



Centro de Ciências Exatas, Naturais e Tecnologias

Programa de Pós-Graduação em Biotecnologia

JÉSSICA RENATA NOGUEIRA

**TRIAGEM DE CHALCONAS COM ATIVIDADE INIBITÓRIA
DA PROTEÍNA CARREADORA CRM1**

**RIBEIRÃO PRETO, SP
2021**

JÉSSICA RENATA NOGUEIRA

**TRIAGEM DE CHALCONAS COM ATIVIDADE INIBITÓRIA
DA PROTEÍNA CARREADORA CRM1**

Tese apresentada como requisito parcial para obtenção do título de Doutora pelo programa de Pós-Graduação em Biotecnologia do Centro de Ciências Exatas, Naturais e Tecnologias da Universidade de Ribeirão Preto.

Área de concentração: Biotecnologia Aplicada à Saúde

Orientador: Prof. Dr. Mozart A. Marins
Coorientador: Dr. Ícaro Puntinhon Caruso

RIBEIRÃO PRETO, SP
2021

Ficha catalográfica preparada pelo Centro de Processamento
Técnico da Biblioteca Central da UNAERP
- Universidade de Ribeirão Preto -

V778t Nogueira, Jéssica Renata, 1991-
Triagem de chalconas com atividade inibitória da proteína
carreadora CRM1 / Jéssica Renata Nogueira. - - Ribeirão Preto,
2021.
137 f.: il. color.

Orientador: Prof. Dr. Mozart de Azevedo Marins.

Tese (doutorado) - Universidade de Ribeirão Preto, UNAERP,
Biotecnologia. Ribeirão Preto, 2021.

1. CRM1. 2. Simulações computacionais. 3. Molecular
Docking. 4. Dinâmica molecular. I. Título.

CDD 660.6

JÉSSICA RENATA NOGUEIRA
TRIAGEM DE CHALCONAS COM ATIVIDADE INIBITÓRIA DA PROTEÍNA
CARREADORA CRM1

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Biotecnologia da Universidade de Ribeirão Preto, para obtenção do título de Doutora em Biotecnologia.

Área de Concentração: Biotecnologia Aplicada à Saúde

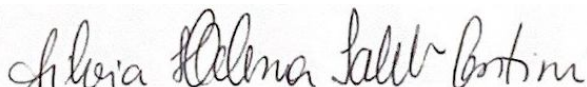
Data da defesa: 29 de março de 2021

Resultado: Aprovada

BANCA EXAMINADORA



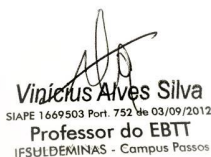
Prof. Dr. Mozart de Azevedo Marins
Universidade de Ribeirão Preto – UNAERP



Profa. Dra. Sílvia Helena Taleb Contini
Universidade de Ribeirão Preto – UNAERP

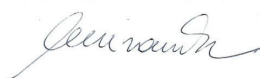


Prof. Dr. Rene de Oliveira Beleboni
Universidade de Ribeirão Preto – UNAERP



Vinicius Alves Silva
SIAPE 1.669.503 Port. 752 de 03/09/2012
Professor do EBT
IFSULDEMINAS - Campus Passos

Prof. Dr. Vinicius Alves Silva
Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais – IFSul de Minas



Prof. Dr. Carlos Eduardo Saraiva Miranda
Universidade de Ribeirão Preto – UNAERP

RIBEIRÃO PRETO

2021

AGRADECIMENTOS

Agradeço a mim pela coragem e a essa força que rege minha vida em todos os momentos, me fazendo sempre seguir em frente (se preferir, chame-a de Deus). A meu orientador, Mozart, e a meu coorientador, Ícaro, por toda ajuda e disponibilidade. Admiro muito o trabalho de ambos e aprendi muito nesse espaço de tempo. Agradeço meus pais e minha irmã pelo apoio de sempre. Ao meu Pretinho, meu cãopanheiro e às minhas amigas Lully e Tati, por representarem tanto para mim. A minha querida amiga Isa, por compartilhar parte do conhecimento comigo e me ajudar no desenvolvimento deste trabalho. Agradeço à CAPES pelo apoio financeiro, que foi fundamental para o desenvolvimento deste trabalho. Agradeço também à Fátima, pelos ensinamentos e por me permitir realizar os experimentos que resultaram neste trabalho na UNESP – São José do Rio Preto. Meu carinho e agradecimento aos professores da UNAERP, por me ensinarem tanto de uma área que pouco entendia. Valeu muito essa caminhada!

RESUMO

O uso de métodos computacionais para verificar a interação entre moléculas possibilita reduzir o tempo de produção de novos fármacos, uma vez que esse processo pode ser utilizado para excluir compostos não-reagentes. Nesse sentido, neste trabalho são analisadas, computacionalmente, interações entre a proteína CRM1 e chalconas. A CRM1 é a principal proteína de exportação presente em mamíferos e atua no transporte de macromoléculas da membrana nuclear para o citoplasma. Nesse estudo, em uma fase inicial, foram realizadas simulações de *docking* molecular para a triagem de chalconas que apresentam maior afinidade de ligação com a proteína de interesse. Em seguida, os ligantes selecionados na fase anterior foram utilizados em simulações de dinâmica computacional, em que foram verificados valores de RMSD, número de pontes de hidrogênio e número de contatos. Como a atividade da proteína CRM1 possui relação direta com a descrita pelas proteínas p53 e MDM2, ambas foram empregadas em simulações computacionais de *docking* e de dinâmica molecular com chalconas. Na triagem computacional realizada, foram selecionadas oito chalconas de interesse específico, que são: G14, G15, G19, D14, D15, D19, A6 e A20, que apresentaram melhores afinidades de ligação para a CRM1. Em experimentos de dinâmica molecular, foi verificado que, dentre as oito chalconas selecionadas, algumas delas são comparáveis ou apresentaram melhores resultados de RMSD, número de contatos e de pontes de hidrogênio, quando confrontadas ao ligante de referência. Nesse sentido, através de simulações computacionais, foram selecionadas moléculas de interesse, cuja atividade pode auxiliar na regulação da proteína CRM1, p53 e MDM2.

Palavras-chave: CRM1; Simulações Computacionais; Molecular *Docking*; Dinâmica Molecular.

LISTA DE ILUSTRAÇÕES

Figura 1: Número estimado de novos casos e de mortes no Brasil e no mundo decorrentes do câncer.	17
Figura 2: Isômeros de chalconas.	21
Figura 3: O tratamento com transchalcona diminuiu os níveis de CRM1 e, conseqüentemente, reduziu a quantidade de proteína p53 ligada à CRM1, reduzindo a exportação de p53 para o citoplasma.	24
Figura 4: Estratégias consideradas em programas de <i>docking</i> molecular.	30
Figura 5: Estrutura da proteína CRM1 utilizada nas simulações computacionais de <i>docking</i> e de dinâmica molecular. A molécula está na cor cinza e está representada em <i>cartoon</i>	37
Figura 6: Estrutura da proteína p53 utilizada nas simulações computacionais de <i>docking</i> e dinâmica molecular. Sua estrutura é apresentada em cinza e é representada em <i>cartoon</i>	38
Figura 7: Estrutura da proteína MDM2 utilizada nas simulações computacionais de <i>docking</i> e de dinâmica molecular. Sua estrutura é apresentada em cinza e é representada em <i>cartoon</i>	39
Figura 8: Estrutura genérica de chalconas.	42
Figura 9: Resumo da metodologia utilizada nos experimentos com a proteína CRM1.	45
Figura 10: Estrutura da proteína CRM1, em cor cinza, contendo o ligante Leptomicina B, com dados da conformação real, obtida em cristal, na cor azul, e experimental obtida no programa AutoDock Vina, na cor magenta. A proteína CRM1 é representada em <i>cartoon</i> e o ligante de referência, Leptomicina B, em <i>stick</i>	47
Figura 11: Resíduos formados entre CRM1 e ligantes: (a) T10; (b) T33; (c) T37; (d) T38; (e) G2; (f) G3; (g) G4; (h) G5; (i) G6; (j) G14; (k) G15; (l) G18; (m) G19; (n) G20; (o) G21; (p) G23; (q) G24; (r) G25; (s) D1; (t) D2; (u) D3; (v) D4; (w) D6; (x) D14; (y) D15; (z) D18; (aa) D19; (ab) D20; (ac) D21; (ad) D23; (ae) D24; (af) D25; (ag) A2; (ah) A5; (ai) A6; (aj) A8; (ak) A9; (al) A10; (am) A11; (an) A12; (ao) A14; (ap) A15; (aq) A17; (ar) A19; (as) A20 (continua).	50
Figura 12: Resultados de simulação de dinâmica computacional para RMSD utilizando a proteína CRM1 (PDB ID 4HAT) com os ligantes: (a) G14; (b) G15; (c) G19; (d) D14; (e) D15; (f) D19; (g) A6; e (h) A20.	58

Figura 13: Resultados de simulação de dinâmica computacional para pontes de hidrogênio utilizando a proteína CRM1 (PDB ID 4HAT) com os ligantes: (a) G14; (b) G15; (c) G19; (d) D14; (e) D15; (f) D19; (g) A6; (h) A20.....	59
Figura 14: Resultados de simulação de dinâmica computacional para número de contatos da proteína CRM1 com os ligantes: (a) G14; (b) G15; (c) G19; (d) D14; (e) D15; (f) D19; (g) A6; (h) A20.....	60
Figura 15: Resultados de <i>redocking</i> para o ligante K76. A proteína é representada em <i>cartoon</i> , na cor cinza, e o ligante K76, em todas as conformações, em <i>stick</i> . Em magenta, encontra-se a conformação do K76 obtida em <i>redocking</i> no programa AutoDock Vina, em azul no programa AutoDock, em verde o resultado no SwissDock e em amarelo o ligante em sua estrutura no cristal.	62
Figura 16: Resíduos de aminoácidos identificados para os ligantes: (a) G14; (b) G15; (c) G19; (d) D14; (e) D15; (f) D19; (g) A6; (h) A20 e (i) K76 utilizando o arquivo de PDB ID 4HB2. Nessa figura, estão escritos em verde os nomes de resíduos da proteína e em preto estão escritos os nomes de resíduos dos ligantes.	66
Figura 17: Resultados de simulação de dinâmica computacional para RMSD utilizando a proteína CRM1 (PDB ID 4HB2) com os ligantes: (a) G14; (b) G15; (c) G19; (d) D14; (e) D15; (f) D19; (g) A6; e (h) A20.	68
Figura 18: Resultados de simulação de dinâmica computacional para o número de pontes de hidrogênio da proteína CRM1 (PDB id.: 4HB2) com os ligantes: (a) G14; (b) G15; (c) G19; (d) D14; (e) D15; (f) D19; (g) A6; (h) A20.....	69
Figura 19: Resultados de simulação de dinâmica computacional para número de contatos utilizando a proteína CRM1 (PDB ID 4HB2) com os ligantes: (a) G14; (b) G15; (c) G19; (d) D14; (e) D15; (f) D19; (g) A6; e (h) A20.	70
Figura 20: resultados de dinâmica computacional utilizando o ligante de referência K76 nas simulações de: (a) RMSD; (b) número de pontes de hidrogênio; e (c) número de contatos.	71
Figura 21: Nas Figuras (a) e (b), a estrutura da proteína CRM1 utilizada nas simulações computacionais é apresentada em cinza e sua representação está em <i>cartoon</i> . Nessas figuras, as chalconas são apresentadas utilizando o modelo <i>stick</i> e a cada uma delas é associada uma cor, sendo: i) G14: vermelho; ii) G15: verde; iii) G19: azul; iv) D14: amarelo; v) D15: magenta; vi) D19: ciano; vii) A6: laranja; e viii) A20: marrom. Em (a), apresenta-se a estrutura para a CRM1	

de identificador PDB ID 4HAT. Em (b) apresenta-se a proteína de identificador PDB ID 4HB2, que contém também o ligante de referência K76, que é apresentado em roxo.....	73
Figura 22: Resíduos formados entre CRM1 e ligantes: (a) D14; (b) D15; e (c) Rita.	79
Figura 23: Resultados de simulação de dinâmica computacional para RMSD da proteína p53 com os ligantes: (a) D14; (b) D15; (c) Rita.	80
Figura 24: Resultados de simulação de dinâmica computacional para número de contatos da proteína p53 com os ligantes: (a) D14; (b) D15; (c) Rita.	81
Figura 25: Resultados de simulação de dinâmica molecular para número de pontes de hidrogênio da proteína p53 com os ligantes: (a) D14; (b) D15; (c) Rita.	82
Figura 26: Estrutura da MDM2 utilizada para simulações computacionais e os subdomínios gerados com os ligantes D14, D15 e Nutlin-3a. Nas figuras a.1, b.1 e c.1, a proteína está representada em cartoon e a região da MDM2 que contém o ligante avaliado encontra-se circulada. Para a representação das chalconas D14 e D15 e do ligante de referência (Nutlin-3a), foi utilizado o modelo de <i>stick</i> e <i>ball</i> . Nas figuras a.2, a.3, b.2, b.3, c.2 e c.3 podem ser observados os resíduos encontrados entre MDM2 e: (a.2, a.3) D14; (b.2, b.3) D15; (c.2, c.3) Nutlin-3a.	86
Figura 27: Resultados de simulação de dinâmica computacional para RMSD da proteína MDM2 com os ligantes: (a) D14; (b) D15; (c) Nutlin-3a.....	87
Figura 28: Número de contatos da proteína MDM2 com os ligantes: (a) D14; (b) D15; (c) Nutlin-3a.....	88
Figura 29: Resultados de simulação de dinâmica computacional para número de pontes de hidrogênio da proteína MDM2 com os ligantes: (a) D14; (b) D15; (c) Nutlin-3a.	88
Figura 30: Amostras antes de induzir (AI) e depois de induzir (DI) com os clones C1, C2 e C3, com temperatura: 20°C; concentração de IPTG: 1.0 milimol/Litro; tempo de indução: 18h. Linhagem: Ril(DE3).	91
Figura 31: Antes de induzir (AI) e depois de induzir (DI) com os clones C1 e C2 e temperatura 20°C; IPTG: 0.4 milimol/Litro; e tempo de indução: 18h. Linhagem: Ril(DE3). Sobrenadante (Sob) e pellet (Pel) são apresentados.	91
Figura 32: Antes de induzir (AI) e depois de induzir (DI) com os clones 3 (C3, RilDE3) e 1 (C2, BL21DE3) utilizando temperatura: 20°C; concentração de IPTG: 0.4 milimol/Litro; e tempo de indução: 18h.....	92

Figura 33: Antes de induzir (AI) e depois de induzir (DI) com os clones 2 (C2) e 3 (C3) com temperatura: 20°C; concentração de IPTG: 0.4 milimol/Litro; e tempo de indução: 18h. Linhagem: BL21(DE3). Continuação da Figura 32.	92
Figura 34: Antes de induzir (AI) e depois de induzir (DI) com o clone 1 (C1) com temperatura: 37°C; concentração de IPTG: 0.5 milimol/Litro e 1.0 milimol/Litro; e tempo de indução: 4h. Linhagem: Ril(DE3).	93
Figura 35: Antes de induzir (AI) e depois de induzir (DI) com o clone 1 (C1) com temperatura: 37°C; concentração de IPTG: 0.5 milimol/Litro e 1.0 milimol/Litro; e tempo de indução: 4h. Linhagem: BL21(DE3).	93
Figura 36: Três programas de maior citação na base de dados Scopus para as simulações de <i>docking</i> molecular entre proteína e ligante.	136

LISTA DE TABELAS

Tabela 1: Estimativa de casos de câncer e taxas de mortalidade para Brasil e mundo nos anos de 2018 a 2040.....	15
Tabela 2: Condições avaliadas nos ensaios.	44
Tabela 3: Resultados de afinidade de ligação, em kcal/mol, utilizando os programas AutoDock Vina, AutoDock 4, SwissDock e DockThor (continua).	47
Tabela 4: Combinação dos dez melhores resultados de afinidade de ligação obtidos em experimentos de <i>docking</i> molecular para a proteína CRM1 (PDB ID 4HAT) e chalconas.	49
Tabela 5: Resíduos (RES) identificados nos complexos contendo a proteína CRM1 e chalconas e suas respectivas porcentagens (P) de ocorrência.	55
Tabela 6: Resíduos calculados no programa LigPlot+ para os complexos contendo CRM1 e chalconas.	56
Tabela 7: Resultados de afinidade de ligação, em kcal/mol, utilizando os programas AutoDock Vina, AutoDock 4 e SwissDock (continua).....	62
Tabela 8: Combinação dos dez melhores resultados de afinidade de ligação obtidos em experimentos de <i>docking</i> molecular para a proteína CRM1 (PDB ID 4HB2) e chalconas.	64
Tabela 9: Resíduos formados entre a proteína CRM1 (PDB ID 4HB2) e ligantes G14, G15, G19, D14, D15, D19, A6 e A20. Na última coluna, são apresentadas as porcentagens de ocorrência desses resíduos.	65
Tabela 10: Resultados de RMSD, número de contatos e pontes de hidrogênio obtidos para os arquivos 4HAT e 4HB2.	75
Tabela 11: Resultados de afinidade de ligação, em kcal/mol, utilizando os programas AutoDock Vina, e AutoDock 4.	76
Tabela 12: Resultados de RMSD, Número de Contatos e Número de Pontes de Hidrogênio para a proteína p53.	82
Tabela 13: Resultados de docking molecular para a proteína MDM2 e chalconas.	83
Tabela 14: Resultados de RMSD, Número de Contatos e Número de Pontes de Hidrogênio para a proteína MDM2.	89
Tabela 15: Programas para simulações computacionais de <i>docking</i> molecular em complexos contendo proteína-ligante. Nessa tabela, <i>indisp.</i> indicam programas cuja licença encontra-se	

indiponível, NR refere-se ao número de artigos encontrados na revisão sistemática entre os anos de 2015 e 2019 e NH indica que não há citações nesse período..... 112

LISTA DE ABREVIATURAS E SIGLAS

ATF3	Fator de ativação de transcrição 3
Ala549	Alanina 549
Ala552	Alanina 552
Arg110	Arginina 110
Asn131	Asparagina 131
Asn268	Asparagina 268
Asn571	Asparagina 571
CADD	Planejamento de fármacos auxiliado por computadores
CRM1	<i>Chromosomal Maintenance</i> 1
Cys539	Cisteína 539
DANT	Doenças e agravos não transmissíveis
DHFR	<i>Dihydrofolate reductase</i>
DM	Dinâmica molecular
Gln18	Glutamina 18
Gln24	Glutamina 24
Gly58	Glutamina 58
Gln59	Glutamina 59
Gln72	Glutamina 72
Gln144	Glutamina 144
Glu582	Glutamina 582
Glu586	Glutamina 586
Gly112	Glicina 112
GPL	<i>General Public License</i>
GSTP1	<i>Glutathione S-transferase</i> P1
HPV	Papilomavírus Humano
His73	Histidina 73
His96	Histidina 96

ICCC	Classificação Internacional do Câncer Infantil
Ile61	Isoleucina 61
Ile99	Isoleucina 99
Ile532	Isoleucina 532
Ile555	Isoleucina 555
LBDD	Planejamento de fármacos baseado na estrutura do ligante
Leu54	Leucina 54
Leu111	Leucina 111
Leu536	Leucina 536
Leu580	Leucina 580
LGA	<i>Lamarckian Genetic Algorithm</i>
LNCC	Laboratório Nacional de Computação Científica
Lys51	Lisina 51
Lys548	Lisina 548
Lys579	Lisina 579
MDR1	Multidrug-resistant
Met62	Metionina 62
Met556	Metionina 556
mRNA	RNA mensageiro
P ³ M	<i>Particle-Particle-Particle Mesh Ewald</i>
P-GP	<i>Membrane pump P-glycoprotein</i>
Phe113	Fenilalanina 113
Phe572	Fenilalanina 572
Phe583	Fenilalanina 583
PME	<i>Particle Mesh Ewald</i>
OMS	Organização Mundial da Saúde
OS	Osteossarcoma
RMSD	Tolerância do <i>root-square-mean deviation</i>

rRNA	Ácido ribonucleico ribossomal
SBDD	Planejamento de fármacos baseado na estrutura do receptor
Ser269	Serina 269
snRNA	Pequenos RNA nucleares
SPME	<i>Smooth Particle Mesh Ewald</i>
Thr102	Treonina 102
Thr575	Treonina 575
TIP	<i>Transferable Intermolecular Potentials</i>
Top2A	Topo 2A
Trp146	Triptofano 146
Tyr67	Tirosina 67
Tyr100	Tirosina 100
Tyr126	Tirosina 126
Val75	Valina 75
Val93	Valina 93
Val551	Valina 551
Val559	Valina 559
Val576	Valina 576
Val591	Valina 591

SUMÁRIO

1. INTRODUÇÃO	15
1.1. Objetivos.....	34
1.2. Organização do Trabalho.....	35
2. MATERIAIS E MÉTODOS	36
2.1. Materiais	36
2.2. Metodologia.....	36
3. RESULTADOS	46
3.1. Resultados Computacionais	46
3.1.1. Resultados para a CRM1 (PDB ID 4HAT)	46
3.1.2. Resultados para a CRM1 (PDB ID 4HB2).....	61
3.1.3. Considerações sobre os Resultados Computacionais Obtidos para a proteína CRM1 ..	72
3.1.4. Resultados Computacionais para a Proteína p53	76
3.1.4.1. Resultados de <i>Docking</i> Molecular	76
3.1.4.2. Resultados de Dinâmica Molecular	79
3.1.5. Resultados Computacionais para a Proteína MDM2	83
3.1.5.1. Resultados de <i>Docking</i> Molecular	83
3.1.5.2. Resultados de Dinâmica Molecular	85
3.1.6. Considerações sobre os Resultados Computacionais para as Proteínas CRM1, p53 e MDM2	89
3.2. Resultados de Expressão e Purificação do Domínio da Proteína CRM1.....	90
4. CONCLUSÕES.....	95
REFERÊNCIAS BIBLIOGRÁFICAS.....	96
APÊNDICE.....	111
Programas de <i>Docking</i> Molecular entre Proteína-Ligante	111

1. INTRODUÇÃO

Doenças e agravos não transmissíveis (DANT) representam a principal causa de adoecimento e de óbito da população mundial. No ano de 2018, cerca de 36 milhões dos óbitos, o que representa aproximadamente 63% dos casos, ocorreram em consequência das DANT, com destaque para as doenças cardiovasculares, que representam 48% dos casos, e o câncer, que representa 21% delas (INCA, 2018).

Neste trabalho é dado enfoque ao câncer, que é a segunda maior causa de morte no mundo e foi associado a aproximadamente 9,6 milhões dos casos de óbito em 2018 (OMS, 2018). A palavra câncer é utilizada para referenciar um conjunto com mais de 100 doenças diferentes, que possuem em comum o crescimento desordenado de células (INCA, 2019). Segundo a Organização Mundial da Saúde (OMS), mundialmente, cerca de 1 dos 6 casos de morte são decorrentes dessa enfermidade e aproximadamente 70% desses casos ocorrem em países de renda baixa e média (OMS, 2018). Em estudo realizado por da Cruz e Irfi (2014), foi verificado que o efeito renda está relacionado diretamente à facilidade de indivíduos receberem um diagnóstico da doença.

Predições realizadas pela OMS, contendo os números estimados para novos casos e o número de mortes decorrentes do câncer, discriminando valores por gênero (feminino e masculino) e apresentando dados do Brasil e do mundo, podem ser observados na Tabela 1.

Tabela 1: Estimativa de casos de câncer e taxas de mortalidade para Brasil e mundo nos anos de 2018 a 2040.

Brasil						
	Número de Novos Casos			Número de Mortes		
Ano	Masculino	Feminino	Total	Masculino	Feminino	Total
2018	278.607	280.764	559.371	129.134	114.454	243.588
2025	349.007	340.018	689.025	164.418	143.376	307.794
2030	404.577	384.623	789.200	193.571	166.396	359.967
2040	523.281	474.775	998.056	259.626	216.646	476.272

Mundo						
	Número de Novos Casos			Número de Mortes		
Ano	Masculino	Feminino	Total	Masculino	Feminino	Total
2018	9.456.418	8.622.539	18.078.957	5.385.640	4.169.387	9.555.027
2025	11.380.492	10.091.504	21.471.996	6.510.260	4.957.708	11.467.968
2030	12.892.837	11.219.281	24.112.118	7.427.060	5.599.961	13.027.021
2040	16.014.743	13.518.251	29.532.994	9.397.256	6.991.203	16.388.459

Fonte: Compilação da autora¹.

¹ Tabela gerada com dados extraídos de GLOBOCAN (2018).

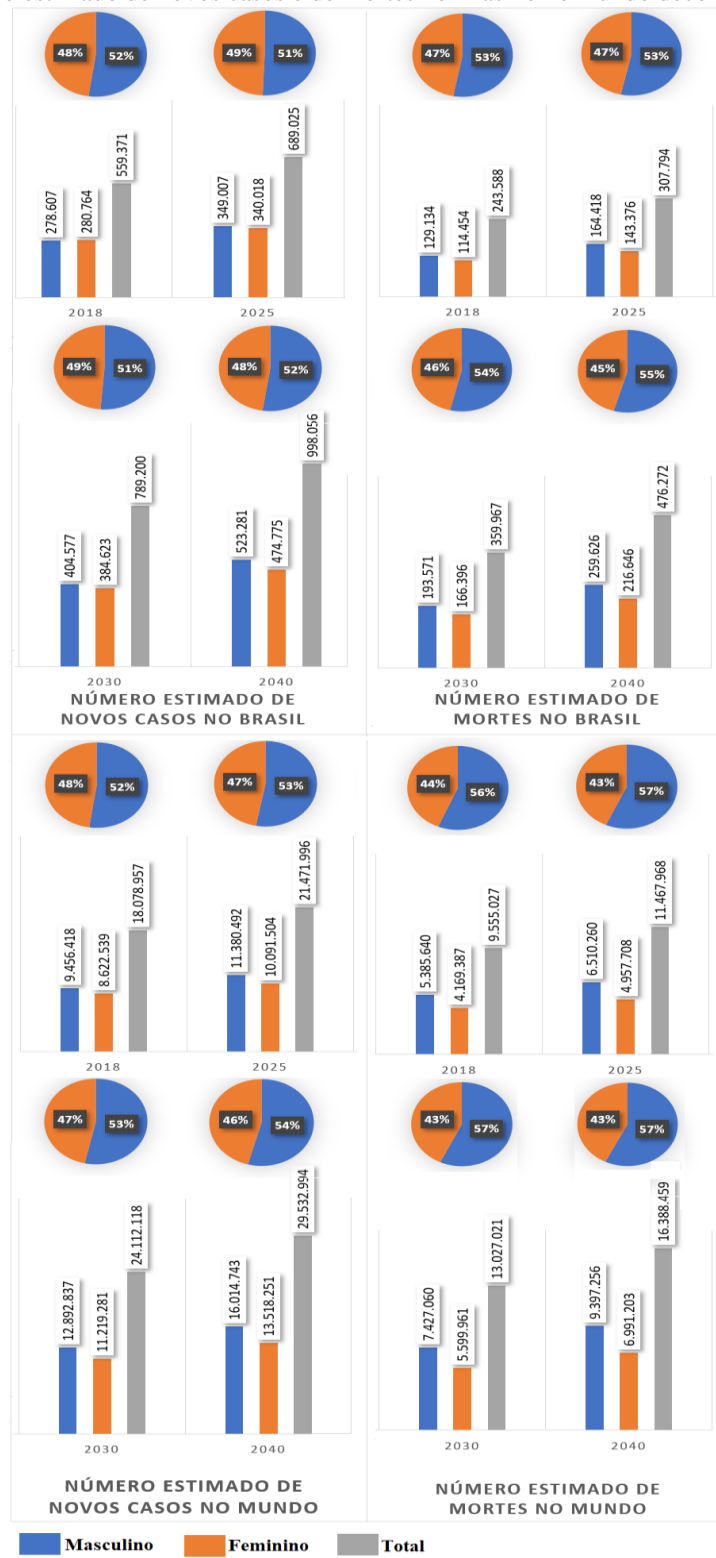
Note na Tabela 1 que, para o Brasil, é estimado que no biênio de 2018-2019, uma ocorrência de cerca de 600 mil novos casos de câncer, para cada ano. Em termos de incidência por região geográfica, as regiões Sul e Sudeste concentram 70% desses novos casos, sendo que a Região Sudeste concentra cerca de 50% da incidência total para o biênio considerado (INCA, 2018). Ainda considerando a Tabela 1, observa-se um aumento do número estimado de casos de incidência e de mortes decorrentes do câncer em homens, uma vez que no ano de 2018, o índice de novos casos é de 49,8% e aumenta para o ano de 2040 para 52,4%. O mesmo pode ser verificado para o número estimado de mortes, que no ano de 2018 é de 53,0% para homens e aumenta para 54,5% em 2040.

Resultados semelhantes aos descritos anteriormente podem ser verificados quando os dados são analisados mundialmente. Nota-se que o número estimado de novos casos entre os homens, no ano de 2018, sobe de 52,3% para 54,2% no ano de 2040. Ainda, em relação a taxa de mortalidade estimada, há um aumento de 56,3%, no ano de 2018, para 57,3%, no ano de 2040. Em todos os cenários, a taxa de mortalidade estimada para os homens sempre foi maior do que a aproximada para as mulheres. Detalhes referentes à Tabela 1 podem ser observados na Figura 1. Considerando os anos discriminados, o número de novos casos no Brasil representa cerca de 3% dos casos do mundo e o número de mortes representa aproximadamente 2% dos casos.

O câncer pode ser desencadeado em diferentes tecidos ou órgãos e de acordo com a célula ou tecido que abriga o ponto de origem, define-se o seu tipo. Ao originarem em tecidos epiteliais, como pele ou mucosas, são denominados carcinomas e se originarem em tecidos conjuntivos, como ossos, músculos ou cartilagens, são chamados sarcomas (INCA, 2019). Ainda, em um mesmo tecido ou órgão, podem ser originados diferentes tipos de câncer. Na pele, por exemplo, podem ser gerados diversos tipos dessa doença, pois essa é formada por mais de um tipo de célula.

As causas que desencadeiam o câncer estão relacionadas a fatores externos e internos. Exemplos de fatores externos referem-se à exposição excessiva à radiação ou a agentes químicos, tabagismo, consumo excessivo de álcool e sedentarismo. Como exemplos de fatores internos, tem-se o sistema imunológico comprometido e a predisposição genética (ONCOGUIA, 2017). Segundo a OMS, atualmente, de 30 a 50% dos cânceres podem ser evitados, pela adoção de estratégias de prevenção, como a redução à exposição à ultravioleta e a vacinação contra Papilomavírus Humano (HPV) e vírus da hepatite B, bem como pela diminuição dos fatores de risco, como a redução do uso do álcool e do tabaco (OMS, 2018).

Figura 1: Número estimado de novos casos e de mortes no Brasil e no mundo decorrentes do câncer.



Fonte: Compilação da autora².

² Imagem gerada a partir de dados extraídos de GLOBOCAN (2018).

Uma característica importante de vários tipos de tumores cancerígenos é a metástase, processo no qual as células tumorais migram do seu ponto de origem para outras partes do corpo e formam os tumores. Esses tumores são denominados secundários ou também metástases. Um tumor é classificado como cancerígeno primário quando ocorre um acúmulo de células cancerígenas em um determinado ponto do organismo, sem que ele parta para outras partes do corpo. Praticamente todos os tipos de câncer podem formar tumores metastáticos, mas os locais em que a ocorrência de metástase é mais comum são ossos, fígado e pulmão (ONCOGUIA, 2013). Embora a maioria dos cânceres tenha a capacidade de se espalhar para diferentes partes do corpo, eles tendem a disseminar com mais frequência para certas partes do organismo que para outras (ONCOGUIA, 2013). Um exemplo de tumor que forma metástase é o osteossarcoma, que é um câncer originário nos ossos.

Os tumores ósseos malignos não estão entre as neoplasias infanto-juvenis de maior incidência, representando apenas 5% dos cânceres que ocorrem na infância e na adolescência, sendo o osteossarcoma (OS) o mais comum (OTTAVIANI; JAFFE, 2009; SAVAGE; MIRABELLO, 2011). O OS é um tumor primário que afeta principalmente jovens com idade entre 10 e 16 anos (ABARRATEGI *et al.*, 2016), por desenvolver preferencialmente em tecidos ósseos em fase de rápido crescimento, o que justifica o pico de incidência na adolescência durante o surto de crescimento puberal (MIRABELLO; TROISI; SAVAGE, 2009). Porém, esse tipo de câncer apresenta distribuição de idade bimodal, com 30% dos casos afetando pacientes com mais de 40 anos (ABARRATEGI *et al.*, 2016).

Em estudo realizado, considerando-se os tipos de câncer definidos na Classificação Internacional do Câncer Infantil (ICCC), foi verificado pelos autores que o OS representa 2% do total dos tipos possíveis de câncer desencadeados em pacientes americanos com até 14 anos (SIEGEL; MILLER; JEMAL, 2019). Embora considerado raro, com índices abaixo de 1% quando comparado à outras neoplasias (MIRABELLO; TROISI; SAVAGE, 2009), o osteossarcoma chama a atenção pelo caráter extremamente agressivo e metastático, com elevada capacidade de invadir e destruir tecidos moles (SAVAGE; MIRABELLO, 2011).

Embora a taxa de sobrevivência dos pacientes submetidos à cirurgia e à terapia multidrogas seja estimada em cerca de 75%, a taxa de sobrevivência para os doentes metastáticos cai para 19% a 30% (KAGER *et al.*, 2003). Para pessoas com até 14 anos, em um período de 5 anos, o osteossarcoma é o segundo tipo de câncer com maior taxa de mortalidade (SIEGEL; MILLER;

JEMAL, 2019). Os autores apontam que em 30,4% dos casos, os pacientes vêm a óbito e esse valor é inferior apenas aos casos de leucemia mielóide aguda, cujo índice de mortalidade é de 33,6%.

A origem mais provável do osteossarcoma é a partir de células mesenquimais, sendo caracterizado por formação irregular de osso imaturo, produção de matriz osteoide e células fusiformes estromais malignas (MUTSAERS; WALKLEY, 2014). Aproximadamente 42% dos casos ocorrem nas regiões acima do fêmur que estão próximas ao joelho, 19% dos casos ocorrem na região da tíbia e 10% na região úmero (OTTAVIANI; JAFFE, 2009).

Considerando o potencial de crescimento e disseminação, existem três diferentes tipos de osteossarcoma, que são os sarcomas de grau baixo, intermediário e alto (SOCIETY, s.d.), sendo o OS de alto grau é o mais prevalente. Ele apresenta elevada taxa de crescimento e desdiferenciação e seus subtipos mais comuns são o osteoblástico, condroblástico e fibroblástico. O osteosarcoma de alto grau espalha-se pelo organismo predominantemente por meio da via hematogênica, atingindo, preferencialmente, os pulmões, que é o sítio mais comum para metástase (85-90% dos casos) derivado desse sarcoma (KUIJER; HOGENDOORN; CLETON-JANSEN, 2013).

O protocolo usual de tratamento do osteossarcoma inclui quimioterapia pré-operatória (neoadjuvante), cirurgia e quimioterapia pós-operatória (adjuvante) (ABARRATEGI *et al.*, 2016). As drogas antitumorais utilizadas comumente para o tratamento incluem Doxorubicina, Cisplatina, Metotrexato, Epirubicina, Ifosfamida, Ciclofosfamida, Etoposídeo, Gemcitabina e Topotecano. O regime de tratamento mais utilizado é a associação de Doxorubicina, Cisplatina e Metotrexato, tanto na quimioterapia neoadjuvante como adjuvante. Apesar da eficácia da associação dessas drogas, elas causam sérios efeitos adversos. A Doxorubicina induz cardiotoxicidade (COLOMBO *et al.*, 2013) e a Cisplatina provoca nefrotoxicidade, neurotoxicidade e pode levar ao surgimento de neoplasia secundárias (SASTRY; KELLIE, 2005). O Metotrexato, além de problemas renais, hepáticos e neurológicos, também pode provocar mielossupressão (HOWARD *et al.*, 2016).

Outro problema com o protocolo de quimioterapia é o desenvolvimento de resistência a drogas, bastante comum em doentes com recorrência da doença devido às metástases, fato que é um obstáculo no sucesso do tratamento (YE *et al.*, 2016). A resistência à cisplatina tem sido relacionada ao aumento nos mecanismos de reparo do DNA e superexpressão da enzima GSTP1 (*glutathione S-transferase P1*), que metaboliza e inativa drogas anticâncer. Os casos de resistência à doxorubicina estão relacionados a alta expressão da GSTP1 e do gene MDR1 (*multidrug-*

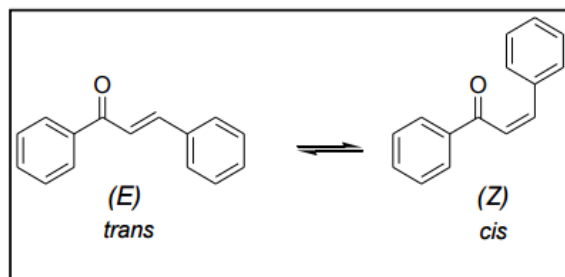
resistant) que codifica a proteína P-GP (*membrane pump P-glycoprotein*) responsável por exportar substâncias para fora da célula (HE; NI; HUANG, 2014). Por sua vez, os mecanismos de resistência ao metotrexato envolvem redução na absorção da droga e o aumento dos níveis da enzima DHFR (*dihydrofolate reductase*) através da amplificação do seu gene correspondente (JAFFE, 2014). Desta forma, considerando que as terapias atuais para tratamento do osteossarcoma são ainda insatisfatórias, especialmente em casos de metástases, e que muitas vezes ocorre uma resistência às drogas comumente usadas, é imperativo continuar os esforços para desenvolver novas terapias e drogas para o tratamento desta doença.

O uso de polifenóis é apontado como efetivo para a prevenção de doenças, tais como as doenças cardíacas, diabetes e doenças neurodegenerativas e apresentam potencial aplicação no tratamento de casos de câncer (SCALBERT *et al.*, 2005). Polifenóis representam um grupo de substâncias que são abundantes, podendo também ser encontrados em bebidas, em que exemplos incluem chás e vinho tinto.

Um tipo específico de polifenol são os flavonoides, que são polifenóis encontrados, por exemplo, em frutas cítricas, legumes e verduras. As chalconas, flavonoides de cadeia aberta e precursoras comuns da biossíntese de outros flavonoides e isoflavonoides, são compostos que vêm recebendo atenção especial por apresentarem diversas atividades biológicas (LEON-GONZALEZ *et al.*, 2015; MAHAPATRA; BHARTI; ASATI, 2015).

No esqueleto básico das chalconas (Figura 2), existem duas formas isoméricas, *cis* e *trans*. A forma *trans*chalcona é considerada termodinamicamente mais estável e exerce atividade citotóxica contra diferentes linhagens de células tumorais através de mecanismos variados que incluem a interrupção do ciclo celular, inibição da angiogênese, inibição da polimerização da tubulina, indução de apoptose e a inibição de quinases reguladoras do ciclo celular (MAHAPATRA; BHARTI; ASATI, 2015). Além disso, nosso grupo de pesquisa identificou a *trans*chalcona como um inibidor transcricional da topo 2A (Top2A), um alvo clássico de muitas drogas antitumorais (SILVA *et al.*, 2013).

Figura 2: Isômeros de chalconas.



Fonte: Grupo de Pesquisa da Autora.

Moléculas como a transchalcona possuem potenciais redox relativamente baixos e com uma maior probabilidade de sofrerem reações de transferência de elétrons por conta da presença de ligações duplas conjugadas em sua estrutura. Esta propriedade da transchalcona, juntamente com sua lipofilicidade, aumentam sua atividade biológica e ainda é notada uma característica interessante: o baixo efeito genotóxico, associado ao grupo amino de ácidos nucleicos, que a maioria das drogas antitumorais em uso possuem (DIMMOCK *et al.*, 1999). Dessa forma, a transchalcona é considerada um composto apropriado para a síntese de vários derivados heterocíclicos, incluindo pirimidina, imidazol, tiazina, tiazol, quinolona, oxazol, benzopirano e imidazolina. Esses híbridos de chalcona com anel heterocíclico exibem várias propriedades biológicas, que os destacam como agentes antibacterianos, antifúngicos, antioxidantes, antitumorais e anti-HIV (ZHOU E XING, 2015).

De interesse específico deste projeto, a atividade inibidora da transchalcona e de seus derivados sobre várias linhagens celulares tumorais já foi descrita em diversos trabalhos da literatura, incluindo trabalhos de nosso grupo de pesquisa (BORTOLOTTI *et al.*, 2016; SIKANDER *et al.*, 2011; SILVA *et al.*, 2013; SILVA *et al.*, 2016). Também, nosso grupo de pesquisa demonstrou que a atividade anticancerígena da transchalcona é mediada por aumento da estabilidade da proteína supressora de tumores p53 e redução nos níveis do fator de transcrição Sp1, parceiro comum de p53 (SILVA *et al.*, 2016).

A proteína p53 atua como um fator de transcrição o qual é ativado em resposta a diferentes tipos de estresse celular como danos ao DNA, ativação de oncogênese e hipóxia. No genoma humano, ela foi um dos primeiros genes supressores tumorais descobertos. Exemplos de suas funções incluem o controle do ciclo celular, apoptose, manutenção da integridade do genoma, metabolismo, fertilidade, reprogramação celular e autofagia (GU; ZHU, 2012).

Gu e Zhu (2012) verificaram que a p53 coordena uma rede de regulação que supervisiona e responde a uma variedade de sinais de estresse. Segundo esses autores, exemplos dessas regulações correspondem a danos ao DNA, ativação oncogênica, estresse ribossômico, perda da adesão célula-célula e outras. Gu e Zhu (2012) afirmam que mutações da p53 ou perturbações em sua coordenação, em menor escala, podem representar um desequilíbrio fisiológico normal. Caso esse desarranjo torne-se crítico, ocorrem os casos de câncer.

Embora diferentes mecanismos subjacentes possíveis para a regulação do p53 tenham sido propostos por décadas, nenhum deles é conclusivo. Além disso, a grande maioria das drogas anticancerígenas utilizadas clinicamente possuem mecanismos de ação baseados na indução de danos ao DNA com respostas dependentes da atividade da p53. As mutações que ocorrem no gene p53 são as alterações genéticas mais frequentes em tumores malignos. Segundo Cavalcanti Júnior, Klumb e Maia (2002), essa mutação na p53 ocorre em cerca de 60% das neoplasias. Além disso, em sua forma ativa, a p53 tem um tempo de vida médio curto, que está em torno de 6 minutos, o que dificulta sua detecção (CAVALCANTI JÚNIOR; KLUMB; MAIA, 2002). Os autores afirmam que na forma mutada ou inativa, a p53 tende a acumular-se no núcleo celular, permitindo sua detecção através de métodos imunológicos, como a imunocitoquímica, imunohistoquímica, *western blot* ou citometria de fluxo.

A regulação de p53 ocorre principalmente a nível de estabilidade da proteína dentro de uma rede de regulação, em que a p53 sofre poliubiquitinação mediada pelo seu principal inibidor, a proteína MDM2 e subsequentemente degradação pelo proteassoma (SHADFAN; LOPEZ-PAJARES; YUAN, 2012). Outros estudos têm mostrado que a estabilidade e a ativação de p53 pode ser reforçada pelo bloqueio da sua ubiquitinação mediada por MDM2 através da interação de MDM2 com proteínas, tais como DNAJB1, um membro da família Hsp40, ou a ligação direta de p53 com o fator de ativação de transcrição 3 (ATF3), um fator de transcrição da família ATF/CREB (QI *et al.*, 2014; YAN *et al.*, 2005).

Ao formar complexos com a p53, a ATF3 também pode atuar como um cofator de transcrição na ativação de genes alvos da p53, como no caso do gene DR5 (*Death receptor 5*) (TAKETANI *et al.*, 2012). Além disso, a ATF3 é capaz de ligar-se a formas mutantes da p53 (R175H e R273H) e de suprimir as funções oncogênicas, reduzindo assim a migração e a invasão celular de células que carregam essas versões mutantes da p53 (WEI *et al.*, 2014). A proteína Hsp40, por sua vez, também tem demonstrado capacidade de se ligar a formas mutantes da p53,

especialmente naquelas carregando a mutação R175H, e provocar o restabelecimento da conformação selvagem da p53 (HIRAKI *et al.*, 2015).

Um elemento importante na regulação de p53, mediada por MDM2, é o seu transporte para o citoplasma. Em relação a p53, esse transporte é realizado pela proteína CRM1 (do inglês *Chromosomal Maintenance 1*), também conhecida como *Exportin 1* ou XPO1. A proteína CRM1 facilita o transporte de grandes macromoléculas, incluindo o rRNA (ácido ribonucleico ribossomal), snRNA (pequenos RNA nucleares), alguns mRNA (RNA mensageiro) e de proteínas, através da membrana nuclear para o citoplasma (NGUYEN; HOLLOWAY; ALTURA, 2012). Em sua forma livre, a proteína CRM1 volta para o núcleo para que realize a próxima rodada de exportação (MONECKE *et al.*, 2013).

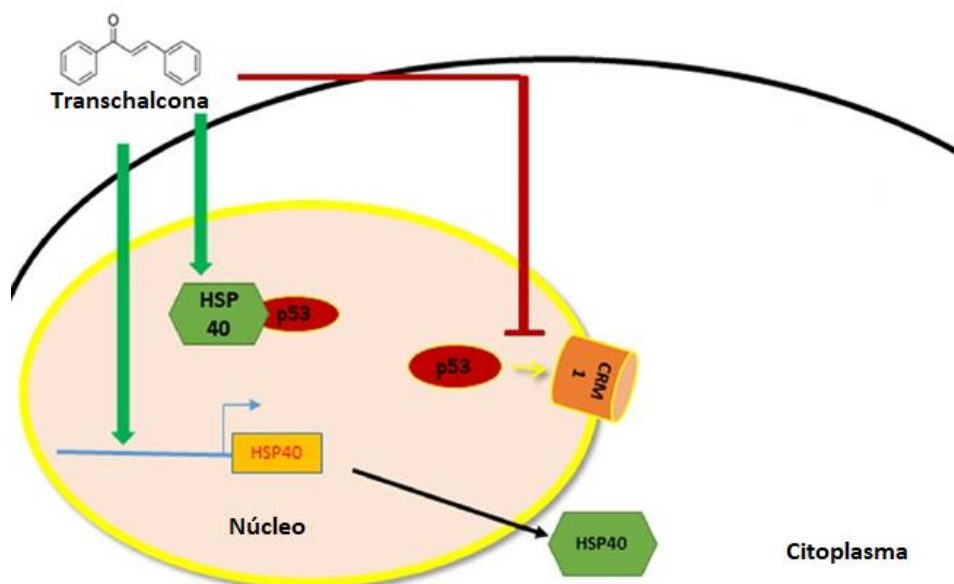
Segundo Niu *et al.* (2013), as proteínas de carga podem se ligar à CRM1 por meio de sinais de sequência específicos, que, em geral, são chamados de sinais de exportação nuclear. Um sinal de exportação nuclear é um peptídeo alvo pequeno, que contém quatro resíduos hidrofóbicos e que marcam a proteína para que ele seja transportado do núcleo celular até o citoplasma da célula. Segundo Niu *et al.* (2013), fatores de transcrição e proteínas supressoras de tumores são transportadas pelo CRM1 do núcleo para o compartimento citoplasmático. As proteínas de carga da CRM1 incluem numerosas proteínas de supressão tumoral e proteínas reguladoras do crescimento, tais como a p53, FoxO3a e o inibidor NF- κ B, I κ B (LAPALOMBELLA *et al.*, 2012). Foi relatado que um desbalanço no nível citosólico dessas proteínas tem sido observado nas células cancerígenas. Segundo Lapalombella *et al.*, (2012), mais de 14 proteínas de supressão tumoral e proteínas reguladoras do crescimento foram identificadas como sendo exportadas pela CRM1 e muitas delas coexistem em diferentes tipos de câncer.

A CRM1 é um alvo potencial para o tratamento de diversas doenças, sendo uma proteína alvo no tratamento do câncer e de inflamações. Segundo Lapalombella *et al.*, (2012), índices elevados de CRM1 estão relacionados a um prognóstico ruim da doença e de resistência à terapia anticancerígena. Estuda-se também a relação da CRM1 e outras doenças. Em Mahboobi, Javanpour, Mofrad (2015), por exemplo, estuda-se a relação entre a CRM1 e o HIV. Além disso, existem estudos em que compostos são testados na tentativa de inibir a atividade dessa proteína, uma vez que seu desequilíbrio, sendo uma expressão elevada ou disfunção, têm sido relatadas em vários casos de tumores sólidos e hematológicos (LAPALOMBELLA *et al.*, 2012). Em Niu *et al.* (2013), por exemplo, utiliza-se a curcumina para inibir a atividade da proteína CRM1.

