

**UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO**  
**INSTITUTO DE QUÍMICA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA**

**DISSERTAÇÃO**

**ESTUDO PELO MÉTODO SEMI-EMPÍRICO PM7 DE COMPLEXOS  
ENTRE O RBD DA PROTEÍNA SPIKE DO SARS-CoV-2 COM  
PEPTÍDEOS DERIVADOS DA hACE2, CANDIDATOS A INIBIDORES  
DA INFECÇÃO VIRAL**

**DANDARA DE PAULA CANDIDO**

**2023**



**UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO  
INSTITUTO DE QUÍMICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA**

**ESTUDO PELO MÉTODO SEMI-EMPÍRICO PM7 DE COMPLEXOS  
ENTRE O RBD DA PROTEÍNA SPIKE DO SARS-CoV-2 COM  
PEPTÍDEOS DERIVADOS DA hACE2, CANDIDATOS A INIBIDORES  
DA INFECÇÃO VIRAL**

**DANDARA DE PAULA CANDIDO**

*Sob Orientação do Professor*  
**Dr. Carlos Maurício R. Sant'Anna**

*e Coorientação do Professor*  
**Dr. Marco Edilson Freire de Lima**

Dissertação submetida como  
requisito parcial para a obtenção do  
grau de **Mestra em Química**, no  
Programa de Pós-graduação em  
Química, Área de concentração  
Química.

Seropédica, RJ  
Junho/2023

Universidade Federal Rural do Rio de Janeiro  
Biblioteca Central / Seção de Processamento Técnico

Ficha catalográfica elaborada  
com os dados fornecidos pelo(a) autor(a)

C217e Candido, Dandara de Paula, 1991-  
Estudo pelo Método Semi-empírico PM7 de Complexos  
entre o RBD da Proteína Spike do SARS-CoV-2 com  
Peptídeos Derivados da hACE2, Candidatos a Inibidores  
da Infecção Viral / Dandara de Paula Candido. -  
Seropédica, 2023.  
97 f.: il.

Orientador: Carlos Maurício Rabello de Sant'Anna.  
Coorientador: Marco Edilson Freire de Lima.  
Dissertação(Mestrado). -- Universidade Federal  
Rural do Rio de Janeiro, Programa de Pós-graduação em  
Química, 2023.

1. COVID-19. 2. Spike. 3. Peptídeos derivados da  
hACE2. 4. PM7. 5. Inibição covalente. I. Sant'Anna,  
Carlos Maurício Rabello de, 1965-, orient. II. Lima,  
Marco Edilson Freire de, 1965-, coorient. III  
Universidade Federal Rural do Rio de Janeiro.  
Programa de Pós-graduação em Química. IV. Título.

**UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO**  
**INSTITUTO DE QUÍMICA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA**

**DANDARA DE PAULA CANDIDO**

Dissertação submetida como requisito parcial para a obtenção do grau de Mestra em Química,  
no Programa de Pós-graduação em Química, Área de concentração Química.

**DISSERTAÇÃO APROVADA EM 21/06/2023**

Carlos Maurício Rabello de Sant'Anna. Dr. UFRRJ  
(Orientador)

Aurea Echevarria Aznar Neves Lima. Dr<sup>a</sup>. UFRRJ

Nelilma Correia Romeiro. Dr<sup>a</sup>. UFRJ

---

Emitido em 2023

**TERMO N° 716/2023 - PPGQ (12.28.01.00.00.00.60)**

**(Nº do Protocolo: NÃO PROTOCOLADO)**

*(Assinado digitalmente em 22/06/2023 08:54 )*  
AUREA ECHEVARRIA AZNAR NEVES LIMA  
*PROFESSOR DO MAGISTERIO SUPERIOR*  
*DQO (11.39.00.23)*  
*Matrícula: ###71#9*

*(Assinado digitalmente em 22/06/2023 10:48 )*  
CARLOS MAURICIO RABELLO DE SANT ANNA  
*PROFESSOR DO MAGISTERIO SUPERIOR*  
*DQF (11.39.00.25)*  
*Matrícula: ###204##*

*(Assinado digitalmente em 22/06/2023 11:08 )*  
NELILMA CORREIA ROMEIRO  
*ASSINANTE EXTERNO*  
*CPF: ####.###.877-##*

Visualize o documento original em <https://sipac.ufrrj.br/documentos/> informando seu número: **716**, ano: **2023**, tipo: **TERMO**, data de emissão: **22/06/2023** e o código de verificação: **740a65c885**

## **AGRADECIMENTOS**

À Deus, por ter me abençoado com saúde e sabedoria, guiado a minha trajetória e me dado força nos momentos difíceis.