Micromoléculas de origem natural, como a leptomicina B, foram identificadas como inibidores de CRM1. Porém, o desenvolvimento de drogas anticancerígenas utilizando essas micromoléculas foi comprometido pela toxicidade sistêmica significativa que apresentaram em estudos com animais e humanos (LAPALOMBELLA *et al.*, 2012). Dessa maneira, a CRM1 tem sido considerado um alvo promissor para o desenvolvimento de novas drogas anticancerígenas (NIU *et al.*, 2013).

Experimentos realizados pelo grupo de pesquisa do Prof. Marins, da Unidade de Biotecnologia da UNAERP, indicaram que a transchalcona pode se ligar à CRM1 e inibir sua atividade. Em Silva *et al.* (2018), foi verificado que o tratamento com transchalcona diminuiu os níveis de CRM1 e, conseqüentemente, reduziu a quantidade de proteína p53 ligada à CRM1 (Figura 3). Nesse sentido, a p53 fica por mais tempo no núcleo, não se deslocando para o citoplasma, o que favorece a atividade anticancerígena. Cumulativamente, a transchalcona não apenas degradou a expressão CRM1, mas também interferiu na interação entre p53 e CRM1. Porém, os mecanismos decorrentes dessa inibição não foram elucidados, sendo este um dos objetivos do presente estudo. Considerando a atividade anticancerígena da transchalcona com probabilidade de baixa toxicidade, a atividade inibidora sobre CRM1 pode auxiliar no desenvolvimento de novos inibidores de CRM1 baseados em chalconas.

Figura 3: O tratamento com transchalcona diminuiu os níveis de CRM1 e, conseqüentemente, reduziu a quantidade de proteína p53 ligada à CRM1, reduzindo a exportação de p53 para o citoplasma.



Fonte: Silva *et al.*, 2018.

Simulações computacionais podem ser empregadas para verificar a interação entre dois ou mais compostos. De interesse específico para este trabalho, esses experimentos podem ser realizados entre proteína e ligantes, abordagem essa que será adotada nesse estudo para verificar a ligação entre a proteína CRM1 e chalconas. A utilização da análise computacional para verificar a interação entre proteínas e ligantes é uma abordagem viável para o desenvolvimento de novos fármacos, frente ao tempo e ao custo para o desenvolvimento de uma nova droga que são altos. Levando-se em consideração que, em média, gasta-se de 10 a 15 anos e de 500 a 800 milhões de dólares para introduzir um novo medicamento no mercado (BASAK, 2012), o uso de ferramentas computacionais permite que esse tempo seja reduzido.

Na produção de um fármaco, nas fases iniciais de planejamento do mesmo, são identificadas moléculas de baixa potência e afinidade e que devem ser otimizadas em relação a uma série de propriedades, como potência, afinidade, absorção e metabolismo (GUIDO, 2010). Esses compostos, após a otimização, são selecionados e tornam-se candidatos para o desenvolvimento de um novo fármaco. Nesse sentido, uma triagem computacional de compostos de interesse pode ser realizada como uma fase inicial do processo de seleção de compostos. Como exemplo, as técnicas *in silico* foram utilizadas para o desenvolvimento do inibidor da integrase do HIV, Raltegravir, do anticoagulante Tiroban e do antiviral Zanamivir (OPENZICA, 2016).

A criação de fármacos empregando ferramentas computacionais é denominada planejamento de fármacos auxiliado por computadores (do inglês *Computer Assisted Drug Design*, CADD). Essa área é dividida em duas subáreas principais, que são: i) planejamento de fármacos baseado na estrutura do receptor (do inglês *structure-based drug design*, SBDD); e ii) planejamento de fármacos baseado na estrutura do ligante (do inglês *ligand-based drug design*, LBDD).

O planejamento de fármacos baseado na estrutura do receptor (SBDD) é um método em que se tem conhecimento das estruturas dos alvos moleculares, que são obtidas, por exemplo, através de cristalografia por raio-X, NMR e microscopia eletrônica. Essas técnicas permitem modelar computacionalmente as estruturas de biomoléculas tridimensionais em sua forma de associação, identificando resíduos formados entre diferentes moléculas, as forças de interação e as energias envolvidas. As representações computacionais das estruturas biológicas são realizadas o

mais próximo do que é encontrado na natureza e as biomoléculas são, geralmente, proteínas ou RNAs.

Para o planejamento de fármacos baseado na estrutura do ligante (LBDD), não é necessário o conhecimento do receptor-alvo. Nesse método, são investigadas as propriedades e características de ligantes conhecidos para uma molécula-alvo, para que se estabeleçam relações entre as propriedades físico-químicas e atividades antibióticas. Essas informações podem ser utilizadas na otimização de drogas já conhecidas ou para o desenvolvimento de novos fármacos, com atividade otimizada (WENBO; MACKERELL Jr., 2017). O SBDD e o LBDD são métodos úteis na identificação, seleção e otimização de moléculas bioativas candidatas. Ainda, por conta de sua complementaridade, quando possível, essas técnicas podem ser empregadas em conjunto.

O enfoque deste trabalho é utilizar as simulações computacionais para experimentos de *docking* molecular, técnica em que é empregada a estrutura do receptor, o que a caracteriza como SBDD. Estudos relacionados ao *docking molecular* iniciaram na década de 1980, com o trabalho de Kuntz *et al.*, (1982), que descreveram a técnica e a empregaram na investigação da interação entre heme-mioglobina e da ligação de hormônios análogos da tireóide à prealbumina. Os primeiros métodos automáticos de *docking* molecular surgiram na década de 1990, com os programas AutoDock, GOLD, HammerHead e FlexX, que são detalhados no apêndice 1 deste trabalho.

Desde a proposta desses programas, essa técnica possui aplicações em áreas como a química medicinal e a biologia estrutural. Exemplos de trabalhos em que programas de *docking* molecular são utilizados incluem Bora, Suzuki e Funar-Timofei (2019), em que os métodos são empregados na formulação de inseticidas e de Ahmed e Alkali (2019), em que são descritos inibidores para a enzima caspase-3, cuja atividade fica desregulada em casos de câncer. Ainda, simulações de *docking* molecular podem ser utilizadas na filtragem de grandes bibliotecas de compostos, na otimização dos mesmos e na concepção de novos fármacos.

O estudo da interação entre ligante e proteína é, em geral, o mais realizado ao empregar o *docking* molecular. Porém, outras moléculas podem ser empregadas nessas simulações, o que permite, por exemplo, que simulações entre ligante e proteína, proteína e proteína e proteína e DNA sejam realizadas. Para essas simulações, são disponibilizados bancos de dados contendo informações estruturais de moléculas de interesse. Exemplos para bancos de dados são: i) de

enzima: Brenda³; e ii) de proteína: RCSB *Protein Data Bank*⁴ e UniProt⁵ (*Universal Protein Resource*).

De posse das estruturas de interesse, ao realizar simulações de *docking* molecular, é possível prever computacionalmente a orientação preferencial que uma molécula apresenta em relação a uma segunda quando elas estão ligadas, a afinidade de ligação entre ambas e a estrutura tridimensional do complexo gerado, o que possibilita identificar o local de interação entre essas moléculas.

Em programas de *docking* molecular, é possível prever diferentes conformações formadas entre moléculas, que são denominados modos de ligação, e a cada complexo gerado associa-se uma afinidade de ligação, que é expressa em quilocaloria/mol (kcal/mol). Nos programas de *docking* molecular, para verificar diferentes conformações formadas e selecionar as que apresentam melhor afinidade de ligação, são utilizadas duas fases principais, que são: i) a fase de busca, em que é determinado o ponto de interação entre as moléculas selecionadas; e ii) o cálculo da afinidade de ligação entre essas moléculas, na posição determinada na fase i.

Na fase i, em que ocorre a busca, é identificada uma posição adequada, com energia baixa no complexo, em que é interessante incluir, quando possível, modos de ligação determinados experimentalmente. Caso a termodinâmica do sistema, considerando tanto os efeitos entálpicos quanto os entrópicos, seja modelada de maneira correta pela função de energia, a solução gerada pelo programa de *docking* será a mesma que a obtida experimentalmente (DE PAULA; HORTA, 2014). Porém, devido à complexidade desse procedimento, em que o espaço de busca na macromolécula pode ser grande, para otimizar o tempo de resposta para o problema, são empregados algoritmos que adotam abordagens distintas, que são: sistemática, determinística ou estocástica. Essa classificação considera a flexibilidade das moléculas utilizadas e o espaço de busca selecionado. Na abordagem sistemática, há uma convergência para a melhor conformação, que é o ponto de energia mínima. Ao empregá-la, são explorados todos os graus de liberdade do ligante e a busca ocorre de maneira exaustiva, em toda a região do receptor. Porém, o custo computacional dessa abordagem é alto, uma vez que está relacionado com os graus de liberdades do sistema, bem como ao tamanho do receptor utilizado. Desse modo, para ligantes e receptores maiores, com muitos graus de liberdade, essa abordagem é proibitiva. Na abordagem

³ Banco de Dados Brenda: disponível em <http://www.brenda-enzymes.info/>

⁴ Banco de Dados RCSB Protein Data Bank: disponível em <https://www.rcsb.org/>

⁵ Banco de Dados Uniprot: disponível em <http://www.uniprot.org/>

determinística, são considerados métodos de minimização de energia e de dinâmica molecular. Segundo de Paula e Horta (2014), são utilizadas ferramentas como o gradiente, em que é verificada a direção de menor crescimento de energia no sistema e a movimentação do sistema ao longo do tempo. Ao utilizar essa abordagem, em geral, ocorre a convergência para um mínimo de energia do sistema, que não é, obrigatoriamente, o mínimo global (que possui o melhor resultado), pois há a possibilidade de mínimos locais. Dessa forma, a solução final para o sistema ao utilizar a abordagem determinística está diretamente relacionada às suas condições iniciais. Na abordagem estocástica, são utilizadas heurísticas, que são métodos computacionais em que são fornecidas respostas aproximadas para um problema de solução computacional demorada. Em geral, seu uso retorna resultados aproximados em intervalo de tempo finito, cujo critério de parada do programa é definido pelo usuário do programa. Exemplos de heurísticas empregadas em programas de *docking* molecular são:

- Algoritmos Genéticos: são utilizados os conceitos genéticos de evolução, cruzamento e mutação para a proposta de algoritmos que empregam essa abordagem. Através dessas operações, evoluções (ou seja, melhorias) na solução (denominada população) atual são alcançadas através de operações de cruzamentos e mutações, que são realizados em populações cuja afinidade já foi avaliada.
- Busca Tabu: dado um ponto inicial, são realizadas buscas locais (pequenas alterações conformacionais) e é utilizada uma lista para impedir que certas posições sejam avaliadas, uma vez que elas já foram verificadas previamente.
- Recozimento Simulado (*Simulated Annealing*): nessa classe de algoritmos, são empregados os conceitos da física relacionados à temperatura e à estrutura molecular dos sólidos, para determinar uma nova conformação para o sistema.
- Otimização de Colônia de Formigas: essa classe de algoritmos baseia-se no comportamento real das formigas, que buscam o menor caminho entre o formigueiro e a fonte de comida. Durante esse processo, as formigas comunicam entre si através da deposição de feromônio para marcar os caminhos de maior interesse. Nesse caso, é provável que as melhores soluções apresentem maior deposição de feromônio.

No algoritmo de busca, é possível realizar ou não uma alteração na conformação do receptor e do ligante. Geralmente, o receptor é mantido rígido e alterações no ligante são realizadas (chamado de *docking* semi-rígido). Exemplos de programas de *docking* que empregam o *docking*

semi-rígido são o FlexX e o Dock. Em outros programas, ligante e receptor são rígidos (*docking* rígido) ou ambos são flexíveis (*docking* flexível). Ainda, existem programas que utilizam a abordagem *ensemble*, em que a afinidade de ligação para múltiplas conformações da molécula-alvo são avaliadas.

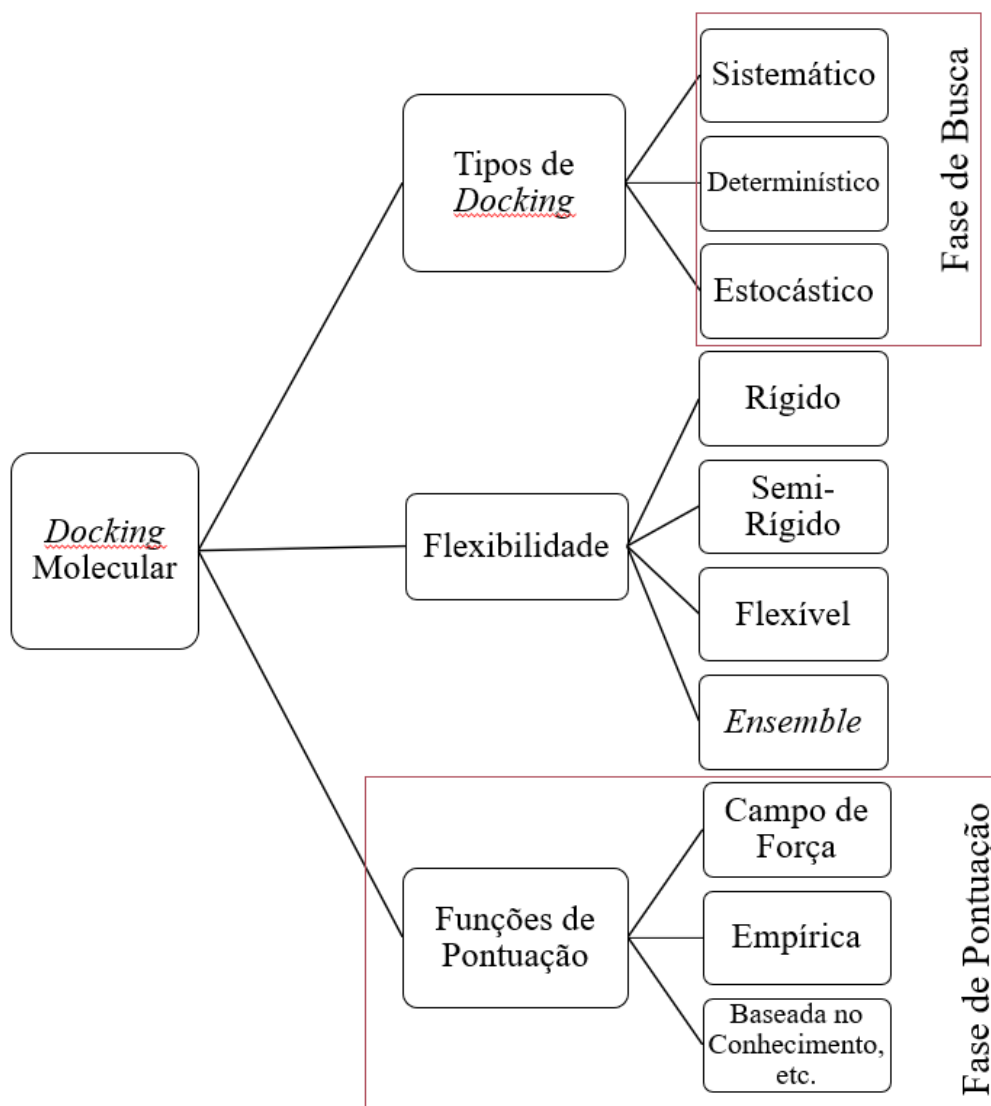
Em uma segunda fase, após a seleção de um novo sítio de interesse, é aplicada uma função de avaliação do sistema formado. Conforme descrito por de Paula e Horta (2014), para esse cálculo, são utilizadas funções de campo de força clássicas, que são resultantes do somatório de diferentes fontes de energia. Segundo os mesmos autores, para realizar essa avaliação, deve-se considerar a interação entre o receptor-ligante e a energia interna do ligante, considerando os termos ligados e não ligados, de modo que são utilizadas funções de aproximação para calcular a afinidade de interação. Nessa fase, são aplicadas funções de pontuação como:

- Funções de Campo de Força, em que as afinidades são estimadas pela soma da força de van der Waals intermolecular e interações eletrostáticas entre todos os átomos das moléculas contidas no complexo formado.
- Funções empíricas, em que são consideradas interações como: hidrofóbico-hidrofóbico, hidrofóbico-hidrofílico, número de pontes de hidrogênio e número de ligações que podem ser rotacionadas e estão imobilizadas no complexo formado.
- Funções baseadas no conhecimento, em que são empregadas informações já disponíveis em bases de dados. Nessa categoria, são consideradas estatísticas de contatos intermoleculares de grandes bancos de dados tridimensionais para definir a função de pontuação.

De acordo com de Paula e Horta (2014), não existe um modelo de função que possa ser aplicado a qualquer tipo de caso e que expresse, de maneira eficaz, todos os tipos de ligação. Porém, ao realizar a simulação computacional da interação de uma proteína com um receptor, devem-se tratar as diversas forças envolvidas, em que exemplos de forças incluem as interações eletrostáticas, eletrodinâmicas e os graus de flexibilidade das moléculas.

Conforme descrito anteriormente, diferentes estratégias são adotadas em programas de *docking* molecular. Fatores como tipo de *docking*, flexibilidade do ligante e funções de pontuação são considerados para avaliar a afinidade de ligação presente no sistema. Nesse sentido, essas estratégias são resumidas na Figura 4.

Figura 4: Estratégias consideradas em programas de *docking* molecular.



Fonte: Autora.

Os programas de *docking* molecular diferem no algoritmo de busca, nas alterações conformacionais do ligante e do receptor e na função de avaliação, o que tem relação direta com o método implementado nesse programa. Na literatura, são disponibilizados diversos programas para simulações de *docking* molecular. Desse modo, no apêndice 1 deste trabalho, encontra-se uma revisão sistemática contendo 122 programas de *docking* molecular para ligações do tipo proteína-ligante e suas descrições.

De interesse para esse trabalho, devido a sua utilização, os programas AutoDock 4, AutoDock Vina, SwissDock e DockThor são detalhados a seguir. Dentre os programas de *docking*

molecular, o AutoDock (MORRIS *et al.*, 1998) é o mais citado na literatura, considerando a revisão sistemática disponível no apêndice deste trabalho. Esse programa é disponível sob a *General Public License*, sendo que seu uso comercial e não comercial e redistribuição são permitidos. No AutoDock 4, é utilizada a heurística Algoritmos Genéticos como algoritmo de busca.

O programa de *docking* molecular AutoDock Vina (TROTT; OLSON, 2010) é uma variante para o AutoDock e foi proposto de forma a permitir um número maior que torções nos ligantes. Ainda, em geral, fornece respostas de maneira mais eficiente, ou seja, em menor espaço de tempo que o AutoDock. O AutoDock Vina é utilizado nas predições de conformações do tipo proteína-ligante e permite a programação *multithread*, para acelerar o tempo de execução das simulações. No AutoDock Vina, são utilizados Algoritmos Genéticos e é adotado um passo de otimização, em que são selecionados ligantes com resultados de afinidade descrevendo mínimos locais, que são resultados de interesse, que são armazenados e refinados, a fim de melhorar as predições nas saídas geradas pelo programa. Os programas AutoDock Vina e AutoDock são disponibilizados pelo Laboratório *Molecular Graphics*, no *The Scripps Research Institute*⁶.

O programa de *docking* molecular para proteína e pequenos ligantes SwissDock (GROSDIDIER; ZOETE; MICHIELIN, 2011) é *webservice, open source* (Apache *webserver*), gratuito e segundo seus autores, foi proposto para permitir uma interface de uso agradável ao usuário. Após o envio do receptor e do ligante nesse programa, para a realização das simulações de *docking*, as execuções são realizadas em um servidor remoto de alto desempenho. Em seguida, após a finalização das simulações de *docking*, o complexo resultante e a energia do sistema são enviados por e-mail.

O programa de *docking* molecular DockThor (ALMEIDA, 2011; DE MAGALHÃES, 2006) é *webservice*, gratuito, utiliza Algoritmos Genéticos em sua implementação e foi desenvolvido pelo grupo de Modelagem Molecular em Sistemas Biológicos do Laboratório Nacional de Computação Científica (LNCC). O portal contendo esse programa foi disponibilizado em 2013 e, segundo seus criadores, o DockThor é o primeiro programa brasileiro, e possivelmente de todo o hemisfério Sul, para realização de *docking* molecular receptor-ligante. De modo similar ao SwissDock, após o envio do receptor e do ligante no portal do DockThor, ocorre a realização das simulações de *docking* em um servidor remoto de alto desempenho.

⁶ Disponível em <http://autodock.scripps.edu/>

Computacionalmente, também é possível verificar os resíduos de aminoácidos presentes no complexo formado entre proteína e ligante. Um resíduo de aminoácido ocorre quando um aminoácido se liga a outro e nessa reação química é liberada água. Essa ligação se estabelece pela ligação covalente do grupo amina de um aminoácido com o grupo carboxila do outro.

Simulações computacionais de *docking* apresentam, em geral, resultados aproximados ao calcular a afinidade de ligação entre moléculas. Nos programas para essa finalidade, após determinar a conformação final gerada para o complexo, não são consideradas as interações que ocorrem após essa ligação. Considerando os sistemas reais, em que átomos e moléculas podem interagir após essa união inicial, o uso de métodos que considerem alterações na conformação do sistema durante um intervalo de tempo é de interesse. Nesse sentido, estudos de dinâmica molecular (DM) são úteis para investigar se as interações se mantêm após de ligação inicial, considerando também se a ligação se mantém estável ao longo da trajetória e se novas interações ocorrem nesse complexo.

Nos programas de DM, consideram-se os princípios de mecânica clássica e são simulados comportamentos microscópicos descritos pelo complexo formado. Nessas avaliações, também é analisado o comportamento individual dos átomos presentes no sistema. Nesses estudos, fatores como pressão, energia interna, volume, temperatura, entropia, energia livre e outros são considerados.

Em simulações de dinâmica molecular, a partir da estrutura tridimensional do sistema e de uma posição inicial das partículas, calcula-se o potencial de interação entre as partículas, que é denominado campo de força. Segundo Namba, Silva, Silva (2008), esse cálculo é definido através da soma de vários termos de energia, considerando os termos para átomos ligados, que são comprimentos e ângulos de ligação e ângulos diedros, e os termos para átomos não-ligados, obtidos pelas interações de van der Waals e de Coulomb. Exemplos de campos de força disponíveis na literatura incluem CHARMM (BROOKS *et al.*, 1983), GROMOS (GUNSTEREN E BERENDSEN, 1987; GUNSTEREN *et al.*, 1996) e AMBER (WEINER *et al.*, 1984) e a seleção do campo de força deve considerar o sistema estudado, bem como as informações que se tem interesse relacionadas a esse complexo. Sabendo-se que em estudos de DM são verificados intervalos de tempo, a cada passo são calculadas as novas posições e velocidades associadas a cada partícula.

Em simulações de DM, é utilizada uma caixa de simulação que contém as moléculas empregadas nos experimentos, que pode adotar diferentes geometrias. Exemplos para essas geometrias incluem a cúbica, a paralelepípedica e a octaédrica truncada. Segundo Nimba, Silva, Silva (2008), é comprovado que os resultados gerados nas simulações são praticamente independentes da geometria utilizada para as caixas de simulação.

Antes de iniciar as simulações de dinâmica, é necessário eliminar contatos ruins entre os átomos, o que ocorre através do processo de minimização de energia, em que são verificadas as coordenadas que permitem que a energia potencial do sistema seja minimizada. Nesse procedimento, a superfície de potencial é percorrida em direção ao ponto em que a energia decresce, de modo que o sistema é levado a um mínimo de energia local próximo, através do ajuste nas ligações químicas, nos ângulos entre ligações e nos contatos de van der Waals (NIMBA, SILVA, SILVA, 2008). Exemplos de algoritmos de minimização incluem o método *steepest descent* (ARFKEN, 1985; MENZEL, 1960, MORSE, FESHACH, 1953), o método dos gradientes conjugados (HESTENES, STIEFEL, 1952) e o método de Newton-Raphson (FLETCHER, 1980; KREYSZIG, 1993). Após o passo de minimização, o sistema é aquecido de modo gradual até que atinja a temperatura de interesse.

Em simulações de DM, como são utilizados sistemas microscópicos que imitam o comportamento de sistemas macroscópicos, devem ser consideradas as condições de contorno presentes no sistema. Isso ocorre pois o número relativo de partículas na superfície do sistema microscópico é muito maior do que para um sistema macroscópico, o que causa efeitos de superfície indesejáveis (NIMBA, SILVA, SILVA, 2008) e deve ser evitado, para obtenção de resultados mais realísticos para o modelo. Desse modo, através das condições de contorno, os efeitos que ocorrem na superfície são eliminados e as partículas são consideradas como se sempre estivessem internas a esse sistema. Para tratar interações de curto alcance, utiliza-se o procedimento denominado convenção da imagem mínima, em que é considerado um raio de corte esférico para impedir a sobreposição de uma partícula com ela mesma ou para impossibilitar interações entre a mesma partícula (NIMBA, SILVA, SILVA, 2008). Para as interações de longo alcance, métodos como P³M (*Particle-Particle-Particle Mesh Ewald*), PME (*Particle Mesh Ewald*), descrito por Darden, York, Pedersen (1993) e SPME (*Smooth Particle Mesh Ewald*), proposto por Essmann *et al.* (1995), são utilizados para produzir trajetórias estáveis.

Além disso, nas simulações de DM, devem ser considerados os modelos de solvatação utilizados nos sistemas. Exemplos de solvatação incluem TIP (do inglês *Transferable Intermolecular Potentials*, que possui as variações TIP3P, TIP4P e TIP5PSPC) e que foi descrito por Jorgensen *et al.* (1983) e SPC/E, proposto por Berendsen, Grigera, Straatsma (1987).

Devido a aplicabilidade de simulações de dinâmica molecular na área de planejamento de fármacos assistido por computador, essa metodologia computacional também será utilizada neste trabalho, utilizando o programa GROMACS 5.0.1 (PRONK *et al.*, 2013). O desenvolvimento do GROMACS iniciou em 1991 e nesse programa são realizadas simulações contendo moléculas que são flexíveis ao longo de sua trajetória. Além disso, utiliza-se da dinâmica molecular clássica, em que são empregadas equações de movimento de Newton para o cálculo das trajetórias de partículas. A cada partícula no sistema é calculada uma força total, que é resultante de sua interação com outras partículas.

O GROMACS permite realizar, em especial, simulações em proteínas, lipídios e ácidos nucleicos. Nesse programa, é possível que as simulações sejam realizadas de maneira sequencial, utilizando apenas um processador, ou utilizando-se de paralelismo, o que permite que o programa gere respostas mais rapidamente. O GROMACS funciona de maneira nativa em sistemas do tipo Unix (como Linux ou Mac OS X). Porém, é possível adaptá-lo para ser executado no sistema operacional Windows. Nesse caso, segundo seus autores, o uso desse programa é mais pesado e sua execução é lenta. O GROMACS é um programa de código aberto e está sob a GNU *General Public License* (GPL).

No GROMACS, selecionando-se as moléculas de interesse, o programa, ao final de sua execução, produz um arquivo que contém a trajetória em que são descritos os movimentos dos átomos em estudo ao longo do tempo. Através desse arquivo, a partir da trajetória gerada, é possível analisar, por exemplo, se há estabilidade ou não na ligação entre duas ou mais moléculas, o número de contatos e o número de pontes de hidrogênio.

1.1. Objetivos

O objetivo geral deste projeto é a triagem de chalconas com atividade inibidora potencial da proteína CRM1. Trabalhamos com a hipótese de que através dos estudos de interação, identificaremos o sítio de ligação da transchalcona em CRM1. Ainda, tem-se como objetivos

específicos:

- Verificar a atividade de inibição das chalconas para a proteína CRM1 através de simulação computacional de *docking* molecular.
- Avaliar as chalconas selecionadas na triagem do passo anterior através de simulações de dinâmica computacional e correlacionar a estrutura e a atividade desses ligantes.
- Realizar simulações de *docking* e dinâmica computacional com as proteínas MDM2 e p53.

1.2. Organização do Trabalho

Este trabalho possui as seções de Introdução, Materiais e Métodos, Resultados e Conclusão, cujos conteúdos são apresentados nas seções 1, 2, 3 e 4, respectivamente. Na seção 1, são apresentadas as motivações para realização desse estudo e o conteúdo necessário para compreensão deste trabalho. Na seção 2, descreve-se a metodologia aplicada ao realizar os experimentos, bem como os materiais necessários para sua elaboração. Na seção 3, são apresentados os resultados obtidos nas simulações computacionais e nos ensaios laboratoriais. Finalmente, na seção 4, são apresentadas as conclusões obtidas através dos experimentos realizados.

2. MATERIAIS E MÉTODOS

Nas subseções 2.1 e 2.2, são apresentados, respectivamente, os materiais necessários para a realização deste trabalho e a metodologia a ser aplicada para a obtenção dos resultados esperados.

2.1. Materiais

Para realização dessa pesquisa, os seguintes materiais foram utilizados:

- Programas AutoDock Vina (TROTT; OLSON, 2010), AutoDock 4 (MORRIS *et al.*, 1998), SwissDock (GROSDIDIER; ZOETE; MICHIELIN, 2011) e DockThor (ALMEIDA, 2011; DE MAGALHÃES, 2006), utilizados nas simulações de *docking* molecular;
- Programa AutoDock Tools (SANNER, 1999), que contém ferramentas para a manipulação das estruturas das proteínas e ligantes empregadas nas simulações computacionais;
- Programa GROMACS 5.0.1 (PRONK *et al.*, 2013), para realizar simulações de dinâmica computacional e para cálculo de RMSD;
- Programa PyMol (DELANO, 2002), utilizado para visualização e manipulação de compostos;
- Banco de dados RCSB (RCSB, 2020), para obtenção da estrutura das proteínas CRM1, p53 e MDM2.
- Programa Gaussian 09 (FRISCH *et al.*, 2009), utilizado para desenho dos ligantes;
- Programa LigPlot+, v. 2.2 (WALLACE, LASKOWSKI e THORNTON, 1996), para cálculo de resíduos presentes nos complexos;
- Programa VMD, versão 1.9.3, para cálculo de número de pontes de hidrogênio e número de contatos.

2.2. Metodologia

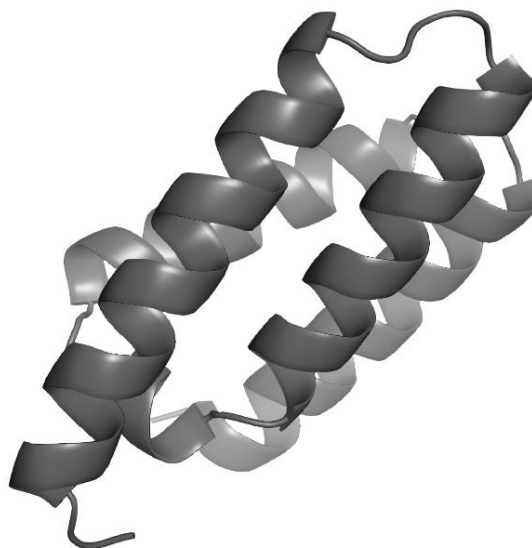
Inicialmente, foi realizada uma revisão da literatura envolvendo aspectos teóricos e práticos necessários para o desenvolvimento do projeto. Dentre os temas estudados, destacam-se: i) levantamento de referencial abordando as proteínas CRM1, p53, MDM2 e chalconas; ii) programas necessários para as simulações de *docking* molecular e dinâmica computacional.

Em seguida, foram realizados os desenhos das chalconas utilizadas nas simulações computacionais no programa Gaussian 09 (FRISCH *et al.*, 2009), através de licença disponibilizada pela Universidade Estadual de São Paulo – UNESP. Os ligantes foram otimizados

com os parâmetros GROUND STATE, semi-empírico e PM6 (STEWART, 2007).

Logo após, foram selecionadas as estruturas computacionais da proteína CRM1 para as simulações, utilizando trabalhos correlatos que as empregam, seguida da verificação dos desenhos que apresentam melhores resoluções. Como exemplo, em Niu *et al.* (2013), utiliza-se a estrutura da CRM1 (Figura 5) com identificador 4HAT (SUN *et al.*, 2013). No presente trabalho, optou-se por utilizar a estrutura da CRM1 de mesmo identificador (resolução: 1.78 Å) e a estrutura de identificador 4HB2 (SUN *et al.*, 2013), cuja resolução é de 1.8 Å. Ambas foram obtidas no banco de dados de proteínas (RSCB PDB), em que os autores empregaram difração de raio-X e utilizaram sistema de expressão *Escherichia coli*. Nessas simulações, foi selecionado apenas o domínio de interesse da CRM1. Foi realizado um *redocking* com os ligantes Leptomicina B, disponível no arquivo de PBD ID 4HAT, e K76, disponível no arquivo de PBD ID 4HB2. Um *redocking* é a comparação da conformação de dados reais de ligação, obtidos na estrutura cristalina, com os dados obtidos em simulações computacionais. A Leptomicina B foi identificada como um inibidor da atividade da CRM1 em Kudo, Matsumori e Taoka (1999). Uma vez que o *redocking* com a Leptomicina B não foi realístico, utilizou-se o ligante, K76 como ligante de referência nas simulações de dinâmica computacional. Um ligante de referência é uma molécula já descrita na literatura cuja atividade é conhecida para a proteína de interesse.

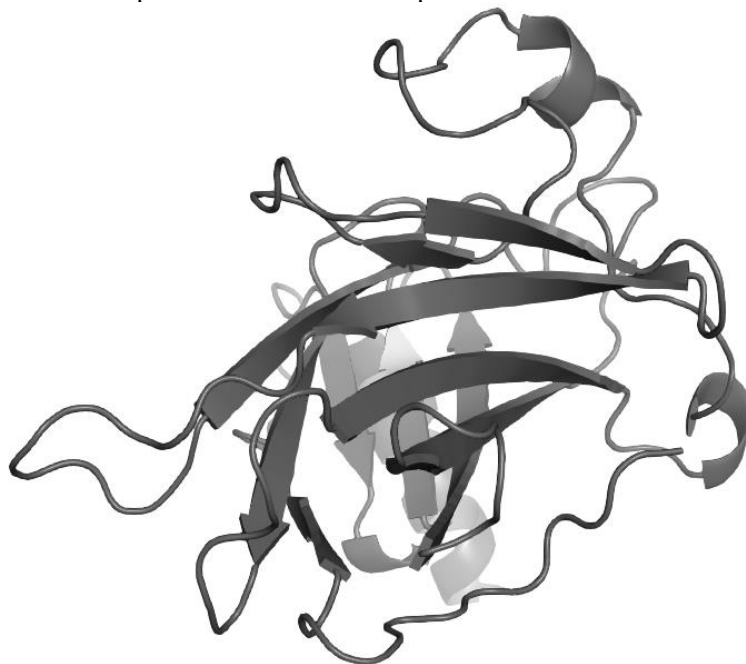
Figura 5: Estrutura da proteína CRM1 utilizada nas simulações computacionais de *docking* e de dinâmica molecular. A molécula está na cor cinza e está representada em *cartoon*.



Fonte: Autora. A representação estrutural foi preparada utilizando o programa PyMOL (DELANO, 2002).

A proteína p53 (Figura 6) também foi utilizada em simulações de *docking* e de dinâmica molecular. Os programas AutoDock 4 e AutoDock Vina, foram utilizados para as simulações de *docking* molecular para as 45 transchalconas avaliadas. As simulações de dinâmica foram realizadas com as chalconas D14 e D15, pois foi verificado elas possuem atividade de inibir a migração de células de osteossarcoma, baseada na atividade da p53 sobre a regulação de genes associados ao processo de transição epitélio mesenquimal (SILVA *et al.*, 2018). Essas chalconas também foram selecionadas pois, em outro trabalho, foi demonstrado que o aumento da atividade de p53 em células de osteossarcoma tratadas com transchalcona, a molécula qual a D14 e D15 são derivadas, é baseado na estabilização de p53 mediada pela chaperona Hsp40 e no bloqueio de sua exportação para o citoplasma via inibição de CRM1 (SILVA *et al.*, 2018). A estrutura da proteína p53 utilizada para simulações computacionais foi retirada do banco de dados de proteínas (RSCB PDB) através do identificador 2PCX e que possui resolução de 1.54 Å. O ligante Rita foi utilizado como referência para a p53. Esse ligante reativa a p53, induz a apoptose tumoral e é descrito como um ativador do tipo selvagem e reativador da função do p53 mutante (WIEGERING *et al.*, 2017) e o desenho de sua estrutura foi realizado no programa Gaussian 09 (FRISCH *et al.*, 2009).

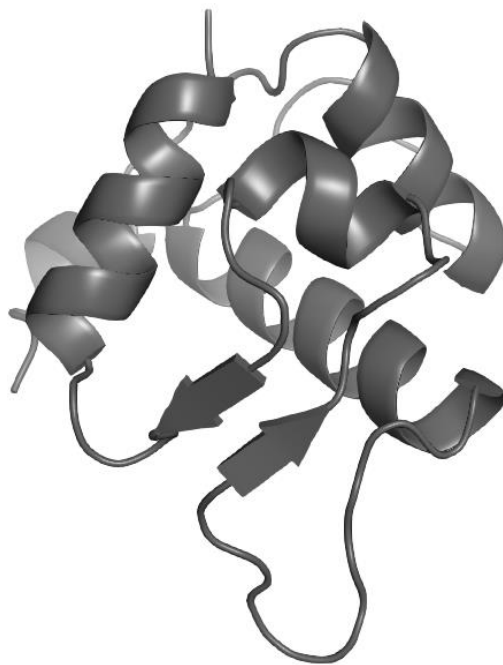
Figura 6: Estrutura da proteína p53 utilizada nas simulações computacionais de *docking* e dinâmica molecular. Sua estrutura é apresentada em cinza e é representada em *cartoon*.



Fonte: Autora. A representação estrutural foi preparada utilizando o programa PyMOL (DELANO, 2002).

Ainda, simulações computacionais de *docking* e de dinâmica molecular, foram realizadas com a proteína MDM2 (Figura 7). As células tumorais da p53 podem ficar inativas através da superexpressão da MDM2, que é o principal regulador negativo da p53. A MDM2 liga-se à proteína supressora da p53 com grande afinidade e modula negativamente sua atividade e estabilidade transcricional (VASSILEV et al., 2004). Desse modo, inibir a interação MDM2-p53 pode estabilizar a p53 e, por esse motivo, simulações computacionais foram realizadas com a MDM2. A estrutura da proteína MDM2 utilizada para simulações computacionais foi retirada do banco de dados de proteínas (RSCB PDB) através do identificador 4HG7 (ANIL et al., 2013). O ligante Nutlin-3a, disponível no PDB ID 4HG7 da proteína MDM2 foi utilizado para *redocking* com a proteína MDM2, pois foi descrito como inibidor para a MDM2 em Kojima *et al.* (2006).

Figura 7: Estrutura da proteína MDM2 utilizada nas simulações computacionais de *docking* e de dinâmica molecular. Sua estrutura é apresentada em cinza e é representada em *cartoon*.



Fonte: Autora. A representação estrutural foi preparada utilizando o programa PyMOL (DELANO, 2002).

O programa AutoDockTools (SANNER, 1999) foi utilizado para preparar as proteínas e as chalconas utilizadas para as simulações de *docking* molecular. Os hidrogênios da proteína e das chalconas ligados a átomos não polares foram removidos. O estado de protonação dos resíduos de histidina da proteína foram definidos de acordo com a programa PROPKA 3.0 (OLSON *et al.*, 2011) do server PDB2PQR versão 2.0.0 (DOLINSKY *et al.*, 2004), considerando pH 7,0. As

estruturas da Leptomicina B e do K76 foram removidas do arquivo da CRM1 (PDB ID 4HAT e PDB ID 4HB2, respectivamente), bem como o composto Nutlin-3a foi removido da estrutura da MDM2. A raiz rígida dos ligantes foi identificada automaticamente, definindo todas as possíveis ligações covalentes rotáveis como ativas para a realização de torções na busca conformacional no sítio de ligação.

2.2.1. Parâmetros Utilizados nos Programas de *Docking* Molecular

Os programas AutoDock 4, AutoDock Vina, SwissDock e DockThor foram utilizados em simulações de *docking* molecular. Para esses programas, foram utilizadas as seguintes configurações:

- **AutoDock 4:** em todas as simulações que o utiliza, foi adotado um espaçamento de 0,375 Å, 100 corridas, população de 250, número máximo de avaliações igual a 25.000.000. Os demais parâmetros utilizados foram padrões do programa. Nesse programa, foram utilizados *Lamarckian Genetic Algorithm* (LGA), tolerância do *root-square-mean deviation* (RMSD) para a análise de *cluster* de 2.0 Å e foi calculada a energia eletrostática interna. Para a CRM1, p53 e MDM2 foram utilizadas, respectivamente, caixas delimitadoras de tamanho 80 x 80 x 80 Å, 100 x 100 x 100 Å e 70 x 70 x 70 Å.
- **AutoDock Vina:** empregou-se uma caixa delimitadora de tamanho 30 x 30 x 30 Å, para CRM1 e p53, e 24 x 26 x 24 Å para a MDM2. Foi utilizada uma exaustão igual a 16 para CRM1 e p53 e 32 para a MDM2. Em todas as simulações, foi adotado um espaçamento de 1 Å.
- **SwissDock:** foi utilizado apenas para as simulações com a proteína CRM1, em que foi empregada uma caixa delimitadora de tamanho 30 x 30 x 30 Å e uma precisão acurada.
- **DockThor:** esse programa foi utilizado apenas para as simulações com a CRM1 de PDB ID 4HAT, pois seus resultados apresentaram grande divergência aos encontrados nos programas AutoDock 4, AutoDock Vina e SwissDock. Nessa simulação, utilizou-se uma caixa delimitadora 8 x 8 x 8 Å, espaçamento de 0.25 Å, 1.000.000 avaliações, população de tamanho 1.000 e 30 corridas.

2.2.2. Parâmetros Utilizados no programa GROMACS para Simulação de Dinâmica Molecular

O GROMACS versão 5.0.1 (PRONK et al., 2013) foi utilizado para as simulações com a CRM1 e MDM2. As simulações de dinâmica molecular (DM) foram realizadas no programa GROMACS utilizando o campo de força CHARMM36, para CRM1 e MDM2, e GROMOS54a7, para p53. O modelo de água utilizado foi TIP3P (JORGENSEN *et al.*, 1983).

Foram utilizados em simulações de dinâmica computacional os ligantes Rita, K76 e Nutlin-3a como referência, para fins de comparação com os resultados obtidos com chalconas. A posição inicial do composto Rita, utilizado como ligante de referência nas simulações de DM com a p53, foi obtida a partir da saída do programa de *docking* molecular do programa AutoDock Vina. A estrutura desse ligante desenhada no programa Gaussian 09 e otimizada com o método semi-empírico PM6 (STEWART, 2007). Para os demais ligantes de referência utilizados, foi utilizada a estrutura do composto obtida em cristal. Para as simulações, foi considerada uma caixa cúbica de 74 Å e solvatação em solução de 100 milimol/Litro de NaCl em água. O estado de protonação dos resíduos foi definido considerando-se os resultados obtidos no PROPKA 3.0 (OLSON *et al.*, 2011) utilizando um pH 7,0. Foram empregadas condições de contorno periódicas e NPT *ensemble*, em todas as simulações. Para o sistema, foi utilizada uma temperatura de 298 K (25 °C) e pressão de 1.0 bar, através do termostato de Berendsej modificado ($\tau_t = 0.1$ ps) e do barostato Parrinello-Rahman ($\tau_p = 2.0$ ps e compressibilidade $4.5 \times 10^{-5} \text{bar}^{-1}$). Um ponto de corte de 14 Å para os potenciais de Lennard-Jones e Coulomb foi utilizado. Para medir as interações eletrostáticas, utilizou-se o algoritmo *particle mesh* Ewald (PME). Ainda, para as simulações de DM, utilizou-se um intervalo de tempo de 2.0 fs e todas as ligações covalentes envolvendo átomos de hidrogênio foram restritas às suas distâncias de equilíbrio. O algoritmo de minimização do gradiente conjugado foi utilizado para relaxar os átomos, a fim de evitar as sobreposições que ocorrem no início do processo de construção da caixa. Para minimizar a energia do sistema, utilizou-se o algoritmo do gradiente conjugado e *steepest descent integrator*, com critério de força máxima igual a $1000 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-1}$.

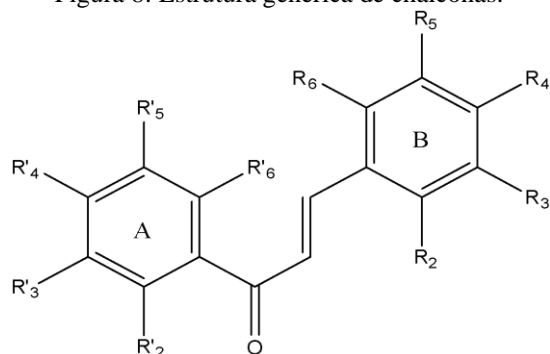
Após a execução dos testes, as trajetórias obtidas foram analisadas de acordo com os seguintes parâmetros: a) RMSD; b) número de pontes de hidrogênio (distância de corte = 3.5 Å e ângulo máximo = 45°); e c) número de contatos (< 0.6 nm). Para obtenção dos valores de RMSD, foi utilizado o programa VMD v. 1.9.3. O RMSD é dado em Ångström (do inglês *Root Mean Square deviation*, ou desvio médio quadrático), em que é calculada a mudança de posição do

ligante de interesse ao longo de sua trajetória. Para verificar o número de pontes de hidrogênio, utilizou-se o programa VMD, versão 1.9.3. Definiu-se como 45° o ângulo máximo formado entre as moléculas e 3,5 a distância entre ligante e receptor. Para cálculo do número de contatos, utilizou-se o comando gmx mindist, que é um programa é disponibilizado pelo GROMACS e é utilizado no cálculo da distância entre o ligante e a proteína. O contato de um átomo com um outro átomo ou grupo de átomos é contado com um contato. Ao final, o gráfico gerado contém o número de contatos que ocorreram entre a proteína e o ligante ao longo da trajetória.

2.2.3. Chalconas

As chalconas (Tabela 1, Figura 8) modeladas computacionalmente neste trabalho, foram cedidas pelo Prof. Luis Octávio Regasini, do grupo de química medicinal do Laboratório de Química Verde e Medicinal, Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista (UNESP - São José do Rio Preto) e pelo Prof. Dr. Valdir Cechinel Filho do Núcleo de Investigações Químico-Farmacêuticas, da Universidade do Vale do Itajaí. As chalconas utilizadas nos testes originais foram obtidas pelo método de condensação de Claisen-Schmidt e caracterizadas quanto à sua atividade citotóxica em linhagens de células tumorais em trabalhos anteriores de nosso grupo de pesquisa (SANTOS *et al.*, 2017; SILVA *et al.*, 2016).

Figura 8: Estrutura genérica de chalconas.



Fonte: Grupo de Pesquisa da Autora.

Chalcona		Anel A					Anel B					MW
		R'2	R'3	R'4	R'5	R'6	R2	R3	R4	R5	R6	
1	T10	H	H	H	H	H	H	OH	H	H	H	224.26
2	T33	H	H	H	H	H	H	OH	OH	H	H	240.26
3	T37	H	H	H	H	H	OH	H	H	F	H	224.25
4	T38	H	H	H	H	H	OH	H	H	CH ₃	H	238.29

5	G2	NH ₂	H	H	H	H	H	H	CH ₃	H	H	237.3
6	G3	NH ₂	H	H	H	H	H	H		H	H	241.26
7	G4	NH ₂	H	H	H	H	H	H	Cl	H	H	257.72
8	G5	NH ₂	H	H	H	H	H	H	Br	H	H	302.17
9	G6	NH ₂	H	H	H	H	H	H	OCH ₃	H	H	253.3
10	G14	NH ₂	H	H	H	H	Phenyl		H	H	H	273.33
11	G15	NH ₂	H	H	H	H	Phenyl		CH ₃	H	H	287.36
12	G18	NH ₂	H	H	H	H	H	H	C-F ₃	H	H	291.27
13	G19	NH ₂	H	H	H	H	H	Phenyl		H	H	273.33
14	G20	NH ₂	H	H	H	H		R ₁ – Thiophene				229.3
15	G21	NH ₂	H	H	H	H		R ₁ – Furan				213.24
16	G23	NH ₂	H	H	H	H	H	H	CN	H	H	236.27
17	G24	NH ₂	H	H	H	H	H	H	NO ₂	H	H	222.27
18	G25	NH ₂	H	H	H	H	H	NO ₂	H	H	H	222.27
19	D1	H	H	NH ₂	H	H	H	H	H	H	H	223.38
20	D2	H	H	NH ₂	H	H	H	H	CH ₃	H	H	237.3
21	D3	H	H	NH ₂	H	H	H	H	F	H	H	241.26
22	D4	H	H	NH ₂	H	H	H	H	Cl	H	H	257.72
23	D6	H	H	NH ₂	H	H	H	H	OCH ₃	H	H	253.3
24	D14	H	H	NH ₂	H	H	Phenyl		H	H	H	273.33
25	D15	H	H	NH ₂	H	H	Phenyl		CH ₃	H	H	287.36
26	D18	H	H	NH ₂	H	H	H	H	CF ₃	H	H	291.27
27	D19	H	H	NH ₂	H	H	H	Phenyl		H	H	273.33
28	D20	H	H	NH ₂	H	H		R ₁ – Thiophene				229.3
29	D21	H	H	NH ₂	H	H		R ₁ – Furan				213.24
30	D23	H	H	NH ₂	H	H	H	H	CN	H	H	236.27
31	D24	H	H	NH ₂	H	H	H	H	NO ₂	H	H	222.27
32	D25	H	H	NH ₂	H	H	H	NO ₂	H	H	H	222.27
33	A2	H	OCH ₃	OH	H	H	H	H	H	H	H	254.28
34	A5	H	OCH ₃	OH	H	H	H	N	H	H	H	255.27
35	A6	H	OCH ₃	OH	H	H	H	Phenyl		H	H	304.35
36	A8	H	OCH ₃	OH	H	H	H	NO ₂	H	H	H	253.28
37	A9	H	OCH ₃	OH	H	H	H	H	CH ₃	H	H	268.31
38	A10	H	OCH ₃	OH	H	H	H	H	F	H	H	272.27
39	A11	H	OCH ₃	OH	H	H	H	OH	H	H	H	270.28
40	A12	H	OCH ₃	OH	H	H	H	H	OH	H	H	270.28
41	A14	H	OCH ₃	OH	H	H	H	H	Cl	H	H	288.73
42	A15	H	OCH ₃	OH	H	H	H	H	OCH ₃	H	H	284.31
43	A17	H	OCH ₃	OH	H	H	H	H	C-F ₃	H	H	322.28
44	A19	H	OCH ₃	OH	H	H	H	NO ₂	H	H	OH	269.28
45	A20	H	OCH ₃	OH	H	H	Phenyl		H	H	OH	320.34

2.2.4. Descrição dos Ensaios de Expressão e Purificação

Os ensaios utilizando o plasmídeo contendo o domínio da CRM1 foram obtidos na Empresa FastBio. Nos experimentos, foram avaliadas as linhagens de bactéria *E. coli* Ril(DE3) e *E. coli* BL21(DE3), que foram transformadas através do método de choque térmico utilizando o plasmídeo solicitado à empresa FastBio. A sequência utilizada para os ensaios foi:

MSEDTEKRFVVTVIKDLLDLKCVKKRGKDNKAVVASDIMYVVGQYPRFLKAHWNFLRTVILKLFEEFMHETH
EGVQDMACDTFIKIVQKCKYHFVIQ

Nos experimentos, as colônias utilizadas foram escolhidas e cresceram em meio LB com $50 \mu\text{g mL}^{-1}$ de kanamicina, para a linhagem BL21(DE3) e $50 \mu\text{g mL}^{-1}$ de kanamicina e $50 \mu\text{g mL}^{-1}$ de cloranfenicol, para a linhagem Ril(DE3). Para crescimento, em ambas as linhagens, foi utilizada a temperatura de 37°C , até obter a densidade ótica 0.6.

Em seguida, para indução da expressão proteica, foram avaliadas, de forma independente, diferentes temperaturas, que variaram entre 20°C , 28°C e 37°C e diferentes concentrações de Isopropyl b-D-1-thiogalactopyranoside (IPTG). Nos experimentos, a expressão da proteína de interesse foi induzida com concentrações de IPTG de 0.4, 0.5 e 1.0 milimol/Litro. O tempo de indução também foi alterado, empregando 4h ou 18h. De forma específica, nesse passo, foram realizadas as combinações de concentração de IPTG, tempo de indução e temperatura conforme apresentado na Tabela 2.

Tabela 2: Condições avaliadas nos ensaios.

Concentração de IPTG	Tempo de Indução	Temperatura
0.5 milimol/Litro	4h	37°C
1.0 milimol/Litro	4h	37°C
0.4 milimol/Litro	18h	28°C
0.4 milimol/Litro	18h	20°C
0.5 milimol/Litro	18h	20°C
1.0 milimol/Litro	18h	20°C

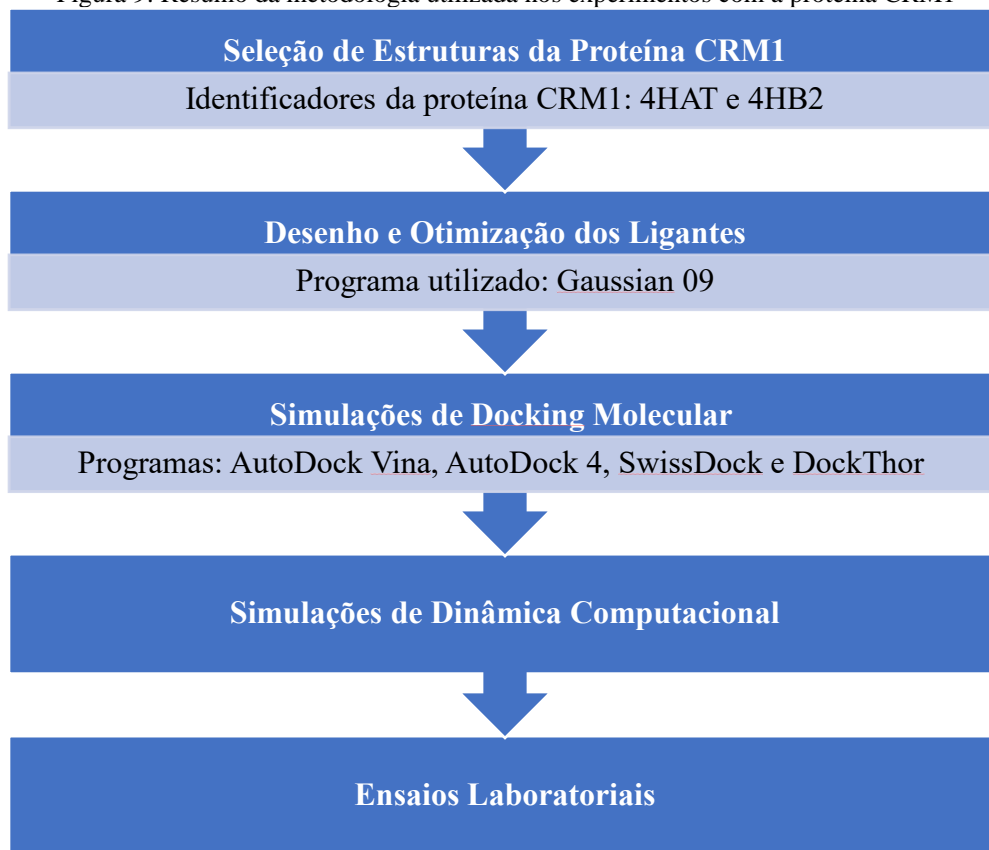
Fonte: Autora.

Após o tempo de indução, as amostras foram centrifugadas a $4,800 \times g$ por 50 minutos, a 4°C e o pellet foi ressuspendido utilizando tampão de lise (50 milimol/Litro Tris-HCl pH 8.0, 500 milimol/Litro NaCl). As células foram lisadas através de sonicação com um pulso de 30 s. Em

seguida, as células foram centrifugadas a 25,000 x g por 50 minutos, a 4°C. O sobrenadante foi coletado e a amostra foi ressuspensa em tampão. Em seguida, foi coletado uma amostra do pellet.

Selecionadas as amostras, utilizou-se o gel SDS-PAGE de 20% por aproximadamente 3 horas (90 volts) para avaliar expressão e solubilidade da proteína de interesse. O domínio de interesse da proteína possui ~12 kDa. Na Figura 9, pode-se observar um fluxograma contendo os experimentos computacionais e de laboratório utilizando a proteína CRM1.

Figura 9: Resumo da metodologia utilizada nos experimentos com a proteína CRM1



Fonte: Autora.

3. RESULTADOS

Os resultados obtidos no presente trabalho são descritos a seguir. Na subseção 3.1, serão apresentados os resultados de *docking* e de dinâmica molecular obtidos utilizando a CRM1 (PDB ID 4HAT e ID 4HB2), p53 e MDM2. Na subseção 3.2, os dados obtidos através dos ensaios laboratoriais utilizando o domínio da CRM1 foram descritos.

3.1. Resultados Computacionais

Nas subseções 3.1.1 e 3.1.2, são apresentados, respectivamente, os resultados de *docking* e dinâmica molecular obtidos ao utilizar os arquivos da proteína CRM1 com PDB de identificador 4HAT e 4HB2. As considerações obtidas através dessas simulações encontram-se na subseção 3.1.3. Nas subseções 3.1.4 e 3.1.5 são apresentados, respectivamente, os resultados de *docking* e dinâmica molecular utilizando a proteína p53 e MDM2.

3.1.1. Resultados para a CRM1 (PDB ID 4HAT)

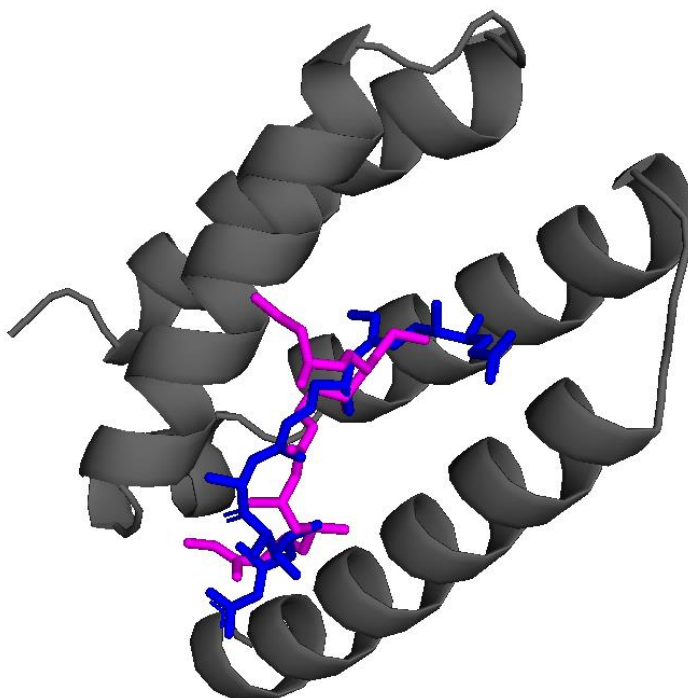
Na seção 3.1.1.1, são apresentados os resultados de *docking* molecular obtidos com a proteína CRM1, empregando o PDB ID 4HAT. Na seção 3.1.1.2, são apresentados os resultados de dinâmica molecular, para essa mesma estrutura.

3.1.1.1. Resultados de *Docking* Molecular

O PDB de identificador 4HAT possui em sua estrutura de cristal o composto Leptomicina B, que foi utilizado para *redocking* e é o ligante de referência nessa simulação. No *redocking*, foi obtida uma afinidade de ligação igual a -7,3 kcal/mol, no programa AutoDock Vina. Na Figura 10, podem ser observadas a conformação real (na estrutura cristalina) em azul e a computacional (obtidas em simulações de *docking* molecular) em magenta. Nessas simulações, no *redocking*, é obtida uma conformação que diverge com os dados da estrutura em cristal, o que pode ser explicado, em partes, devido ao tamanho da estrutura Leptomicina B. Em programas de *docking*, quando o ligante apresenta muitos pontos de torção, como ocorre na Leptomicina B, em que os pontos totalizam 19, os resultados podem não refletir os dados reais. Por esse motivo, a Leptomicina B não foi utilizada nas simulações de dinâmica computacional. Ainda, é de interesse que o ligante de referência empregado possua estrutura semelhante à apresentada pelos demais ligantes avaliados. Em relação

ao trabalho proposto, note que as chalconas utilizadas possuem menos pontos de torção que a Leptomicina B, o que induz a um aumento na confiabilidade dos resultados de *docking* e de dinâmica molecular.

Figura 10: Estrutura da proteína CRM1, em cor cinza, contendo o ligante Leptomicina B, com dados da conformação real, obtida em cristal, na cor azul, e experimental obtida no programa AutoDock Vina, na cor magenta. A proteína CRM1 é representada em *cartoon* e o ligante de referência, Leptomicina B, em *stick*.



Fonte: Autora. A representação estrutural foi preparada utilizando o programa PyMOL (DELANO, 2002).

Os resultados de *docking* molecular para os programas AutoDock 4, AutoDock Vina, DockThor e SwissDock, empregando as 45 chalconas, podem ser observados na Tabela 3. Nessa mesma tabela, é discriminada a posição dos ligantes de acordo com a afinidade de ligação encontrada nesses programas.

Tabela 3: Resultados de afinidade de ligação, em kcal/mol, utilizando os programas AutoDock Vina, AutoDock 4, SwissDock e DockThor (continua).

Chalcona	AutoDock Vina		AutoDock 4		DockThor		SwissDock	
	Afinidade de Ligação (kcal/mol)	Posição	Afinidade de Ligação (kcal/mol)	Posição	Afinidade de Ligação (kcal/mol)	Posição	Afinidade de Ligação (kcal/mol)	Posição
T10	-7,6	14	-5,81	16	-8,052	21	-6,55	43
T33	-7,1	30	-5,45	27	-7,832	32	-6,67	34
T37	-7,6	15	-5,53	21	-8,059	19	-6,66	37
T38	-7,7	12	-5,88	15	-8,189	6	-6,74	29
G2	-7,0	33	-5,42	28	-8,019	24	-6,69	31

G3	-7,2	25	-5,16	37	-8,275	1	-6,69	32
G4	-6,9	36	-5,49	24	-8,244	2	-6,75	28
G5	-6,9	37	-5,60	18	-8,226	3	-6,77	23
G6	-6,7	41	-5,16	38	-7,448	39	-7,15	4
G14	-8,5	4	-6,54	5	-7,984	27	-7,35	1
G15	-9,0	2	-6,52	6	-7,731	35	-7,03	8
G18	-7,7	13	-5,18	36	-8,023	23	-7,32	2
G19	-8,3	6	-6,03	12	-8,135	14	-6,89	13
G20	-6,2	45	-5,01	42	-8,182	7	-6,99	10
G21	-6,3	44	-4,76	44	-8,219	4	-6,54	44
G23	-6,9	38	-5,54	20	-8,017	25	-6,67	35
G24	-7,1	31	-6,14	10	-8,149	10	-6,77	24
G25	-7,5	16	-7,47	1	-8,205	5	-6,86	16
D1	-7,5	17	-5,30	31	-8,148	11	-6,60	41
D2	-7,2	26	-5,59	19	-8,142	13	-6,63	40
D3	-7,3	22	-5,22	32	-8,157	9	-6,59	42
D4	-7,0	34	-5,48	25	-8,158	8	-6,88	15
D6	-7,0	35	-5,20	33	-7,684	37	-6,68	33
D14	-8,6	3	-6,63	4	-7,811	34	-6,85	18
D15	-9,4	1	-7,03	2	-7,388	41	-7,12	5
D18	-7,8	10	-5,19	34	-8,058	20	-6,76	25
D19	-8,3	7	-6,51	7	-7,001	45	-6,91	12
D20	-6,5	43	-5,05	41	-7,837	30	-6,67	36
D21	-6,6	42	-4,70	45	-8,130	16	-6,48	45
D23	-7,2	27	-5,35	30	-7,818	33	-6,66	38
D24	-7,2	28	-5,97	13	-8,107	17	-6,72	30
D25	-8,0	9	-6,1	11	-7,879	29	-6,81	20
A2	-7,5	18	-5,48	26	-8,134	15	-6,65	39
A5	-6,9	39	-5,09	39	-8,086	18	-6,80	21
A6	-8,2	8	-6,48	8	-7,306	43	-7,09	7
A8	-7,8	11	-5,91	14	-7,945	28	-7,02	9
A9	-7,4	20	-5,68	17	-7,372	42	-6,76	26
A10	-7,3	23	-5,38	29	-8,048	22	-6,79	22
A11	-7,4	21	-5,51	23	-8,145	12	-6,76	27
A12	-6,9	40	-5,08	40	-7,834	31	-6,84	19
A14	-7,2	29	-5,52	22	-7,715	36	-6,89	14
A15	-7,1	32	-5,19	35	-7,443	40	-6,86	17
A17	-7,5	19	-4,88	43	-7,199	44	-7,18	3
A19	-7,3	24	-6,22	9	-7,988	26	-6,95	11
A20	-8,4	5	-6,94	3	-7,683	38	-7,11	6

Fonte: Resultados compilados pela autora, com dados de afinidade de ligação obtidos nos programas AutoDock Vina, AutoDock 4, DockThor e SwissDock.

Considerando a afinidade de ligação obtida para o ligante de referência no programa AutoDock Vina, que foi igual a -7,3 kcal/mol, é possível verificar que os dados obtidos no mesmo programa para 21 chalconas foi superior. As chalconas D15, G15, D14, G14, A20, G19, D19, A6, D25, D18, A8, T38, G18, T10, T37, G25, D1, A2, A17, A9 e A11 apresentaram melhores afinidades de ligação que a Leptomicina B. A ordenação das dez melhores afinidades de ligação obtidas nos programas AutoDock Vina AutoDock 4, DockThor e SwissDock pode ser observada na Tabela 4.

Tabela 4: Combinação dos dez melhores resultados de afinidade de ligação obtidos em experimentos de *docking* molecular para a proteína CRM1 (PDB ID 4HAT) e chalconas.

AutoDock Vina		AutoDock		SwissDock		DockThor	
Chalcona	Afinidade de Ligação (kcal/mol)	Chalcona	Afinidade de Ligação (kcal/mol)	Chalcona	Afinidade de Ligação (kcal/mol)	Chalcona	Afinidade de Ligação (kcal/mol)
D15	-9,4	G25	-7,47	G14	-7,35	G3	-8,275
G15	-9,0	D15	-7,03	G18	-7,32	G4	-8,244
D14	-8,6	A20	-6,94	A17	-7,18	G5	-8,226
G14	-8,5	D14	-6,63	G6	-7,15	G21	-8,219
A20	-8,4	G14	-6,54	D15	-7,12	T38	-8,189
G19	-8,3	G15	-6,52	A20	-7,11	G20	-8,182
D19	-8,3	D19	-6,51	A6	-7,09	D4	-8,158
A6	-8,2	A6	-6,48	G15	-7,03	D3	-8,157
D25	-8,0	A19	-6,22	A8	-7,02	G24	-8,149
D18	-7,8	G24	-6,14	G20	-6,99	D1	-8,148

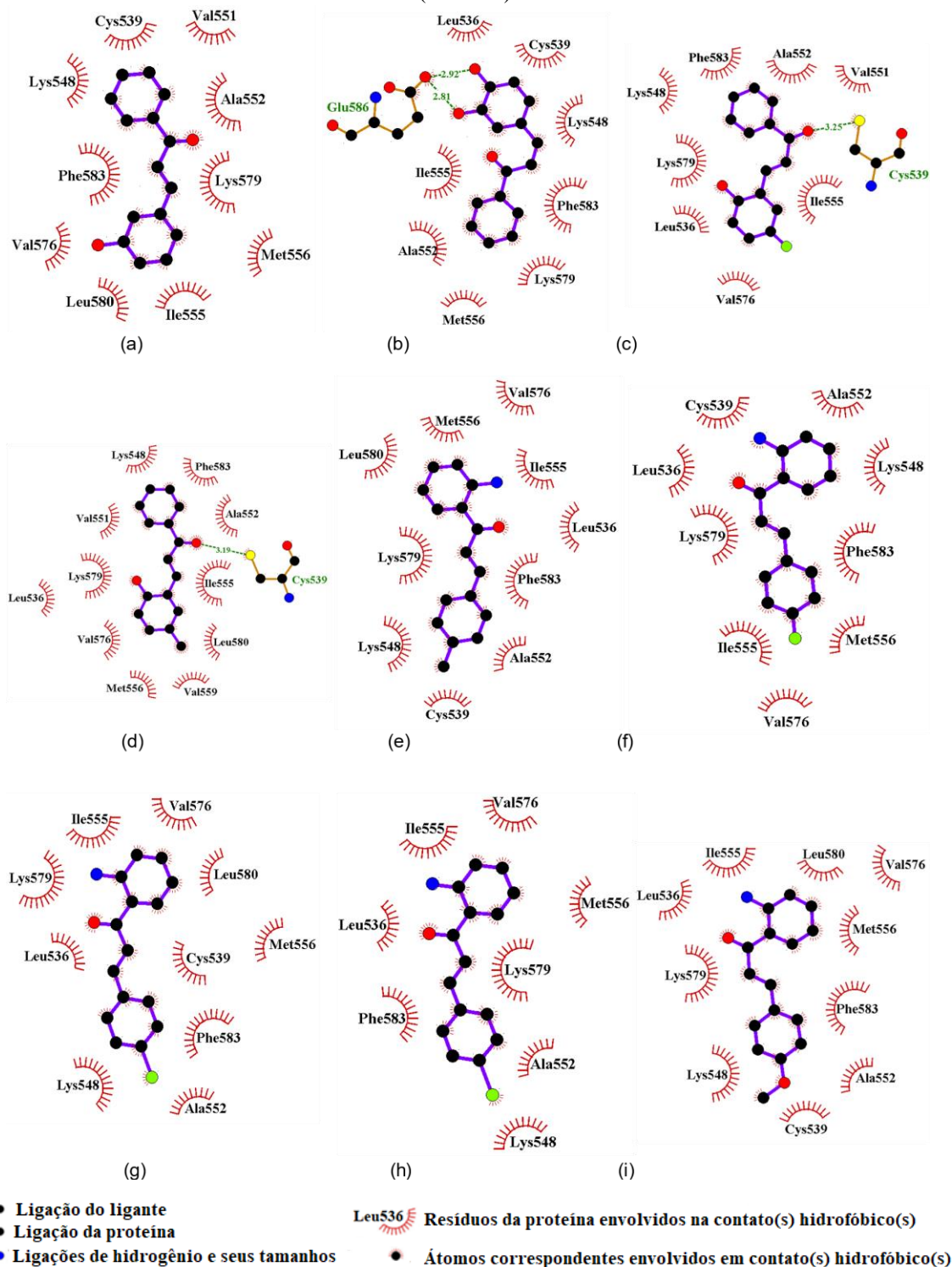
Fonte: Resultados compilados pela autora, com dados de afinidade de ligação obtidos nos programas AutoDock Vina AutoDock 4, DockThor e SwissDock.

É possível observar que o uso do programa DockThor apresentou resultados consideravelmente divergentes quando comparados aos obtidos pelos demais programas de *docking* molecular. As afinidades de ligação identificadas por esse programa para os dez primeiros ligantes, quando ordenadas, são coincidentes somente para o ligante G20 (também posicionado entre os dez primeiros para o programa SwissDock) e G24 (verificado entre os dez primeiros pelo programa AutoDock). É possível que esse programa não seja adequado para simular o comportamento da interação entre a proteína CRM1 e as 45 chalconas. Por esse motivo, esse programa não foi utilizado nos experimentos com a proteína CRM1, empregando o PDB ID 4HB2. Desconsiderando os valores obtidos no programa DockThor, em relação aos dados apresentados na Tabela 4, é possível verificar a seguinte ocorrência das chalconas avaliadas nos dez primeiros posicionamentos em relação à afinidade de ligação:

- Ocorrência em todos os programas: D15, G15, G14, A20 e A6.
- Ocorrência em dois programas: D14 e D19, ambos verificados nos programas AutoDock e AutoDock Vina.
- Ocorrência em apenas um programa: G19, D25 e D18, para o AutoDock Vina; G25, A19 e G24, para o AutoDock 4; e G18, A17, G6, A8 e G20, para o SwissDock.

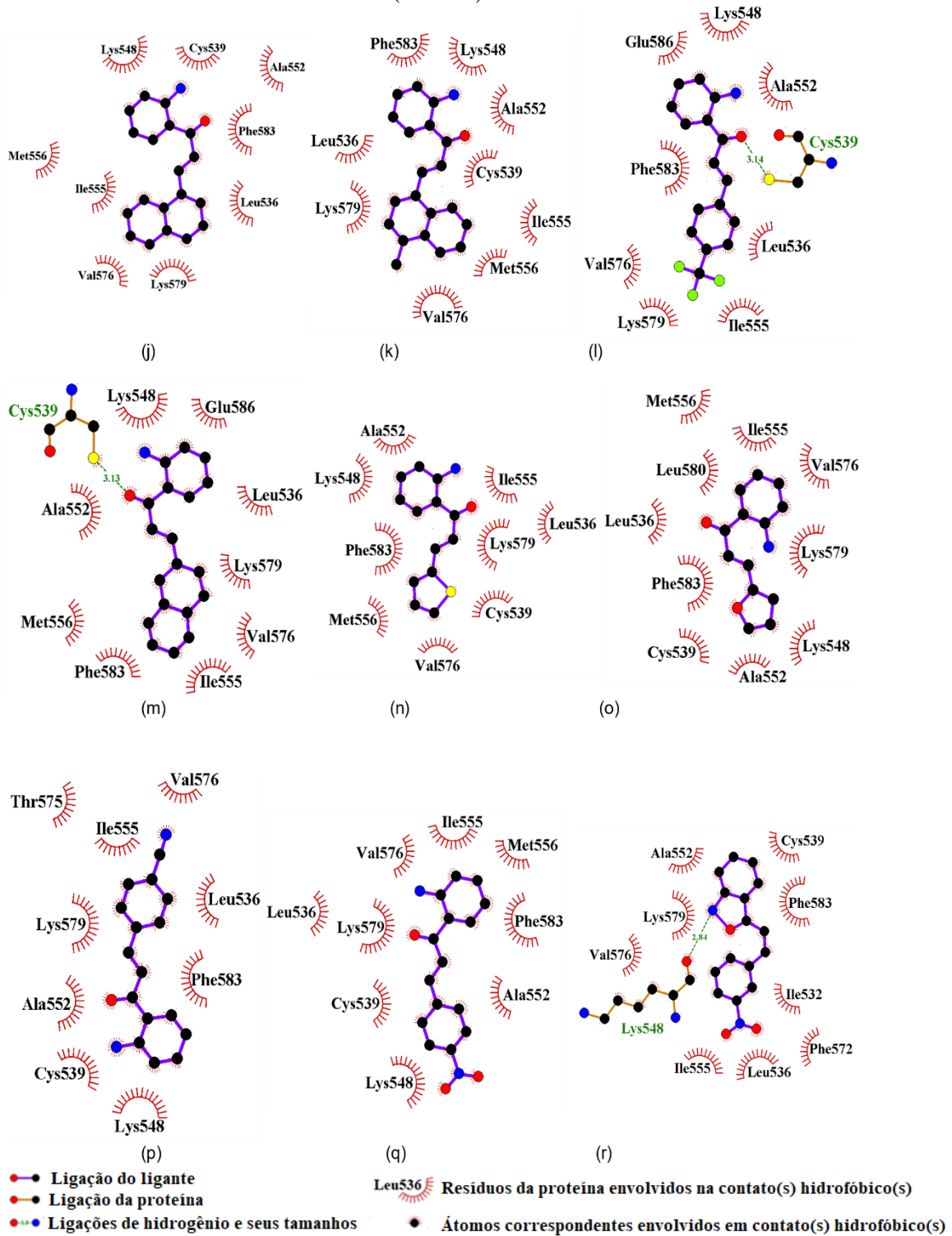
Os resíduos gerados entre chalconas e a proteína CRM1, obtidos no programa LigPlot+ com a saída do programa AutoDock Vina, são apresentados a seguir, na Figura 11. Essas informações também podem ser verificadas nas Tabelas 5 e 6.

Figura 11: Resíduos formados entre CRM1 e ligantes: (a) T10; (b) T33; (c) T37; (d) T38; (e) G2; (f) G3; (g) G4; (h) G5; (i) G6; (j) G14; (k) G15; (l) G18; (m) G19; (n) G20; (o) G21; (p) G23; (q) G24; (r) G25; (s) D1; (t) D2; (u) D3; (v) D4; (w) D6; (x) D14; (y) D15; (z) D18; (aa) D19; (ab) D20; (ac) D21; (ad) D23; (ae) D24; (af) D25; (ag) A2; (ah) A5; (ai) A6; (aj) A8; (ak) A9; (al) A10; (am) A11; (an) A12; (ao) A14; (ap) A15; (aq) A17; (ar) A19; (as) A20 (continua).



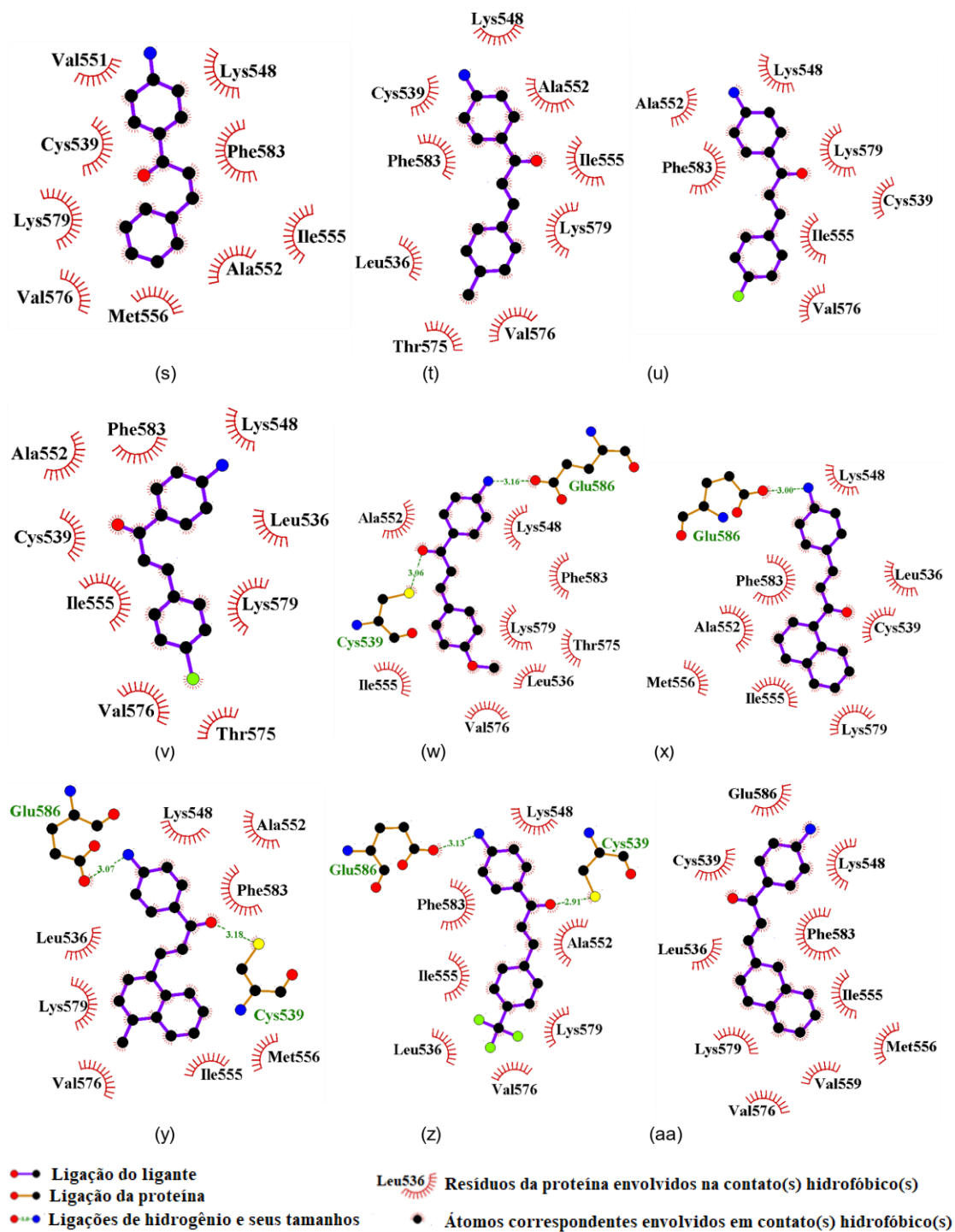
Fonte: Autora, com dados de resíduos obtidos no programa LigPlot+.

Figura 11: Resíduos formados entre CRM1 e ligantes: (a) T10; (b) T33; (c) T37; (d) T38; (e) G2; (f) G3; (g) G4; (h) G5; (i) G6; (j) G14; (k) G15; (l) G18; (m) G19; (n) G20; (o) G21; (p) G23; (q) G24; (r) G25; (s) D1; (t) D2; (u) D3; (v) D4; (w) D6; (x) D14; (y) D15; (z) D18; (aa) D19; (ab) D20; (ac) D21; (ad) D23; (ae) D24; (af) D25; (ag) A2; (ah) A5; (ai) A6; (aj) A8; (ak) A9; (al) A10; (am) A11; (an) A12; (ao) A14; (ap) A15; (aq) A17; (ar) A19; (as) A20 (continua).



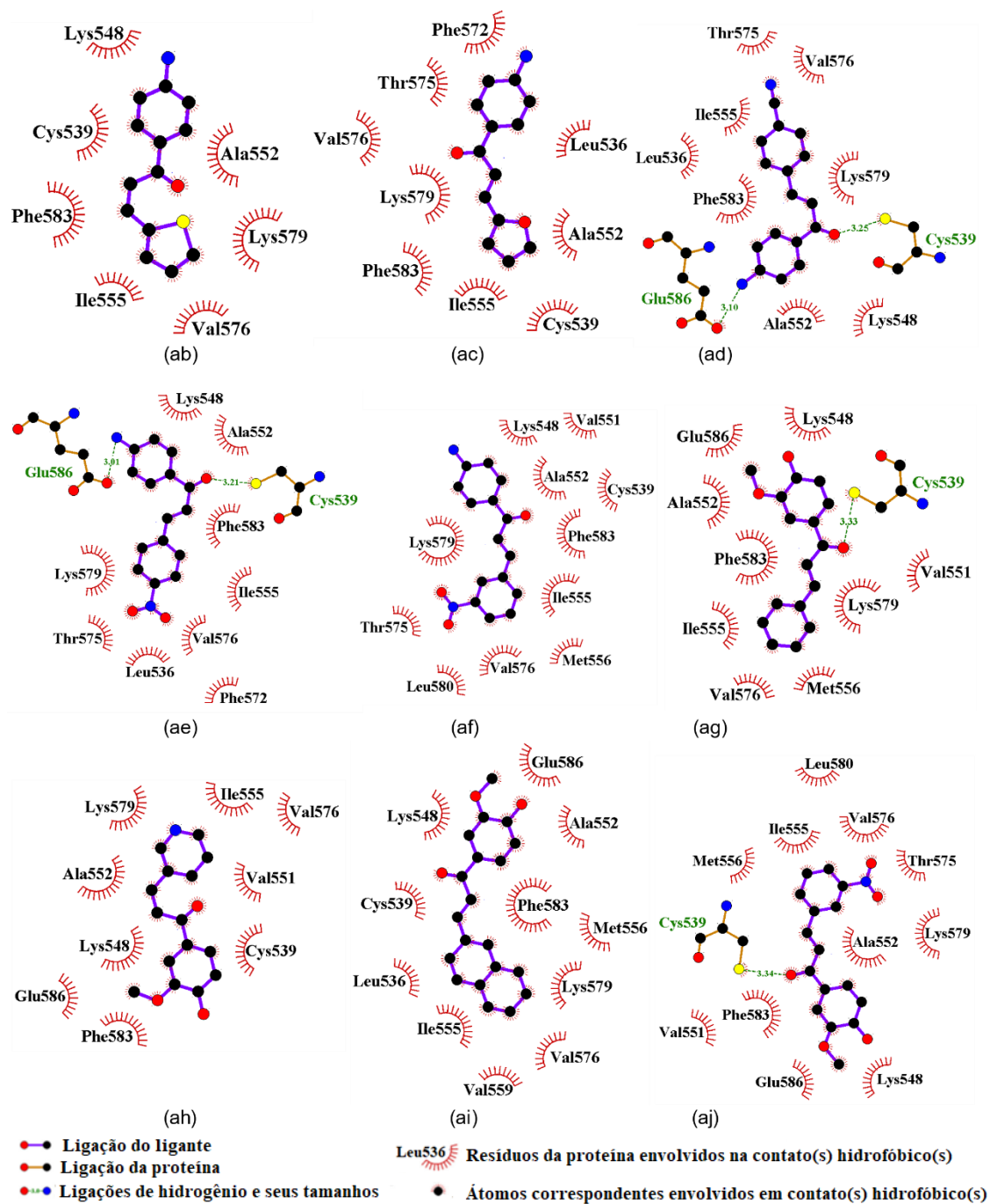
Fonte: Autora, com dados de resíduos obtidos no programa LigPlot+.

Figura 11: Resíduos formados entre CRM1 e ligantes: (a) T10; (b) T33; (c) T37; (d) T38; (e) G2; (f) G3; (g) G4; (h) G5; (i) G6; (j) G14; (k) G15; (l) G18; (m) G19; (n) G20; (o) G21; (p) G23; (q) G24; (r) G25; (s) D1; (t) D2; (u) D3; (v) D4; (w) D6; (x) D14; (y) D15; (z) D18; (aa) D19; (ab) D20; (ac) D21; (ad) D23; (ae) D24; (af) D25; (ag) A2; (ah) A5; (ai) A6; (aj) A8; (ak) A9; (al) A10; (am) A11; (an) A12; (ao) A14; (ap) A15; (aq) A17; (ar) A19; (as) A20 (continua).



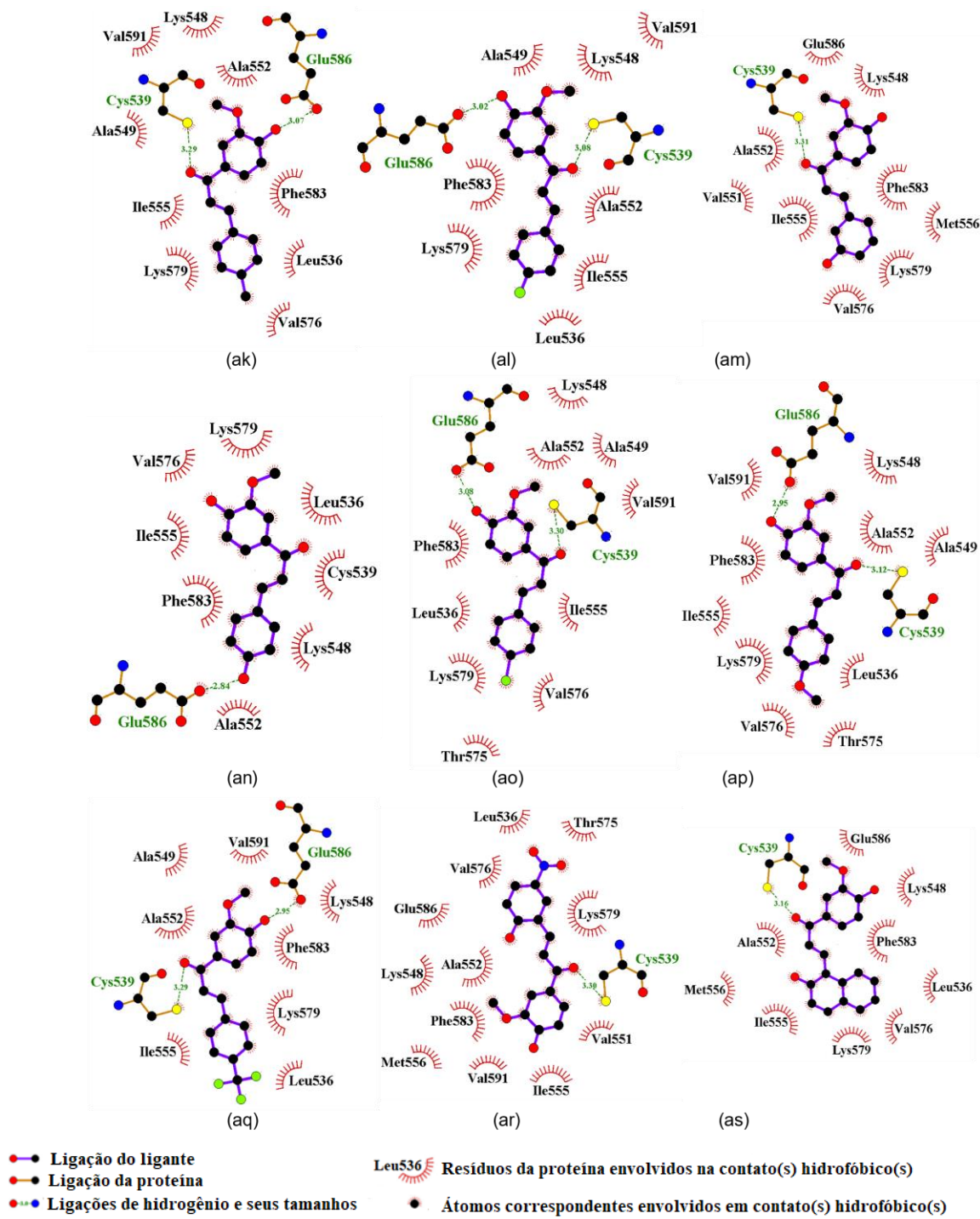
Fonte: Autora, com dados de resíduos obtidos no programa LigPlot+.

Figura 11: Resíduos formados entre CRM1 e ligantes: (a) T10; (b) T33; (c) T37; (d) T38; (e) G2; (f) G3; (g) G4; (h) G5; (i) G6; (j) G14; (k) G15; (l) G18; (m) G19; (n) G20; (o) G21; (p) G23; (q) G24; (r) G25; (s) D1; (t) D2; (u) D3; (v) D4; (w) D6; (x) D14; (y) D15; (z) D18; (aa) D19; (ab) D20; (ac) D21; (ad) D23; (ae) D24; (af) D25; (ag) A2; (ah) A5; (ai) A6; (aj) A8; (ak) A9; (al) A10; (am) A11; (an) A12; (ao) A14; (ap) A15; (aq) A17; (ar) A19; (as) A20 (continua).



Fonte: Autora, com dados de resíduos obtidos no programa LigPlot+.

Figura 11: Resíduos formados entre CRM1 e ligantes: (a) T10; (b) T33; (c) T37; (d) T38; (e) G2; (f) G3; (g) G4; (h) G5; (i) G6; (j) G14; (k) G15; (l) G18; (m) G19; (n) G20; (o) G21; (p) G23; (q) G24; (r) G25; (s) D1; (t) D2; (u) D3; (v) D4; (w) D6; (x) D14; (y) D15; (z) D18; (aa) D19; (ab) D20; (ac) D21; (ad) D23; (ae) D24; (af) D25; (ag) A2; (ah) A5; (ai) A6; (aj) A8; (ak) A9; (al) A10; (am) A11; (an) A12; (ao) A14; (ap) A15; (aq) A17; (ar) A19; (as) A20.



Fonte: Autora, com dados de resíduos obtidos no programa LigPlot+.

Na Tabela 5, pode-se observar que os resíduos de aminoácidos Ile555, Lys579 e Phe583 são comuns a todos os ligantes considerados. Ainda, os resíduos Ala552, Cys539 e Lys548 são encontrados em 97,7% das ligações formadas entre as chalconas consideradas e a proteína CRM1. Esses são os resíduos identificados com maior frequência para os complexos avaliados. Na Tabela 6, pode-se observar essas informações de forma detalhada, explicitando os ligantes e os respectivos resíduos.

Tabela 6: Resíduos calculados no programa LigPlot+ para os complexos contendo CRM1 e 45 chalconas.

Chalcona	Resíduos de aminoácidos identificados
T10	Ala552, Cys539, Ile555, Leu580, Lys548, Lys579, Met556, Phe583, Val551 e Val576.
T33	Ala552, Cys539, Glu586, Ile555, Leu536, Lys548, Lys579, Met556, Phe583.
T37	Ala552, Cys539, Ile555, Leu536, Lys548, Lys579, Phe583, Val551 e Val576.
T38	Ala552, Cys539, Ile555, Leu536, Leu580, Lys548, Lys579, Met556, Phe583, Val551, Val559 e Val576.
G2	Ala552, Cys539, Ile555, Leu536, Leu580, Lys548, Lys579, Met556, Phe583 e Val576.
G3	Ala552, Cys539, Ile555, Leu536, Lys548, Lys579, Met556, Phe583 e Val576.
G4	Ala552, Cys539, Ile555, Leu536, Leu580, Lys548, Lys579, Met556, Phe583 e Val576.
G5	Ala552, Ile555, Leu536, Lys548, Lys579, Met556, Phe583 e Val576.
G6	Ala552, Cys539, Ile555, Leu536, Leu580, Lys548, Lys579, Met556, Phe583 e Val576.
G14	Ala552, Cys539, Ile555, Leu536, Lys548, Lys579, Met556, Phe583 e Val576.
G15	Ala552, Cys539, Ile555, Leu536, Lys548, Lys579, Met556, Phe583 e Val576.
G18	Ala552, Cys539, Glu586, Ile555, Leu536, Lys548, Lys579, Phe583 e Val576.
G19	Ala552, Cys539, Glu586, Ile555, Leu536, Lys548, Lys579, Met556, Phe583 e Val576.
G20	Ala552, Cys539, Ile555, Leu536, Lys548, Lys579, Met556, Phe583 e Val576.
G21	Ala552, Cys539, Ile555, Leu536, Leu580, Lys548, Lys579, Met556, Phe583 e Val576.
G23	Ala552, Cys539, Ile555, Leu536, Lys548, Lys579, Phe583, Thr575 e Val576.
G24	Ala552, Cys539, Ile555, Leu536, Lys548, Lys579, Met556, Phe583 e Val576.
G25	Ala552, Cys539, Ile532, Ile555, Leu536, Lys548, Lys579, Phe572, Phe583 e Val576.
D1	Ala552, Cys539, Ile555, Lys548, Lys579, Met556, Phe583, Val551 e Val576.
D2	Ala552, Cys539, Ile555, Leu536, Lys548, Lys579, Phe583, Thr575 e Val576.
D3	Ala552, Cys539, Ile555, Lys548, Lys579, Phe583 e Val576.
D4	Ala552, Cys539, Ile555, Leu536, Lys548, Lys579, Phe583, Thr575 e Val576.
D6	Ala552, Cys539, Glu586, Ile555, Leu536, Lys548, Lys579, Phe583, Thr575 e Val576.
D14	Ala552, Cys539, Glu586, Ile555, Leu536, Lys548, Lys579, Met556 e Phe583.
D15	Ala552, Cys539, Glu586, Ile555, Leu536, Lys548, Lys579, Met556, Phe583 e Val576.
D18	Ala552, Cys539, Glu586, Ile555, Leu536, Lys548, Lys579, Phe583 e Val576.
D19	Cys539, Glu586, Ile555, Leu536, Lys548, Lys579, Met556, Phe583, Val559 e Val576.
D20	Ala552, Cys539, Ile555, Lys548, Lys579, Phe583 e Val576.
D21	Ala552, Cys539, Ile555, Leu536, Lys579, Phe572, Phe583, Thr575 e Val576.

D23	Ala552, Cys539, Glu586, Ile555, Leu536, Lys548, Lys579, Phe583, Thr575 e Val576.
D24	Ala552, Cys539, Glu586, Ile555, Leu536, Lys548, Lys579, Phe572, Phe583, Thr575 e Val576.
D25	Ala552, Cys539, Ile555, Leu580, Lys548, Lys579, Met556, Phe583, Thr575, Val551 e Val576.
A2	Ala552, Cys539, Glu586, Ile555, Lys548, Lys579, Met556, Phe583, Val551 e Val576.
A5	Ala552, Cys539, Glu586, Ile555, Lys548, Lys579, Phe583, Val551 e Val576.
A6	Ala552, Cys539, Glu586, Ile555, Leu536, Lys548, Lys579, Met556, Phe583, Val559 e Val576.
A8	Ala552, Cys539, Glu586, Ile555, Leu580, Lys548, Lys579, Met556, Phe583, Thr575, Val551 e Val576.
A9	Ala549, Ala552, Cys539, Glu586, Ile555, Leu536, Lys548, Lys579, Phe583, Val576 e Val591.
A10	Ala549, Ala552, Cys539, Glu586, Ile555, Leu536, Lys548, Lys579, Phe583 e Val591.
A11	Ala552, Cys539, Glu586, Ile555, Lys548, Lys579, Met556, Phe583, Val551 e Val576.
A12	Ala552, Cys539, Glu586, Ile555, Leu536, Lys548, Lys579, Phe583, Val576.
A14	Ala549, Ala552, Cys539, Glu586, Ile555, Leu536, Lys548, Lys579, Phe583, Thr575, Val576 e Val591.
A15	Ala549, Ala552, Cys539, Glu586, Ile555, Leu536, Lys548, Lys579, Phe583, Thr575, Val576 e Val591.
A17	Ala549, Ala552, Cys539, Glu586, Ile555, Leu536, Lys548, Lys579, Phe583 e Val591.
A19	Ala552, Cys539, Glu586, Ile555, Leu536, Lys548, Lys579, Met556, Phe583, Thr575, Val551, Val576 e Val591.
A20	Ala552, Cys539, Glu586, Ile555, Leu536, Lys548, Lys579, Met556, Phe583 e Val576.