À minha família, pelas orações e apoio financeiro nos meses seguinte ao término da bolsa.

Ao meu orientador, com quem tenho aprendido, pela mentoria, apoio, incentivo, empatia e palavras de conforto quando estive insegura. Você inspira os seus alunos e, com toda a certeza, ser sua orientanda foi um presente.

Ao meu coorientador, pelas direções em algumas tomadas de decisão do nosso trabalho e pelo acolhimento na minha primeira participação num evento científico, em sessão coordenada. A conversa que tivemos minutos antes da apresentação me deixou mais confiante.

Ao Dr. Paulo Pitasse, que conheci ainda durante o seu doutorado, pela disponibilidade em esclarecer as minhas dúvidas sobre peptídeos, pelo envio de alguns artigos, acompanhamento e discussão no desenvolvimento inicial do trabalho.

Ao Marcos Vinicius Braz, por me comunicar da vaga de mestrado para o desenvolvimento deste projeto.

Às amigas e irmãs desde o ensino médio: Thamires Gomes e Thamires Oliveira, por sempre torcerem e vibrarem com as minhas conquistas.

Aos amigos desde a graduação: Rafael de Lima e Talita de Paiva, pelo carinho e por partilharem um pouco das suas experiências na pós-graduação, que contribuíram para a minha caminhada no mestrado. Em especial, aquela que é irmã, confidente, conselheira e se tornou afilhada: Geicy Kelly, por sempre me encorajar a me desafiar e estar lado a lado comigo nas minhas lutas.

Aos colegas do Laboratório de Modelagem Molecular - LabMol: Carol Gurski, Layla Dorbação e Bruno Sena, por todas as conversas e trocas de conhecimentos.

Ao Prof. Dr. Glauco Favilla, pela atenção ao vir conversar comigo sobre este trabalho e me incentivar a me inscrever no doutorado.

Aos membros da banca, por aceitarem o convite para ler, avaliar e contribuírem para a qualidade deste documento.

À CAPES pela concessão da bolsa de incentivo à pesquisa.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

## RESUMO

CANDIDO, Dandara de Paula. **Estudo pelo método semi-empírico PM7 de complexos entre o RBD da proteína Spike do SARS-CoV-2 com peptídeos derivados da hACE2, candidatos a inibidores da infecção viral.** 2023. 97p. Dissertação (Mestrado em Química, Química Medicinal e Biológica). Instituto de Química, Departamento de Química Fundamental, Universidade Federal Rural do Rio de Janeiro, Seropédica, RJ.

No final de 2019, o mundo conheceu uma nova doença viral, atualmente chamada de COVID-19, causada pelo coronavírus SARS-CoV-2. A COVID-19 se espalhou rapidamente pelo mundo, tendo causado mais de 6,95 milhões de mortes até hoje. Dentre as proteínas do SARS-CoV-2, a *Spike* tem despertado o interesse especial, visto que o coronavírus a utiliza para se ligar através do domínio de ligação ao receptor (RBD) à enzima conversora de angiotensina 2 humana (hACE2), iniciando, assim, a invasão na célula hospedeira. Um estudo *in vitro* recente demonstrou que peptídeos derivados da hACE2 contendo a sequência de aminoácidos **EDLFYQ** (SAP1, SAP2 e SAP6) são capazes de inibir a infecção viral mediada pela *Spike*. O objetivo deste projeto é empregar métodos teóricos para se investigar o modo de interação desses peptídeos com o RBD da *Spike*, de forma a se obter informações para o planejamento racional de novos peptídeos, candidatos a inibidores da infecção pelo SARS-CoV-2, explorando a estratégia de inibição covalente. Por envolver a formação de ligações covalentes, o estudo foi conduzido com o uso do método quântico semi-empírico PM7, que é rápido o suficiente para ser aplicado a proteínas inteiras e seus complexos. Todos os cálculos envolveram o RBD completo, usando a aproximação do contínuo para incluir o efeito do meio externo na minimização de energia, sem restrições estruturais. A proposta dos novos peptídeos foi baseada na química da troca enxofre-triazol (SuTEx), que permite seletividade para formação de ligações covalentes com resíduos de tirosina. Os SAPs foram modificados pela incorporação de grupos ogiva sulfonil-triazol, além de um novo grupo ogiva proposto fosfonil-triazol, de modo a elevar a eficiência da interação por meio de ligações covalentes. Dados de  $\Delta H_{int}$  calculados revelaram a mesma tendência que a  $K_d$  determinada experimentalmente para SAP1, SAP2 e SAP6, e confirmam o peptídeo SAP1 tanto nativo quanto com modificações do tipo SuTEx como melhor candidato a ligante do RBD, seguido pelos SAPs 6 e 2, nesta ordem. Também foi observado que os complexos covalentes formados pelo SAP1 com grupos ogiva sulfonil-triazol demonstraram ser mais estáveis entalpicamente, indicando uma maior tendência de serem formados e persistirem no meio em que foram estudados do que aqueles contendo grupos ogiva fosfonil-triazol.