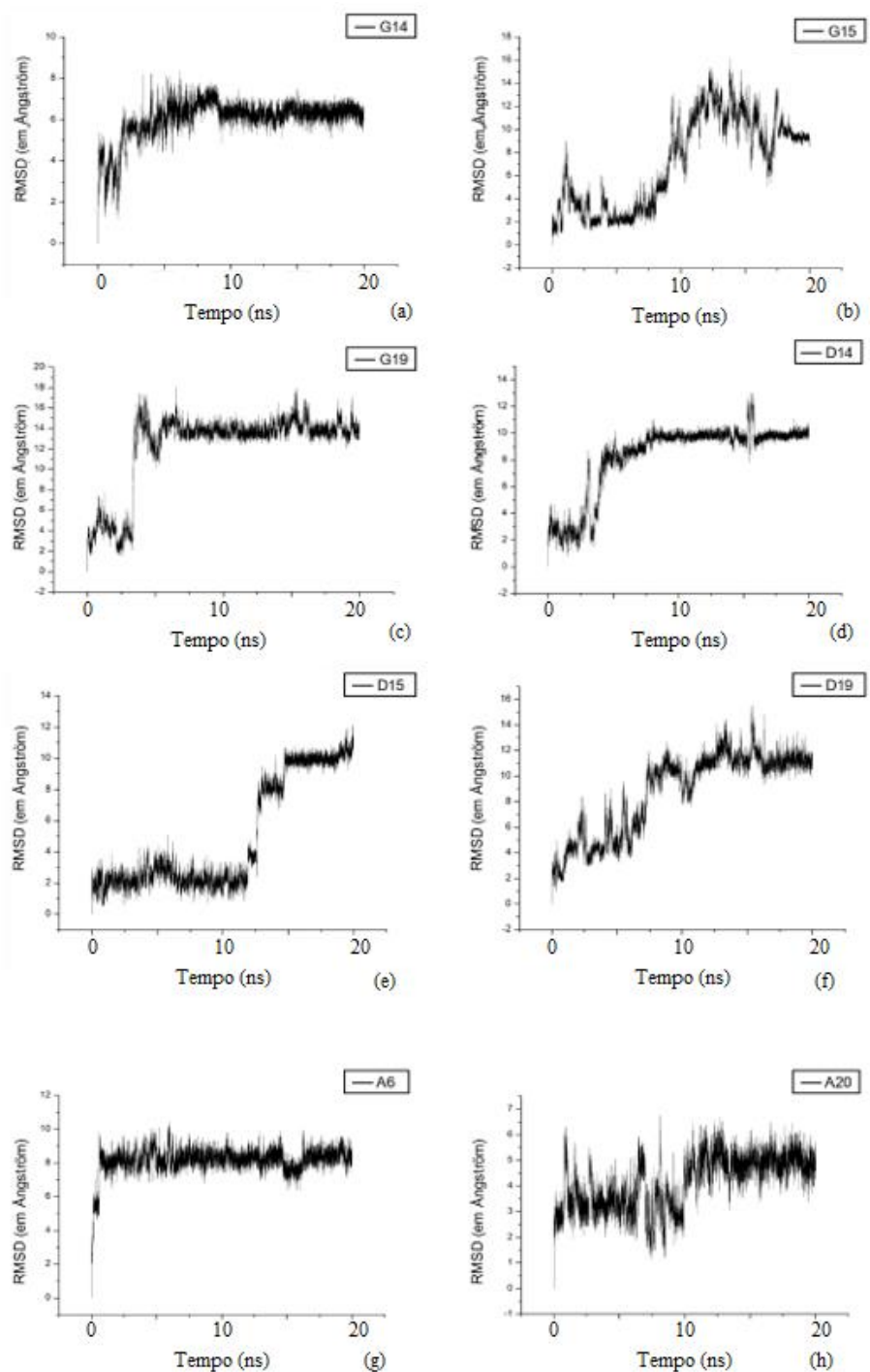
Fonte: Autora, com dados de resíduos obtidos no programa LigPlot+.

As simulações de *docking* molecular ocorreram em duplicata, de modo que o PDB ID 4HB2 também foi utilizado nas avaliações, em que resultados são detalhados na seção 3.1.2. Os ligantes foram avaliados conforme a afinidade de ligação encontrada nos dois arquivos, considerando os programas AutoDock, AutoDock Vina e SwissDock e foram selecionadas as chalconas com os melhores resultados para os experimentos de *docking*. As chalconas selecionadas foram: i) G14; ii) G15; iii) G19; iv) D14; v) D15; vi) D19; vii) A6 e; viii) A20 e esses ligantes foram utilizados em simulações de dinâmica molecular. Detalhes sobre como foi realizada essa seleção das 8 chalconas são apresentados na seção 3.1.4.1.

3.1.1.2. Resultados de Dinâmica Molecular

Nas simulações de DM, foram avaliados valores de RMSD, número de pontes de hidrogênio e número de contatos, em trajetórias de 20 ns. Na Figura 12, são observados resultados de RMSD de todos os átomos que não são hidrogênio para os 8 ligantes selecionados, com o PDB ID 4HAT. Para esse arquivo, não foi utilizado o ligante de referência, uma vez que a Leptomicina B não possui estrutura molecular semelhante à das chalconas selecionadas.

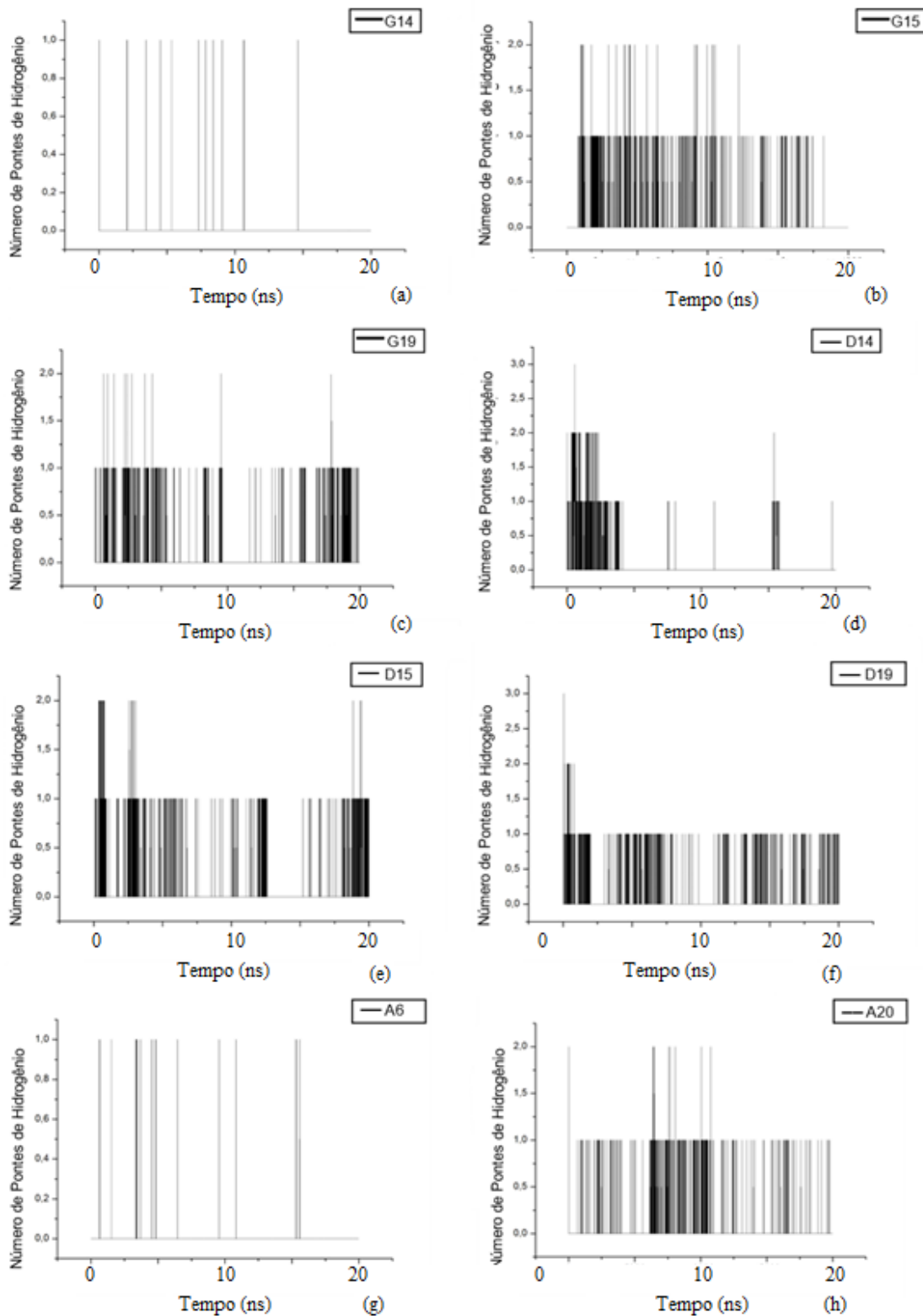
Figura 12: Resultados de simulação de dinâmica computacional para RMSD utilizando a proteína CRM1 (PDB ID 4HAT) com os ligantes: (a) G14; (b) G15; (c) G19; (d) D14; (e) D15; (f) D19; (g) A6; e (h) A20.



Fonte: Autora, com resultados de RMSD obtidos no programa GROMACS.

Na Figura 13, são verificados os resultados obtidos em simulações de dinâmica molecular referentes ao número de pontes de hidrogênio formados entre CRM1 e ligantes.

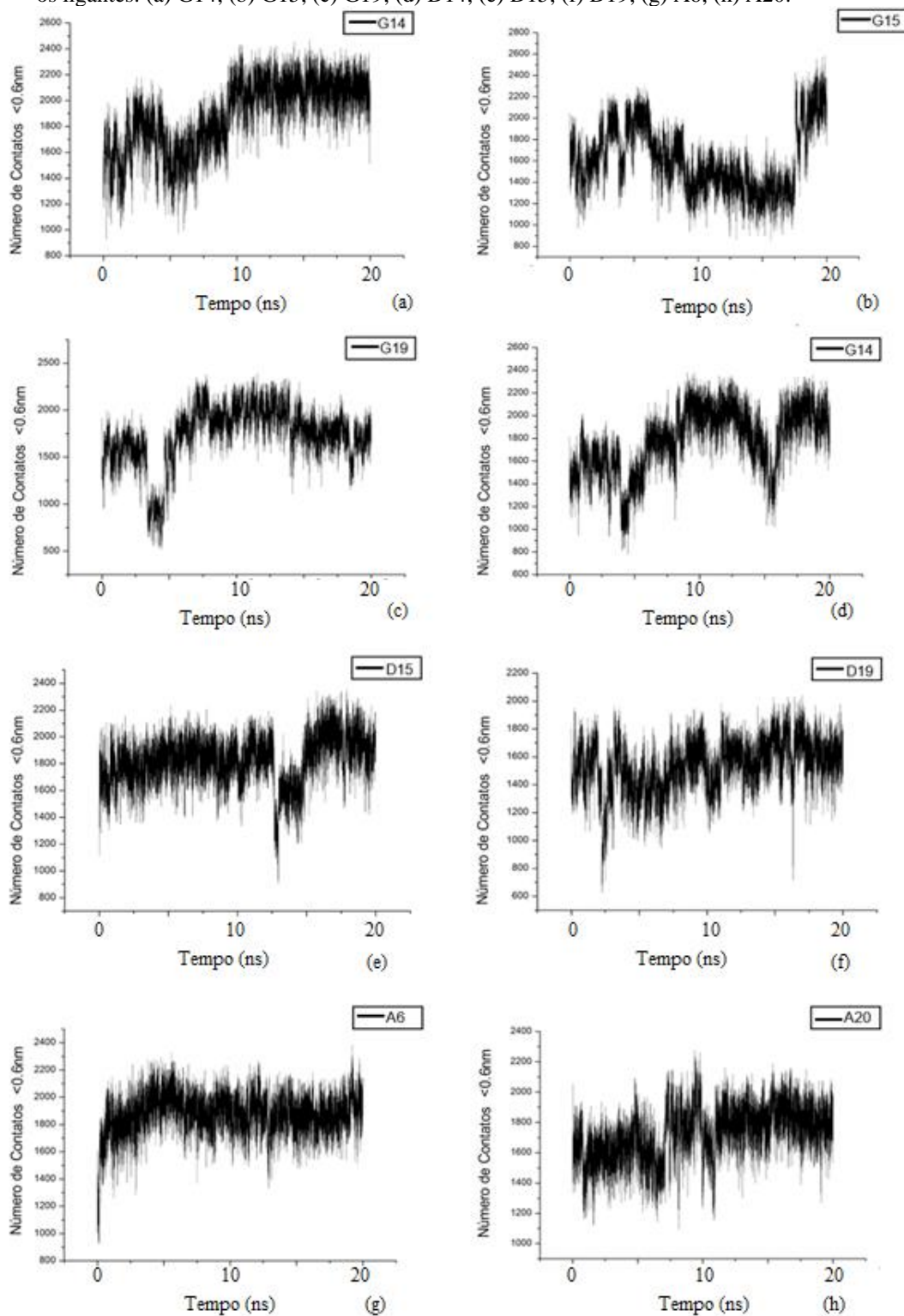
Figura 13: Resultados de simulação de dinâmica computacional para pontes de hidrogênio utilizando a proteína CRM1 (PDB ID 4HAT) com os ligantes: (a) G14; (b) G15; (c) G19; (d) D14; (e) D15; (f) D19; (g) A6; (h) A20.



Fonte: Autora, com resultados de RMSD obtidos no programa GROMACS.

Uma terceira informação obtida em simulações de dinâmica molecular foi o número de contatos, cujos resultados obtidos para a proteína CRM1 são observados na Figura 14.

Figura 14: Resultados de simulação de dinâmica computacional para número de contatos da proteína CRM1 com os ligantes: (a) G14; (b) G15; (c) G19; (d) D14; (e) D15; (f) D19; (g) A6; (h) A20.



Fonte: Autora, com resultados de RMSD obtidos no programa GROMACS.

Nas simulações de RMSD, cujos resultados são apresentados na Figura 12, pode-se observar, para o ligante G14 (Figura 12a), é obtida uma estabilidade na ligação a partir de 10ns, que é mantida ao longo da trajetória, que possui 20ns. Essa estabilidade também é observada para o ligante G19 (Figura 12c), cuja estabilidade tem início aos 3ns de trajetória e que se mantém até o final da simulação. Para a chalcona D14 (Figura 12d), a estabilidade é iniciada próximo aos 8ns, sendo que aos 15ns há uma alteração na conformação, mas essa estabilidade é retomada. Para a D15 (Figura 12e), há uma alteração na conformação gerada no complexo aos 12ns, que é mantida até o final da simulação computacional. Para o ligante A6 (Figura 12g), essa estabilidade tem início aos 2ns e é mantida ao longo de toda a trajetória. Para os demais ligantes, que são G15 (Figura 12b), D19 (Figura 12f) e A20 (Figura 12h), é observada uma menor estabilidade na ligação formada com a proteína CRM1.

Tratando-se do número de pontes de hidrogênio, em que os dados para o arquivo de PDB de identificador 4HAT são observados na Figura 13, o valor obtido foi 0 ± 0 ponte. Para os resultados referentes ao número de contatos para o arquivo de PDB ID 4HAT, cujos comportamentos podem ser verificados na Figura 14, nota-se que os ligantes G14 e A6 apresentaram os maiores números de contatos, sendo que a chalcona G14 (Figura 14a) apresenta 1889 ± 263 contatos e a A6 (Figura 14g) apresenta 1880 ± 147 contatos. Os menores valores foram verificados para os ligantes D19 (Figura 14f) e G15 (Figura 14b), que formam 1542 ± 180 e 1618 ± 301 contatos, respectivamente. Os resultados obtidos para as demais chalconas, que são G19 (Figura 14c), D14 (Figura 14d), D15 (Figura 14e) e A20 (Figura 14h) foram 1738 ± 307 , 1785 ± 277 , 1827 ± 175 e 1729 ± 168 , respectivamente.

3.1.2. Resultados para a CRM1 (PDB ID 4HB2)

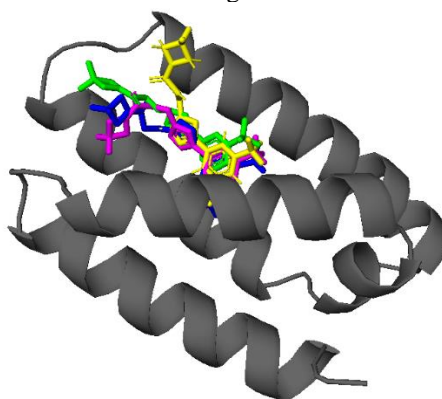
Na seção 3.1.2.1, são apresentados os resultados de *docking* molecular obtidos com a proteína CRM1, empregando o arquivo de PDB ID 4HB2. Na seção 3.1.2.2, são apresentados os resultados de dinâmica molecular, contendo dados de RMSD, número de pontes de hidrogênio e número de contatos, utilizando essa mesma estrutura.

3.1.2.1. Resultados de *Docking* Molecular

A fim de confrontar os resultados de *docking* molecular obtidos com a proteína de identificador 4HAT e assegurar a escolha dos melhores ligantes para as simulações de dinâmica, utilizou-se o arquivo de PDB ID 4HB2, que também contém a estrutura da proteína CRM1. Definiu-se o ligante K76 como referência, uma vez que a estrutura em cristal dessa molécula está disponível

no PDB de mesmo identificador. Para esse ligante, foi verificada uma afinidade de ligação igual a -9,3 kcal/mol, no programa AutoDock Vina, -4,82 kcal/mol no programa AutoDock 4 e -7,02 kcal/mol no programa SwissDock. Os resultados obtidos em *redocking* nos programas AutoDock, AutoDock Vina e SwissDock para o K76 podem ser observados na Figura 15. Em magenta, encontra-se a conformação obtida no programa AutoDock Vina, em azul no AutoDock, em verde no SwissDock e em amarelo está o ligante em sua estrutura no cristal. Note que as conformações obtidas no AutoDock e AutoDock Vina são semelhantes à encontrada na estrutura em cristal. Já o programa SwissDock gera uma conformação semelhante à do cristal apenas na 15ª conformação, que é a pose apresentada na Figura 15. Conforme descrito anteriormente, essa concordância em conformações não foi verificada no *redocking* para a Leptomicina B (PDB ID 4HAT). Assim, somente o K76 é utilizado como ligante de referência nas simulações de DM.

Figura 15: Resultados de *redocking* para o ligante K76. A proteína é representada em *cartoon*, na cor cinza, e o ligante K76, em todas as conformações, em *stick*. Em magenta, encontra-se a conformação do K76 obtida em *redocking* no programa AutoDock Vina, em azul no programa AutoDock, em verde o resultado no SwissDock e em amarelo o ligante em sua estrutura no cristal.



Fonte: Autora. A representação estrutural foi preparada no programa PyMOL (DELANO, 2002).

Para as simulações de *docking* molecular com o arquivo PDB ID 4HB2, foram utilizados os programas AutoDock Vina, AutoDock 4 e SwissDock, para as 45 chalconas. Na Tabela 7, podem ser observados os resultados obtidos ao utilizar o programa AutoDock Vina, AutoDock e SwissDock.

Tabela 7: Resultados de afinidade de ligação, em kcal/mol, utilizando os programas AutoDock Vina, AutoDock 4 e SwissDock (continua).

Ligante	AutoDock Vina		AutoDock 4		SwissDock	
	Afinidade de Ligação (kcal/mol)	Posição	Afinidade de Ligação (kcal/mol)	Posição	Afinidade de Ligação (kcal/mol)	Posição

Ligante de Referência	-9,3 kcal/mol	-	-4,82 kcal/mol	-	-7,02 kcal/mol	-
T10	-8,0	16	-5,80	25	-6,80	25
T33	-7,7	26	-5,73	26	-6,98	11
T37	-7,6	31	-5,60	30	-6,55	38
T38	-7,8	21	-5,94	20	-6,69	31
G2	-7,8	22	-5,55	33	-7,36	3
G3	-7,6	32	-5,02	42	-6,55	39
G4	-7,6	33	-5,70	27	-6,54	40
G5	-7,6	34	-5,94	21	-7,45	1
G6	-7,4	38	-5,32	38	-6,64	35
G14	-8,9	8	-6,69	11	-6,84	22
G15	-9,4	5	-7,01	7	-7,12	6
G18	-8,7	10	-5,27	39	-6,79	26
G19	-9,5	3	-6,75	9	-6,70	30
G20	-6,9	42	-4,83	44	-6,84	23
G21	-6,9	43	-4,73	45	-6,91	14
G23	-7,3	39	-5,91	23	-6,78	28
G24	-7,9	18	-6,52	13	-6,53	41
G25	-8,3	13	-7,34	1	-6,60	36
D1	-7,7	27	-5,68	28	-6,68	33
D2	-8,0	17	-6,14	16	-6,50	42
D3	-7,8	23	-5,58	32	-6,91	13
D4	-7,8	24	-6,26	15	-6,49	43
D6	-7,7	28	-5,92	22	-7,42	2
D14	-9,6	2	-7,06	4	-6,69	32
D15	-10,2	1	-7,21	2	-7,09	7
D18	-8,8	9	-5,59	31	-6,85	20
D19	-9,5	4	-7,05	5	-6,72	29
D20	-7,0	41	-5,44	36	-6,91	15
D21	-6,9	44	-4,95	43	-6,41	44
D23	-7,2	40	-6,40	14	-6,66	34
D24	-8,1	15	-6,88	8	-6,36	45
D25	-8,4	12	-6,57	12	-6,90	16
A2	-7,5	35	-5,54	34	-6,85	21
A5	-6,9	45	-5,25	40	-6,89	18
A6	-9,3	7	-7,02	6	-6,91	12
A8	-8,2	14	-6,13	17	-7,06	8
A9	-7,9	19	-6,02	18	-6,79	27
A10	-7,8	25	-5,45	35	-6,86	19
A11	-7,5	36	-5,37	37	-6,82	24
A12	-7,5	37	-5,18	41	-7,26	4
A14	-7,7	29	-6,0	19	-6,90	17
A15	-7,7	30	-5,82	24	-6,57	37
A17	-8,7	11	-5,66	29	-7,00	10
A19	-7,9	20	-6,70	10	-7,24	5
A20	-9,4	6	-7,09	3	-7,02	9

Fonte: Resultados compilados pela autora, com dados de afinidade de ligação obtidos nos programas AutoDock Vina AutoDock 4 e SwissDock.

Considerando os experimentos de *docking* molecular realizados para o ligante de referência, que ocorreram para o arquivo de PDB ID 4HB2, nas simulações no AutoDock Vina, verifica-se que sete, das oito chalconas selecionadas, possuem afinidade de ligação maior ou igual à apresentada pelo K76, que foi -9,3 kcal/mol. Somente o ligante G14 apresenta afinidade de ligação inferior, que é -8,9 kcal/mol. Nas simulações realizadas no AutoDock 4, verifica-se que todas as chalconas selecionadas apresentam afinidade de ligação melhores que a do ligante de referência. No programa SwissDock, os ligantes selecionados G15, D15, A6 e A20 apresentam afinidades de ligação melhores, quando comparadas às descritas pelo ligante de referência. A combinação dos dez melhores resultados de afinidade de ligação obtidos em experimentos de *docking*, de acordo com os dados disponíveis na Tabela 7, pode ser observada na Tabela 8.

Tabela 8: Combinação dos dez melhores resultados de afinidade de ligação obtidos em experimentos de *docking* molecular para a proteína CRM1 (PDB ID 4HB2) e chalconas.

AutoDock Vina		AutoDock 4		SwissDock	
Chalcona	Afinidade de Ligação (kcal/mol)	Chalcona	Afinidade de Ligação (kcal/mol)	Chalcona	Afinidade de Ligação (kcal/mol)
D15	-10,2	G25	-7,34	G5	-7,45
D14	-9,6	D15	-7,21	D6	-7,42
G19	-9,5	A20	-7,09	G2	-7,36
D19	-9,5	D14	-7,06	A12	-7,26
G15	-9,4	D19	-7,05	A19	-7,24
A20	-9,4	A6	-7,02	G15	-7,12
A6	-9,3	G15	-7,01	D15	-7,09
G14	-8,9	D24	-6,88	A8	-7,06
D18	-8,8	G19	-6,75	A20	-7,02
G18	-8,7	A19	-6,70	A17	-7,00

Fonte: Resultados compilados pela autora, com dados de afinidade de ligação obtidos nos programas AutoDock Vina AutoDock 4 e SwissDock.

Considerando os valores obtidos no programa AutoDock 4, AutoDock Vina e SwissDock, de acordo com os dados da Tabela 8, é possível verificar a seguinte ocorrência das chalconas avaliadas nos dez primeiros posicionamentos em relação à afinidade de ligação:

- Ocorrência em todos os programas de *docking* molecular: D15, G15 e A20.
- Ocorrência em dois programas de *docking* molecular: D14, G19, D19 e A6 (verificados nos programas AutoDock e AutoDock Vina).
- Ocorrência em apenas um programa de *docking* molecular: G14, D18 e G18, para o programa AutoDock Vina; D14, G19, D19 e A6, para o programa AutoDock 4; e G14, D18, G18, G25, D24 e A19 para o programa SwissDock.

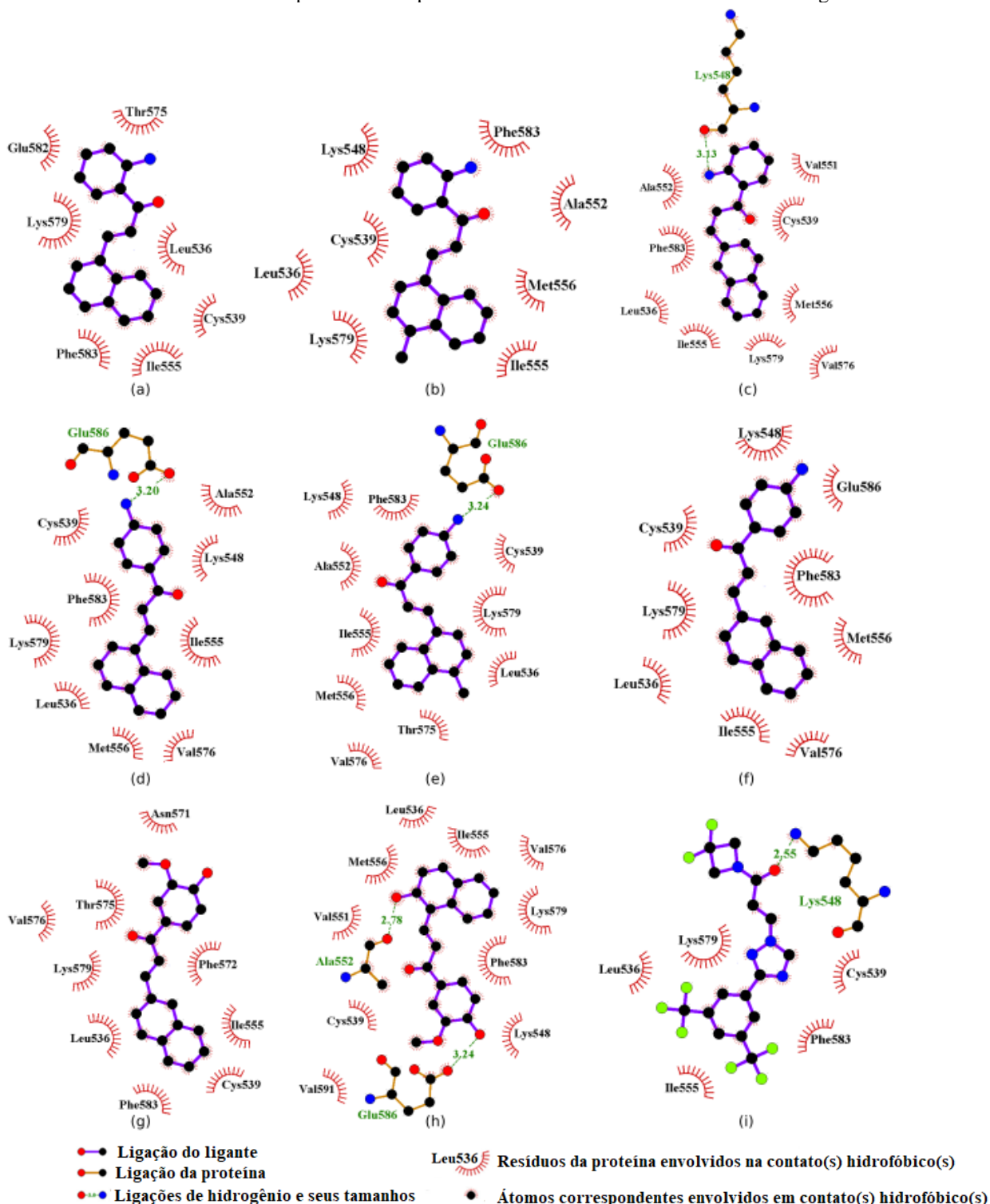
Considerando as informações apresentadas acima e dados verificados para as melhores afinidades de ligação no arquivo de PDB ID 4HAT, foram combinados esses resultados e foi observado que as chalconas G14, G15, G19, D14, D15, D19, A6 e A20 apresentaram afinidades de ligação de interesse, com melhores resultados nas simulações computacionais de *docking* molecular. Como critério de desempate para a seleção das chalconas, foi dado enfoque nos resultados obtidos para o PDB ID 4HAT, uma vez que essa estrutura possui maior resolução, quando comparada à apresentada pelo arquivo de PDB ID 4HB2. Desse modo, como descrito anteriormente, essas oito chalconas foram utilizadas nas simulações de dinâmica computacional e os resultados obtidos para o PDB ID 4HB2 são apresentados a seguir, na subseção 3.1.2.2. Em relação aos resíduos gerados para esse arquivo, serão discriminados apenas os resultados obtidos para as oito chalconas selecionadas e pode-se observar esses dados obtidos no programa LigPlot+ na Tabela 9 e na Figura 16.

Tabela 9: Resíduos formados entre a proteína CRM1 (PDB ID 4HB2) e ligantes G14, G15, G19, D14, D15, D19, A6 e A20. Na última coluna, são apresentadas as porcentagens de ocorrência desses resíduos.

Resíduo	G14	G15	G19	D14	D15	D19	A6	A20	Porcentagem
Ala552		▪	▪	▪	▪			▪	62,5%
Asn571							▪		12,5%
Cys539	▪	▪	▪	▪	▪	▪	▪	▪	100,0%
Glu582	▪								12,5%
Glu586				▪	▪	▪		▪	50,0%
Ile555	▪	▪	▪	▪	▪	▪	▪	▪	100,0%
Leu536	▪	▪	▪	▪	▪	▪	▪	▪	100,0%
Lys548		▪	▪	▪	▪	▪		▪	75,0%
Lys579	▪	▪	▪	▪	▪	▪	▪	▪	100,0%
Met556		▪	▪	▪	▪	▪		▪	75,0%
Phe572							▪		12,5%
Phe583	▪	▪	▪	▪	▪	▪	▪	▪	100,0%
Thr575	▪				▪		▪		37,5%
Val551			▪					▪	25,0%
Val576			▪	▪	▪	▪	▪	▪	75,0%
Val591								▪	12,5%

Fonte: Autora, com dados de resíduos obtidos no programa LigPlot+.

Figura 16: Resíduos de aminoácidos identificados para os ligantes: (a) G14; (b) G15; (c) G19; (d) D14; (e) D15; (f) D19; (g) A6; (h) A20 e (i) K76 utilizando o arquivo de PDB ID 4HB2. Nessa figura, estão escritos em verde os nomes de resíduos da proteína e em preto estão escritos os nomes de resíduos dos ligantes.



Fonte: Autora, com dados de resíduos obtidos no programa LigPlot+.

Em relação ao ligante K76 (Figura 16i), para o PDB ID 4HB2, foram identificados os seguintes resíduos: Cys539, Ile555, Leu536, Lys548, Lys579 e Phe583. Combinados os

resultados, é possível verificar que os resíduos comuns a todas as oito chalconas selecionadas são: Cys539, Ile555, Leu536, Lys579 e Phe583. Esses mesmos resíduos são comuns ao ligante de referência K76.

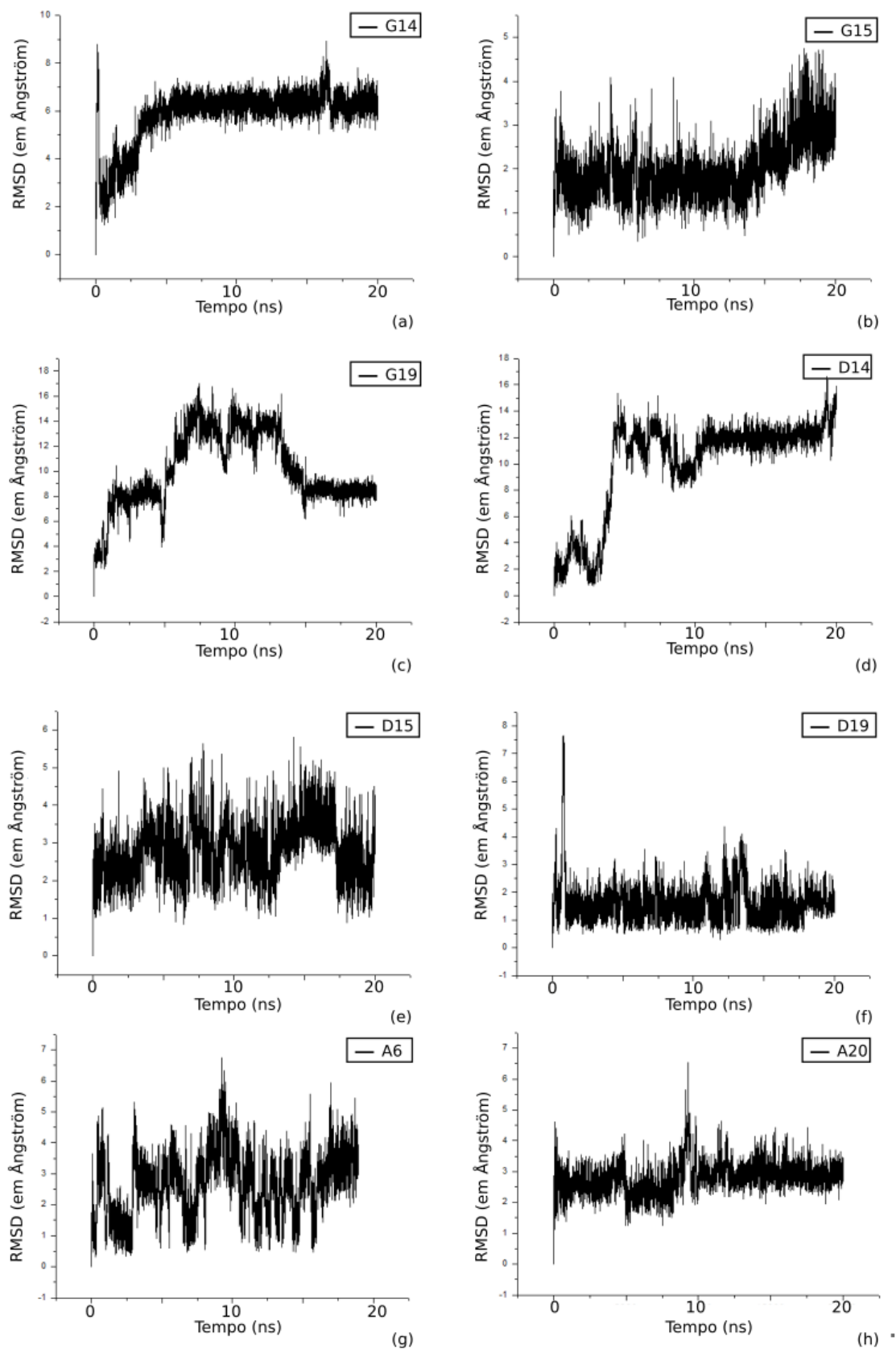
3.1.2.2. Resultados de Dinâmica Molecular

Nas simulações de DM para o arquivo da CRM1 de identificador 4HB2, também foram avaliados os valores de RMSD (Figura 17), número de pontes de hidrogênio (Figura 18) e número de contatos (Figura 18), utilizando as 8 chalconas selecionadas e o ligante K76 (Figura 20) como referência, em trajetórias de 20 ns. Na Figura 17, são apresentados os valores de RMSD obtidos para o arquivo considerado. Verificando essa figura, tratando-se dos valores de RMSD obtidos para o arquivo de identificador 4HB2, é observada uma menor estabilidade na ligação para os ligantes avaliados, se comparado aos resultados obtidos com o arquivo de PDB 4HAT. Observa-se que o ligante G14 (Figura 17a) apresenta uma estabilidade na ligação a partir de 3ns, que se mantém até o final da simulação. Observa-se também que o ligante D14 (Figura 17d) apresenta estabilidade a partir dos 10 ns de trajetória. Para os ligantes D19 (Figura 17f) e A20 (Figura 17h), observa-se uma menor alteração na conformação, se comparada às trajetórias descritas pelos demais ligantes, que são G15 (Figura 17b), G19 (Figura 17c), D15 (Figura 17e) e A6 (Figura 17g).

Considerando o arquivo de PDB de identificador 4HB2, cujos resultados são apresentados na Figura 18, as chalconas que apresentaram 0 ± 1 pontes de hidrogênio foram G14 (Figura 18a), D19 (Figura 18f) e A20 (Figura 18h).

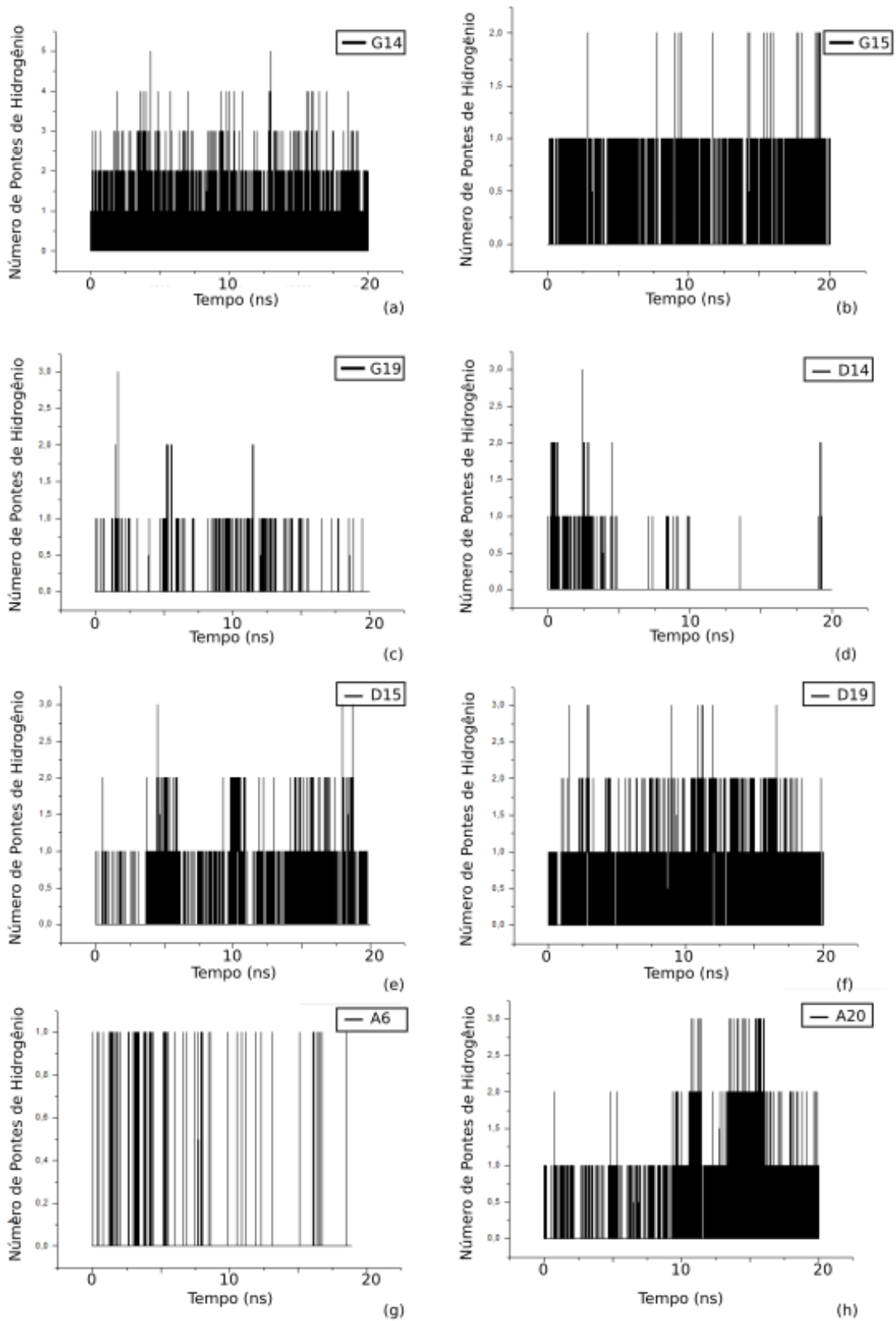
Para o arquivo de PDB ID 4HB2, cujos resultados são apresentados na Figura 19, os ligantes que apresentaram maiores números de contatos foram G14 e G15. Para a G14 (Figura 19a), foram obtidos 1915 ± 230 contatos e para a chalcona G15 (Figura 19b), foram observados 1865 ± 136 contatos. Em relação aos menores números de contatos, esses valores foram obtidos para os ligantes D14 (Figura 19d) e A6 (Figura 19g), para os quais foram obtidos 1367 ± 194 e 1587 ± 159 contatos, respectivamente. Os valores obtidos para os demais ligantes, que são G19 (Figura 19c), D15 (Figura 19e), D19 (Figura 19f) e A20 (Figura 19h), foram, respectivamente, 1611 ± 357 , 1698 ± 159 , 1715 ± 136 e 1760 ± 146 .

Figura 17: Resultados de simulação de dinâmica computacional para RMSD utilizando a proteína CRM1 (PDB ID 4HB2) com os ligantes: (a) G14; (b) G15; (c) G19; (d) D14; (e) D15; (f) D19; (g) A6; e (h) A20.



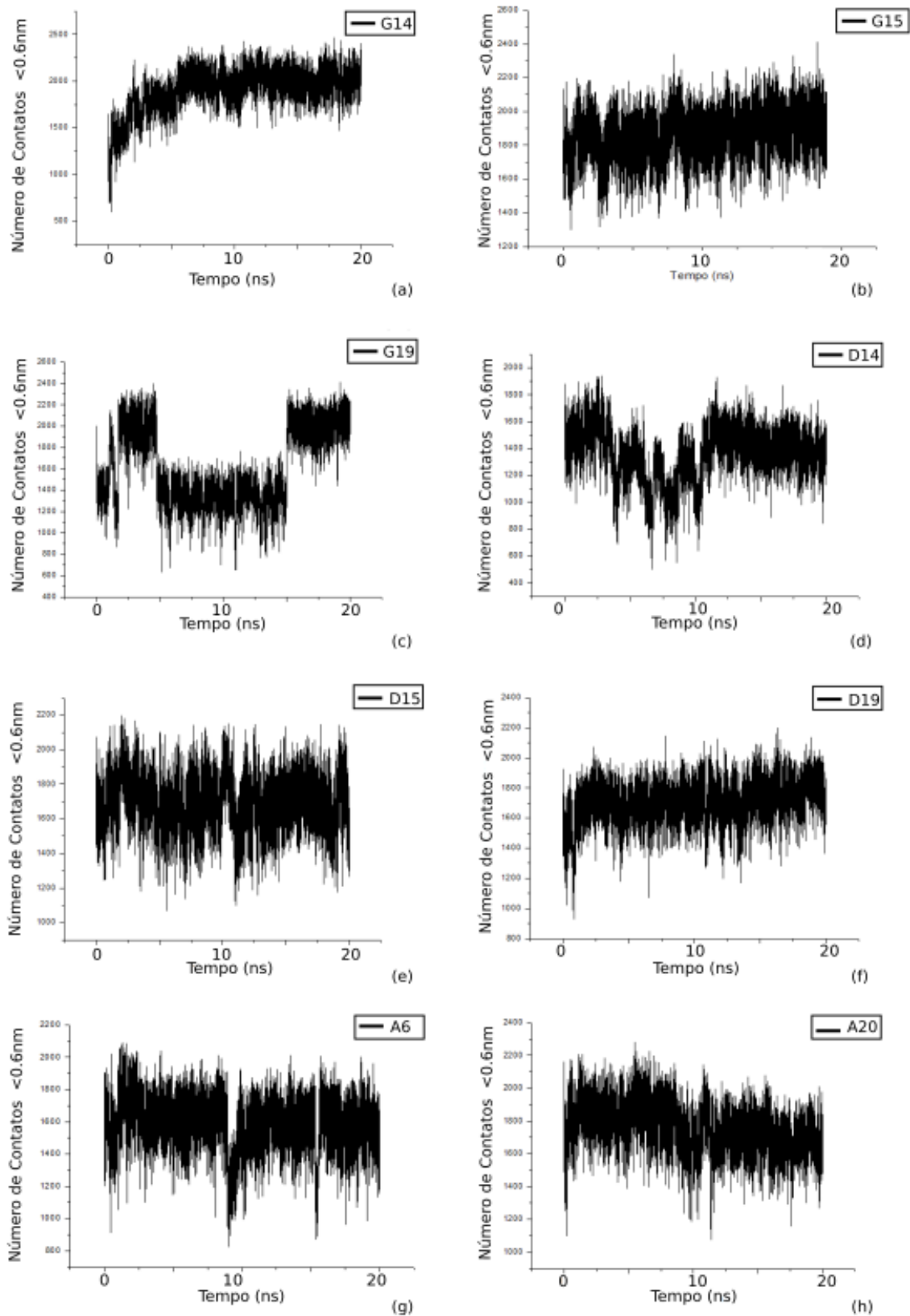
Fonte: Autora, com resultados de RMSD obtidos no programa GROMACS.

Figura 18: Resultados de simulação de dinâmica computacional para o número de pontes de hidrogênio da proteína CRM1 (PDB id.: 4HB2) com os ligantes: (a) G14; (b) G15; (c) G19; (d) D14; (e) D15; (f) D19; (g) A6; (h) A20.



Fonte: Autora, com resultados de RMSD obtidos no programa GROMACS.

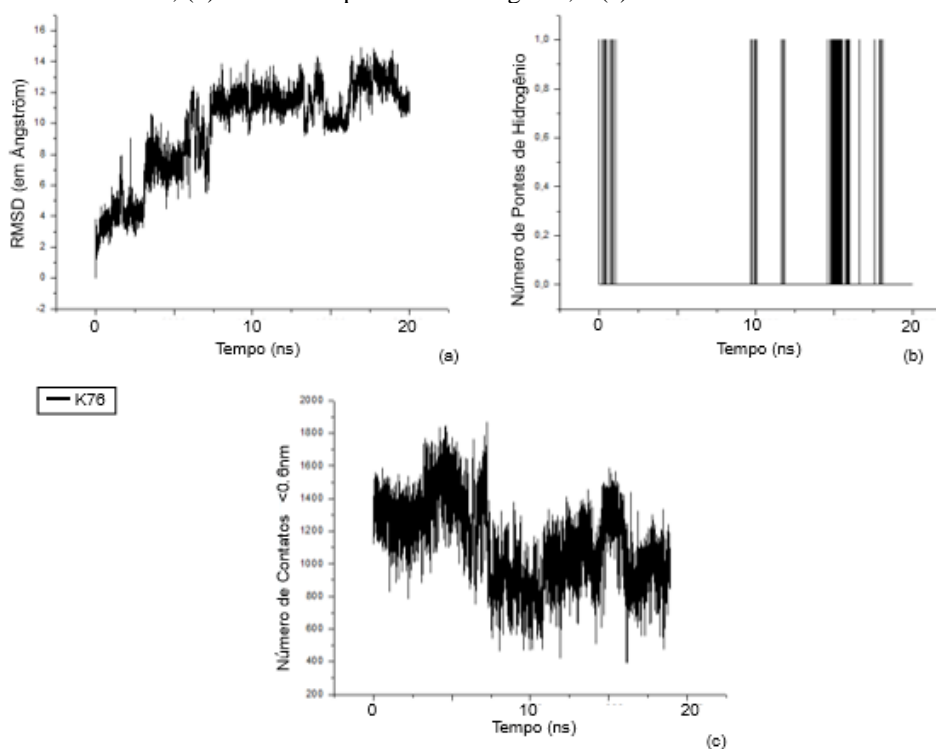
Figura 19: Resultados de simulação de dinâmica computacional para número de contatos utilizando a proteína CRM1 (PDB ID 4HB2) com os ligantes: (a) G14; (b) G15; (c) G19; (d) D14; (e) D15; (f) D19; (g) A6; e (h) A20.



Fonte: Autora, com resultados de RMSD obtidos no programa GROMACS.

Os resultados obtidos para o ligante de referência K76, utilizando o arquivo de identificador 4HB2, podem ser observados na Figura 20. Para esse ligante, foram calculados valores de RMSD (Figura 20a), pontes de hidrogênio (Figura 20b) e número de contatos (Figura 20c).

Figura 20: resultados de dinâmica computacional utilizando o ligante de referência K76 nas simulações de: (a) RMSD; (b) número de pontes de hidrogênio; e (c) número de contatos.



Fonte: Autora, com resultados de RMSD obtidos no programa GROMACS.

Em uma avaliação dos cálculos de dinâmica molecular para o ligante de controle K76, considerando valores de RMSD, observa-se na Figura 20a indicativos de uma interação estável, uma vez que os valores de RMSD são praticamente constantes ao longo de toda a trajetória, após 7ns. Porém, observa-se que há uma instabilidade maior na ligação formada para esse ligante se comparada à apresentada nos resultados de RMSD obtidos por outras chalconas. Note que a estabilidade desse ligante não é comparável, por exemplo, à apresentada por chalconas como a G14 e a D14. Ainda, para esse mesmo ligante, foi obtido um resultado igual a 0 para o número de pontes de hidrogênio, que é comparável aos resultados dados pelos ligantes G15, G19, D14, D15 e A6 e inferior ao apresentado pelos ligantes G14, D19 e A20, que foi 0 ± 1 . Considerando os dados presentes na Figura 20c, observa-se que o ligante utilizado como referência, K76, apresentou 1133 ± 241 contatos, que é o menor valor quando comparado ao calculado para as oito chalconas selecionadas. Desse modo, o K76 não apresentou resultados

superiores quando comparado às chalconas selecionadas nesse trabalho. Nesse sentido, é possível verificar que o uso de chalconas, se comparado ao uso do ligante de referência, pode gerar melhores valores de RMSD, pontes de hidrogênio e número de contatos.

3.1.3. Considerações sobre os Resultados Computacionais Obtidos para a proteína CRM1

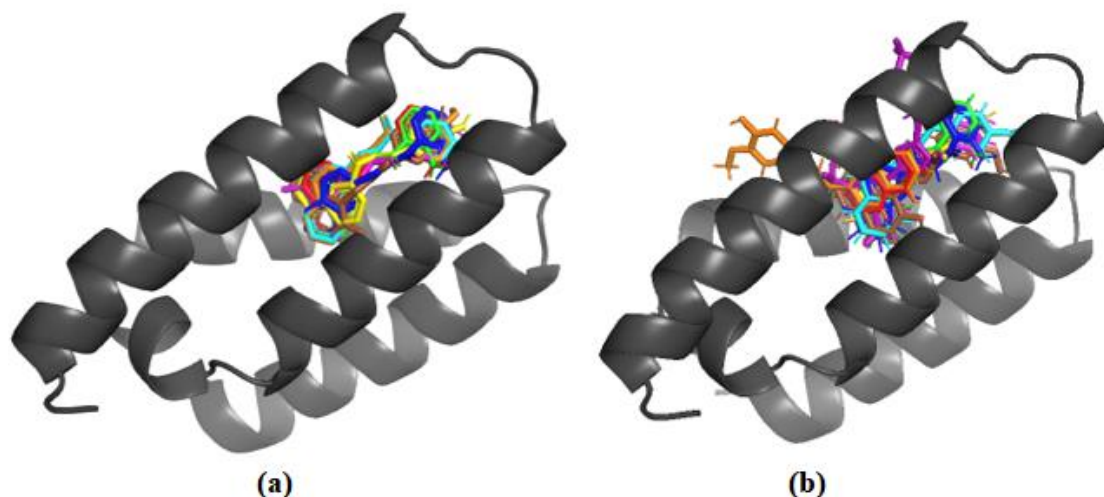
Nas simulações de *docking* molecular, considerando os dez ligantes que apresentaram maior afinidade de ligação, somente as chalconas G15, D15 e A20 foram comuns na ordenação de acordo com os programas AutoDock, AutoDock Vina e SwissDock, nos dois arquivos de PDB considerados. A seleção dos outros cinco ligantes (G14, D14, D19, A6 e A20) ocorreu por essas chalconas estarem entre os melhores resultados, em grande parte das simulações de *docking* molecular. Os resultados obtidos ao utilizar o programa DockThor foram desconsiderados nessa análise, uma vez que as afinidades de ligação obtidas nesse programa não foram condizentes com as verificadas para os demais programas.

Para o PDB ID 4HAT, a Leptomicina B não foi utilizada como ligante de referência nas simulações de DM, uma vez que essa molécula não possui estrutura molecular similar às chalconas avaliadas, bem como por não apresentar um *redocking* com resultados realísticos. Nas simulações de *docking* molecular e de dinâmica computacional para a CRM1 com o arquivo de PDB ID 4HB2, utilizou-se o ligante K76 como referência, devido a sua similaridade estrutural com as chalconas avaliadas e pelos resultados condizentes no *redocking*.

Na Figura 21, pode-se observar a localização dessas chalconas selecionadas, em que é possível verificar a conformação dessas moléculas na estrutura da proteína CRM1. Para gerar essas poses, foram selecionados os melhores resultados de afinidade de ligação, em simulações de *docking* molecular nos programas AutoDock Vina, AutoDock 4 e SwissDock. Na Figura 21a, são apresentados os resultados para o arquivo de PDB ID 4HAT. Os mesmos resultados podem ser verificados na Figura 21b, para o arquivo de PDB ID 4HB2, contendo também a estrutura em cristal do ligante de referência K76, que é apresentada em roxo. Observe que essa localização é similar a grande parte dos ligantes considerados, exceto para o ligante A6 na

segunda simulação (Figura 21b).

Figura 21: Nas Figuras (a) e (b), a estrutura da proteína CRM1 utilizada nas simulações computacionais é apresentada em cinza e sua representação está em *cartoon*. Nessas figuras, as chalconas são apresentadas utilizando o modelo *stick* e a cada uma delas é associada uma cor, sendo: i) G14: vermelho; ii) G15: verde; iii) G19: azul; iv) D14: amarelo; v) D15: magenta; vi) D19: ciano; vii) A6: laranja; e viii) A20: marrom. Em (a), apresenta-se a estrutura para a CRM1 de identificador PDB ID 4HAT. Em (b) apresenta-se a proteína de identificador PDB ID 4HB2, que contém também o ligante de referência K76, que é apresentado em roxo.



Fonte: Autora, com resultados de RMSD obtidos no programa AutoDock Vina, AutoDock 4 e SwissDock. A representação estrutural foi preparada utilizando o programa PyMOL (DELANO, 2002).

Considerando os dados das estruturas que são apresentados na Tabela 1 e na Figura 8, dentre as oito chalconas selecionadas, é possível observar a ocorrência do grupo fenil no anel B dessas moléculas. Nas chalconas G14, G15, D14, D15 e A20, esse grupo encontra-se nas posições R₂ e R₃ do anel B. Para os ligantes G19, D19 e A6, o fenil encontra-se nas posições R₃ e R₄ desse mesmo anel. Compostos que contêm em sua estrutura o grupo fenil e que são aplicados no tratamento do câncer são descritos na literatura. Como exemplo, no trabalho de Abid *et al.* (2012), foram analisados análogos de isocumarina e foi verificado que o composto 3-(3',4'-difluorophenyl) isocoumarin 1 foi mais potente para exercer atividade antimetastática, quando comparado ao controle. Em outro trabalho, realizado por Tseng *et al.* (2015), em que foram incluídas diferentes linhagens de células cancerígenas de pulmão e de mama, foi verificado que o composto (E)-6-methoxy-3-(4-methoxyphenyl)-2-[2-(5-nitrofuranyl)vinyl]quinoline 22 apresentou baixa toxicidade em todas as linhagens celulares normais e cancerígenas avaliadas e que possuiu capacidade de inibir a migração e invasão celular.

Ainda em relação ao anel B, nos ensaios de *docking* molecular, foi possível observar que moléculas que apresentam Fluor (F), Cloro (Cl), Bromo (Br), Metoxilo (OCH₃), Fluorcarbono (C-F₃), Thiophene, Furan, Nitrila (CN) e Dióxido de Nitrogênio (NO₂) em sua estrutura, não estão presentes entre as oito chalconas selecionadas. Essas restrições não foram

identificadas para o anel A, uma vez que as variações que apresentam NH₂, Grupo Metoxi (OCH₃) e Hidroxila (OH) em sua estrutura, são verificadas nas oito chalconas de interesse. Em estudos realizados por Liew *et al.* (2020), foi descrito que o grupo metoxi (OCH₃) contribuiu de modo potencial para efeitos antimetastáticos, uma vez que sua ocorrência nos compostos avaliados acarretou em um aumento em sua atividade. Nesse mesmo estudo, também foi verificado um aumento na atividade de compostos ao utilizar o NH₂. No trabalho de Jiang *et al.* (2011), em que foram avaliados análogos de brartemicina, que é um inibidor de invasão de células tumorais à base de trealose, foi verificado que a atividade anti-invasiva foi moderadamente ativa quando o grupo hidroxila (OH) foi substituído na posição 2 ou 4 do anel de ácido benzóico.

Considerando as chalconas selecionadas, os ligantes G14 e G15 diferem em sua conformação apenas no elemento presente no anel B, na posição R₄, sendo que o G14 apresenta H e o G15 apresenta CH₃ nessa posição. Nas simulações de *docking* molecular com o PDB ID 4HAT, considerando os três programas adotados, não foi possível apontar qual desses ligantes apresentou melhor afinidade de ligação. Para o PDB ID 4HB2, o ligante G15 apresentou melhores afinidades, nos três programas. Nesse sentido, não é possível concluir, computacionalmente, se o uso do H ou do CH₃ na posição R₄ do anel B gera melhores compostos.

De modo similar, os ligantes G14 e G19 divergem em sua conformação apenas na posição do fenil, sendo que o G14 apresenta o fenil em R₂ e R₃ e o G19 em R₃ e R₄. Comparando esses ligantes, nas simulações de *docking* para o PDB ID 4HAT, o ligante G14 apresentou melhores afinidades, nos três programas. Em relação ao PDB ID 4HB2, houve divergência em relação a esse dado. Desse modo, não é possível inferir, computacionalmente, se a posição do fenil, nas posições avaliadas, gera melhores compostos com melhores afinidades de ligação.

Os ligantes D14 e D15 diferem em sua conformação apenas no elemento presente no anel B, na posição R₄, sendo que o G14 apresenta H e o G15 apresenta CH₃ nessa posição. Considerando os programas AutoDock, AutoDock Vina e SwissDock, nos dois arquivos de PDB utilizados, o ligante D15 apresenta melhores afinidades de ligação. Nesse sentido, o uso do CH₃ na posição R₄ do anel B melhora os resultados computacionais de *docking* molecular.

De forma análoga, os ligantes D14 e D19 divergem em sua conformação apenas na posição do fenil, sendo que o G14 apresenta o fenil em R₂ e R₃ e o G19 em R₃ e R₄. Para os dois arquivos, não foi encontrado um ligante cuja afinidade de ligação sempre convergisse para os melhores resultados, considerando os três programas empregados.

Os ligantes G14 e D14 diferem apenas na disposição do NH₂, no anel A, uma vez que G14 apresenta na posição R'₂ e D14 em R'₄. Considerando esses ligantes, nos arquivos de PDB ID 4HAT e 4HB2, não foi encontrado um ligante cuja afinidade de ligação sempre foi superior, nos três programas empregados.

Os ligantes G15 e D15 diferem apenas na disposição do NH₂, no anel A, uma vez que G14 apresenta na posição R'₂ e D14 em R'₄. Comparando esses ligantes, nas simulações de *docking* molecular para o PDB ID 4HAT, o ligante D15 apresentou melhores afinidades, nos programas AutoDock, AutoDock Vina e SwissDock. Em relação ao PDB ID 4HB2, houve divergência em relação a essa melhor afinidade. Desse modo, não é possível inferir, computacionalmente, se alterar a posição do NH₂ pode gerar melhores compostos, considerando as conformações avaliadas.

Os ligantes G19 e D19 diferem apenas na disposição do NH₂, no anel A, uma vez que G19 apresenta na posição R'₂ e D19 em R'₄. Considerando esses ligantes, para os arquivos de PDB ID 4HAT e 4HB2, não foi encontrado um ligante cuja afinidade de ligação sempre convergisse para os melhores resultados, considerando os três programas.

Nesse sentido, somente as alterações realizadas para os ligantes D14 e D15, que diferem em sua conformação apenas no elemento presente no anel B, na posição R₄, sendo H e CH₃, respectivamente, apresentaram melhorias nos resultados de afinidade de ligação. A substituição do H por NH₂ gerou melhores resultados de *docking* molecular, nas posições consideradas.

Em relação aos resultados de dinâmica molecular, que estão descritos nas seções 3.1.1.2 e 3.1.2.2, na Tabela 10 é verificada a compilação desses resultados. Essa tabela contém dados de RMSD, número de contatos e número de pontes de hidrogênio para experimentos com chalconas e ligante K76.

Tabela 10: Resultados de RMSD, número de contatos e pontes de hidrogênio obtidos para os arquivos 4HAT e 4HB2.

Ligante	RMSD		Número de Contatos		Pontes de Hidrogênio
	PDB: 4HAT	PDB: 4HB2	PDB: 4HAT	PDB: 4HB2	PDB: 4HB2
G14	6,03 ± 0,94	5,86 ± 1,14	1889 ± 263	1915 ± 230	0 ± 1
G15	7,13 ± 3,91	1,96 ± 0,61	1618 ± 301	1865 ± 136	0
G19	10,04 ± 2,92	10,04 ± 2,92	1738 ± 307	1611 ± 357	0
D14	8,24 ± 2,70	10,02 ± 3,70	1785 ± 277	1367 ± 194	0
D15	4,94 ± 3,54	2,78 ± 0,73	1827 ± 175	1698 ± 159	0
D19	8,71 ± 3,20	1,57 ± 0,69	1542 ± 180	1715 ± 136	0 ± 1
A6	8,16 ± 0,76	2,65 ± 1,01	1880 ± 147	1587 ± 159	0

A20	4,11 ± 1,01	2,78 ± 0,48	1729 ± 168	1760 ± 146	0 ± 1
K76	-	9,69 ± 3,03	-	1133 ± 241	0

Fonte: Autora, com resultados de RMSD obtidos no programa

Considerando os resultados de valores de RMSD, os ligantes que apresentaram maior estabilidade na ligação, por descreverem trajetórias com menores oscilações, considerando os arquivos de PDB ID 4HAT e 4HB2, foram G14 e D14. Em relação ao número de pontes de hidrogênio, 0 ± 1 pontes foram identificadas para os ligantes G14, D19 e A20. Para as demais chalconas e para o ligante de referência, a média encontrada foi zero. Considerando valores de números de contatos para o arquivo de PDB ID 4HB2, todas as 8 chalconas apresentadas melhores resultados, quando comparadas ao ligante de referência. Nesse sentido, é possível verificar que o uso de chalconas, se comparado ao uso do ligante de referência, pode gerar melhores valores de RMSD, pontes de hidrogênio e número de contatos. Porém, os resultados obtidos computacionalmente não indicam qual é a melhor chalcona, dentre as oito selecionadas neste trabalho.

3.1.4. Resultados Computacionais para a Proteína p53

Na seção 3.1.4.1, são apresentados os resultados de *docking* molecular obtidos ao utilizar a proteína p53, empregando o arquivo de PDB ID 2PCX. Na seção 3.1.4.2, são apresentados os resultados de dinâmica molecular, utilizando essa mesma estrutura. Para as simulações de dinâmica molecular apresentados nesta subseção, utilizou-se o ligante Rita como referência.

3.1.4.1. Resultados de *Docking* Molecular

Para as simulações de *docking* molecular, foram utilizados os programas AutoDock Vina e AutoDock 4 e os resultados de afinidade de ligação obtidos para todas as chalconas consideradas, utilizando a proteína p53, podem ser observados na Tabela 11. Os dez melhores resultados de afinidade de ligação encontram-se em sublinhado.

Tabela 11: Resultados de afinidade de ligação, em kcal/mol, utilizando os programas AutoDock Vina, e AutoDock 4.

Chalcona	AutoDock Vina		AutoDock 4	
	Afinidade de Ligação (kcal/mol)	Posição	Afinidade de Ligação (kcal/mol)	Posição
T10	-6,0	30	-4,62	27
T33	-6,2	25	-4,42	34
T37	-5,9	34	-4,55	30

T38	-5,9	35	-4,89	18
G2	-6,0	31	-4,39	36
G3	-5,9	36	-4,16	41
G4	-5,9	37	-4,47	33
G5	-5,9	38	-4,62	28
G6	-5,9	39	-4,13	43
G14	<u>-6,7</u>	8	-5,24	12
G15	<u>-6,7</u>	9	<u>-5,35</u>	9
G18	-6,5	13	-3,94	45
G19	<u>-6,8</u>	7	<u>-5,33</u>	10
G20	-5,2	45	-4,14	42
G21	-5,3	44	-4,77	23
G23	-6,1	27	-4,80	21
G24	-6,6	12	-4,97	15
G25	-6,5	14	<u>-6,12</u>	1
D1	-5,7	41	-4,79	22
D2	-6,1	28	-4,95	16
D3	-6,0	32	-4,66	26
D4	-5,9	40	-5,05	14
D6	-6,0	33	-4,71	24
D14	<u>-6,9</u>	4	<u>-5,45</u>	8
D15	<u>-7,1</u>	3	<u>-5,66</u>	5
D18	-6,4	19	-4,52	32
D19	<u>-6,9</u>	5	-5,92	2
D20	-5,4	42	-4,62	29
D21	-5,4	43	-5,30	11
D23	-6,4	20	-4,90	17
D24	-6,4	21	<u>-5,58</u>	6
D25	-6,5	15	<u>-5,85</u>	3
A2	-6,4	22	-4,41	35
A5	-6,1	29	-4,29	38
A6	<u>-7,2</u>	2	-5,24	13
A8	-6,5	16	<u>-5,54</u>	7
A9	<u>-6,7</u>	10	-4,53	31
A10	-6,5	17	-4,11	44
A11	-6,2	26	-4,84	20
A12	-6,3	24	-4,87	19
A14	-6,4	23	-4,67	25
A15	-6,5	18	-4,38	37
A17	<u>-6,9</u>	6	-4,24	40
A19	-6,7	11	-4,29	39
A20	<u>-7,3</u>	1	<u>-5,75</u>	4

Fonte: Resultados compilados pela autora, com dados de afinidade de ligação obtidos nos programas

AutoDock Vina e AutoDock 4.

Considerando os resultados apresentados na Tabela 11, é possível verificar que os ligantes G15, G19, D14, D15, D19 e A20 posicionam-se entre as dez melhores afinidade de ligação, considerando os programas AutoDock e AutoDock Vina. Os ligantes G14, A6, A17 e

A9 apresentam-se entre os dez ligantes com maior afinidade de ligação somente no programa AutoDock Vina e os ligantes G25, D25, D24 e A8 somente no programa AutoDock.

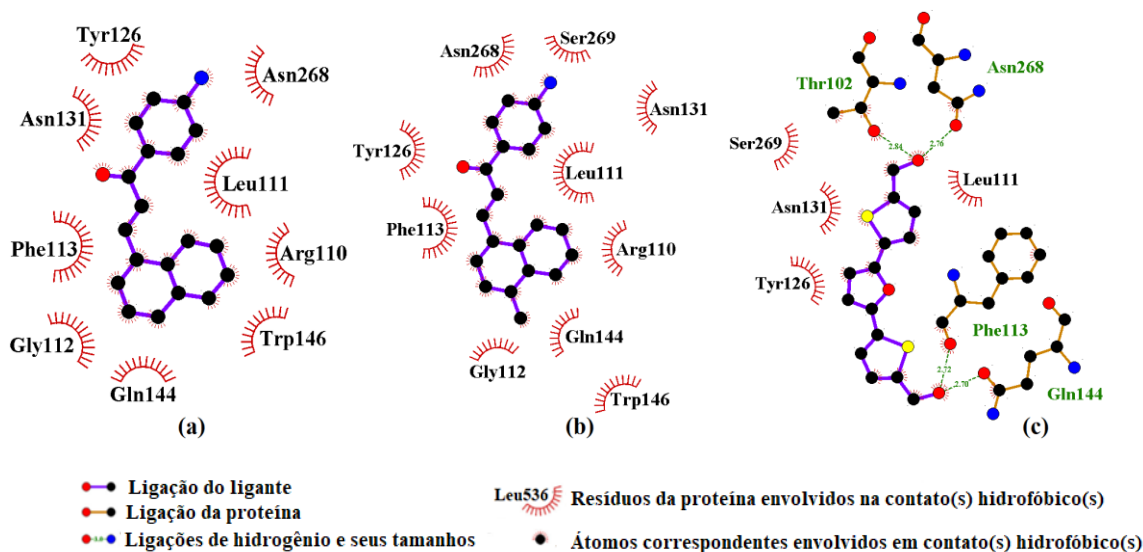
Considerando os oito ligantes que apresentaram os melhores resultados de afinidade de ligação para a proteína CRM1 e que foram selecionados para a simulação de dinâmica molecular com essa proteína (G14, G15, D14, D15, D19, A6, A20 e G19), verifica-se que os ligantes G15, D14, D15, D19, G19 e A20 também configuram como os que possuem as dez melhores afinidades de ligação para a proteína p53, considerando os dados de *docking* molecular com os programas AutoDock e AutoDock Vina. Já os ligantes G14 e A6 figuram entre os dez compostos com melhores afinidades de ligação apenas no programa AutoDock Vina. Desse modo, todos os ligantes selecionados para as simulações de dinâmica molecular com a proteína CRM1 também foram encontrados como os que possuem as dez melhores afinidades de ligação, considerando ao menos um dos programas de *docking* molecular utilizado, indicando que as soluções são convergentes para a CRM1 e para a p53.

Para realizar as demais simulações, nos cálculos de resíduos formados e de dinâmica molecular com a p53, foram utilizadas apenas as chalconas D14 e D15. Essas chalconas foram selecionadas por apresentarem um efeito inibitório na migração e invasão das linhas celulares de osteossarcoma U2OS (p53 +/+) e SAOS-2 (p53 -/-) (SEBA *et al.*, 2018). Como ligante de referência, utilizou-se o composto Rita, que é descrito na literatura como inibidor da atividade da proteína p53. Porém, a estrutura desse ligante para as simulações computacionais foi obtida através do programa Gaussian, uma vez que sua estrutura cristalina não é disponibilizada no PDB ID 2PCX e, por esse motivo, não foi realizado o *redocking*.

Com o programa LigPlot+, foram criados diagramas dos resíduos gerados entre ligante e proteína, de acordo com convenções químicas. Os resultados são apresentados na Figura 22, para os ligantes D14, D15 e Rita. Essas conformações também foram as poses iniciais para as simulações de dinâmica molecular. A pose do composto Rita e das chalconas D14 e D15 utilizadas para cálculo de resíduo foram definidas como a saída gerada pelo programa AutoDock Vina.

Os resíduos identificados para esses ligantes foram: i) D14: Arg110, Asn131, Asn268, Gln144, Gly112, Leu111, Phe113, Trp146 e Tyr126; ii) D15: Arg110, Asn131, Asn268, Gln144, Gly112, Leu111, Phe113, Ser269 e Trp146; e iii) Rita: Asn131, Asn268, Gln144, Leu111, Phe113, Phe113, Ser269, Thr102 e Tyr126.

Figura 22: Resíduos formados entre CRM1 e ligantes: (a) D14; (b) D15; e (c) Rita.



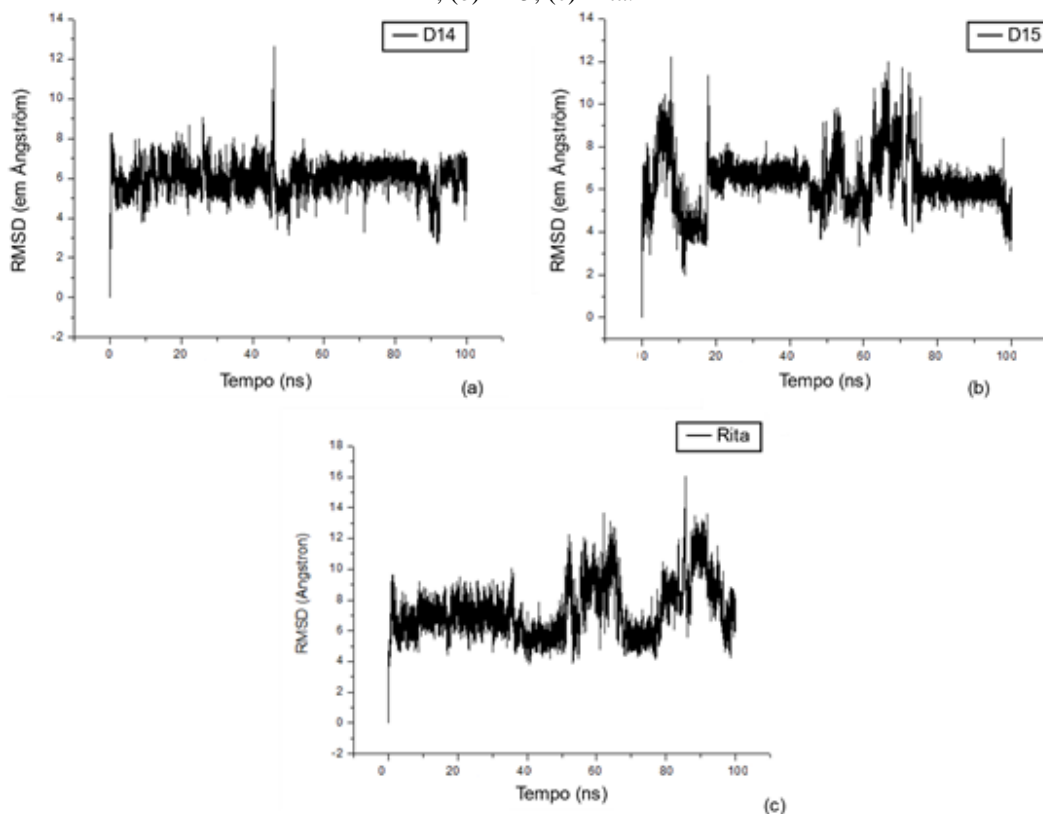
Fonte: Autora, com dados de resíduos obtidos no programa LigPlot+.

3.1.4.2. Resultados de Dinâmica Molecular

Conforme descrito anteriormente, as simulações de dinâmica molecular para a proteína p53 foram realizadas somente para os ligantes D14 e D15. Os resultados obtidos nas simulações de DM, com trajetórias de 100 ns, com cálculos de RMSD, número de contatos e pontes de hidrogênio, podem ser observados nas Figuras 23, 24 e 25, respectivamente.

Na Figura 23, podem ser observados, em Ångström, os valores de RMSD de todos os átomos que não são hidrogênio para os ligantes avaliados, de modo que valores para a D14 constam na Figura 23a, para a D15 na Figura 23b e para o Rita na Figura 23c. Com esses dados de RMSD, é possível verificar uma maior estabilidade na trajetória descrita pelo ligante D14 (RMSD: 6 ± 3), quando comparada à apresentada pelos ligantes D15 (RMSD: 6 ± 3) e Rita (RMSD: 7 ± 3). Para o D14, é possível identificar uma mudança na trajetória (~50 ns), que é seguida, em um breve espaço de tempo, de um retorno à sua posição anterior. Em relação ao ligante D15, é possível identificar alterações em sua posição nos intervalos entre 10ns a 20ns, 40ns a 80ns e após os 90 ns. Já para o ligante de referência Rita, variações na posição podem ser verificadas após 50 ns, até o final da trajetória. Porém, em relação à média de RMSD, esse ligante (Rita) é o que apresenta maior valor nas simulações computacionais.

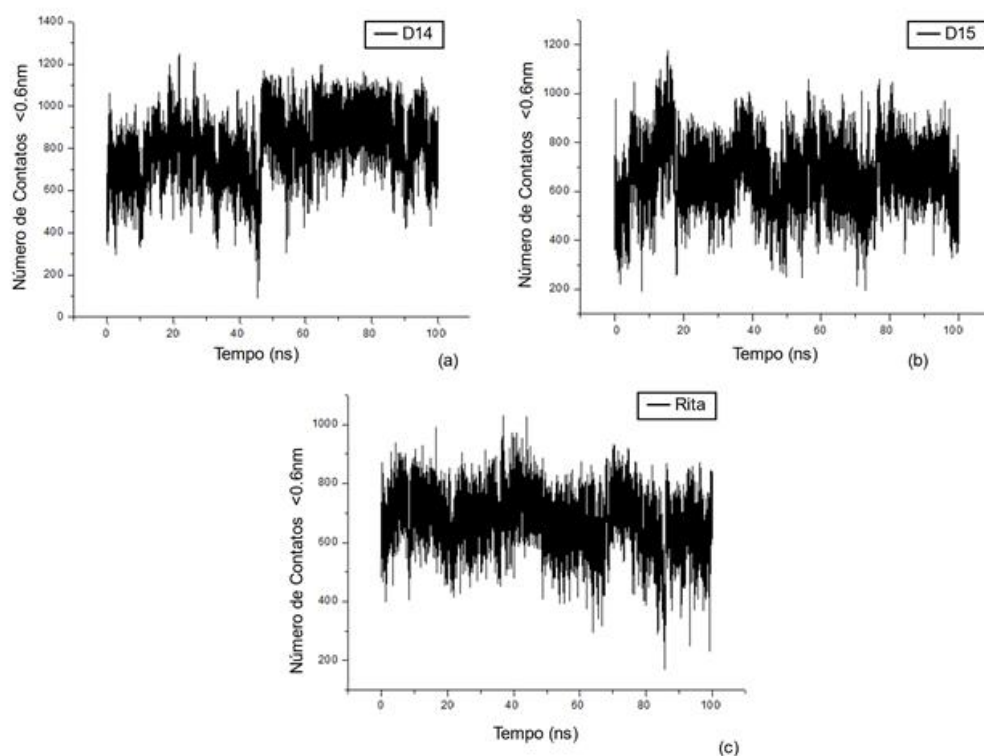
Figura 23: Resultados de simulação de dinâmica computacional para RMSD da proteína p53 com os ligantes: (a) D14; (b) D15; (c) Rita.



Fonte: Autora, com resultados de RMSD obtidos no programa GROMACS.

A seguir, na Figura 24, podem ser verificados os resultados referentes ao número de contatos (<0.6 nm) formados entre p53 e os ligantes D14 (Figura 24a), D15 (Figura 24b) e Rita (Figura 24c), em simulações de 100ns. Considerando os dados dessa tabela, o ligante D14 é o que apresenta um maior número de contatos (813 ± 136), seguido do ligante Rita (678 ± 85) e D15 (677 ± 117). Os resultados para número de contatos obtidos para o ligante de referência e D15 foram competitivos nessa simulação.

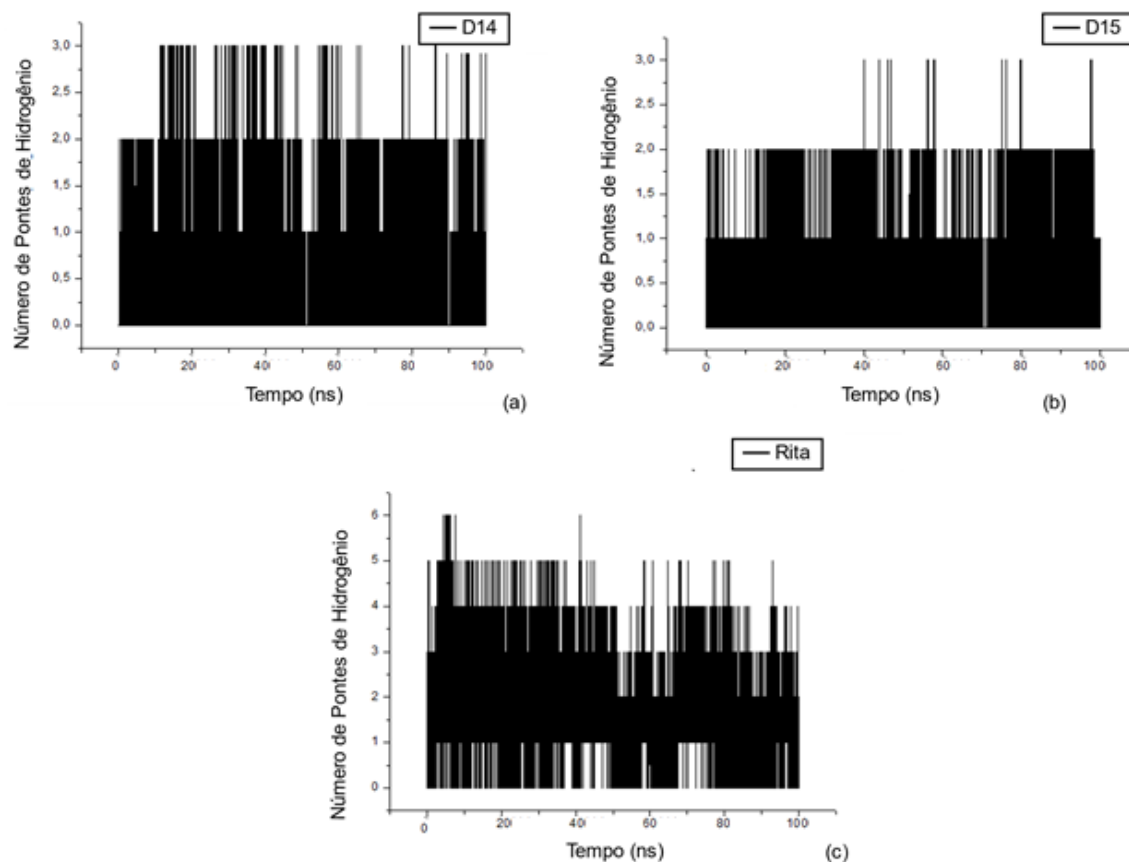
Figura 24: Resultados de simulação de dinâmica computacional para número de contatos da proteína p53 com os ligantes: (a) D14; (b) D15; (c) Rita.



Fonte: Autora, com resultados de número de contatos obtidos no programa GROMACS.

Dados referentes ao número de pontes de hidrogênio formadas entre p53 e ligantes D14 (Figura 25a), D15 (Figura 19b) e Rita (Figura 19c) podem ser verificados na Figura 19. O número médio de pontes de hidrogênio verificado para os ligantes D14 e D15 é 1 ± 1 e para o ligante de referência é 2 ± 1 . Desse modo, em relação ao número de pontes de hidrogênio, o ligante Rita apresenta a maior média.

Figura 25: Resultados de simulação de dinâmica molecular para número de pontes de hidrogênio da proteína p53 com os ligantes: (a) D14; (b) D15; (c) Rita.



Fonte: Autora, com resultados de pontes de hidrogênio obtidos no programa GROMACS.

Sumarizando os dados obtidos para RMSD de todos os átomos que não são hidrogênio, número de pontes de hidrogênio e número de contatos (<0.6 nm), com os ligantes avaliados e p53, os resultados obtidos podem ser verificados na Tabela 12.

Tabela 12: Resultados de RMSD, Número de Contatos e Número de Pontes de Hidrogênio para a proteína p53.

Ligante	RMSD	Pontes de Hidrogênio	Número de Contatos
D14	6 ± 3	1 ± 1	813 ± 136
D15	6 ± 3	1 ± 1	677 ± 117
Rita	7 ± 3	2 ± 1	678 ± 85

Fonte: Autora, com resultados obtidos no programa GROMACS.

Conforme descrito anteriormente e que pode ser verificado na Tabela 12, o ligante Rita apresenta maiores médias de RMSD e pontes de hidrogênio. Já o ligante D14 apresenta maior média em relação ao número de contatos.

3.1.5. Resultados Computacionais para a Proteína MDM2

Nesta subseção, são apresentados os resultados obtidos para a proteína MDM2. Na subseção 3.1.5.1, são apresentados os resultados de *docking* molecular obtidos para essa proteína, empregando o PDB ID 4HG7. Na subseção 3.1.5.2, são apresentados os resultados de dinâmica molecular, empregando essa mesma estrutura. Para as simulações de *docking* e dinâmica molecular, utilizou-se o Nutlin-3a como ligante de referência, cuja conformação é disponibilizada no PDB de mesmo identificador.

3.1.5.1. Resultados de *Docking* Molecular

Os resultados obtidos para as simulações de *docking* molecular são apresentados na Tabela 13. As dez melhores afinidades de ligação obtidas nos programas AutoDock Vina e AutoDock encontram-se sublinhadas nesta tabela e a afinidade obtida para o ligante de referência Nutlin-3a também é apresentada.

Tabela 13: Resultados de docking molecular para a proteína MDM2 e chalconas.

Chalcona	AutoDock Vina		AutoDock	
	Afinidade de Ligação (kcal/mol)	Posição	Afinidade de Ligação (kcal/mol)	Posição
Nutlin-3a	-6,40	-	-6,19 kcal/mol	-
T10	-6,7	24	-4,75	27
T33	-6,6	30	-4,86	23
T37	-6,6	31	-5,11	15
T38	-6,8	20	-5,36	11
G2	-6,6	32	-4,69	28
G3	-6,6	33	-4,34	39
G4	-6,6	34	-4,80	26
G5	-6,6	35	-4,97	18
G6	-6,6	36	-4,46	36
G14	<u>-7,7</u>	2	-4,91	20
G15	<u>-7,7</u>	3	-5,20	13
G18	-7,1	11	-4,46	37
G19	<u>-7,4</u>	6	-5,43	7
G20	-6,0	45	-4,19	43
G21	-6,1	42	-4,01	45
G23	-6,8	21	-4,91	21
G24	-6,7	25	-7,32	1
G25	-7,1	12	-7,30	3
D1	-6,6	37	-4,62	30

D2	-6,9	15	-4,87	22
D3	-6,9	16	-4,52	32
D4	-6,8	22	-4,83	25
D6	-6,7	26	-4,63	29
D14	<u>-7,4</u>	7	-5,20	14
D15	<u>-7,5</u>	5	-5,38	10
D18	<u>-7,4</u>	8	-4,46	38
D19	<u>-7,7</u>	4	-5,39	8
D20	-6,1	43	-4,47	35
D21	-6,1	44	-4,22	42
D23	-7,1	13	-4,98	17
D24	<u>-7,2</u>	9	-7,24	5
D25	-6,9	17	-7,27	4
A2	-6,6	38	-4,55	31
A5	-6,3	41	-4,24	40
A6	<u>-7,8</u>	1	-5,21	12
A8	-6,8	23	-7,31	2
A9	-7,0	14	-4,96	19
A10	-6,9	18	-4,50	33
A11	-6,5	40	-4,86	24
A12	-6,6	39	-4,24	41
A14	-6,9	19	-4,99	16
A15	-6,7	27	-4,11	44
A17	<u>-7,2</u>	10	-4,48	34
A19	-6,7	28	-7,01	6
A20	-6,7	29	-5,39	9

Fonte: Resultados compilados pela autora, com dados de afinidade de ligação obtidos nos programas AutoDock Vina e AutoDock.

Conforme pode ser observado na Tabela 13, os ligantes D15, D19, G19 e D24 são comuns aos programas AutoDock e AutoDock Vina, entre os dez ligantes com maior afinidade de ligação. Os ligantes G14, G15, A6, D19, D14 e A17 posicionam-se entre os dez ligantes com maior afinidade de ligação somente para o programa AutoDock Vina e os ligantes G24, G25, D25, A19, A20 e A8 somente para o programa AutoDock.

Considerando os oito ligantes selecionados para simulações de dinâmica molecular (G14, G15, D14, D15, D19, A6, A20 e G19) com a CRM1, verifica-se que os ligantes D15, D19 e G19 também configuram como os que posicionam-se entre os dez melhores ligantes em relação à afinidade de ligação para a proteína MDM2, considerando os programas AutoDock e AutoDock Vina. Os ligantes G14, G15, D14 e A6 são coincidentes apenas para o programa AutoDock Vina e o ligante A20 somente para o programa AutoDock.

Nesse sentido, todos os ligantes selecionados para simulações de dinâmica em *docking*

molecular com a CRM1 também foram encontrados como os que possuem as dez melhores afinidades de ligação com a proteína MDM2, nos programas AutoDock e/ou AutoDock Vina. Compilando os resultados, observam-se que os dados obtidos nas simulações utilizando a CRM1 e MDM2 convergem em experimentos de *docking* molecular.

De maneira similar à realizada para a proteína p53, nos cálculos de resíduos formados e de dinâmica computacional com a proteína MDM2, foram utilizadas apenas as chalconas D14 e D15. O ligante Nutlin-3a, que é um inibidor descrito na literatura para a MDM2 (KOJIMA *et al.*, 2006), foi utilizado como ligante de referência. Com o programa LigPlot+ (WALLACE; LASKOWSKI; THOMTON, 1995), foram criados diagramas dos resíduos gerados entre ligante e proteína, de acordo com convenções químicas. e detalhes podem ser verificados na Figura 26. As afinidades de ligação obtidas para esse composto foram -6,19 kcal/mol, no programa AutoDock 4 e -6,40 no programa AutoDock Vina.

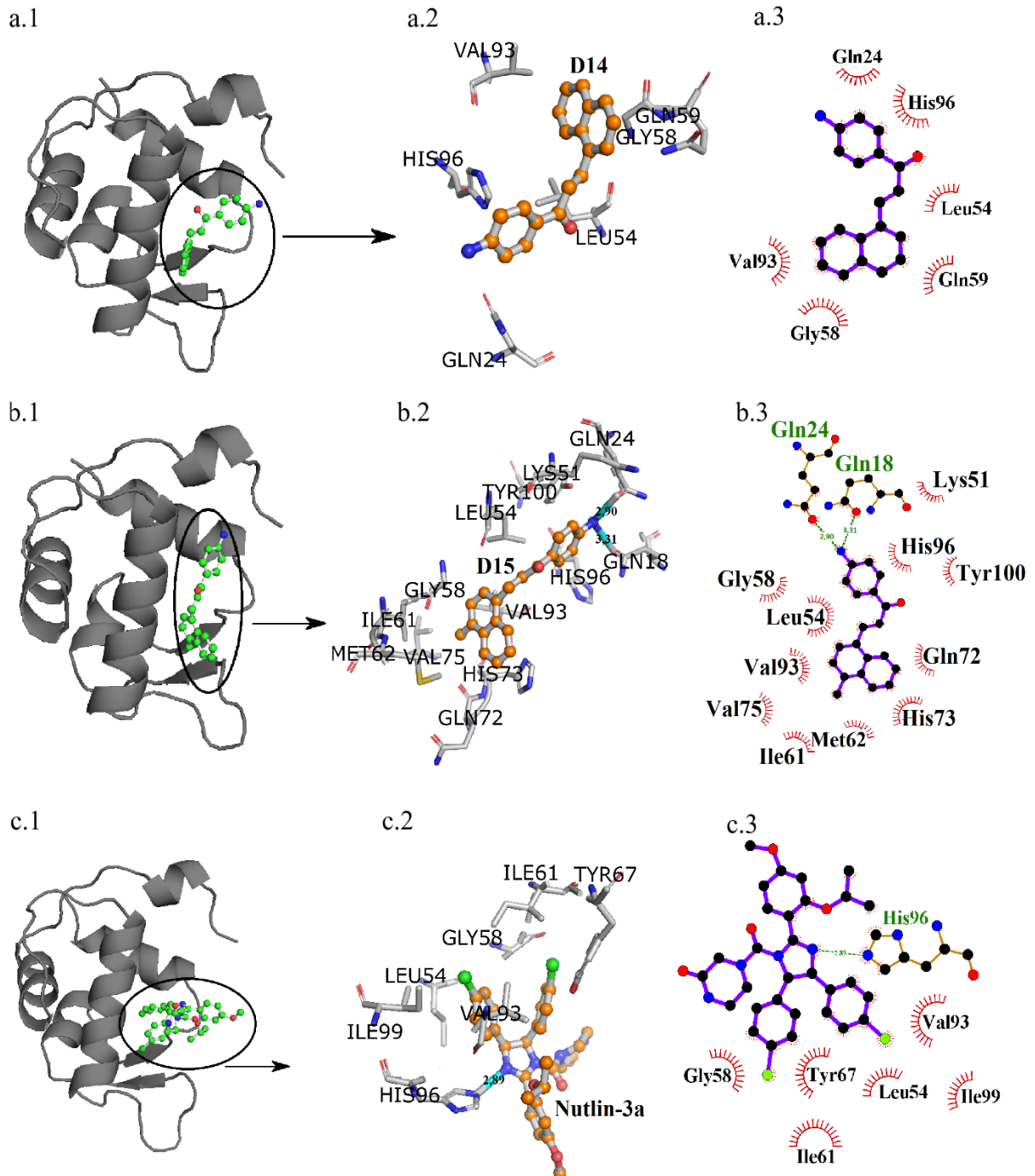
Considerando as informações observadas na Figura 26, os resíduos encontrados para o ligante D14 (Figura 26a), foram Gln24, Gln58, Gln59, His96, Leu54, Val93. Para o ligante D15 (Figura 26b), foram identificados os resíduos Gln18, Gln24, Gln72, Gly58, Ile61, Leu54, Lys51, His73, His96, Met62, Tyr100, Val75 e Val93. Para o ligante de referência Nutlin-3a (Figura 26c), foram encontrados os resíduos His96, Val93, Gly58, Tyr67, Leu54, Ile61 e Ile99. Resíduos comuns a esses ligantes são Val93, Leu54 e His96.