**Palavras-chave:** COVID-19. *Spike*. Peptídeos derivados da hACE2. PM7. Inibição covalente.

## ABSTRACT

CANDIDO, Dandara de Paula. **Study by the semi-empirical PM7 method of complexes between RBD of SARS-CoV-2 Spike protein with hACE2 peptides, candidates for inhibitors of viral infection.** 2023. 97p. Dissertation (Master in Chemistry, Medicinal and Biological Chemistry). Institute of Chemistry, Department of Fundamental Chemistry, Federal Rural University of Rio de Janeiro, Seropédica, RJ.

At the end of 2019, the world experienced a new viral disease, currently called COVID-19, caused by the coronavirus SARS-CoV-2. COVID-19 has spread rapidly around the world, causing more than 6.95 million deaths to date. Among the proteins of SARS-CoV-2, *Spike* has aroused special interest, since the coronavirus uses it to bind through the receptor binding domain (RBD) to the converting enzyme of human angiotensin 2 (hACE2), thus initiating invasion into the host cell. A recent *in vitro* study demonstrated that hACE2-derived peptides containing the amino acid sequence **EDLFYQ** (SAP1, SAP2 and SAP6) are able to inhibit Spike-mediated viral infection. The objective of this project is to employ theoretical methods to investigate the interaction mode of these peptides with *Spike*'s RBD, in order to obtain information for the rational design of new peptides, candidates for SARS-CoV-2 infection inhibitors, exploring the covalent inhibition strategy. Because it involves the formation of covalent bonds, the study was conducted using the semi-empirical quantum method PM7, which is fast enough to be applied to whole proteins and their complexes. All calculations involved the complete RBD, using the continuum approach to include the effect of the external environment on energy minimization, without structural constraints. The proposal of the new peptides was based on the chemistry of sulfur-triazole exchange (SuTEx), which allows selectivity for the formation of covalent bonds with tyrosine residues. The SAPs were modified by the incorporation of warhead groups sulfonyl-triazole, and a new proposed warhead group, phosphonyl-triazole, in order to increase the efficiency of the interaction through covalent bond formation.  $\Delta H_{int}$  calculated revealed the same trend as the experimentally determined  $K_d$  for SAP1, SAP2 and SAP6, and confirm the peptide SAP1 both in native and in SuTEx-type modifications as the best candidate for RBD ligand, followed by SAPs 6 and 2, in this order. It was also observed that the covalent complexes formed by SAP1 derivatives containing sulfonyl-triazole warhead groups have a greater tendency to be formed in the medium in which they were studied than those containing phosphonyl-triazole warheads.