3.1.5.2. Resultados de Dinâmica Molecular

As simulações de dinâmica molecular foram realizadas somente para a D14 e D15 e utilizou-se o ligante de referência Nutlin-3a, presente na estrutura da proteína MDM2. Os resultados das simulações de dinâmica molecular relativos a RMSD, número de contatos e número de pontes de hidrogênio para os ligantes D14, D15 e Nutlin-3a (ligante de referência) podem ser observados, respectivamente, nas Figuras 27, 28 e 29. Na Figura 27, visualizam-se os resultados de RMSD de todos os átomos que não são hidrogênio para os ligantes D14 (Figura 27a), D15 (Figura 27b) e de referência (Figura 27c), respectivamente. Note que, nesses resultados, o RMSD do ligante D14 ($7,43 \pm 0,5$, Figura 27a) é estável a partir de 5ns e mantém-se estável até o final da trajetória. O RMSD do ligante D15 ($5,07 \pm 1,48$, Figura 27b) é estável entre 30ns e 50 ns. Com 50 ns, o ligante D15 realiza uma pequena alteração em sua trajetória e retorna à posição anterior imediatamente, mantendo essa posição até o final da simulação. Para o ligante de referência ($4,54 \pm 0,73$, Figura 27c), a estabilidade é obtida a partir de 25ns e mantém-se até o final da execução. Esses resultados indicam que os ligantes avaliados

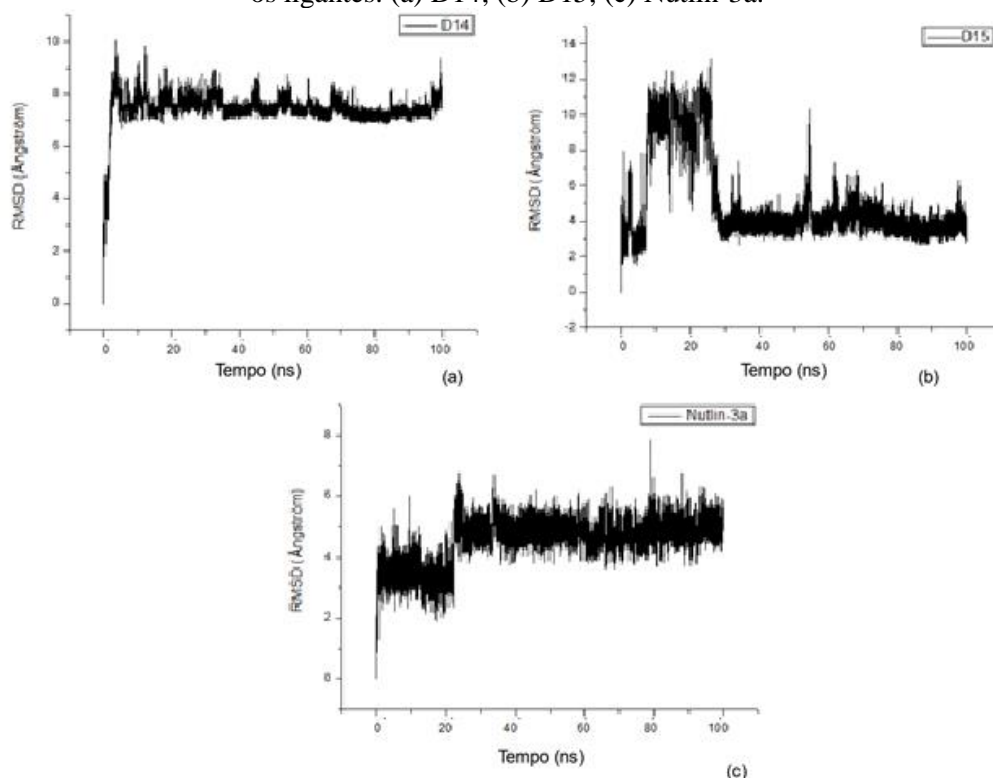
apresentam estabilidade nas trajetórias.

Figura 26: Estrutura da MDM2 utilizada para simulações computacionais e os subdomínios gerados com os ligantes D14, D15 e Nutlin-3a. Nas figuras a.1, b.1 e c.1, a proteína está representada em cartoon e a região da MDM2 que contém o ligante avaliado encontra-se circunscrita. Para a representação das chalconas D14 e D15 e do ligante de referência (Nutlin-3a), foi utilizado o modelo de *stick* e *ball*. Nas figuras a.2, a.3, b.2, b.3, c.2 e c.3 podem ser observados os resíduos encontrados entre MDM2 e: (a.2, a.3) D14; (b.2, b.3) D15; (c.2, c.3) Nutlin-3a.



Fonte: Resultados compilados pela autora, com dados obtidos no programa LigPlot+. A representação estrutural foi preparada utilizando o programa PyMOL (DELANO, 2002).

Figura 27: Resultados de simulação de dinâmica computacional para RMSD da proteína MDM2 com os ligantes: (a) D14; (b) D15; (c) Nutlin-3a.



Fonte: Autora, com resultados de RMSD obtidos no programa GROMACS.

Na Figura 28, observa-se o número de contatos (<0.6 nm) dos ligantes avaliados. Nessa tabela, observa-se que o número de contatos para o ligante D14 (Figura 28a), D15 (Figura 28b) e de referência (Figura 28c) são, respectivamente, 1487 ± 186 , 1505 ± 265 e 1499 ± 171 . Desse modo, dentre os ligantes avaliados, o que apresentou maiores números de contatos foi o ligante D15, seguido do ligante de referência. Embora o ligante D14 tenha apresentado números menores de contatos, esse ligante ainda é competitivo com os demais.

Na Figura 29, observa-se o número de pontes de hidrogênio formado entre os três ligantes avaliados. Na Figura 29a e 29b, observa-se o número de pontes de hidrogênio entre a MDM2 e os ligantes D14 e D15, respectivamente, que é igual a 1 ± 1 . Na Tabela 29c, pode-se observar que a média dos valores do número de pontes de hidrogênio é zero. Assim, o ligante que apresenta os menores valores para o número de pontes de hidrogênio é o Nutlin-3a.

Considerando os valores de RMSD, número de contatos e número de pontes de hidrogênio, computacionalmente, o ligante D15 indica possuir resultados de maior interesse se comparado às demais moléculas avaliadas. Ainda, por apresentar um maior número de pontes de hidrogênio e estabilidade na ligação, que pode ser verificada através de RMSD, o ligante D14 parece possuir desempenho superior quando comparado ao ligante de referência. Os dados obtidos são sumarizados na Tabela 14.

Tabela 14: Resultados de RMSD, Número de Contatos e Número de Pontes de Hidrogênio para a proteína MDM2.

Ligante	RMSD	Pontes de Hidrogênio	Número de Contatos
D14	7,43 ± 0,5	1487±186	1±1
D15	5,07 ± 1,48	1505±265	1±1
Nutlin-3a	4,54 ± 0,73	1499±171	0

Fonte: Autora, com resultados obtidos no programa GROMACS.

3.1.6. Considerações sobre os Resultados Computacionais para as Proteínas CRM1, p53 e MDM2

Simulações de *docking* molecular são modelos computacionais, cujos resultados gerados para a afinidade de ligação são dependentes da abordagem adotada no programa selecionado. Nesse sentido, neste trabalho, considerando a proteína CRM1, foram utilizados os programas AutoDock, AutoDock Vina, SwissDock e DockThor para essas simulações, a fim de confrontar as saídas geradas e selecionar as chalconas com melhores afinidades de ligação. Combinados os resultados obtidos pelos programas de *docking* molecular, os ligantes com melhores resultados foram G14, G15, D14, D15, D19, A6, A20 e G19. Esses ligantes também foram submetidos a simulações de dinâmica computacional. Porém, com os dados computacionais obtidos em simulações de dinâmica molecular, não foi possível determinar o melhor ligante para essa proteína, pois os dados para melhores RMSD, pontes de hidrogênio, número de contatos e menor média de RMSD divergem entre si. Porém, as oito chalconas selecionadas apresentam resultados competitivos aos obtidos pelo ligante de referência, K76.

Uma vez que a atividade da CRM1 possui relação direta com a apresentada por outras proteínas, as simulações de *docking* molecular foram refeitas com a p53 e MDM2, a fim de confrontar as chalconas selecionadas em testes anteriores com a CRM1. Foi verificado nessas simulações que todas as chalconas (G14, G15, D14, D15, D19, A6, A20 e G19) também apresentam bons resultados nos programas de *docking* empregando as proteínas MDM2 e p53. Ainda, quando comparados os resultados de dinâmica molecular obtidos com as chalconas D14

e D15 e ligantes de referência, o uso das chalconas mostrou-se competitivo, para as proteínas MDM2 e p53.

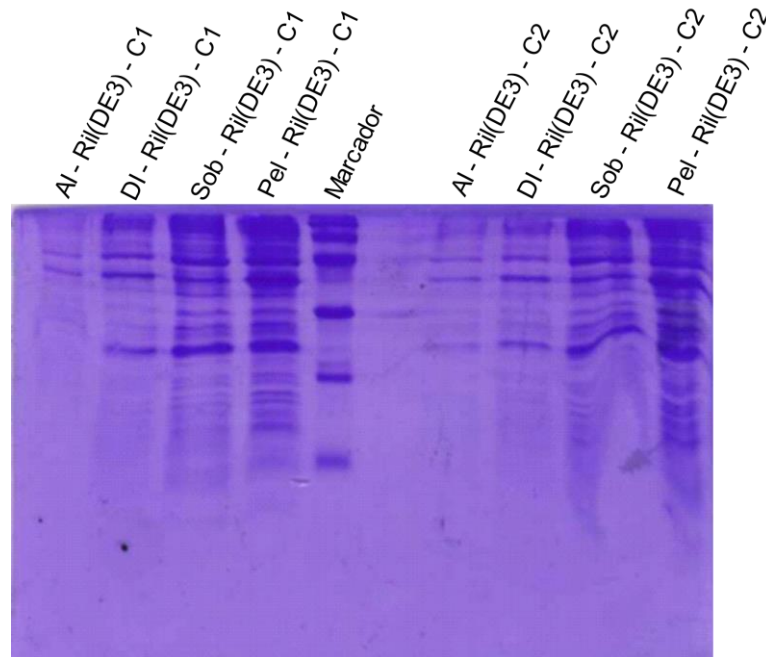
A fim de corroborar os resultados obtidos computacionalmente, ensaios laboratoriais foram realizados com o domínio da proteína CRM1. A realização desses ensaios teve por objetivo verificar, através de experimentos de ressonância magnética, se os ligantes selecionados apresentariam interação no domínio de interesse da CRM1. Os resultados alcançados são apresentados a seguir, na seção 3.2 deste trabalho.

3.2. Resultados de Expressão e Purificação do Domínio da Proteína CRM1

Os resultados dos ensaios de expressão e purificação do domínio da CRM1 são apresentados nesta subseção, cujas amostras foram obtidas conforme metodologia descrita na seção 2.2.4 deste trabalho. Os géis de SDS-PAGE de 20% obtidos podem ser observados nas Figuras 30 a 35, em que se deve considerar que o domínio da CRM1 possui tamanho ~12 kDa, que está abaixo do último marcador dos géis.

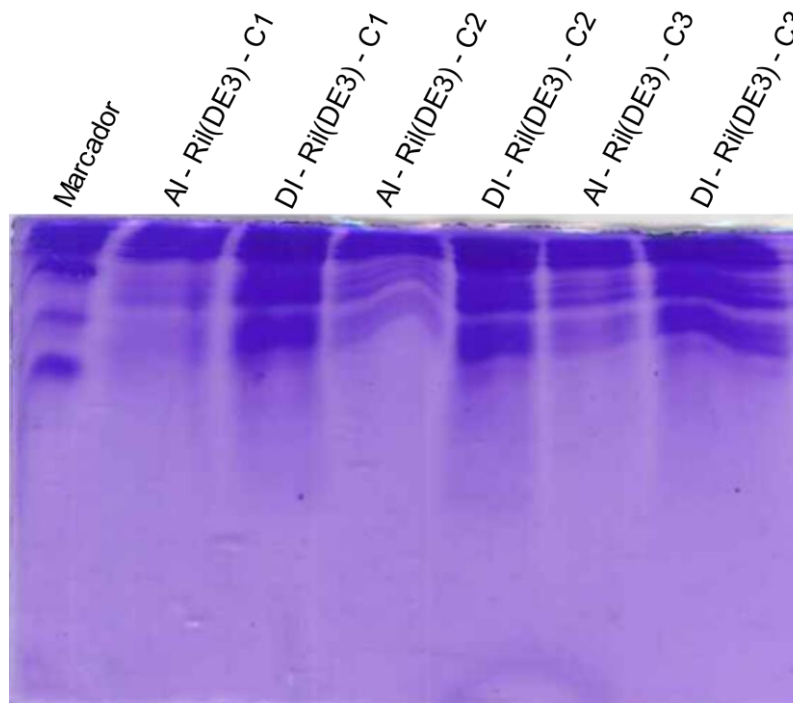
Nas Figuras 30, 31, 32, 33, 34 e 35, apresentadas a seguir, são apresentados os resultados obtidos com experimentos utilizando os clones 1 (C1), clone 2 (C2) e clone 3 (C3), em que também foram coletadas amostras de pellet (Pel) e sobrenadante (Sob) e foram empregadas as seguintes condições de expressão: Figura 30) temperatura: 20°C; concentração de IPTG: 1.0 milimol/Litro; tempo de indução: 18h. Linhagem: Ril(DE3); Figura 31) temperatura: 20°C; concentração de IPTG: 0.4 milimol/Litro; tempo de indução: 18h. Linhagem: Ril(DE3). Na Figura 31, resultados para sobrenadante e pellet também são apresentados; Figuras 32 e 33) temperatura: 20°C; concentração de IPTG: 0.4 milimol/Litro; e tempo de indução: 18h. Linhagens: Ril(DE3) e BL21(De3); Figura 34) temperatura: 37°C; concentração de IPTG: 0.5 milimol/Litro e 1.0 milimol/Litro; tempo de indução: 4h. Linhagem: Ril(DE3); Figura 35) temperatura: 37°C; concentração de IPTG: 0.5 milimol/Litro e 1.0 milimol/Litro; e tempo de indução: 4h. Linhagem: BL21(DE3).

Figura 30: Amostras antes de induzir (AI) e depois de induzir (DI) com os clones C1, C2 e C3, com temperatura: 20°C; concentração de IPTG: 1.0 milimol/Litro; tempo de indução: 18h. Linhagem: Ril(DE3).



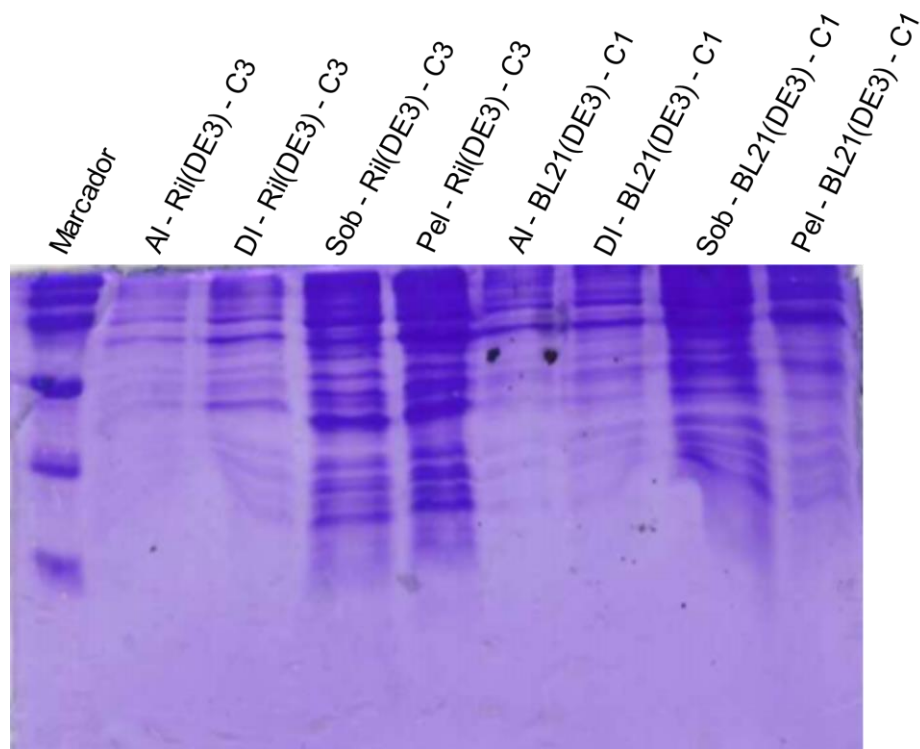
Fonte: Autora.

Figura 31: Antes de induzir (AI) e depois de induzir (DI) com os clones C1 e C2 e temperatura 20°C; IPTG: 0.4 milimol/Litro; e tempo de indução: 18h. Linhagem: Ril(DE3). Sobrenadante (Sob) e pellet (Pel) são apresentados.



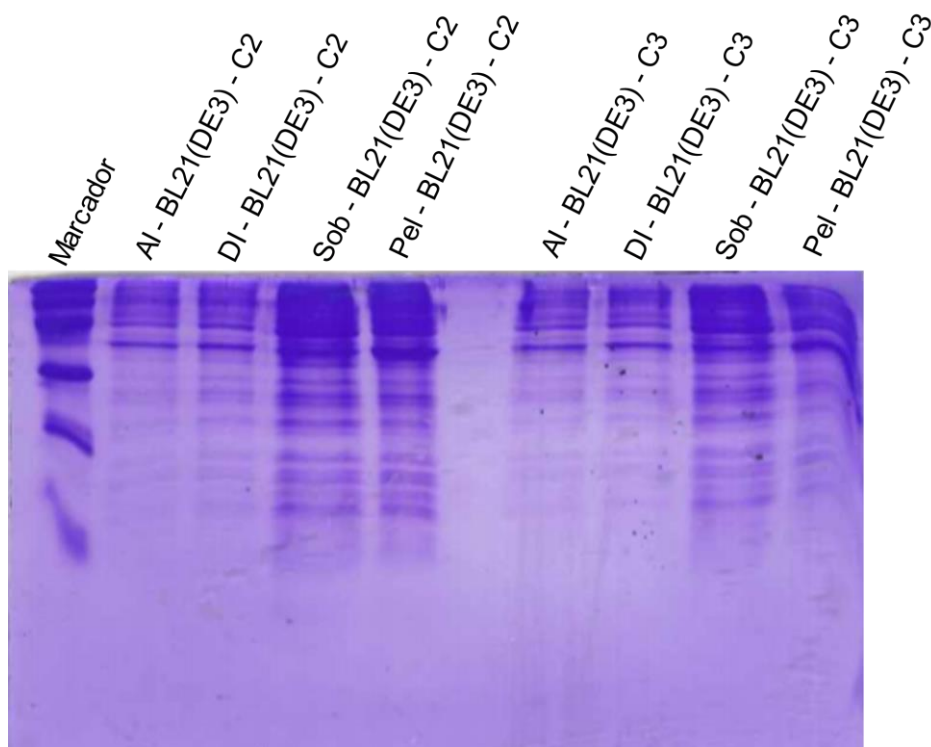
Fonte: Autora

Figura 32: Antes de induzir (AI) e depois de induzir (DI) com os clones 3 (C3, RiIDE3) e 1 (C2, BL21DE3) utilizando temperatura: 20°C; concentração de IPTG: 0.4 milimol/Litro; e tempo de indução: 18h.



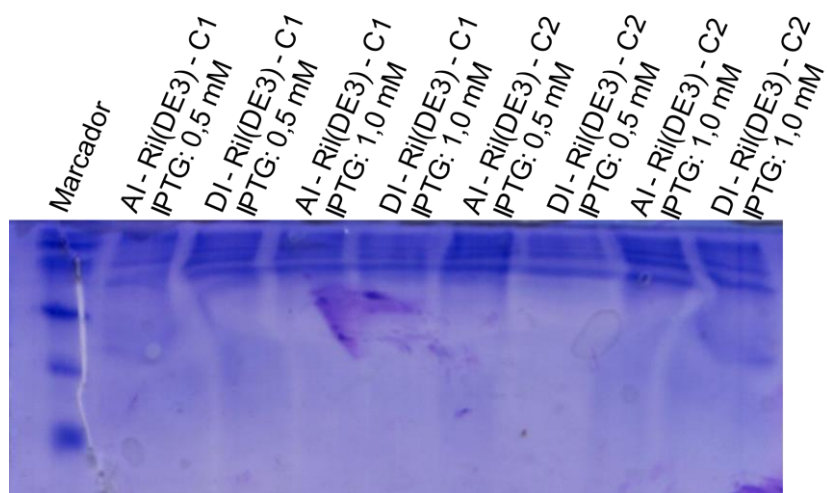
Fonte: Autora

Figura 33: Antes de induzir (AI) e depois de induzir (DI) com os clones 2 (C2) e 3 (C3) com temperatura: 20°C; concentração de IPTG: 0.4 milimol/Litro; e tempo de indução: 18h. Linhagem: BL21(DE3). Continuação da Figura 32.



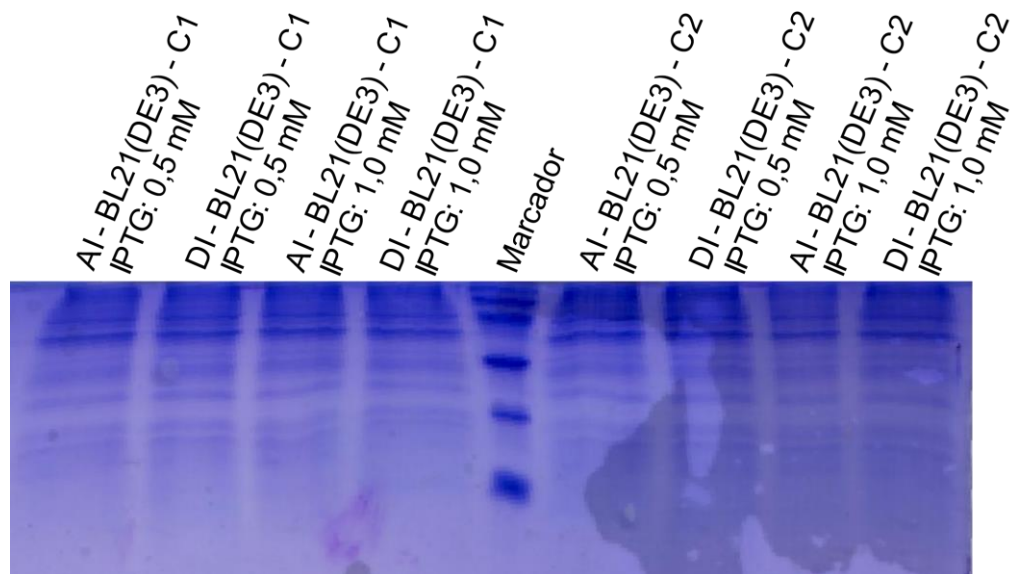
Fonte: Autora

Figura 34: Antes de induzir (AI) e depois de induzir (DI) com o clone 1 (C1) com temperatura: 37°C; concentração de IPTG: 0.5 milimol/Litro e 1.0 milimol/Litro; e tempo de indução: 4h. Linhagem: Ril(DE3).



Fonte: Autora

Figura 35: Antes de induzir (AI) e depois de induzir (DI) com o clone 1 (C1) com temperatura: 37°C; concentração de IPTG: 0.5 milimol/Litro e 1.0 milimol/Litro; e tempo de indução: 4h. Linhagem: BL21(DE3).



Fonte: Autora

Conforme pode ser observado nos géis SDS-PAGE, não houve expressão da CRM1 nas condições avaliadas. Uma justificativa é o plasmídeo apresentar caráter altamente hidrofóbico, especialmente no C-terminal. Aminoácidos hidrofóbicos evitam estar em contato com água e, para isso, tendem a ficar no interior da proteína. A cadeia do alvo é apresentada abaixo e podem ser verificados, em cor cinza, resíduos hidrofóbicos na porção C-terminal. Nas demais cores, outros resíduos que podem aumentar a hidrofobicidade da cadeia também estão destacados. Apenas cerca de 35% dos resíduos do plasmídeo da CRM1 não apresentam condições de hidrofobicidade.

MSEDTEKRFVVTVIKDLLDLCVKKRGKDNKAVVASDIMYVVGQYPRFLKAHWNFLRIVILKLFEEFMHE
THEGVQDMACDIFIKIVQCKYHFVIQ

Ainda, a sequência utilizada para o alvo é pequena, possuindo cerca de 12 kDa (95 aminoácidos), o que dificulta sua superexpressão. Por esse motivo, é possível que caso sejam encontradas as condições ideais para expressão da proteína, ela seja expressa em pequenas quantidades. Desse modo, no presente trabalho, não foram obtidos resultados utilizando o domínio da proteína CRM1. Uma abordagem possível é a utilização da proteína CRM1, utilizando as condições descritas na literatura, como no trabalho de BOOTH, CHENG, FRANKEL (2014).

4. CONCLUSÕES

Neste trabalho, foram realizados experimentos de *docking* e de dinâmica molecular empregando 45 chalconas e a proteína CRM1. Nas simulações de *docking*, foram selecionadas 8 chalconas de interesse, que são: G14, G15, D14, D15, D19, A6, A20 e G19. Em simulações de dinâmica molecular, considerando o ligante de referência K76, pode-se observar que algumas das chalconas selecionadas apresentaram resultados superiores de RMSD, número de contatos e pontes de hidrogênio.

Ainda, as proteínas p53 e MDM2 também foram submetidas a experimentos computacionais. Pode-se observar um desempenho superior das chalconas D14 e D15 se comparadas aos ligantes de referência, em experimentos de dinâmica molecular para valores de RMSD, número de contatos e pontes de hidrogênio.

Nos ensaios de expressão e de purificação do domínio de interesse da CRM1, após o uso de diferentes concentrações de IPTG e de variações de temperatura e de tempo de indução, não foram obtidas as condições ideais para sua expressão. Ainda, foi verificado que esse domínio é hidrofóbico e devido ao seu tamanho (~12 kDa), é provável que seja difícil haver uma superexpressão dessa proteína nos ensaios. Desse modo, trabalhos futuros referem-se à realização de ensaios contendo o plasmídeo da proteína CRM1 completa, bem utilizar anticorpos específicos para experimentos de *western blot*, para verificar a presença dessa proteína nas amostras.

Nesse sentido, através de simulações computacionais, neste trabalho foram selecionadas moléculas de interesse, cuja atividade pode auxiliar na regulação da proteína CRM1, p53 e MDM2.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABARRATEGI, A.; TORNIN, J.; MARTINEZ-CRUZADO, L.; HAMILTON, A. *et al.* Osteosarcoma: Cells-of-Origin, Cancer Stem Cells, and Targeted Therapies. **Stem Cells Int**, 2016, p. 3631764, 2016.
- ABID, O-U-R; KHALID, M; HUSSAIN, M. T.; HANIF, M. *et al.* Synthesis and anti-cancer, anti-metastatic evaluation of some new fluorinated isocoumarins and 3,4-dihydroisocoumarins. **Journal of Fluorine Chemistry**, 135, p. 240-245, 2012.
- AHMED, A.; ALKALI, Y. In silico Pharmacokinetics and Molecular Docking Studies of Lead Compounds Derived from Diospyros Mespiliformis. **PharmaTutor**, 7, n. 3, p. 31-37, 2019.
- ALMEIDA, D. M. **Dockthor: Implementação, Aprimoramento e Validação de um Programa de Docking Receptor-Ligante**. 2011. (Dissertação de Mestrado) - Laboratório Nacional de Computação Científica, Brasil. Disponível em: oi 10. Acesso em: 09/03/2021.
- ANIL, B.; RIEDINGER, C.; ENDICOTT, J. A.; NOBLE, M. E. M. The structure of an MDM2–Nutlin-3a complex solved by the use of a validated MDM2 surface-entropy reduction mutant. **Acta Cryst**, D69, p. 1358-1366, 2013.
- BASAK, S. C. Chemobioinformatics: the advancing frontier of computer-aided drug design in the post-genomic era. **Curr Comput Aided Drug Des**, 8, n. 1, p. 1-2, Mar 2012.
- BIKADI, Z.; HAZAI, E. Application of the PM6 semi-empirical method to modeling proteins enhances docking accuracy of AutoDock. **J Cheminform**, 1, p. 15, Sep 11 2009.
- BIO-HPC. **Achilles Blind Docking Server, Bioinformatics and High Performance Computing (BIO-HPC) Research group** 2017. Acesso em: <https://bio-hpc.ucam.edu/achilles/>
- BIOVIA. **Discovery Studio Modeling Environment**. Dassault Systèmes. San Diego 2017.
- BOLIA, A.; OZKAN, S. B. Adaptive BP-Dock: An Induced Fit Docking Approach for Full Receptor Flexibility. **J Chem Inf Model**, 56, n. 4, p. 734-746, Apr 25 2016.
- BOOTH, D. S.; CHENG, Y.; FRANKEL, A. D. The export receptor Crm1 forms a dimer to promote nuclear export of HIV RNA. **Elife**, 2014, 3, p. 04121.
- BORA, A.; SUZUKI, T.; FUNAR-TIMOFEI, S. Neonicotinoid insecticide design: molecular docking, multiple chemometric approaches, and toxicity relationship with Cowpea aphids. **Environ Sci Pollut Res Int**, 26, n. 14, p. 14547-14561, May 2019.
- BORTOLOTTI, L. F.; BARBOSA, F. R.; SILVA, G.; BITENCOURT, T. A. *et al.* Cytotoxicity of trans-chalcone and licochalcone A against breast cancer cells is due to apoptosis induction and cell cycle arrest. **Biomed Pharmacother**, Nov 27 2016.

- BOSTROM, J.; GREENWOOD, J. R.; GOTTFRIES, J. Assessing the performance of OMEGA with respect to retrieving bioactive conformations. **J Mol Graph Model**, 21, n. 5, p. 449-462, Mar 2003.
- BROOKS, B. R.; BRUCCOLERI, R. E.; OLAFSON, B. D.; STATES, D. J. *et al.* CHARMM: a program for macromolecular energy, minimization and dynamics calculations. . **J Comp Chem** 4, p. 187–217, 1983.
- BRYLINSKI, M.; SKOLNICK, J. Q-Dock: Low-resolution flexible ligand docking with pocket-specific threading restraints. **J Comput Chem**, 29, n. 10, p. 1574-1588, Jul 30 2008.
- BRYLINSKI, M.; SKOLNICK, J. FINDSITE: a threading-based approach to ligand homology modeling. **PLoS Comput Biol**, 5, n. 6, p. e1000405, Jun 2009.
- BURKHARD, P.; TAYLOR, P.; WALKINSHAW, M. An example of a protein ligand found by database mining: description of the docking method and its verification by a 2.3 Å X-ray structure of a Thrombin-Ligand complex. **Journal of Molecular Biology**, 277, n. 2, p. 449-466, 1998.
- CAO, Y.; DAI, W.; MIAO, Z. Evaluation of Protein-Ligand Docking by Cyscore. **Methods Mol Biol**, 1762, p. 233-243, 2018.
- CAVALCANTI JÚNIOR, G. B.; KLUMB, C. E.; MAIA, R. C. p53 e as hemopatias malignas **Revista Brasileira de Cancerologia**, 48, n. 3, p. 419-427, 2002.
- CERA, T. P.; PANCOTE, C. G. Planejamento de fármacos. **Revista Científica UNILAGO**, p. 137-148, 2012.
- CHANG, D. T.; OYANG, Y. J.; LIN, J. H. MEDock: a web server for efficient prediction of ligand binding sites based on a novel optimization algorithm. **Nucleic Acids Res**, 33, n. Web Server issue, p. W233-238, Jul 1 2005.
- CHEN, H. M.; LIU, B. F.; HUANG, H. L.; HWANG, S. F. *et al.* SODOCK: swarm optimization for highly flexible protein-ligand docking. **J Comput Chem**, 28, n. 2, p. 612-623, Jan 30 2007.
- CHOI, V. Yucca: an efficient algorithm for small-molecule docking. **Chem Biodivers**, 2, n. 11, p. 1517-1524, Nov 2005.
- CHUNG, J. Y.; CHO, S. J.; HAH, J. M. A python-based docking program utilizing a receptor bound ligand shape: PythDock. **Arch Pharm Res**, 34, n. 9, p. 1451-1458, Sep 2011.
- CLARK, K. P.; AJAY. Flexible ligand docking without parameter adjustment across four ligand-receptor complexes. **Journal of Computational Chemistry**, 16, n. 10, p. 1210–1226, 1995.
- CLAUSSEN, H.; BUNING, C.; RAREY, M.; LENGAUER, T. FlexE: efficient molecular docking considering protein structure variations. **J Mol Biol**, 308, n. 2, p. 377-395, Apr 27 2001.

COLOMBO, A.; CIPOLLA, C.; BEGGIATO, M.; CARDINALE, D. Cardiac toxicity of anticancer agents. **Curr Cardiol Rep**, 15, n. 5, p. 362, May 2013.

da CRUZ, M. S.; IRFFI, G. **A relação entre diagnóstico de câncer e aspectos socioeconômicos, ambientais e de estilo de vida no Brasil.**

https://coreconpe.gov.br/eventos/venpecon/teo_aplic/A%20rela%C3%A7%C3%A3o%20entre%20diagn%C3%B3stico.pdf, 2017. Acesso em 08/04/2021.

DALLAKYAN, S. **PyRx**. <https://pyrx.sourceforge.io/>, 2009-2019. Acesso em: 17/07/2019.

DARDENN, L. E. Dockthor: Development and Validation of a New Docking Program. . *In*: XLI Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular - SBBq, 2012, Foz do Iguaçu, Brazil.

DE MAGALHÃES, C. S. **Algoritmos Genéticos para o Problema de Docking Proteína-Ligante**. 2006. (Tese de Doutorado) - Laboratório Nacional de Computação Científica, Brasil.

DE PAULA, G. C.; HORTA, B. A. C. **Métodos de multi-docking**. http://www.puc-rio.br/pibic/relatorio_resumo2014/relatorios_pdf/ctc/ELE/ELE-Gabriel%20Coutinho%20de%20Paula.pdf, 2014. Acesso em: 27/12/2017.

DELANO, W. L. **The PyMOL molecular graphics system**. San Carlos, CA, USA 2002.

DIMMOCK, J. R.; ELIAS, D. W.; BEAZELY, M. A.; KANDEPU, N. M. Bioactivities of chalcones. **Curr Med Chem**, 6, n. 12, p. 1125-1149, Dec 1999.

DOLINSKY, T. J.; NIELSEN, J. E.; MCCAMMON, J. A.; BAKER, N. A. PDB2PQR: an automated pipeline for the setup, execution, and analysis of Poisson-Boltzmann electrostatics calculations. **Nucleic Acids Res** 32, p. 665-667, 2004.

ELLINGSON, S. R.; SMITH, J. C.; BAUDRY, J. VinaMPI: facilitating multiple receptor high-throughput virtual docking on high-performance computers. **J Comput Chem**, 34, n. 25, p. 2212-2221, Sep 30 2013.

FRIESNER, R. A.; BANKS, J. L.; MURPHY, R. B.; HALGREN, T. A. *et al.* Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. **J Med Chem**, 47, n. 7, p. 1739-1749, Mar 25 2004.

FRISCH, M. J.; TRUCKS, G. W.; SCHLEGEL, H. B.; SCUSERIA, G. E. *et al.* **Gaussian 09, revision A.02**. . Gaussian Inc. Wallingford, CT. 2009.

FU, D. Y.; MEILER, J. RosettaLigandEnsemble: A Small-Molecule Ensemble-Driven Docking Approach. **ACS Omega**, 3, n. 4, p. 3655-3664, Apr 30 2018.

GABB, H. A.; JACKSON, R. M.; STERNBERG, M. J. E. Modelling protein docking using shape complementarity, electrostatics and biochemical information **Journal of Molecular Biology**, Edited by J. Thornton, 272, n. 1, 1997.

GAUDREAU, F.; NAJMANOVICH, R. J. FlexAID: Revisiting Docking on Non-Native-Complex Structures. **J Chem Inf Model**, 55, n. 7, p. 1323-1336, Jul 27 2015.

GLOBOCAN. **Cancer Tomorrow**. <http://gco.iarc.fr/tomorrow/home>, 2018. Acesso em: 21/01/2020.

GOODSELL, D. S.; OLSON, A. J. Automated docking of substrates to proteins by simulated annealing. **Proteins**, 8, n. 3, p. 195-202, 1990.

GORELIK, B.; GOLDBLUM, A. High quality binding modes in docking ligands to proteins. **Proteins**, 71, n. 3, p. 1373-1386, May 15 2008.

GOTO, J.; KATAOKA, R.; MUTA, H.; HIRAYAMA, N. ASEDock-docking based on alpha spheres and excluded volumes. **J Chem Inf Model**, 48, n. 3, p. 583-590, Mar 2008.

GROSDIDIER, A.; ZOETE, V.; MICHIELIN, O. EADock: docking of small molecules into protein active sites with a multiobjective evolutionary optimization. **Proteins**, 67, n. 4, p. 1010-1025, Jun 1 2007.

GROSDIDIER, A.; ZOETE, V.; MICHIELIN, O. SwissDock, a protein-small molecule docking web service based on EADock DSS. **Nucleic Acids Res**, 39, n. Web Server issue, p. W270-277, Jul 2011.

GU, B.; ZHU, W.-G. Surf the Post-translational Modification Network of p53 Regulation. **International Journal of Biological Sciences**, 8, n. 5, p. 672-684, 2012.

GUPTA, A.; GANDHIMATHI, A.; SHARMA, P.; JAYARAM, B. ParDOCK: an all atom energy based Monte Carlo docking protocol for protein-ligand complexes. **Protein Pept Lett**, 14, n. 7, p. 632-646, 2007.

HART, T. N.; READ, R. J. A multiple-start Monte Carlo docking method. **Proteins**, 13, n. 3, p. 206-222, Jul 1992.

HE, H.; NI, J.; HUANG, J. Molecular mechanisms of chemoresistance in osteosarcoma (Review). **Oncol Lett**, 7, n. 5, p. 1352-1362, May 2014.

HINDLE, S. A.; RAREY, M.; BUNING, C.; LENGAE, T. Flexible docking under pharmacophore type constraints. **J Comput Aided Mol Des**, 16, n. 2, p. 129-149, Feb 2002.

HIRAKI, M.; HWANG, S. Y.; CAO, S.; RAMADHAR, T. R. *et al.* Small-Molecule Reactivation of Mutant p53 to Wild-Type-like p53 through the p53-Hsp40 Regulatory Axis. **Chem Biol** 22, n. 9, p. 1206-1216, 2015.

HOWARD, S. C.; MCCORMICK, J.; PUI, C. H.; BUDDINGTON, R. K. *et al.* Preventing and Managing Toxicities of High-Dose Methotrexate. **Oncologist**, 21, n. 12, p. 1471-1482, Dec 2016.

HSU, K. C.; CHEN, Y. F.; LIN, S. R.; YANG, J. M. iGEMDOCK: a graphical environment of enhancing GEMDOCK using pharmacological interactions and post-screening analysis. **BMC Bioinformatics**, 12 Suppl 1, p. S33, Feb 15 2011.

HU, B.; LILL, M. A. PharmDock: a pharmacophore-based docking program. **J Cheminform**, 6, p. 14, 2014.

HUANG, S. Y.; LI, M.; WANG, J.; PAN, Y. HybridDock: A Hybrid Protein-Ligand Docking Protocol Integrating Protein- and Ligand-Based Approaches. **J Chem Inf Model**, 56, n. 6, p. 1078-1087, Jun 27 2016.

HYPERCHEM. **Docking Study with HyperChem**. <https://www.molfunction.com/>, 2004. Acesso em: 07/07/2019.

IMBERNÓN, B.; CECILIA, J. M.; PÉREZ-SÁNCHEZ, H.; GIMÉNEZ, D. METADOCK: A parallel metaheuristic schema for virtual screening methods. **The International Journal of High Performance Computing Applications**, 32, n. 6, 2017.

IMBERNÓN, B.; LLANES, A.; CUTILLAS, J.-M.; GIMÉNEZ, D. HYPERDOCK: Improving virtual screening through parallel hyperheuristics. **The International Journal of High Performance Computing Applications**, 34, n. 1, p. 1-12, 2019.

IMMD. **Adam**. http://www.immd.co.jp/en/product_2.html, 2011. Acesso em: 25/07/2019.

INCA. **Estimativa**. <http://www1.inca.gov.br/estimativa/2018/estimativa-2018.pdf>, 2018. Acesso em: 15/02/2020.

INCA. **O que é o câncer?** <https://www.inca.gov.br/o-que-e-cancer>, 2019. Acesso em: 18/01/2020.

ITZAMNA. **Mind the Byte**. <https://www.mindthebyte.com/>, 2018. Acesso em: 15/07/2019.

JAFFE, N. Historical Perspective on the Introduction and Use of Chemotherapy for the Treatment of Osteosarcoma. *In*: KLEINERMAN, E. (Ed.). **Current Advances in Osteosarcoma**. Switzerland: Springer International Publishing, 2014.

JAIN, A. N. Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine. **J Med Chem**, 46, n. 4, p. 499-511, Feb 13 2003.

JIANG, Y.-L.; TANG, L.-Q.; MIYANAGA, S.; IGARASHI, *et al.* Synthesis and evaluation of trehalose-based compounds as anti-invasive agents. **Bioorganic & Medicinal Chemistry Letters**, 21, n. 4, p. 1089-1091, 2011.

JONES, G.; WILLETT, P.; GLEN, R. C. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. **Journal of Molecular Biology**, 245, n. 1, p. 43-53, 1995.

JORGENSEN, W. L.; CHANDRASEKHAR, J.; MADURA, J. D.; IMPEY, R. W. *et al.* Comparison of simple potential functions for simulating liquid water. **Journal of Chemical Physics**, 79, p. 926-935, 1983.

KAGER, L.; ZOUBEK, A.; POTSCHEGER, U.; KASTNER, U. *et al.* Primary metastatic osteosarcoma: presentation and outcome of patients treated on neoadjuvant Cooperative Osteosarcoma Study Group protocols. **J Clin Oncol**, 21, n. 10, p. 2011-2018, May 15 2003.

KAINRAD, T.; HUNOLD, S.; SEIDEL, T.; LANGER, T. LigandScout Remote: A New User-Friendly Interface for HPC and Cloud Resources. **J Chem Inf Model**, 59, n. 1, p. 31-37, Jan 28 2019.

KELLEY, B. P.; BROWN, S. P.; WARREN, G. L.; MUCHMORE, S. W. POSIT: Flexible Shape-Guided Docking For Pose Prediction. **J Chem Inf Model**, 55, n. 8, p. 1771-1780, Aug 24 2015.

KIM, D. S.; KIM, C. M.; WON, C. I.; KIM, J. K. *et al.* BetaDock: shape-priority docking method based on beta-complex. **J Biomol Struct Dyn**, 29, n. 1, p. 219-242, Aug 2011.

KIN. **Mind the Byte** <https://www.mindthebyte.com/>, 2018. Acesso em: 15/07/2019.

KOJIMA, K.; KONOPLEVA, M.; MCQUEEN, T.; O'BRIEN, S. *et al.* Mdm2 inhibitor Nutlin-3a induces p53-mediated apoptosis by transcription-dependent and transcription-independent mechanisms and may overcome Atm-mediated resistance to fludarabine in chronic lymphocytic leukemia. **Blood**, 108, n. 3, p. 993-1000, Aug 1 2006.

KOLB, P.; CAFLISCH, A. Automatic and efficient decomposition of two-dimensional structures of small molecules for fragment-based high-throughput docking. **J Med Chem**, 49, n. 25, p. 7384-7392, Dec 14 2006.

KORB, O.; STUTZLE, T.; EXNER, T. E. PLANTS: Application of Ant Colony Optimization to Structure-Based Drug Design. **Lecture Notes in Computer Science**, p. 247-258, 2006.

KUDO, N.; MATSUMORI, N.; TAOKA, H., 1999, **Leptomycin B inactivates CRM1/exportin 1 by covalent modification at a cysteine residue in the central conserved region**. 9112-9117.

KUIJER, M. L.; HOGENDOORN, P. C.; CLETON-JANSEN, A. M. Genome-wide analyses on high-grade osteosarcoma: making sense of a genomically most unstable tumor. **Int J Cancer**, 133, n. 11, p. 2512-2521, Dec 1 2013.

KUNTZ, I. D.; BLANEY, J. M.; OATLEY, S. J.; LANGRIDGE, R. *et al.* A geometric approach to macromolecule-ligand interactions. **J Mol Biol**, 161, n. 2, p. 269-288, Oct 25 1982.

LAM, P. C.; ABAGYAN, R.; TOTROV, M. Ligand-biased ensemble receptor docking (LigBEnD): a hybrid ligand/receptor structure-based approach. **J Comput Aided Mol Des**, 32, n. 1, p. 187-198, Jan 2018.

LAPALOMBELLA, R.; SUN, Q.; WILLIAMS, K.; TANGEMAN, L. *et al.* Selective inhibitors of nuclear export show that CRM1/XPO1 is a target in chronic lymphocytic leukemia. **Blood**, 120, n. 23, p. 4621-4634, Nov 29 2012.

LAUNAY, G.; TELETSCHEA, S.; WADE, F.; PAJOT-AUGY, E. *et al.* Automatic modeling of mammalian olfactory receptors and docking of odorants. **Protein Eng Des Sel**, 25, n. 8, p. 377-386, Aug 2012.

LEADIT. **LeadIT version 2.3.2, BioSolveIT GmbH**. Sankt Augustin, Germany, www.biosolveit.de/LeadIT, 2017. Acesso em: 23/07/2019.

LEDOCK. **Lephar Research, Molecular Docking** <http://www.lephar.com/software.htm>, 2014 - 2015. Acesso em: 15/07/2019.

LEE, G. R.; SEOK, C. Galaxy7TM: flexible GPCR-ligand docking by structure refinement. **Nucleic Acids Res**, 44, n. W1, p. W502-506, Jul 8 2016.

LEE, H. S.; ZHANG, Y. BSP-SLIM: a blind low-resolution ligand-protein docking approach using predicted protein structures. **Proteins**, 80, n. 1, p. 93-110, Jan 2012.

LEON-GONZALEZ, A. J.; ACERO, N.; MUNOZ-MINGARRO, D.; NAVARRO, I. *et al.* Chalcones as Promising Lead Compounds on Cancer Therapy. **Curr Med Chem** 22, n. 30, p. 3407-3425, 2015.

LI, H.; LEUNG, K.-S.; WONG, M.-H. idock: A multithreaded virtual screening tool for flexible ligand docking. **2012 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)**, 2012.

LI, L.; BUM-ERDENE, K.; BAENZIGER, P. H.; ROSEN, J. J. *et al.* BioDrugScreen: a computational drug design resource for ranking molecules docked to the human proteome. **Nucleic Acids Res**, 38, n. Database issue, p. D765-773, Jan 2010.

LIEW S. K.; MALAGOBADAN, S; ARSHAD, N. M.; NAGOOR, N. H. A Review of the Structure-Activity Relationship of Natural and Synthetic Antimetastatic Compounds. **Biomolecules**, 10, n. 1, p. 138, Jan 14 2020.

LIN, H.; SIU, S. W. I. A Hybrid Cuckoo Search and Differential Evolution Approach to Protein(-)Ligand Docking. **Int J Mol Sci**, 19, n. 10, Oct 15 2018.

LIN, T.-H.; LIN, G.-L. An Anchor-Dependent Molecular Docking Process for Docking Small Flexible Molecules into Rigid Protein Receptors. **Journal of Chemical Information and Modeling**, 48, n. 8, p. 1638-1655, 2008.

LIU, M.; WANG, S. MCDOCK: a Monte Carlo simulation approach to the molecular docking problem. **J Comput Aided Mol Des**, 13, n. 5, p. 435-451, Sep 1999.

LIU, Y.; ZHAO, L.; LI, W.; ZHAO, D. *et al.* FIPSDock: a new molecular docking technique driven by fully informed swarm optimization algorithm. **J Comput Chem**, 34, n. 1, p. 67-75, Jan 5 2013.

LONDON, N.; MILLER, R. M.; KRISHNAN, S.; UCHIDA, K. *et al.* Covalent docking of large libraries for the discovery of chemical probes. **Nat Chem Biol**, 10, n. 12, p. 1066-1072, Dec 2014.

LYSKOV, S.; CHOU, F.-C.; CONCHÚIR, S. Ó.; DER, B. S. *et al.* Serverification of Molecular Modeling Applications: The Rosetta Online Server That Includes Everyone (ROSIE). **PLoS ONE**, 8, n. 5, 2013.

- MACINDOE, G.; MAVRIDIS, L.; VENKATRAMAN, V.; DEVIGNES, M. D. *et al.* HexServer: an FFT-based protein docking server powered by graphics processors. **Nucleic Acids Res**, 38, n. Web Server issue, p. W445-449, Jul 2010.
- MAHAPATRA, D. K.; BHARTI, S. K.; ASATI, V. Anti-cancer chalcones: Structural and molecular target perspectives. **Eur J Med Chem**, 98, p. 69-114, Jun 15 2015.
- MAHBOOBI, S. H.; JAVANPOUR, A. A.; MOFRAD, M. R. The interaction of RNA helicase DDX3 with HIV-1 Rev-CRM1-RanGTP complex during the HIV replication cycle. **PLoS One**, 10, n. 2, p. e0112969, 2015.
- MAJEUX, N.; SCARSI, M.; APOSTOLAKIS, J.; EHRHARDT, C. *et al.* Exhaustive docking of molecular fragments with electrostatic solvation. **Proteins**, 37, n. 1, p. 88-105, Oct 1 1999.
- MARIALKE, J.; TIETZE, S.; APOSTOLAKIS, J. Similarity based docking. **J Chem Inf Model**, 48, n. 1, p. 186-196, Jan 2008.
- MATTHEWS, N.; KITAO, A.; LAYCOCK, S.; HAYWARD, S. Haptic-Assisted Interactive Molecular Docking Incorporating Receptor Flexibility. **J Chem Inf Model**, 59, n. 6, p. 2900-2912, Jun 24 2019.
- MCGANN, M. FRED pose prediction and virtual screening accuracy. **J Chem Inf Model**, 51, n. 3, p. 578-596, Mar 28 2011.
- MCGANN, M. FRED and HYBRID docking performance on standardized datasets. **J Comput Aided Mol Des**, 26, n. 8, p. 897-906, Aug 2012.
- MCULE. **1-Click Docking**. <https://mcule.com/apps/1-click-docking/>, 2011. Acesso em: 25/06/2019.
- MEIER, R.; PIPPEL, M.; BRANDT, F.; SIPPL, W. *et al.* ParaDockS: a framework for molecular docking with population-based metaheuristics. **J Chem Inf Model**, 50, n. 5, p. 879-889, May 24 2010.
- MEILER, J.; BAKER, D. ROSETTALIGAND: protein-small molecule docking with full side-chain flexibility. **Proteins**, 65, n. 3, p. 538-548, Nov 15 2006.
- MENG, E. C.; SHOICHET, B. K.; KUNTZ, I. D. Automated docking with grid-based energy evaluation. **J. Comput. Chem.** , 13, p. 505-524, 1992.
- MIRABELLO, L.; TROISI, R. J.; SAVAGE, S. A. Osteosarcoma incidence and survival rates from 1973 to 2004: data from the Surveillance, Epidemiology, and End Results Program. **Cancer**, 115, n. 7, p. 1531-1543, Apr 1 2009.
- MOE. **Chemical Computing Group, MOE**. MoE Software, Montreal, Quebec, Canada, 2003.
- MOITESSIER, N.; POTTEL, J.; THERRIEN, E.; ENGLEBIENNE, P. *et al.* Medicinal Chemistry Projects Requiring Imaginative Structure-Based Drug Design Methods. **Acc Chem Res**, 49, n. 9, p. 1646-1657, Sep 20 2016.

MONECKE, T.; HASELBACH, D.; VOSS, B.; RUSSEK, A. *et al.* Structural basis for cooperativity of CRM1 export complex formation. **Proc Natl Acad Sci U S A**, 110, n. 3, p. 960-965, Jan 15 2013.

MORRIS, G. M.; GOODSELL, D. S.; HALLIDAY, R. S.; HUEY, R. *et al.* Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. **Journal of Computational Chemistry**, 19, p. 1639-1662, 1998.

MUNIZ, H. S.; NASCIMENTO, A. S. Ligand- and receptor-based docking with LiBELa. **J Comput Aided Mol Des**, 29, n. 8, p. 713-723, Aug 2015.

MUTSAERS, A. J.; WALKLEY, C. R. Cells of origin in osteosarcoma: mesenchymal stem cells or osteoblast committed cells? **Bone**, 62, p. 56-63, May 2014.

NAMASIVAYAM, V.; GUNTHER, R. *ps@autodock*: a fast flexible molecular docking program based on Swarm intelligence. **Chem Biol Drug Des**, 70, n. 6, p. 475-484, Dec 2007.

NG, M. C.; FONG, S.; SIU, S. W. *PSOVina*: The hybrid particle swarm optimization algorithm for protein-ligand docking. **J Bioinform Comput Biol**, 13, n. 3, p. 1541007, Jun 2015.

NGUYEN, K. T.; HOLLOWAY, M. P.; ALTURA, R. A. The CRM1 nuclear export protein in normal development and disease. **Int J Biochem Mol Biol**, 3, n. 2, p. 137-151, 2012.

NIU, M.; WU, S.; MAO, L.; YANG, Y. *Crm1* is a cellular target of curcumin: new insights for the myriad of biological effects of an ancient spice **Traffic**, 14, n. 10, p. 1042-1052, 2013.

OMS. **Cancer**. p. <https://www.who.int/news-room/fact-sheets/detail/cancer>, 2018. Acesso em: 20/01/2020.

ONCOGUIA. **Entenda o que é câncer metastático**. <http://www.oncoguia.org.br/conteudo/entenda-o-que-e-cancer-metastatico/3186/357/>, 2013. Acesso em: 15/09/2017.

ONCOGUIA. **Causas do Câncer**. <http://www.oncoguia.org.br/conteudo/causas-do-cancer/80/1/>, 2017. Acesso em: 20/01/2020.

OPENZICA. **Triagem virtual**. <http://openzika.ufg.br/vs/?lang=pt-br>, 2016. Acesso em: 28/04/2017.

OTTAVIANI, G.; JAFFE, N. The epidemiology of osteosarcoma. **Cancer Treat Res**, 152, p. 3-13, 2009.

PANG, Y. P.; PEROLA, E.; XU, K.; PRENDERGAST, F. G. *EUDOC*: a computer program for identification of drug interaction sites in macromolecules and drug leads from chemical databases. **J Comput Chem**, 22, n. 15, p. 1750-1771, Nov 30 2001.

PATEL, H.; BRINKJOST, T.; KOCH, O. PyGOLD: a python based API for docking based virtual screening workflow generation. **Bioinformatics**, 33, n. 16, p. 2589-2590, Aug 15 2017.

PEDRETTI, A.; VISTOLI, G. **Vega online**. <https://nova.disfarm.unimi.it/index.ebd>, 2004-2019. Acesso em: 25/07/2019.

PEI, J.; WANG, Q.; LIU, Z.; LI, Q. *et al.* PSI-DOCK: towards highly efficient and accurate flexible ligand docking. **Proteins**, 62, n. 4, p. 934-946, Mar 1 2006.

PHAM, T. A.; JAIN, A. N. Parameter Estimation for Scoring Protein-Ligand Interactions Using Negative Training Data. **Journal of Medicinal Chemistry**, 49, n. 20, p. 5856–5868, 2006.

PLEWCZYNSKI, D.; LAZNIEWSKI, M.; VON GROTHUSS, M.; RYCHLEWSKI, L. *et al.* VoteDock: consensus docking method for prediction of protein-ligand interactions. **J Comput Chem**, 32, n. 4, p. 568-581, Mar 2011.

PRONK, S.; PALL, S.; SCHULZ, R.; LARSSON, P. *et al.* GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. **Bioinformatics**, 29, n. 7, p. 845-854, Apr 1 2013.

QI, M.; ZHANG, J.; ZENG, W.; CHEN, X. DNAJB1 stabilizes MDM2 and contributes to cancer cell proliferation in a p53-dependent manner. **Biochim Biophys Acta**, 1839, n. 1, p. 62-69, Jan 2014.

RAREY, M.; KRAMER, B.; LENGAUER, T.; KLEBE, G. A Fast Flexible Docking Method using an Incremental Construction Algorithm. **Journal of Molecular Biology**, 261, n. 3, p. 470–489, 1996.

RAVINDRANATH, P. A.; FORLI, S.; GOODSSELL, D. S.; OLSON, A. J. *et al.* AutoDockFR: Advances in Protein-Ligand Docking with Explicitly Specified Binding Site Flexibility. **PLoS Comput Biol**, 11, n. 12, p. e1004586, Dec 2015.

RCSB. **RCSB PDB, Protein Data Bank**. p. <https://www.rcsb.org/>, 2020. Acesso em: 29/01/2020.

RITCHIE, D. W.; KEMP, G. J. L. Protein docking using spherical polar Fourier correlations. **Proteins: Structure, Function, and Genetics**, 39, n. 2, p. 178–194, 2000.

RITCHIE, D. W.; VENKATRAMAN, V. Ultra-fast FFT protein docking on graphics processors. **Bioinformatics**, 26, n. 19, p. 2398-2405, Oct 1 2010.

SALENTIN, S.; SCHREIBER, S.; HAUPT, V. J.; ADASME, M. F. *et al.* PLIP: fully automated protein-ligand interaction profiler. **Nucleic Acids Res**, 43, n. W1, p. W443-447, Jul 1 2015.

SANNER, M. F. Python: A programming language for software integration and development. **Journal of Molecular Graphics and Modelling**, 17, p. 57-61, 1999.

SANTOS-MARTINS, D.; FORLI, S.; RAMOS, M. J.; OLSON, A. J. AutoDock4(Zn): an improved AutoDock force field for small-molecule docking to zinc metalloproteins. **J Chem Inf Model**, 54, n. 8, p. 2371-2379, Aug 25 2014.

SANTOS, M. B.; PINHANELLI, V. C.; GARCIA, M. A. R.; SILVA, G. *et al.* Antiproliferative and pro-apoptotic activities of 2'- and 4'-aminochalcones against tumor canine cells. **Eur J Med Chem**, 138, p. 884-889, Sep 29 2017.

SASTRY, J.; KELLIE, S. J. Severe neurotoxicity, ototoxicity and nephrotoxicity following high-dose cisplatin and amifostine. **Pediatr Hematol Oncol**, 22, n. 5, p. 441-445, Jul-Aug 2005.

SAUTON, N.; LAGORCE, D.; VILLOUTREIX, B. O.; MITEVA, M. A. MS-DOCK: accurate multiple conformation generator and rigid docking protocol for multi-step virtual ligand screening. **BMC Bioinformatics**, 9, p. 184, Apr 10 2008.

SAVAGE, S. A.; MIRABELLO, L. Using epidemiology and genomics to understand osteosarcoma etiology. **Sarcoma**, 2011, p. 548151, 2011.

SCALBERT, A.; MANACH, C.; MORAND, C.; REMESY, C. *et al.* Dietary polyphenols and the prevention of diseases. **Crit Rev Food Sci Nutr**, 45, n. 4, p. 287-306, 2005.

SCHNECKE, V.; KUHN, L. A. Database screening for HIV protease ligands: the influence of binding-site conformation and representation on ligand selectivity. **Proc Int Conf Intell Syst Mol Biol**, p. 242-251, 1999.

SCHNECKE, V.; KUHN, L. A. Virtual screening with solvation and ligand-induced complementarity. **Perspect. Drug Discov. Design**, 20, p. 171-190, 2000.

SCHNEIDMAN-DUHOVNY, D.; INBAR, Y.; NUSSINOV, R.; WOLFSON, H. J. PatchDock and SymmDock: servers for rigid and symmetric docking. **Nucleic Acids Res**, 33, n. Web Server issue, p. W363-367, Jul 1 2005.

SCORE. **Drug Design Laboratory**. <https://nova.disfarm.unimi.it/cms/>, 2019. Acesso em: 25/07/2019.

SEBA, V.; SILVA, G.; SANTOS, M. B. D.; BAEK, S. J. *et al.* Chalcone Derivatives 4'-Amino-1-Naphthyl-Chalcone (D14) and 4'-Amino-4-Methyl-1-Naphthyl-Chalcone (D15) Suppress Migration and Invasion of Osteosarcoma Cells Mediated by p53 Regulating EMT-Related Genes. **Int J Mol Sci**, 19, n. 9, Sep 19 2018.

SHADFAN, M.; LOPEZ-PAJARES, V.; YUAN, Z. M. MDM2 and MDMX: Alone and together in regulation of p53. **Transl Cancer Res**, 1, n. 2, p. 88-89, Aug 2012.

SHIN, W. H.; HEO, L.; LEE, J.; KO, J. *et al.* LigDockCSA: protein-ligand docking using conformational space annealing. **J Comput Chem**, 32, n. 15, p. 3226-3232, Nov 30 2011.

SHIN, W. H.; SEOK, C. GalaxyDock: protein-ligand docking with flexible protein side-chains. **J Chem Inf Model**, 52, n. 12, p. 3225-3232, Dec 21 2012.

- SHOLZ, C.; KNORR, S.; HAMACHER, K.; SCHMIDT, B. DOCKTITE-a highly versatile step-by-step workflow for covalent docking and virtual screening in the molecular operating environment. **J Chem Inf Model**, 55, n. 2, p. 398-406, Feb 23 2015.
- SIEGEL, R. L.; MILLER, K. D.; JEMAL, A. Cancer statistics, 2019. **CA Cancer J Clin**, 69, n. 1, p. 7-34, Jan 2019.
- SIKANDER, M.; MALIK, S.; YADAV, D.; BISWAS, S. *et al.* Cytoprotective activity of a trans-chalcone against hydrogen peroxide induced toxicity in hepatocellular carcinoma (HepG2) cells. **Asian Pac J Cancer Prev**, 12, n. 10, p. 2513-2516, 2011.
- SILVA, G.; FACHIN, A. L.; BELEBONI, R. O.; FRANCA, S. C. *et al.* In vitro action of flavonoids in the canine malignant histiocytic cell line DH82. **Molecules**, 18, n. 12, p. 15448-15463, Dec 12 2013.
- SILVA, G.; MARINS, M.; CHAICHANASAK, N.; YOON, Y. *et al.* Trans-chalcone increases p53 activity via DNAJB1/HSP40 induction and CRM1 inhibition. **PLoS One** 13, n. 8, p. e0202263, 2018.
- SILVA, G.; MARINS, M.; FACHIN, A. L.; LEE, S. H. *et al.* Anti-cancer activity of trans-chalcone in osteosarcoma: Involvement of Sp1 and p53. **Mol Carcinog**, 55, n. 10, p. 1438-1448, Oct 2016.
- SINGH, T.; BISWAS, D.; JAYARAM, B. AADS--an automated active site identification, docking, and scoring protocol for protein targets based on physicochemical descriptors. **J Chem Inf Model**, 51, n. 10, p. 2515-2527, Oct 24 2011.
- SMIESKO, M. DOLINA - docking based on a local induced-fit algorithm: application toward small-molecule binding to nuclear receptors. **J Chem Inf Model**, 53, n. 6, p. 1415-1423, Jun 24 2013.
- SOBOLEV, V.; WADE, R. C.; VRIEND, G.; EDELMAN, M. Molecular docking using surface complementarity. **Proteins** 25, p. 120-129, 1996.
- SOCIETY, A. C. <https://www.cancer.org/cancer/osteosarcoma/about/what-is-osteosarcoma.html>.
- SOUSA, S. F.; FERNANDES, P. A.; RAMOS, M. J. Protein-ligand docking: current status and future challenges. **Proteins**, 65, n. 1, p. 15-26, Oct 1 2006.
- SPITZER, R.; JAIN, A. N. Surflex-Dock: Docking benchmarks and real-world application. **J Comput Aided Mol Des**, 26, n. 6, p. 687-699, Jun 2012.
- STEWART, J. J. Optimization of parameters for semiempirical methods V: modification of NDDO approximations and application to 70 elements. **J Mol Model**, 13, n. 12, p. 1173-1213, Dec 2007.
- STIERAND, K.; MAASS, P. C.; RAREY, M. Molecular complexes at a glance: automated generation of two-dimensional complex diagrams. **Bioinformatics**, 22, n. 14, p. 1710-1716, 2006.

STIERAND, K.; RAREY, M. From modeling to medicinal chemistry: automatic generation of two-dimensional complex diagrams. **ChemMedChem**, 2, n. 6, p. 853-860, Jun 2007.

STROGANOV, O. V.; NOVIKOV, F. N.; STROYLOV, V. S.; KULKOV, V. *et al.* Lead finder: an approach to improve accuracy of protein-ligand docking, binding energy estimation, and virtual screening. **J Chem Inf Model**, 48, n. 12, p. 2371-2385, Dec 2008.

TAKETANI, K.; KAWAUCHI, J.; TANAKA-OKAMOTO, M.; ISHIZAKI, H. *et al.* Key role of ATF3 in p53-dependent DR5 induction upon DNA damage of human colon cancer cells. **Oncogene** 31, n. 17, p. 2210-2221, 2012.

TAYLOR, J. S.; BURNETT, R. M. DARWIN: a program for docking flexible molecules. **Proteins**, 41, n. 2, p. 173-191, Nov 1 2000.

TAYLOR, R. D.; JEWSBURY, P. J.; ESSEX, J. W. FDS: flexible ligand and receptor docking with a continuum solvent model and soft-core energy function. **J Comput Chem**, 24, n. 13, p. 1637-1656, Oct 2003.

THOMPSON, M. A. **Molecular docking using ArgusLab, an efficient shape-based search algorithm and the AScore scoring function.** ACS meeting, Philadelphia, 172, CINF 42, PA. 2004.

THOMSEN, R.; CHRISTENSEN, M. H. MolDock: a new technique for high-accuracy molecular docking. **J Med Chem**, 49, n. 11, p. 3315-3321, Jun 1 2006.

TIETZE, S.; APOSTOLAKIS, J. GlamDock: development and validation of a new docking tool on several thousand protein-ligand complexes. **J Chem Inf Model**, 47, n. 4, p. 1657-1672, Jul-Aug 2007.

TROSSET, J.-Y.; SCHERAGA, H. A. Prodock: Software package for protein modeling and docking. **Journal of Computational Chemistry**, 20, n. 4, p. 412-427, 1999.

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

TSAI, T. Y.; CHANG, K. W.; CHEN, C. Y. iScreen: world's first cloud-computing web server for virtual screening and de novo drug design based on TCM database@Taiwan. **J Comput Aided Mol Des**, 25, n. 6, p. 525-531, Jun 2011.

TSENG, C. H.; TZENG, C. C.; CHIU, C. C.; HSU, C. Y. *et al.* Discovery of 2-[2-(5-nitrofuranyl)vinyl]quinoline derivatives as a novel type of antimetastatic agents. **Bioorg Med Chem**, 23, n. 1, p. 141-8, Jan 2015.

VAQUE, M.; AROLA, A.; ALIAGAS, C.; PUJADAS, G. BDT: an easy-to-use front-end application for automation of massive docking tasks and complex docking strategies with AutoDock. **Bioinformatics**, 22, n. 14, p. 1803-1804, Jul 15 2006.

VASSILEV, L. T.; VU, B. T.; GRAVES, B.; CARVAJAL, D. *et al.* In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. **Science**, 303, n. 5659, p. 844-848, Feb 6 2004.

VENKATACHALAM, C. M.; JIANG, X.; OLDFIELD, T.; WALDMAN, M. LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. **J Mol Graph Model**, 21, n. 4, p. 289-307, Jan 2003.

VISTOLI, G.; PEDRETTI, A.; MAZZOLARI, A.; TESTA, B. Homology modeling and metabolism prediction of human carboxylesterase-2 using docking analyses by GriDock: a parallelized tool based on AutoDock 4.0. **J Comput Aided Mol Des**, 24, n. 9, p. 771-787, Sep 2010.

VLIFEMDS. **VLifeDock, VLife Sciences Technologies Pvt. Ltd.** p. <https://www.vlifesciences.com/>, 2013. Acesso em: 07/08/2019.

WAGENER, M.; DE VLIEG, J.; NABUURS, S. B. Flexible protein-ligand docking using the Fleksy protocol. **Journal of Computational Chemistry**, 33, p. 1215–1217, 2012.

WALLACE, A. C.; LASKOWSKI, R. A.; THOMTON, J. M. LIGPLOT: A program to generate schematic diagrams of protein-ligand interaction. **Protein Engineering**, 8, p. 127-134, 1995.

WANG, J. C.; CHU, P. Y.; CHEN, C. M.; LIN, J. H. idTarget: a web server for identifying protein targets of small chemical molecules with robust scoring functions and a divide-and-conquer docking approach. **Nucleic Acids Res**, 40, n. Web Server issue, p. W393-399, Jul 2012.

WANG, Z.; SUN, H.; YAO, X.; LI, D. *et al.* Comprehensive evaluation of ten docking programs on a diverse set of protein-ligand complexes: the prediction accuracy of sampling power and scoring power. **Phys Chem Chem Phys**, 18, n. 18, p. 12964-12975, May 14 2016.

WEI, S.; WANG, H.; LU, C.; MALMUT, S. *et al.* The activating transcription factor 3 protein suppresses the oncogenic function of mutant p53 proteins. **J Biol Chem**, 289, n. 13, p. 8947-8959, Mar 28 2014.

WELCH, W.; RUPPERT, J.; JAIN, A. N. Hammerhead: fast, fully automated docking of flexible ligands to protein binding sites. **Chem Biol**, 3, n. 6, p. 449-462, Jun 1996.

WESTHEAD, D. R.; CLARK, D. E.; MURRAY, C. W. Flexible docking using Tabu search and an empirical estimate of binding affinity. **Proteins**, p. 367 - 382, 1998.

WOLBER, G.; LANGER, T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. **J Chem Inf Model**, 45, n. 1, p. 160-169, Jan-Feb 2005.

WU, G.; ROBERTSON, D. H.; BROOKS, C. L., 3rd; VIETH, M. Detailed analysis of grid-based molecular docking: A case study of CDOCKER-A CHARMM-based MD docking algorithm. **J Comput Chem**, 24, n. 13, p. 1549-1562, Oct 2003.

XAVIER, M. M.; HECK, G. S.; AVILA, M. B.; LEVIN, N. M. B. *et al.* SAnDReS a Computational Tool for Statistical Analysis of Docking Results and Development of Scoring Functions. **Comb Chem High Throughput Screen**, 19, n. 10, p. 801-812, 2016.

YAN, C.; LU, D.; HAI, T.; BOYD, D. D. Activating transcription factor 3, a stress sensor, activates p53 by blocking its ubiquitination. **EMBO J**, 24, n. 13, p. 2425-2435, Jul 6 2005.

YANG, J. M.; CHEN, C. C. GEMDOCK: a generic evolutionary method for molecular docking. **Proteins**, 55, n. 2, p. 288-304, May 1 2004.

YE, S.; SHEN, J.; CHOY, E.; YANG, C. *et al.* p53 overexpression increases chemosensitivity in multidrug resistant osteosarcoma cell lines. **Cancer chemotherapy and pharmacology**, 77, n. 2, p. 349-356, 12/23 2016.

ZACHARIAS, M. Rapid protein-ligand docking using soft modes from molecular dynamics simulations to account for protein deformability: binding of FK506 to FKBP. **Proteins**, 54, n. 4, p. 759-767, Mar 1 2004.

ZHAO, Y.; SANNER, M. F. FLIPDock: docking flexible ligands into flexible receptors. **Proteins**, 68, n. 3, p. 726-737, Aug 15 2007.

ZHOU, B.; XING, C. Diverse Molecular Targets for Chalcones with Varied Bioactivities. **Med Chem (Los Angeles)**, 5, n. 8, p. 388-404, Aug 2015.

ZHU, K.; BORRELLI, K. W.; GREENWOOD, J. R.; DAY, T. *et al.* Docking covalent inhibitors: a parameter free approach to pose prediction and scoring. **J Chem Inf Model**, 54, n. 7, p. 1932-1940, Jul 28 2014.

ZSOLDOS, Z.; REID, D.; SIMON, A.; SADJAD, S. B. *et al.* eHiTS: a new fast, exhaustive flexible ligand docking system. **J Mol Graph Model**, 26, n. 1, p. 198-212, Jul 2007.

APÊNDICE

Programas de *Docking* Molecular entre Proteína-Ligante

Na Tabela 15, são apresentados programas de *docking* molecular para simulações entre proteína e ligante encontrados na literatura. Nessa tabela, *indisp.* indicam programas cuja licença encontra-se indisponível, *NR* refere-se ao número de artigos encontrados na revisão sistemática entre os anos de 2015 e 2019 e *NH* indica que não há citações nesse período. Para definir os programas incluídos nessa tabela, foi realizada uma busca na base de dados Scopus com os termos TITLE-ABS-KEY (molecular AND docking AND programs) AND (LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2017) OR LIMIT-TO (PUBYEAR , 2016) OR LIMIT-TO (PUBYEAR , 2015)) AND (LIMIT-TO (LANGUAGE , "English")). A busca se restringiu a título, resumo e palavras-chave e o resultado dessa busca gerou 959 artigos. Os resumos foram lidos e todos os artigos que citavam ou descreviam um novo programa de *docking* molecular foram selecionados e constam nessa tabela. Ainda, como forma de incluir outros programas não disponibilizados nessa base, foram utilizados os resultados encontrados no Google Acadêmico e as informações de programas de *docking* disponibilizadas no Click2Drug⁷. Ao término dessa busca, foram selecionados 122 programas de *docking* molecular. Embora alguns programas apresentem diferentes citações, neste trabalho é informada a primeira versão encontrada na literatura, que consta após o nome do programa, na primeira coluna da Tabela 15, no campo Nome do Programa.

Na segunda coluna da Tabela 15 (Campo Descrição), constam informações relevantes sobre o programa considerado. Ainda, quando ocorrem, descrições sobre aplicações específicas também são discriminadas nessa mesma tabela. Na terceira coluna (Campo Licença), é informado se esses programas são *webservice* ou para *download* e se são gratuitos ou pagos. Quando não encontrados para *download* ou em sua versão *webservice*, essa coluna contém a informação “Indisponível”.

Para preenchimento da quarta coluna da Tabela 15 (campo NA), na seleção do número de usos dos programas de *docking* disponíveis na primeira coluna, fez-se uma filtragem de uso entre os anos de 2015 e 2019 e foram considerados apenas trabalhos em inglês. A busca foi realizada na base de dados Scopus e foi utilizada a seguinte chave de busca: TITLE-ABS-KEY (Nome do programa) AND (LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR ,

⁷ Click2Drug: disponível em <http://click2drug.org/>

2018) OR LIMIT-TO (PUBYEAR , 2017) OR LIMIT-TO (PUBYEAR , 2016) OR LIMIT-TO (PUBYEAR , 2015)) AND (LIMIT-TO (LANGUAGE , "English")). O campo “Nome do programa” utilizado para essa busca segue conforme apresentado na primeira coluna da Tabela 15. Programas cujo retorno na base de dados Scopus excedeu 1000 artigos, para realizar a filtragem dos trabalhos de interesse, utilizou-se a chave de busca "docking". Caso o retorno ainda excedesse 1000 respostas, uma nova filtragem foi realizada com o nome completo do programa, caso existisse. Caso contrário, o refinamento foi realizado de acordo com o(s) sobrenome(s) do(s) autor(es) do programa. As buscas consideraram somente trabalhos em que a busca apresentou resultados no título, nas palavras-chave e/ou no *abstract*.

Tabela 15: Programas para simulações computacionais de *docking* molecular em complexos contendo proteína-ligante. Nessa tabela, *indisp.* indicam programas cuja licença encontra-se indisponível, *NR* refere-se ao número de artigos encontrados na revisão sistemática entre os anos de 2015 e 2019 e *NH* indica que não há citações nesse período.

Nome do Programa	Descrição	Licença	NR
1-Click Docking (MCULE, 2011)	Aplicativo <i>web</i> disponibilizado pela Mcule, Inc., que permite realizar o desenho do ligante no próprio programa ou fazer seu <i>upload</i> . A estrutura bidimensional do ligante é convertida para uma tridimensional equivalente de forma automática e calcula-se a interação do ligante com a molécula-alvo.	<i>Webservice</i> gratuito, disponível em mcule.com/apps/1-click-docking/	2
AADS (SINGH; BISWAS; JAYARAM, 2011)	Programa proposto para identificação de sítio ativo da proteína, para cálculo de <i>docking</i> molecular e para descrever protocolos de pontuação (<i>scoring</i>) para proteínas que possuem a estrutura conhecida. Para o cálculo de <i>docking</i> molecular, é empregado o método de Monte Carlo e são avaliadas as dez principais cavidades da proteína, de acordo com seu sítio ativo.	<i>Webservice</i> gratuito, disponível em www.scfbio-iitd.res.in/dock/ActiveSite_new.jsp	NH
ADAM (IMMD, 2011)	ADAM é uma ferramenta de <i>docking</i> molecular disponibilizada pela IMMD Inc. e que prediz a interação entre ligantes flexíveis e a macromoléculas de interesse. A IMMD Inc. possui outras ferramentas computacionais que auxiliam na descoberta de fármacos, tais como a Key3D, que é utilizada para converter estruturas bidimensionais, na sua fórmula estrutural, para estruturas tridimensionais precisas, e a LEGEND, que pode gerar novos ligantes candidatos, com base na estrutura tridimensional da macromolécula-alvo.	Programa pago, disponível em http://www.immd.co.jp/en/product_2.html	NH
Adaptive BP-Dock (BOLIA; OZKAN, 2016)	No programa Adaptive BP-Dock, são realizadas perturbações no sítio de ligação do receptor e é obtida uma nova conformação para o mesmo, que é baseada na resposta de	<i>Indisp.</i>	1

flutuação dos resíduos. Utilizando-se dessa nova conformação, é medida a interação entre proteína e ligante, empregando-se o programa RosettaLigand, que é descrito adiante, para esses cálculos.

Achilles Blind Docking Server (BIO-HPC, 2017)	Programa proposto para realização de <i>docking molecular</i> com interface simples, sem que haja conhecimento inicial sobre o sítio de interação da proteína. Como somente a estrutura do ligante e do receptor são necessários, a caixa de busca inicial é dividida em múltiplas caixas, com a finalidade de reduzir o espaço de busca do sítio de interação.	Webservice gratuito, disponível em https://bio-hpc.ucam.edu/achilles/	NH
ADDock (LIN; LIN, 2008)	Programa de <i>docking molecular</i> indicado para simulações com pequenas moléculas, que são flexíveis no programa, e proteína, que é mantida rígida. No ADDock, utiliza-se o método <i>anchor-and-grow</i> , em que o ligante é dividido em partes rígidas e flexíveis. As partes rígidas são definidas como âncoras e são utilizadas inicialmente para as simulações de <i>docking</i> . Em seguida, as partes flexíveis são adicionadas a ele, observando-se os pontos de torção do ligante. No decorrer da simulação, os pontos definidos como âncoras são alterados e a melhor conformação é retornada.	Programa pago e detalhes estão disponíveis em http://www.biodelight.com.tw/	NH
ArgusDock (THOMPSON, 2004)	O ArgusDock é uma ferramenta para simulações de <i>docking molecular</i> e que é disponibilizado pelo ArgusLab4.0, programa que também possui ferramentas de modelagem molecular, de visualização gráfica e voltadas para o desenvolvimento de fármacos para uso no sistema operacional Windows. O ArgusDock implementa um algoritmo de busca exaustiva aproximada e são permitidos ligantes flexíveis e proteínas rígidas nesse programa.	Gratuito e disponível para <i>download</i> em www.arguslab.com/arguslab.com/ArgusLab.html	2
ASEDock (GOTO <i>et al.</i> , 2008)	Programa de <i>docking molecular</i> que se baseia na similaridade da forma da porção côncava da proteína e do ligante para definir as melhores conformações. Como forma de calcular a interação na ligação, utiliza-se a função de avaliação ASE, em que é combinada uma função do tipo Gaussiana para avaliar sobreposições dessas estruturas e essas sobreposições entre ligante e sítio são otimizadas.	O programa pode ser solicitado por e-mail (disponível no artigo) para seus autores	1
AutoDock (GOODSELL; OLSON, 1990)	Dentre os programas de <i>docking molecular</i> avaliados neste trabalho, o AutoDock foi o que apresentou o maior número de citações na base de dados Scopus. Entre os anos de 2001 e 2006, esse programa também foi o mais utilizado para as simulações de <i>docking molecular</i> (SOUSA; FERNANDES; RAMOS, 2006). Ao utilizar o programa AutoDock, é possível selecionar a heurística adequada para o problema. Heurísticas implementadas neste programa incluem Algoritmos Genéticos, <i>Lamarckian Genetic Algorithms</i> , <i>Simulated Annealing</i> e Busca Local. Atualmente, o programa AutoDock encontra-se na	Gratuito e disponível para <i>download</i> em https://autodock.scripps.edu/	553

versão 4.

AutoDockFR (RAVINDRANATH <i>et al.</i> , 2015)	Programa em que é utilizada uma variação da função de cálculo de interação do programa AutoDock 4. No programa AutoDockFR, os cálculos de ligação ocorrem empregando receptores flexíveis, que não são utilizados no programa AutoDock 4, e há a utilização de algoritmos genéticos.	Gratuito e disponível para <i>download</i> em http://adfr.scripps.edu/	2
AutoDock Vina (TROTT; OLSON, 2010)	Programa que utiliza Algoritmos Genéticos em sua formulação e foi proposto como uma variação para o programa AutoDock. Das variações propostas, o AutoDock Vina possui o tempo de resposta mais eficiente e melhor média de acurácia em predições de ligação que o AutoDock 4, que é a versão mais recente do AutoDock. Ainda, o programa AutoDock Vina permite uma maior rotação no ligante, se comparado ao número de rotações possível de serem realizadas no AutoDock original. Dentre os programas avaliados neste trabalho, o AutoDock Vina é o segundo com maior número de citações na base de dados Scopus.	Gratuito e disponível para <i>download</i> em http://vina.scripps.edu/download.html	445
AutoDock4Zn (SANTOS- MARTINS <i>et al.</i> , 2014)	O AutoDock4Zn utiliza as mesmas medidas de força que as empregadas no programa AutoDock. Porém, no AutoDock4Zn, há uma alteração no cálculo da interação de ligantes com o zinco. Essa proposta foi feita, de acordo com os autores, porque o zinco é importante no metabolismo para a maioria dos organismos, sendo alvo terapêutico em diversas doenças.	Extensão gratuita, que pode ser instalada com os programas AutoDock e MGLTools.	4
BDT (VAQUE <i>et al.</i> , 2006)	Esse programa foi proposto como uma variação para o programa AutoDock. A variação descrita por seus autores possui o intuito de facilitar a triagem computacional quando há uma biblioteca de ligantes, o de permitir a flexibilidade do receptor (que não ocorre no programa AutoDock) e o de realizar um <i>docking</i> cego em toda a superfície do receptor.	O programa pode ser solicitado por e-mail (disponibilizado no artigo) para os autores (não disponível para fins comerciais)	1
BetaDock (KIM <i>et al.</i> , 2011)	O programa BetaDock prioriza a complementaridade de forma entre o receptor e o ligante e baseia-se na teoria do complexo β . Através da forma β , computa-se o <i>pocket</i> que o ligante pode pertencer e utilizando as conformações dos ligantes dentro desses <i>pockets</i> , combinado a algoritmos genéticos, são determinadas as posições de <i>docking</i> molecular.	Indisponível	NH
BioDrugScreen (LI <i>et al.</i> , 2010)	No programa BioDrugScreen foi realizado um <i>ranking</i> de resultados de <i>docking</i> molecular de aproximadamente 1600 moléculas diferentes em 1589 receptores-alvo presentes no	Webservice gratuito, disponível em	1

	<p>proteoma humano e esses resultados são disponibilizados no site do programa, disponível na coluna ao lado. Desses receptores, foram calculadas as energias presentes em 1926 cavidades. Nesse programa, a interação entre esses complexos pode ser verificada por meio das funções de avaliação AutoDock, DockScore, ChemScore, X-Score, GoldScore, DFIRE ou PMF.</p>	<p>nesg.org:9090/HCPIN/index.jsp</p>	
BSP-SLIM (LEE; ZHANG, 2012)	<p>Programa de <i>docking</i> molecular cego, que inicialmente escolhe um suposto sítio de interação e utiliza um padrão apresentado por estruturas similares. A conformação do composto é então construída de maneira local e a interação é calculada através da complementaridade química entre o ligante e das imagens negativas do sítio de ligação.</p>	<p>Webservice gratuito, disponível em https://zhanglab.ccmb.med.umich.edu/BSP-SLIM/</p>	2
CDOCKER (WU <i>et al.</i> , 2003)	<p>No programa CDOCKER, a proteína é mantida rígida e o ligante é rotacionado. Para os cálculos de <i>docking</i> molecular, é utilizada a heurística <i>simulated annealing</i> e é aplicado o campo de força CHARMM nas simulações. Esse programa é disponibilizado pela BIOVIA Discovery Studio.</p>	<p>Versão básica gratuita para fins acadêmicos disponível em www.3dsbiovia.com/resource-center/downloads/freeware/index.html</p>	101
CovDock (ZHU <i>et al.</i> , 2014)	<p>O programa CovDock realiza cálculos utilizando as interações covalentes, sendo que ligações mais covalentes são as de maior interesse. Nesse programa, também é verificada a estabilidade das forças não covalentes no sítio de ligação. Esse programa é disponibilizado pela Schrodinger.</p>	<p>Programa pago, com versão gratuita de teste, disponível em www.schrodinger.com/covdock</p>	2
CuckooVina (LIN; SIU, 2018)	<p>Programa proposto como uma variante para o AutoDock Vina, que é descrito para simulações de <i>docking</i> molecular entre proteína e pequenos ligantes. Em relação às alterações, no CuckooVina é utilizada a busca cuckoo e a evolução diferencial. Segundo seus autores, quando comparado ao programa AutoDock Vina, essa alteração possibilita que o programa CuckooVina melhore os resultados de RMSD, de afinidade de ligação e de taxa de sucesso. Porém, o CuckooVina requer tempo superior entre 9% a 15% para realizar a execução das simulações, quando comparado ao AutoDock Vina.</p>	<p>Indisp.</p>	1

Cyscore (CAO; DAI; MIAO, 2018)	Para os cálculos de molecular <i>docking</i> , o programa Cyscore combina as energias hidrofóbicas livres, energia das interações de van der Waals, energia de pontes de hidrogênio e entropia do ligante. Ainda, nesse programa, rotações na conformação do ligante são permitidas e o receptor é mantido rígido.	Gratuito e disponível para <i>download</i> em http://clab.labshare.cn/software/	2
DAIM-SEED-FFLS (KOLB; CAFLISCH, 2006)	Programa em que o ligante inicial é decomposto em fragmentos, utilizando as formas mais apropriadas. Detalhes sobre essa escolha são apresentados em (KOLB; CAFLISCH, 2006). Em seguida, cada um dos fragmentos é submetido a simulações de <i>docking</i> molecular utilizando o programa SEED, que é descrito adiante neste trabalho. Com base nessas informações, os fragmentos são combinados e a melhor conformação é retornada.	Gratuito e disponível para <i>download</i> em www.biochem-caflisch.uzh.ch/download/	NH
DARWIN (TAYLOR; BURNETT, 2000)	Proposto para realizar cálculos de <i>docking</i> molecular entre proteínas e outras estruturas biológicas, como ligantes e carboidratos. O programa DARWIN utiliza o programa CHARMM (BROOKS; BRUCCOLERI; OLAFSON; STATES <i>et al.</i> , 1983) para os cálculos de energia e algoritmos genéticos como heurística. Ainda, a fim de aumentar a eficiência nos cálculos de <i>docking</i> molecular, o paralelismo é empregado.	Indisp.	NH
DIVALI (CLARK; AJAY, 1995)	No programa DIVALI (<i>Docking with evolutionary Algorithms</i>), o receptor é mantido rígido e torções são realizadas no ligante. Nesse programa, o campo de força AMBER é empregado e algoritmos genéticos são utilizados para busca na proteína.	Indisp.	NH
DOCKTITE (SHOLZ <i>et al.</i> , 2015)	Neste programa, para cálculo de interação entre receptor e ligante, são consideradas as ligações covalentes formadas entre receptor e ligante no Molecular Operating Environment (MOE). Segundo seus autores, a presença de uma ligação covalente no composto é essencial para uma série de medicamentos.	Programa pago, com versão gratuita de teste, solicitada em https://omictools.com/docktite-tool	2
DOCKoalent (LONDON; MILLER; KRISHNAN; UCHIDA <i>et al.</i> , 2014)	O DOCKoalente é uma adaptação do programa de <i>docking</i> molecular DOCK3.6., que emprega cálculos de ligações não-covalentes. No programa DOCKoalent, as simulações computacionais utilizam informações de ligações covalentes de pequenas moléculas eletrofílicas.	Webservice gratuito, disponível em http://covalent.docking.org/	NH
DOCK (MENG; SHOICHET; KUNTZ, 1992)	O programa DOCK encontra-se atualmente na versão 6 e as alterações que ocorreram nesse programa referem-se ao modo de calcular a ligação de energia mínima, o campo de força e,	Disponível para <i>download</i>	DOCK 4: 1 DOCK 6:

	em sua versão inicial, ligante e receptor eram rígidos. Atualmente, o receptor é rígido e o ligante é flexível. As versões 1, 2 e 3 possuem somente o código fonte do programa. A versão 4 é disponível para <i>download</i> para Linux e Unix. As versões 5 e 6 estão disponíveis para Linux, Windows, MacIntosh OSX.	(gratuito para fins acadêmicos e pago para organizações industriais) em http://dock.compbio.ucsf.edu/	7 Demais versões: NH
Docking at UTMB	O programa Docking at UTMB é baseado no programa AutoDock Vina e é mantido pelo laboratório Watowich, da Universidade <i>Texas Medical Branch</i>	<i>Link</i> para acesso encontra-se indisponível	NH
DockingServer (BIKADI; HAZAI, 2009)	Programa proposto para oferecer uma interface de uso fácil, em que é possível realizar simulações de <i>docking</i> molecular com um único ou com múltiplos ligantes para uma única proteína-alvo. Segundo seus autores, por combinar diferentes métodos computacionais voltados para cálculos químicos de interação, esse programa permite realizar o <i>docking</i> molecular de forma eficiente e robusta.	<i>Webservice</i> com licenças gratuitas e pagas, disponível em www.dockingserver.com/web	3
Docking Study with HyperChem (HYPERCHEM, 2004)	No programa Docking Study with HyperChem, proteínas e ácidos nucleicos podem ser utilizados como macromoléculas. Como ligantes, pequenas moléculas e peptídeos são permitidos. Segundo seus autores, por não adotar uma malha (<i>grid</i>) para cálculos, o programa pode determinar a afinidade de ligação de todo o sistema, ao invés de obter a energia aproximada de interação, que é obtida em outros programas de <i>docking</i> que fazem uso do <i>grid</i> .	Programa pago, disponível em https://www.molfunction.com/	NH
DockThor (ALMEIDA, 2011; DE MAGALHÃES, 2006)	No programa DockThor, são aplicados algoritmos genéticos para definir a posição de interesse para cálculo de afinidade entre proteína e ligante. Ainda, em análise realizada por (DARDENN, 2012), esse programa possui resultados comparáveis ao AutoDock Vina, GOLD e GLIDE, que são programas de <i>docking</i> molecular populares na literatura, especialmente para ligantes altamente flexíveis.	<i>Webservice</i> gratuito, disponível em https://dockthor.lncc.br/v2/	3
DockVision (HART; READ, 1992)	Programa para simulações de <i>docking</i> molecular, que se propõe a disponibilizar uma interface amigável ao usuário. No programa DockVision, o ligante é flexível e o receptor é rígido e são empregadas as heurísticas Monte Carlo e Algoritmos Genéticos para busca do sítio de interação.	Gratuito e disponível para <i>download</i> em http://dockvision.sness.net/	NH
DOLINA (SMIESKO,	O algoritmo implementado no programa Dolina emprega a estratégia local <i>induced-fit</i> para cálculo de <i>docking</i> molecular	Pode ser solicitado por	NH

2013)	entre receptor e ligante. Nessa estratégia, alterações estruturais são induzidas na estrutura do ligante e esses rearranjos são realizados de forma combinatória, nas cadeias laterais que revestem o local de ligação entre receptor-ligante.	e-mail para fins acadêmicos (disponível no artigo) para seu autor	
EADock (GROSDIDIER; ZOETE; MICHIELIN, 2007)	O EADock é um programa para calcular a interação entre proteína e pequenos ligantes utilizando a abordagem evolucionária. Para cálculo de energia e para manipulação estrutural, o campo de força CHARMM é empregado. Segundo seu autor, não é esperada acurácia desse programa, uma vez que ainda está em desenvolvimento. Atualmente, o programa EADock possui um sucessor, chamado EADock2.	Executável gratuito para <i>download</i> disponível em aurelien.latitude77.org/projects/eadock/gettingstarted/index.html	2
eHiTS (ZSOLDOS <i>et al.</i> , 2007)	No programa eHiTS, os ligantes são flexíveis e os receptores são rígidos. Nesse programa, é empregada a busca exaustiva em parte do espaço de busca do receptor e para determinar as possíveis conformações do ligante.	Endereço de acesso indisponível atualmente.	2
EUDOC (PANG <i>et al.</i> , 2001)	Programa de <i>docking</i> molecular automatizado, em que os receptores tridimensionais também podem ser metaloproteínas. Ainda, no programa EUDOC, é possível identificar um subconjunto líder que venha a ser utilizado na fabricação de fármacos, de acordo com informações de bases de dados químicas.	Indisp.	NH
FDS (TAYLOR; JEWSEBURY; ESSEX, 2003)	Programa indicado para realização de <i>docking</i> molecular em ligantes pequenos e receptores grandes, em que ambos, nas simulações, possuem a estrutura flexível. Esse programa utiliza a heurística de Monte Carlo para busca e é aplicado o campo de força AMBER.	Indisp.	NH
FINDSITE-LHM (BRYLINSKI; SKOLNICK, 2009)	O programa FINDSITE-LHM utiliza informações estruturais extraídas de proteínas fracamente relacionadas em relação a proteína de interesse para realizar o <i>docking</i> molecular. Desse modo, para esses cálculos, ocorre a modelagem por homologia. Segundo seus autores, proteínas relacionadas de forma remota frequentemente compartilham um local de ligação comum, que é ocupado por ligantes quimicamente semelhantes.	Informações extras encontram-se disponíveis em http://cssb.biology.gatech.edu/skolnick/files/FINDSITE_LHM/	NH
FIPSDock (LIU; ZHAO; LI; ZHAO <i>et al.</i> , 2013)	No FIPSDock, é implementada uma variante para a heurística <i>Fully Informed Particle Swarm</i> (FIPS). Como subrotina para esse programa, é adotada a função de energia utilizada no AutoDock 4.20 para resolver problemas do tipo proteína-	Indisp.	NH

ligante flexíveis.

Fitted (MOITESSIER; <i>et al.</i> , 2016)	No Fitted, proteína e ligante possuem estruturas flexíveis. Segundo (MOITESSIER; POTTEL; THERRIEN; ENGLEBIENNE <i>et al.</i> , 2016), nesse programa é permitido que as moléculas de água sejam deslocáveis e é utilizado o desenvolvimento de farmacóforos baseados no ligante. Segundo seus autores, essas abordagens aumentam a precisão do programa de <i>docking</i> molecular.	Disponível para <i>download</i> (gratuito para fins acadêmicos) em http://fitted.ca	2
Fleksy (WAGENER; DE VLIEG; NABUURS, 2012)	No programa Fleksy, a estrutura da proteína e do ligante são flexíveis nas simulações. Nesse programa, o processo de cálculo de <i>docking</i> molecular é dividido em duas fases: i) todos os arquivos necessários para a simulação são coletados, analisados, gerados, validados e armazenados; ii) é realizada uma etapa de otimização e ocorre a avaliação dos ligantes, selecionando o com melhor resposta.	Indisp.	NH
FlexX (RAREY <i>et al.</i> , 1996)	Programa em que é implementado um algoritmo incremental para <i>docking</i> molecular e é indicado para ligantes pequenos. Após sua primeira publicação (RAREY; KRAMER; LENGAUER; KLEBE, 1996), variações para o FlexX foram propostas. Atualmente, o programa encontra-se na versão 4.	Possui licenças gratuitas e pagas, disponíveis em https://www.biosolveit.de/download/	38
FlexAID (GAUDREAU; NAJMANOVICH, 2015)	Programa de <i>docking</i> molecular indicado para pequenos ligantes e peptídeos (como micromoléculas), e proteínas e ácidos nucleicos (como receptores). Os parâmetros de energia empregados nesse programa foram obtidos em uma classificação de um conjunto de dados com aproximadamente 1500 complexos. Esses dados foram otimizados por sucessivas iterações empregando o método de Monte Carlo.	Código-fonte disponível em https://github.com/NRGLab/FlexAID	4
FlexE (CLAUSSEN <i>et al.</i> , 2001)	O programa FlexE (<i>Flex-Ensemble</i>) é um variante para o programa Flex e também é distribuído pela empresa BioSolveIT. Esse programa também utiliza um algoritmo incremental para as simulações de <i>docking</i> molecular. No FlexE, são consideradas pequenas alterações válidas nas cadeias lateral e principal da proteína, de forma com que ela se adapte ao ligante.	Possui licenças gratuitas e pagas disponíveis em https://www.biosolveit.de/download/	NH
FlexX-Pharm (HINDLE <i>et al.</i> , 2002)	FlexX-Pharm é uma extensão do programa FlexX, em que é permitido inserir informações extras sobre proteína e ligante de interesse, em relação aos sítios de interação entre ambos. Essas	Indisp.	1

informações são utilizadas como guia para as simulações de *docking* molecular.

FLIPDock (ZHAO; SANNER, 2007)	FLIPDock é um programa de <i>docking</i> molecular em que receptor e ligante são flexíveis. Nesse programa, algoritmos genéticos são empregados como heurística para otimizar os movimentos do ligante no campo da proteína. No FLIPDock, podem ser utilizadas diferentes funções de avaliação para determinar a melhor pose para o ligante.	Gratuito para fins acadêmicos e deve ser solicitado conforme descrito em flipdock.scripps.edu/download	1
FLOG (MILLER <i>et al.</i> , 1994)	No programa FLOG (<i>Flexible Ligands Oriented on Grid</i>), é realizada uma busca em um banco de dados em que são armazenadas as coordenadas tridimensionais do ligante para encontrar moléculas complementares ao receptor de estrutura tridimensional conhecida. Nesse banco de dados, múltiplas conformações da estrutura do ligante são armazenadas, como forma de permitir a flexibilidade do mesmo e para realizar buscas rápidas, como forma de acelerar o processo de <i>docking</i> molecular.	Indisp.	NH
FRED (MCGANN, 2011)	O programa FRED emprega um algoritmo de busca exaustiva e sistemática, em que todas as conformações possíveis do ligante, contendo rotações e translações, são avaliadas no sítio ativo da proteína. Os melhores resultados obtidos são otimizados. Esse programa está contido no OEDock, que contém outras ferramentas de <i>docking</i> molecular.	Disponível para <i>download</i> (gratuito para fins acadêmicos) em www.eyesopen.com/oedocking	18
FTDOCK (GABB; JACKSON; STERNBERG, 1997)	Nesse programa ocorre, inicialmente, uma busca global, em que são realizadas rotações e translações no ligante. Em seguida, baseado nas melhores predições encontradas na etapa anterior, é realizado um refinamento dos resultados. O cálculo de energia do sistema é definido de acordo com a complementaridade entre proteína e ligante e com a teoria de correlação de Fourier.	Gratuito e disponível para <i>download</i> em http://www.sbg.bio.ic.ac.uk/docking/ftdock.html	5
GalaxyDock (SHIN; SEOK, 2012)	GalaxyDock é um programa de <i>docking</i> molecular entre proteína e ligante que permite selecionar partes em que as cadeias laterais dos ligantes sejam flexíveis. Atualmente, o programa GalaxyDock encontra-se na versão 3. Segundo seus autores, nas atualizações deste programa, houve um aumento na acurácia da predição na afinidade de ligação, na função de pontuação e um aumento na flexibilidade conformacional do	As versões 1, 2 e 3 estão disponíveis para <i>download</i> em http://galaxy.seoklab.org/s	4

ligante, em relação a sua primeira versão.

oftwares/galaxydock.html

Galaxy7TM (LEE; SEOK, 2016)	Segundo (LEE; SEOK, 2016), receptores acoplados à proteína G (GPCRs) possuem papéis fisiológicos importantes relacionados à transdução de sinal e representam um alvo importante para o desenvolvimento de drogas. Segundo os mesmos autores, o programa Galaxy7TM melhora a estrutura GPCR de entrada pela realização simultânea de <i>docking</i> molecular e pelo refinamento das estruturas flexíveis da GPCR.	Webservice gratuito, disponível em http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=7TM	1
GEMDOCK (YANG; CHEN, 2004)	O programa GEMDOCK (<i>Generic Evolutionary Method for molecular DOCKing</i>) é um programa de <i>docking</i> molecular entre proteína e ligante que emprega uma função de avaliação empírica para calcular a interação entre ambos. No GEMDOCK, algoritmos evolucionários são utilizados como heurística para reduzir o espaço de busca no receptor.	Gratuito e disponível para <i>download</i> em http://gemdock.life.nctu.edu.tw/dock/download.php	3
GlamDock (TIETZE; APOSTOLAKIS, 2007)	No programa GlamDock, os receptores são fixos e os ligantes são flexíveis. Nesse programa, para cálculo de energia, é utilizado o ChillScore, que é uma função de aproximação diferencial por partes proposta no trabalho, e como algoritmo de busca, emprega-se o método de Monte Carlo.	Gratuito e disponível para <i>download</i> em www.chil2.de/Download.html	1
Glide (FRIESNER <i>et al.</i> , 2004)	No programa Glide, em sua versão original (FRIESNER; BANKS; MURPHY; HALGREN <i>et al.</i> , 2004), inicialmente é determinada uma conformação grosseira para o ligante em diferentes posicionamentos na proteína. Para cada conformação, é aplicada uma função de avaliação (<i>score</i>). Essa etapa é seguida por uma otimização das melhores poses disponíveis e por otimizações nas conformações dos ligantes através do método de Monte Carlo. No programa Glide em sua versão atual, são implementadas as versões SP (precisão padrão, <i>standard precision</i>) e XP (precisão extra, <i>extra precision</i>). Os resultados obtidos ao utilizar a versão XP são mais precisos quando comparados aos da versão SP. Atualmente, esse programa é disponibilizado pela Schrödinger e dentre os programas de molecular <i>docking</i> avaliados neste trabalho, é o que possui finalidade comercial com o maior número de citações na base de dados Scopus.	Programa pago, com versão gratuita de teste, disponível em https://www.schrodinger.com/glide	410
GOLD (JONES; WILLETT; GLEN, 1995)	Em sua versão atual, o programa GOLD possui diferentes maneiras para ser configurado, cuja finalidade é adaptar o sistema de interesse a essas configurações, permitindo a maximização dos resultados obtidos. Como exemplo, o programa disponibiliza diferentes funções de avaliação e	Programa pago, com versão gratuita de teste, que	138

protocolos de *docking* molecular customizáveis. Em todas essas simulações, esse programa utiliza receptores com estruturas parcialmente flexíveis e os ligantes com estrutura flexível.

pode ser solicitada em www.ccdc.cam.ac.uk/solutions/csd-discovery/components/gold/

GPCRautomodel (LAUNAY <i>et al.</i> , 2012)	Neste trabalho, são realizadas duas propostas: i) é descrito um programa para automatizar a modelagem de homologia dos receptores olfativos de mamíferos, com base nas seis estruturas tridimensionais de receptores acoplados à proteína G; ii) é proposto um programa para realizar o <i>docking</i> molecular de odorantes empregando os modelos descritos na etapa 1.	Gratuito e disponível para <i>download</i> em http://genome.jouy.inra.fr/GPCRautomodel/cgi-bin/welcome.pl	NH
GriDock (VISTOLI <i>et al.</i> , 2010)	O programa GriDock foi desenvolvido para realizar <i>docking</i> molecular para um grande número de ligantes armazenados em uma única base de dados. Para realizar o processamento de forma eficiente, o programa faz o balanceamento da carga computacional relacionada às simulações de <i>docking</i> entre os processadores. O programa GriDock utiliza o AutoDock 4 como subrotina para os cálculos de interação entre proteína e ligante.	Gratuito e disponível para <i>download</i> em nova.disfarm.unimi.it/cms/index.php?Software_projects:GriDock	NH
HADDOCK (DOMINGUEZ <i>et al.</i> , 2003)	O programa HADDOCK (<i>High Ambiguity Driven protein-protein DOCKing</i>), em sua primeira versão, foi proposto para calcular a interação entre proteína-proteína. Atualmente, em sua versão 2.2, pode ser utilizado na simulação de <i>docking</i> molecular entre proteína-proteína, proteína-ácido nucléico e proteína-ligante. O HADDOCK é um programa de <i>docking</i> molecular que utiliza a abordagem flexível e, segundo seus autores, ele se difere dos demais pelo fato de armazenar informações de interfaces de proteínas identificadas ou preditas em iterações ambíguas, que são utilizadas para direcionar o processo de <i>docking</i> molecular.	Webservice gratuito, disponível em https://milou.science.uu.nl/services/HADDOCK2.2/haddockserver-easy.html	79
Haptimol FlexiDock (MATTHEWS <i>et al.</i> , 2019)	Através da ferramenta Haptimol FlexiDock, o usuário manuseia um dispositivo e, através dele, de forma tátil o usuário sente as forças de interação entre o ligante rígido e a biomolécula receptora flexível. Desse modo, a intuição e o conhecimento do usuário podem ser utilizados para guiar o conhecimento para o <i>docking</i> molecular.	Gratuito e disponível para <i>download</i> em www.haptimol.co.uk/flexidock/haptimolflexidockinst	1

		aller.zip	
Hammerhead (WELCH; RUPPERT; JAIN, 1996)	Programa de <i>docking</i> molecular, em que são utilizados ligantes flexíveis e receptores fixos. Os autores realizaram testes para grandes bases de dados de forma automatizada e garantiram que, pelo emprego de uma função de avaliação eficiente, descrita em (WELCH; RUPPERT; JAIN, 1996), as simulações foram realizadas em poucos segundos para cada ligante.	Indisponível. Atualizações realizadas pela equipe de desenvolvimento ocorrem no Surfex.	NH
Hex (RITCHIE; KEMP, 2000)	Hex é um programa proposto para cálculo e exibição de resultados de <i>docking</i> molecular entre proteína-DNA e proteína-ligante. Ao realizar <i>docking</i> molecular entre proteína-ligante, o ligante é mantido rígido. Atualmente está na versão 8 e possui uma versão <i>webservice</i> (Hex Server) e uma paralelizada (Hex-Cuda), que serão descritos adiante.	Gratuito e disponível para <i>download</i> (fins acadêmicos e governamentais) em hex.loria.fr	55
Hex-Cuda (RITCHIE; VENKATRAMA, 2010)	Hex-Cuda é a versão paralelizada do programa Hex, em que é possível a utilização da GPU para executar as simulações. Essa versão foi implementada no programa Hex através da plataforma de computação paralela CUDA. Na atualização mais recente do programa Hex, o executável possui suporte a essa forma de paralelismo. Segundo seu autor, esse é o primeiro programa de <i>docking</i> molecular capaz de usar GPUs para acelerar os cálculos.	Gratuito e disponível para <i>download</i> (fins acadêmicos e governamentais) em hex.loria.fr	1
Hex Server (MACINDOE <i>et al.</i> , 2010)	O programa HexServer é a versão <i>webservice</i> para o programa Hex. Desse modo, ao utilizá-lo, não é necessária a realização de <i>download</i> do executável do programa Hex. Para acelerar as simulações computacionais de <i>docking</i> molecular, esse programa utiliza GPU no servidor disponibilizado para testes.	<i>Webservice</i> e gratuito, cujo acesso pode ser realizado em http://hexserver.loria.fr/	2
HomDock (MARIALKE; TIETZE; APOSTOLAKIS, 2008)	No artigo de (MARIALKE; TIETZE; APOSTOLAKIS, 2008), são propostas três abordagens de <i>docking</i> molecular. Conforme apresentado pelos autores, essas técnicas são: i) é empregado um alinhamento molecular baseado em grafo (<i>graph based molecular alignment</i> , GMA). Nessa técnica, é construído um subgrafo comum máximo, gerado com base no grafo químico da molécula de interesse. De forma flexível, a estrutura da molécula alvo é alinhada ao ligante, que está rígido; ii) o GMA é construído e há um relaxamento do ligante no campo da proteína; iii) somente é considerada a informação do ligante, não considerando sua conformação em relação ao complexo.	Gratuito e disponível para <i>download</i> (fins acadêmicos e industriais) em www.chil2.de/HomDock.html	NH

HYBRID (MCGANN, 2012)	Com base nas informações sobre o ligante, o programa HYBRID melhora os posicionamentos do mesmo nas simulações de <i>docking</i> molecular. Esse programa é similar ao FRED, descrito anteriormente, em que é realizada uma busca sistemática, exaustiva e não estocástica dos posicionamentos. De forma divergente ao FRED, o HYBRID reduz o espaço de busca com base na forma e na complementaridade de ligantes conhecidos.	Gratuito e disponível para <i>download</i> (fins acadêmicos) em https://www.eyesopen.com/oedocking	2
HybridDock (HUANG; LI; WANG; PAN, 2016)	No programa HybridDock, são utilizadas as informações relacionadas à estrutura molecular do receptor e do ligante para as simulações de <i>docking</i> molecular. Para isso, com base na informação estrutural do alvo, são extraídas informação de ligantes conhecidos para essa molécula e ocorre também a busca baseada na similaridade do ligante.	Indisp.	1
HyperDock (IMBERNÓN <i>et al.</i> , 2019)	O programa HyperDock utiliza a ideia descrita no METADOCK, em são aplicadas metaheurísticas em clusters heterogêneos de computadores com unidade central de processamento (CPU) e unidade gráfica de processamento (GPU). Nesse sentido, o HyperDock descreve o paralelismo para sistemas compostos por CPU e multi-GPU para melhorar o tempo de processamento e obter melhores resultados de tempo quando comparado ao METADOCK para simulações de <i>docking</i> molecular entre ligante-receptor.	Indisp.	2
ICM-Dock (LAM; ABAGYAN; TOTROV, 2018)	O programa ICM-Dock é implementado no <i>software</i> ICM-Pro, da MolSoft. Esse programa propõe a utilização de proteínas e ligantes flexíveis. O método descrito foi desenvolvido com base na combinação de informações do ligante e do receptor, em que é construída uma função de pontuação em que são relacionados os campos de propriedades atômicas dessas duas moléculas.	Programa pago e disponível em http://molsoft.com/download.html	ICM-Dock: 1 ICM-Pro: 7
idock (LI; LEUNG; WONG, 2012)	Programa de <i>docking</i> molecular <i>opensource</i> e <i>webservice</i> . No idock, são reutilizados dados pré-calculados e são realizadas melhorias na implementação e nos modelos numéricos do programa AutoDock Vina, com a finalidade de acelerar o tempo de processamento do mesmo. Os autores descreveram que, em testes realizados, o programa idock atingiu uma aceleração de 3.3 em termos de tempo de CPU e uma aceleração de 7.5 em termos de tempo decorrido em para as execuções, na média, em comparação em relação ao programa AutoDock Vina.	<i>Webservice</i> e gratuito, disponível em http://istar.cse.cuhk.edu.hk/idock/	11
idTarget (WANG <i>et al.</i> , 2012)	O programa idTarget é utilizado para a predição de possíveis alvos moleculares de pequenas moléculas químicas. No idTarget, é empregada a abordagem de programação por divisão e conquista, para obtenção de resultados com maior eficiência em termos de custo computacional, em combinação com a função de avaliação utilizando análise de regressão e	<i>Webservice</i> e gratuito, disponível em http://idtarget.rcas.sinica.edu.tw/	2

modelos quânticos de carga química.

iGEMDOCK (HSU <i>et al.</i> , 2011)	Programa utilizado para prever interações farmacológicas, em que uma das fases é a de <i>docking</i> molecular. Para essa fase, é empregado o programa GEMDOCK, descrito previamente neste trabalho.	Webservice e gratuito, disponível em gemdock.life.nctu.edu.tw/dock/igemdock.php	37
iScreen (TSAI; CHANG; CHEN, 2011)	Programa de interface amigável, proposto inicialmente em (TSAI; CHANG; CHEN, 2011) para o estudo da medicina tradicional chinesa (TCM). Nesse programa, é descrito um modelo de <i>docking</i> molecular através do programa PLANTS, que é descrito adiante neste trabalho, seguido por um <i>de novo drug design</i> customizado, utilizando o programa LEA3D. Através do <i>de novo drug design</i> , com base na estrutura tridimensional do receptor, novas moléculas receptoras são propostas. Em sua versão atual, o programa iScreen possui o modo padrão, em que a proteína é mantida rígida, e o modo flexível, em que os resíduos flexíveis da proteína podem ser definidos pelo usuário.	Webservice e gratuito, em que os modos padrão e flexível estão disponíveis em http://iscreen.cmu.edu.tw/	NH
ISE-Dock (GORELIK; GOLDBLUM, 2008)	Programa de <i>docking</i> molecular que é baseado no algoritmo iterativo de eliminação estocástica (<i>iterative stochastic elimination</i> , ISE). Nesse programa, são eliminados valores que de maneira consistente levarão aos piores resultados. Essa etapa reduz o espaço de busca por poses de <i>docking</i> molecular.	Indisp.	NH
Itzamna (ITZAMNA, 2018)	Os programas de <i>docking</i> implementados na plataforma SaaS, da empresa Mind the Byte, utilizam o protocolo de <i>docking</i> do <i>software</i> proprietário MtBdock como subrotina. Nesse protocolo, tem-se por objetivo reduzir a quantidade de falsos positivos através da incorporação de todos os grupos prostéticos e de minimizações conformacionais. A plataforma SaaS inclui os programas de <i>docking</i> Itzamina (2018) e Kin (2018). Nesse programa, a descrição do sítio de ligação, as amostragens conformacionais dos ligantes e a descrição de poses de energia mínima são armazenadas.	Programa pago e disponível para <i>download</i> em https://www.mindthebyte.com/products/saas-platform/	NH
KIN (KIN, 2018)	O programa KIN, assim como o programa Itzamna, está implementado na plataforma SaaS, pertencente à empresa Mind the Byte e utiliza o protocolo de <i>docking</i> MtB. Nesse programa, é empregado <i>docking</i> cego, em que a busca pelo sítio de ligação ocorre em toda proteína, em todas as suas cavidades. Um conjunto de diferentes conformações da molécula é avaliado em cada cavidade, e essas conformações são submetidas à função de avaliação. Assim como o programa FRED, o programa Kin está contido no OEDock, plataforma que contém	Programa pago e disponível para <i>download</i> em https://www.mindthebyte.com/products/saas-platform/	NH

ferramentas de *docking* molecular.

Lead finder (STROGANOV <i>et al.</i> , 2008)	Programa de <i>docking</i> molecular, em que são combinados algoritmos genéticos clássicos e otimizações locais, para melhoria dos resultados. Para essa otimização, utiliza-se o conhecimento das conformações geradas ao longo da simulação de <i>docking</i> molecular. Ainda, as funções de avaliação desse programa são determinadas de forma empírica.	Programa pago, disponível para <i>download</i> em moltech.ru/. A versão gratuita de teste deve ser solicitada aos autores	NH
LeadIT (LEADIT, 2017)	Plataforma para desenvolvimento de fármacos (<i>drug design</i>), em que uma das funcionalidades é o de <i>docking</i> molecular. O programa LeadIt encontra-se atualmente na versão 2.3.2. e nele é utilizado como subrotina o programa FlexX, descrito previamente, para as simulações de <i>docking</i> molecular.	Programa pago, com versão gratuita para <i>download</i> disponível em https://www.biosolveit.de/LeadIT/	10
LeDock (LEDOCK, 2014 - 2015)	Segundo seus autores, LeDock é um programa de <i>docking</i> molecular rápido, flexível e acurado, com interface simplificada. Nesse programa, são combinadas as técnicas <i>simulated annealing</i> e otimização evolucionária da posição e orientação do ligante para as simulações computacionais. Dentre dez programas de <i>docking</i> molecular avaliados (LigandFit, Glide, GOLD, MOE Dock, Surflex-Dock, AutoDock, AutoDock Vina, LeDock, rDock, and DOCK), em experimentos realizados por (WANG; SUN; YAO; LI <i>et al.</i> , 2016), o LeDock apresentou um bom desempenho em um conjunto com 2002 complexos do tipo proteína-ligante, embora nesse estudo não tenha obtido um único programa com melhores resultados em todas as simulações.	Gratuito e disponível para <i>download</i> em www.lephar.com/download.htm	8
LibDock (BIOVIA, 2017)	Assim como o programa CDOCKER, o LibDock é disponibilizado pela BIOVIA Discovery Studio. Para avaliar as ligações entre ligante e receptor, esse programa baseia-se nas interações polares e apolares que ocorrem no sítio de ativação da proteína.	Disponível para <i>download</i> (versão básica gratuita para fins acadêmicos) em www.3dsbiovia.com/resource-	44

		center/downl oads/freeware /index.html	
LiBELa (MUNIZ; NASCIMENTO, 2015)	O programa LiBELa (<i>Ligand Binding Energy Landscape</i>) utiliza uma abordagem mista, isto é, utilizam-se informações do ligante e do receptor, para as simulações de <i>docking molecular</i> . Nesse programa, são combinadas as características estéricas e eletrostáticas de um ligante de referência, que são utilizadas nos cálculos de similaridade e sobreposição, a fim de obter uma conformação inicial pré-otimizada do ligante testado. Em seguida, a afinidade de interação é minimizada no sítio ativo de receptor a partir de potenciais energéticos.	Indisp.	2
LigandFit (VENKATACHA LAM <i>et al.</i> , 2003)	No programa LigandFit, é construído um procedimento para detecção de cavidades na proteína, passo que indica seus possíveis sítios. Em seguida, é aplicado um método de busca conformacional de Monte Carlo para gerar poses consistentes para o ligante em relação ao sítio selecionado. As poses candidatas obtidas na fase anterior são minimizadas no sítio ativo da proteína através da utilização de método baseado em malhas (<i>grid-based</i>), de forma a construir uma <i>grid</i> contendo o ligante e ocorre a avaliação da afinidade de ligação entre proteína e ligante.	Disponível para <i>download</i> (gratuito para fins acadêmicos) em pacote disponível em www.phenix- online.org/do wnload/	11
LigandScout (WOLBER; LANGER, 2005)	O programa LigandScout, para descrever os modelos aplicáveis aos farmacóforos, utiliza seis tipos de características químicas e restrições de volume, o que segundo seus autores é seletivo o suficiente na identificação do modo de ligação e é útil quando aplicado a grandes bases de dados. Em (KAINRAD; HUNOLD; SEIDEL; LANGER, 2019) é proposto o LigandScout Remote, programa descrito para gerar uma interface amigável para o LigandScout e que é útil para programação de alto desempenho.	Programa pago, (com versão gratuita para teste) disponível em <a href="http://www.in
teligand.com/
ligandscout/">http://www.in teligand.com/ ligandscout/	23
LigDockCSA (SHIN <i>et al.</i> , 2011)	No programa LigDockCSA é utilizada a técnica de recozimento de espaço conformacional (<i>conformational space annealing</i> , CSA) para otimização global, em que reduz-se o espaço de busca na proteína, a fim de encontrar o sítio de ativação forma mais eficiente. A função de avaliação empregada no LigDockCSA combina a energia calculada pelo programa AutoDock e a energia de torção potencial linear por partes (<i>piecewise linear potential</i> , PLP). Segundo (SHIN; HEO; LEE; KO <i>et al.</i> , 2011), o CSA pode encontrar menor energia no sistema quando comparado aos resultados obtidos empregando os algoritmos genéticos <i>lamarckianos</i> (<i>Lamarckian Genetic Algorithms</i>) implementados no AutoDock.	Indisponível	NH
LIGIN (SOBOLEV <i>et al.</i> ,	Nesse programa, é realizada, inicialmente, uma busca local com translações e rotações no espaço. Em seguida, as melhores	Indisp.	NH

1996)	poses são refinadas. O potencial presente nos complexos é calculado de acordo com a complementaridade entre receptor e ligante e as interações eletrostáticas favoráveis, de acordo com a teoria de correlação de Fourier.		
MCDOCK (LIU; WANG, 1999)	Para a descrição do programa de <i>docking</i> molecular MCDOCK, é empregada a técnica de Monte Carlo (MC) adaptada para determinar o sítio da proteína a ser avaliado. Nesse programa, há flexibilidade no ligante e rigidez no receptor nas simulações computacionais. Já para o cálculo da função de avaliação, são somadas as energias de interação entre ligante e receptor e ainda é avaliada a energia conformacional do ligante.	Código-fonte disponível em https://github.com/andersx/mcdock	NH
MEDock (CHANG; OYANG; LIN, 2005)	O programa MEDock (<i>Maximum - Entropy based Docking</i>) descreve e emprega um novo algoritmo de busca global, em que é utilizada a propriedade de máxima entropia da distribuição gaussiana. Segundo seus autores, o MEDock é capaz de convergir para os modos de ligação com número menor de avaliações de energia, quando comparado aos resultados obtidos ao utilizar os algoritmos genéticos lamarckianos (<i>Lamarckian Genetic Algorithms</i>)	Webservice gratuito, disponível em https://medock.k.ee.ncku.edu.tw/	NH
METADOCK (IMBERNÓN <i>et al.</i> , 2017)	Programa em que é empregada uma heurística paralelizável, desenvolvida especialmente para lidar com arquiteturas computacionais heterogêneas. Para isso, o programa METADOCK propõe um agendamento cooperativo das simulações, com a finalidade de otimizar o desempenho da obtenção da solução para o problema de <i>docking</i> molecular.	Indisp.	6
Molecular Operating Environment (MOE, 2003)	A plataforma <i>Molecular Operating Environment</i> (MOE) contém <i>softwares</i> utilizados para a descoberta de fármacos e integra visualização, modelagem e simulações nesse pacote. Nessa plataforma, são descritas diferentes abordagens para as simulações de <i>docking</i> molecular, que podem ser empregadas, por exemplo, entre proteína-proteína, proteína-ligante e <i>docking</i> com peptídeos.	Programa pago, (com versão gratuita para teste) disponível em www.chemcomp.com/Download_Request.htm	Molecular Operating Environment como chave de busca: 91
Molegro Virtual Docker (MVD, MolDock, THOMSEN; CHRISTENSEN, 2006) (THOMSEN; CHRISTENSEN, 2006)	Programa de molecular <i>docking</i> entre proteína e ligante, que associa, de forma automática, cargas e estados de protonação à proteína. Ainda, de forma automatizada, o programa MolDock verifica as cavidades e os sítios de interesse da proteína. Segundo seus autores, esse programa possui funcionalidades como suporte para águas deslocáveis, restrições definidas pelo usuário e alinhamento molecular.	Programa pago, (com versão gratuita para teste) disponível em https://molegrovirtualdocker.weebly.com/	Molegro Virtual Docker: 57 MolDock : NH

MS-DOCK (SAUTON <i>et al.</i> , 2008)	No programa MS-DOCK, são realizadas duas fases para as simulações de <i>docking</i> molecular. Na primeira, são geradas múltiplas conformações para as moléculas de interesse, fase que é similar, segundo (SAUTON; LAGORCE; VILLOUTREIX; MITEVA, 2008), à busca sistemática implementada no programa DOCK 5. Já na segunda fase, é realizado o <i>docking</i> molecular rígido utilizando o programa DOCK 6.	Indisp.	NH
ParaDockS (MEIER <i>et al.</i> , 2010)	O programa de <i>docking</i> molecular ParaDockS utiliza receptores rígidos e foi desenvolvido para combinar algoritmos de otimização e funções objetivo. Nesse programa é implementado a heurística enxame de partículas adaptada. Os autores avaliaram as seguintes funções objetivo nesse trabalho: i) função objetivo empírica p-Score; ii) potencial baseado no conhecimento PMF04 adaptado. Segundo (MEIER; PIPPEL; BRANDT; SIPPL <i>et al.</i> , 2010), o uso do p-Score garante avaliações mais acuradas e o PMF04 é aparentemente mais adequado para avaliações rápidas.	Código-fonte disponível em https://github.com/cbaldauf/paradocks	NH
ParDOCK (GUPTA; GANDHIMATHI; SHARMA; JAYARAM, 2007)	Programa de molecular <i>docking</i> em que a energia de todos os átomos é calculada de acordo com o método de Monte Carlo. Segundo seus autores, trabalhos futuros que são necessários para o ParDock incluem a predição do sítio de ligação e a flexibilidade de torção no ligante e na proteína.	Webservice gratuito, disponível em www.scfbio-iitd.res.in/paradock/	NH
PatchDock (SCHNEIDMAN- DUHOVNY <i>et al.</i> , 2005)	No programa PatchDock, é implementado um algoritmo de <i>docking</i> molecular baseado nos princípios de complementaridade das formas. Esse programa aceita como entrada duas moléculas, que podem ser proteína, DNA, peptídeo e/ou fármaco e gera como saída o potencial do complexo gerado, de acordo com a função de complementaridade das formas.	Webservice e gratuito disponível em bioinfo3d.cs.tau.ac.il/PatchDock/	59
PC-RELAX (ZACHARIAS, 2004)	O programa PC-RELAX descreve uma abordagem para permitir o relaxamento da conformação da proteína para a simulação de <i>docking</i> molecular. Esse algoritmo foi descrito para calcular a interação do imunossupressor FK506 para a proteína FKBP, uma vez que o <i>docking</i> rígido não identificou o sítio de ligação favorável em experimentos.	Indisp.	NH
PharmDock (HU; LILL, 2014)	O programa PharmDock utiliza modelos de farmacóforos baseados em proteínas para descrever uma função de avaliação empírica. Nesse programa, os modelos foram construídos de forma a reproduzir informações de diferentes complexos do tipo proteína-ligante obtidos experimentalmente. Esses modelos baseiam-se em características biológicas comuns às estruturas que possivelmente serão importantes nas atividades biológicas (HU; LILL, 2014). No programa PharmDock, os	Gratuito e disponível para <i>download</i> em https://people.pharmacy.purdue.edu/~mliill/software/ph	NH

	modelos de farmacóforos empregados contêm informações de interações potenciais entre o ligante e a proteína alvo.	armdock/	
PLANTS (KORB; STUTZLE; EXNER, 2006)	O programa PLANTS utiliza o algoritmo de otimização de colônia de formigas (<i>ant colony optimization</i> , ACO) para a resolução do problema de <i>docking</i> molecular. Nesse programa, a ACO é empregada para encontrar a conformação em que há energia mínima do ligante no sítio de ativação da proteína. Nesse contexto, as formigas armazenam as conformações com baixa energia e representam pontos que são avaliados como possíveis soluções ótimas para o problema ou pontos de possível modificação e posterior avaliação em iterações futuras.	Disponível para <i>download</i> (gratuito para fins acadêmicos) em http://www.tcd.uni-konstanz.de/plants_download/	14
POSIT (KELLEY; BROWN; WARREN; MUCHMORE, 2015)	No programa POSIT, (KELLEY; BROWN; WARREN; MUCHMORE, 2015) propõem que, para aumentar a precisão das simulações de <i>docking</i> molecular, sejam utilizadas informações conhecidas sobre o ligante presente na estrutura inicial da proteína. A informação é obtida através de um alinhamento do ligante de teste com o ligante obtido experimentalmente e através da verificação de similaridades químicas entre ambos.	Disponível para <i>download</i> (gratuito para fins acadêmicos) em www.eyesopen.com/oedocking	3
PRODOCK (TROSSET; SCHERAGA, 1999)	Programa utilizado para modelagem de proteínas e para realizar simulações computacionais de <i>docking</i> molecular, com um nível arbitrário de flexibilidade para proteínas e ligantes. Segundo seus autores, nesse trabalho é utilizado um dicionário de resíduos, cuja finalidade é a de permitir que a flexibilidade molecular seja realizada de forma mais acurada. Para otimização da função de energia, é empregado o método de Monte Carlo, combinado à minimização de energia.	Indisp.	NH
ProLeads (WESTHEAD; CLARK; MURRAY, 1998)	No programa ProLeads, é empregado o método de busca Tabu para as simulações de <i>docking</i> molecular entre ligante flexível e receptor rígido. Nesse programa, para calcular a interação entre receptor e ligante, é utilizada uma função de interação determinada de maneira empírica, utilizando energias de interação presentes nos complexos cristalográficos.	Indisp.	NH
PSI-DOCK (PEI <i>et al.</i> , 2006)	Programa utilizado para simulações de <i>docking</i> molecular entre ligantes flexíveis e proteínas rígidas. No PSI-DOCK, são combinados métodos de busca tabu, algoritmos genéticos e função de avaliação baseada na complementaridade de forma para exploração e armazenamento de poses de interesse do ligante. Em seguida, as poses armazenadas no passo anterior são otimizadas e utiliza-se uma função de avaliação de maior precisão que a empregada na etapa anterior, combinada a	Indisp.	NH

algoritmos genéticos, para avaliar a melhor posição de *docking* molecular.

PSO@AUTODOCK (NAMASIVAYA; GUNTHER, 2007)	O programa PSO@AUTODOCK é descrito para <i>docking</i> molecular em moléculas flexíveis e utiliza o programa AutoDock 3 como base para a implementação. No PSO@AUTODOCK, é empregada a otimização de enxame de partículas (<i>Particle Swarm Optimization</i> , PSO) e, segundo seus autores, o método proposto possui sucesso ao realizar o <i>docking</i> molecular para moléculas altamente flexíveis, como peptídeos, e possui resultados suficientemente rápidos para estudos com descrição ligantes totalmente flexível.	Indisp.	2
PSOVina (NG; FONG; SIU, 2015)	No programa PSOVina, são combinadas as técnicas de Otimização por Enxame de Partículas (<i>particle swarm optimization algorithm</i> , PSO), com o método de busca local empregado no programa Vina, denominado Broyden-Fletcher-Goldfarb-Shannon (BFGS). Em testes realizados, os autores identificaram uma redução de tempo entre 51% e 60%, quando comparado aos apresentados pelo programa AutoDock Vina, sem garantias sobre a qualidade da solução apresentada pelo programa.	Gratuito e disponível para <i>download</i> em https://cbbio.cis.um.edu.mo/software.html	5
PythDock (CHUNG; CHO; HAH, 2011)	O programa PythDock é construído em linguagem de programação Python e utiliza as forças eletrostáticas e de dispersão e de repulsão para cálculo da função de avaliação. Como algoritmo de busca do sítio de ativação da proteína, é utilizando o algoritmo de otimização de Enxame de Partículas (<i>particle swarm optimization algorithm</i> , PSO).	Indisp.	NH
PyGOLD (PATEL; BRINKJOST; KOCH, 2017)	O programa PyGOLD é um módulo em linguagem de programação Python que permite o uso automático do programa GOLD, descrito anteriormente, de forma automática. Desse modo, o programa GOLD pode ser incluído como subrotina em triagens virtuais (<i>virtual screening</i>) e que permite realizar alterações nas configurações no arquivo de configuração do programa GOLD.	Gratuito e disponível para <i>download</i> em https://pypi.org/project/pygold/	1
PyRx (DALLAKYAN, 2009-2019)	O programa PyRx é <i>software</i> de código aberto para realização da triagem computacional e uma das aplicações disponíveis é o <i>docking</i> molecular. Nesse programa, ocorre a combinação de diferentes programas, como o AutoDock Vina, o AutoDock 4.2, o Mayavi, utilizado na visualização de dados, e o Open Babel, programa voltado para a área química. Como programas específicos para a realização de <i>docking</i> molecular, são utilizados os programas AutoDock Vina and AutoDock 4.2 no programa PyRx, que foram descritos previamente.	Diferentes licenças são estão disponíveis em https://pyrx.sourceforge.io/downloads . Versões gratuitas são de versões anteriores.	39

Q-Dock (BRYLINSKI; SKOLNICK, 2008)	A qualidade da solução de avaliações computacionais de <i>docking</i> molecular é dependente das estruturas moleculares utilizadas. (BRYLINSKI; SKOLNICK, 2008) descrevem um método para realizar simulações utilizando estruturas das proteínas de baixa qualidade para as simulações computacionais. No programa Q-Dock, são abordados tratamentos para essas imprecisões e são empregados ligantes flexíveis nessas avaliações.	Gratuito e disponível para <i>download</i> em http://cssb.biology.gatech.edu/skolnick/files/Q-Dock/index.html	NH
rDock (RUIZ-CARMONA <i>et al.</i> , 2014)	rDock é um programa de <i>docking</i> molecular <i>open source</i> que podem ser utilizado para simulações de interação entre pequenas moléculas com proteínas ou com ácidos nucleicos. Devido a forma como foi implementado, seus autores afirmam que esse programa pode ser utilizado em um número ilimitado de CPUs, o que reduz o tempo de computação de Triagens Virtuais de Alta Produção (HTVS).	Gratuito e disponível para <i>download</i> em http://rdock.sourceforge.net/download/	9
RosettaLigand (MEILER; BAKER, 2006)	Programa disponibilizado no <i>software</i> Rosetta, que retorna o resultado de <i>docking</i> molecular entre proteína e pequenos ligantes. O procedimento de Monte Carlo é aplicado nesse algoritmo, de forma que as orientações da proteína e do ligante são otimizadas de forma simultânea. A função de energia utiliza as interações de van der Waals, um modelo implícito de solvatação e modelos eletrostáticos.	Disponível para <i>download</i> (gratuito para fins acadêmicos) em https://www.rosettacommons.org/software/license-and-download	8
RosettaLigandEnsemble (FU; MEILER, 2018)	Programa de <i>docking</i> molecular entre proteína e pequenos ligantes, que é uma variação para o programa RosettaLigand. Segundo (FU; MEILER, 2018), em geral, pequenas moléculas que possuem estruturas similares interagem na mesma pose de um sítio de ligação. No programa RosettaLigandEnsemble, os autores utilizam essa informação e o <i>docking</i> de conjuntos sobrepostos com ligantes similares é realizado de maneira simultânea. Desse modo, nesse programa é determinada uma pose genérica bem-pontuada para esse conjunto sobreposto e, em seguida, são realizadas otimizações nessas estruturas.	Indisp.	1
ROSIE (LYSKOV <i>et al.</i> , 2013)	O programa ROSIE é disponibilizado pela empresa ROSETTA para simulações de <i>docking</i> molecular entre proteína e pequenos ligantes. Esse programa é a versão <i>webservice</i> do algoritmo RosettaLigand, descrito anteriormente.	<i>Webservice</i> e gratuito, disponível em https://rosie.gaylab.jhu.edu/	5

SANDOCK (BURKHARD; TAYLOR; WALKINSHAW, 1998)	O programa SANDOCK foi proposto para simulações de <i>docking</i> molecular entre proteínas e pequenos ligantes. Segundo seus autores, a orientação do ligante no sítio ativo é gerada de acordo com a complementaridade química e de forma entre o ligante e o sítio ativo da cavidade da proteína de interesse. Para a função de avaliação, (BURKHARD; TAYLOR; WALKINSHAW, 1998) utilizaram a força de van der Waals e interações nas ligações hidrofóbicas e de hidrogênio.	Indisponível	NH
SAnDReS (XAVIER <i>et al.</i> , 2016)	Ferramenta computacional integrada, para verificar a precisão de <i>docking</i> molecular entre proteína e ligante e para a construção de uma nova função de avaliação, utilizando métodos de aprendizado de máquina, que são baseados nas funções descritas nos programas PLANTS e MolDock. Segundo seus autores, a estratégia de molecular <i>docking</i> descrita neste trabalho gerou poses com precisão maior que as de <i>benchmarks</i> publicados anteriormente.	Gratuito e disponível para <i>download</i> em https://azevedolab.net/sandres.php	2
Score (PEDRETTI; VISTOLI, 2004- 2019; SCORE, 2019)	No programa <i>webservice</i> Score, pode-se escolher uma ou mais funções de avaliação para cálculo de interação entre proteína e ligante. Essas funções são: i) MLP Interaction Score (MLPInS); ii) MLP Interaction Score (MLPInS2); iii) MLP Interaction Score (MLPInS3); iv) MLP Interaction Score (MLPInSF); v) R ⁶ R ¹² avaliado pelo campo de força CHARMM 22; vi) R ⁶ R ¹² avaliado pelo campo de força CHARMM 36; vii) R ⁶ R ¹² avaliado pelo campo de força CVFF; viii) interação eletrostática; e ix) interação eletrostática dependente de distância. No Score, essas avaliações podem ser realizadas de forma independente ou conjunta nesse programa.	<i>Website</i> e gratuito, disponível em https://nova.disfarm.unimi.it/score.htm	NH
SEED (MAJEUX <i>et al.</i> , 1999)	O programa SEED (<i>Solvation Energy for Exhaustive Docking</i>) é um programa de <i>docking</i> molecular descrito para calcular a interação entre proteínas rígidas e ligantes, de tamanho pequeno a médio. Nesse programa, fragmentos polares do ligante são posicionados de forma que no mínimo uma ponte de hidrogênio com distância ótima seja realizada com um grupo polar da proteína. Para os fragmentos apolares, são identificados pontos hidrofóbicos da proteína.	Código-fonte disponível em https://gitlab.com/CaflischLab/SEED	1
SLIDE (SCHNECKE; KUHN, 1999; SCHNECKE; KUHN, 2000)	No programa SLIDE (<i>Screening for Ligands by Induced-Fit Docking, Efficiently</i>) é possível realizar as simulações de <i>docking</i> molecular para grandes bases de dados contendo componentes orgânicos, que serão utilizados interagir com uma proteína-alvo. Para essas simulações, nesse programa, são realizadas alterações induzidas na conformação do ligante e da proteína, considerando-se a complementaridade de ambos.	Gratuito e disponível para <i>download</i> em https://www.kuhnlab.bmb.msu.edu/software/slide/	NH
SODOCK (CHEN <i>et al.</i> , 2007)	No programa SODOCK é empregada a otimização por enxame de partículas (<i>Particle Swarm Optimization, PSO</i>) para medir a	Indisp.	4

interação entre proteína e ligante, de forma especial para ligantes altamente flexíveis. Para isso, os autores utilizam um método de busca local no SODOCK, que é combinado à função de cálculo de energia empregada no programa AutoDock 3.05. Ao realizar essas alterações, (CHEN; LIU; HUANG; HWANG *et al.*, 2007) verificaram que o SODOCK apresentou resultados superiores ao AutoDock 3.05 empregando algoritmos genéticos *lamarckianos* (*Lamarckian Genetic Algorithms*), quando analisados o desempenho de convergência, a robustez e a energia obtida pelos programas, especialmente em simulações realizadas com ligantes altamente flexíveis. Em comparações realizadas por seus autores, o programa SODOCK também apresentou bons resultados quando comparado ao programa GOLD 1.2, DOCK 4.0 e FlexX 1.8.

Surflex (JAIN, 2003)	O Surflex foi proposto para simulações de <i>docking</i> molecular flexível entre proteína e ligante. Esse programa emprega a função de avaliação descrita no programa Hammerhead, descrito anteriormente, e utiliza um algoritmo para busca baseado na superfície de similaridade molecular entre proteína e ligante. De acordo com seus autores, o tempo de execução desse algoritmo é dependente de forma linear ao número de ligações rotacionáveis.	Uma versão paga e otimizada do programa encontra-se disponível em https://www.certara.com/	21
Surflex-Dock (PHAM; JAIN, 2006)	O Surflex-Dock é derivado do programa Surflex, descrito anteriormente, para simulações de <i>docking</i> molecular. No Surflex-Dock, para as predições computacionais, a função de avaliação empregada é determinada de forma empírica e são geradas poses fictícias da molécula de interesse ou parte dela de acordo com a interação entre pequenos ligantes e proteínas de estrutura conhecida. No Surflex-Dock, é empregada a função de avaliação Hammerhead. Em uma simulação apresentada por (SPITZER; JAIN, 2012) utilizando o programa, para a geração da pose do ligante, foram realizadas simulações considerando 85 complexos contendo proteína e ligante, que são disponibilizados pelo Astex Diverse. Nesse mesmo trabalho, para avaliação da ligação, no Surflex-Dock, foi utilizado um conjunto de 40 proteínas-alvo disponibilizados pelo DUD. Os autores verificaram que, como as conformações dos ligantes foram descritas de acordo com a similaridade de estruturas conhecidas, houve um desempenho robusto no programa.	Programa pago e disponível para <i>download</i> em https://www.certara.com/	48
SwissDock (GROSDIDIER; ZOETE; MICHELIN, 2011)	Programa <i>webservice</i> , proposto para realizar simulações de <i>docking</i> molecular entre proteína e pequenas moléculas. Segundo seus autores, esse programa utiliza como base a <i>engine</i> do programa EADock. Ainda, o SwissDock possui <i>scripts</i> de configuração, que resolvem problemas comuns dos usuários do programa, e que são utilizados no preparo dos arquivos de entrada da proteína e do ligante.	<i>Webservice</i> gratuito, disponível em http://www.swissdock.ch/	19

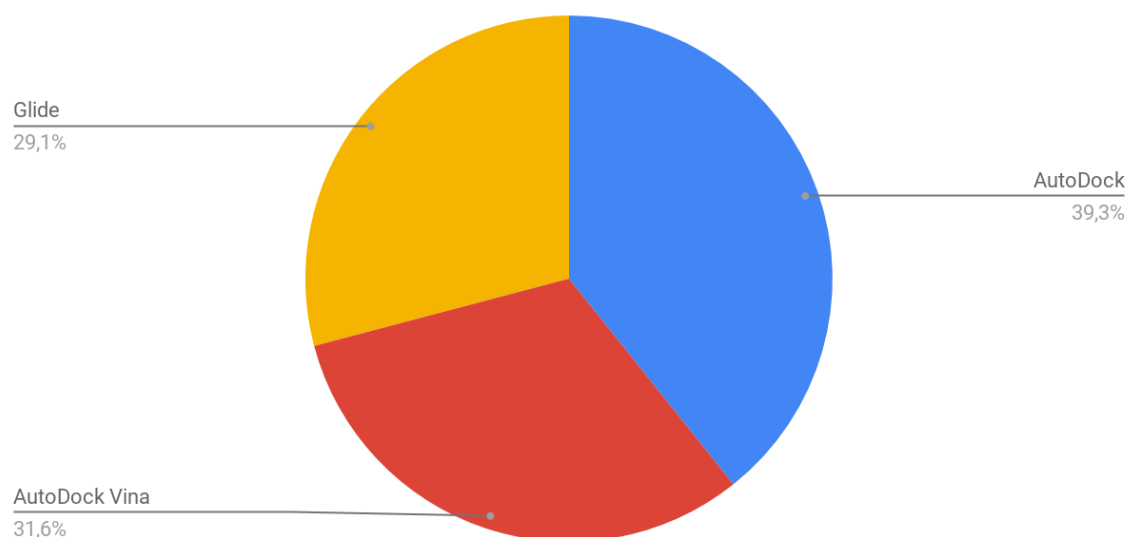
VinaMPI (ELLINGSON; SMITH; BAUDRY, 2013)	O programa VinaMPI é indicado para simulações de <i>docking</i> molecular em grandes bases de dados. Para isso, nesse programa, é empregado paralelismo, pelo uso da Biblioteca de <i>Message Passing (Message Passing Interface, MPI)</i> . Para realizar as simulações de <i>docking</i> molecular de modo eficiente, o programa AutoDock Vina é utilizado de maneira <i>multithread</i> e tarefas individuais também são otimizadas por esse programa.	Código-fonte disponível em https://github.com/mokarrom/mpi-vina	NH
VLifeDock (VLIFEMDS, 2013)	O programa VLifeDock é disponibilizado no pacote de <i>softwares</i> da VLIFEMDS, que possui ferramentas voltadas para desenvolvimento de fármacos assistido por computador e para descoberta molecular e é disponibilizado pela VLife Sciences Technologies Pvt. Ltd. No programa VLifeDock, é possível utilizar algum dos seguintes métodos de <i>docking</i> molecular: <i>docking</i> baseado no uso de malhas (grids), algoritmos genéticos ou um algoritmo próprio descrito pela VLife.	Programa pago, em que versão para teste pode ser obtida em www.vlifesciences.com/support/request_demo.php	NH
VoteDock (PLEWCZYNSKI <i>et al.</i> , 2011)	No VoteDock, são utilizados os resultados de <i>docking</i> molecular obtidos nos programas Surflex, LigandFit, Glide, GOLD, FlexX, eHiTS e AutoDock, já descritos neste trabalho. De posse dos resultados gerados por esses programas, no VoteDock é realizada uma combinação das poses e é descrita uma única predição consensual, que contém a conformação estrutural do ligante e da proteína.	Indisp.	NH
YUCCA (CHOI, 2005)	Programa para simulações de <i>docking</i> molecular entre pequenos ligantes e proteínas, em que ambos são rígidos. No programa YUCCA é implementada uma heurística de busca local, em que é verificada a energia mínima presente em uma vizinhança local, que utiliza as conformações geradas pelo programa OMEGA1.8.1 (BOSTROM; GREENWOOD; GOTTFRIES, 2003). Em seguida, a energia mínima global do sistema é encontrada pelo programa utilizando resultados obtidos na etapa anterior.	Indisp.	NH

Fonte: Compilação dos dados realizada pela autora.

Na Figura 36, podem ser observados os três programas (AutoDock, AutoDock Vina e Glide) com maior número de citações, dentre os programas avaliados, na base de dados Scopus. O AutoDock é o que possui maior número de citações (39,3%), seguido do programa AutoDock Vina (31,6%) e do programa Glide (29,1%). No ano de 2006, em pesquisa realizada por (SOUSA; FERNANDES; RAMOS, 2006), o AutoDock 4 era o programa de *docking* molecular mais utilizado para essa finalidade. Vale ressaltar que nesse ano, o programa AutoDock Vina não era descrito na literatura, pois sua proposta ocorreu no ano de 2010 e que o programa Glide é comercial. Os outros dois são programas disponibilizados de forma gratuita para realizar essas simulações.

Figura 36: Três programas de maior citação na base de dados Scopus para as simulações de *docking* molecular entre proteína e ligante.

Trabalhos utilizando os programas AutoDock 4, AutoDock Vina e Glide



Fonte: Compilação dos dados realizada pela autora.

Conforme pode ser verificado, foram apresentados 122 programas descritos na literatura para simulações de *docking* molecular. A escolha do programa adequado para as simulações permite a obtenção de resultados mais precisos, com poses que caracterizem os resultados obtidos em experimentos laboratoriais. Ainda, o conhecimento sobre esses programas é relevante no sentido de poderem ser utilizados em áreas diversas, para a resolução de problemas de áreas como Biologia e Química.

Dentre os programas encontrados, foram apresentadas características, os números de citações na base de dados Scopus, como forma de verificar a utilização desses programas em outros trabalhos, o tipo de licença disponibilizada e a referência inicial do programa na literatura. Os programas com maior citação no período selecionado foram AutoDock, AutoDock Vina e Glide.