**Keywords:** COVID-19. *Spike*. hACE2-derived peptides. PM7. Covalent inhibition.

## LISTA DE TABELAS

<b>Tabela 1:</b> Variantes de preocupação e mutações de interesse da <i>Spike</i> (continua).....	25
<b>Tabela 2:</b> Erros médios para os conjuntos de dados S22 e S66 (kcal/mol).....	40
<b>Tabela 3:</b> Peptídeos avaliados <i>in vitro</i> como inibidores da infecção viral pelo SARS-CoV-2.....	42
<b>Tabela 4:</b> Parâmetros termodinâmicos dos peptídeos calculados com o método PM7.....	48
<b>Tabela 5:</b> Energias de interação nos complexos nativos calculados pelo método PM7 (continua).....	50
<b>Tabela 6:</b> Resíduos semelhantes na interface de interação proteína-peptídeo.....	51
<b>Tabela 7:</b> Ligações de H semelhantes na interface de interação, calculadas pelo método PM7 (continua).....	58
<b>Tabela 8:</b> Parâmetros calculados pelo método semi-empírico PM7 ( $\varepsilon = 4,0$ ) para os complexos e dados obtidos <i>in vitro</i> <sup>1</sup> para os peptídeos.....	60
<b>Tabela 9:</b> SAPs modificados contendo a sequência EZLFYQ, em que Z é um dos aminoácidos modificados abaixo, contendo grupos sulfonil-triazol (X = S) ou fosfonil-triazol (X = P), propostos neste trabalho para ligação covalente com o RBD.....	62
<b>Tabela 10:</b> Parâmetros calculados pelo método semi-empírico PM7 ( $\varepsilon = 4,0$ ), para os complexos de interação com peptídeos modificados com grupos sulfonil-1,2,4-triazol (Z1S) e com grupo sulfonil-1,2,3-triazol (Z2S).....	63
<b>Tabela 11:</b> Parâmetros calculados pelo método semi-empírico PM7 ( $\varepsilon = 4,0$ ), para os complexos de interação com peptídeos modificados com grupos ogiva fosfonil-1,2,4-triazol (Z1P) e com grupo fosfonil-1,2,3-triazol (Z2P) (continua).....	63
<b>Tabela 12:</b> $\Delta H_{int}$ e energias de interação calculadas pelo método semi-empírico PM7 ( $\varepsilon = 4,0$ ), para os complexos derivados.....	68
<b>Tabela 13:</b> $\Delta H_r$ calculada pelo método semi-empírico PM7 ( $\varepsilon = 4,0$ ), para os complexos covalente.....	69

## LISTA DE FIGURAS

<b>Figura 1:</b> Partículas do vírus SARS-CoV-2, em microfotografia em cores realçadas produzida por microscópio eletrônico (National Geographic, 2022).....	18
<b>Figura 2:</b> Classificação dos Coronavírus (adaptado de MALIK, 2020).....	19
<b>Figura 3:</b> Estrutura e genoma do SARS-CoV-2 (LI <i>et al.</i> , 2020). .....	21
<b>Figura 4:</b> Estruturas obtidas por crio-EM da glicoproteína <i>Spike</i> do SARS CoV-2, mostrando as duas conformações adotadas pelo RBD, que permitem ou não o acesso à hACE2 (adaptado de WALLS <i>et al.</i> , 2020). .....	22
<b>Figura 5:</b> Representação da replicação e morfogênese dos coronavírus (BORGES <i>et al.</i> , 2020). .....	23
<b>Figura 6:</b> (A) Exemplos de regiões de cavidade- $\sigma$ em halogênios; (B) exemplo de interação cátion- $\pi$ (adaptado de FOKUE <i>et al.</i> , 2020).....	28
<b>Figura 7:</b> (A) Possíveis modos de interação entre um ânion e um anel aromático; (B) Intereração amida- $\pi$ (adaptado de FOKUE <i>et al.</i> , 2020).....	29
<b>Figura 8:</b> (A) Interação alquila-arila; (B) Exemplos de empilhamento $\pi$ e suas respectivas orientações (adaptado de FOKUE <i>et al.</i> , 2020).....	30
<b>Figura 9:</b> Mecanismo de interação entre um inibidor covalente e seu alvo (SUTANTO <i>et al.</i> , 2020).....	33
<b>Figura 10:</b> Energia em função da distância de ligação descrita pelo modelo da mecânica molecular em comparação com o comportamento real (SANT'ANNA, 2009).....	36
<b>Figura 11:</b> Estrutura cristalográfica A) da glicoproteína <i>Spike</i> nativa do SARS-CoV-2, no estado aberto, com RBD destacado pelo círculo vermelho (PDB ID: 7FB1); B) do RBD da <i>Spike</i> nativa do SARS-CoV-2 (cor laranja) ligado com a hACE2 (cor verde); em azul estão os quatro N-acetil- $\beta$ -glucosamina glicanos (NAG) ligados a Asn90, Asn322 e Asn546 da hACE2 e a Asn343 do RBD; as esferas vermelhas representam as moléculas de água (PDB ID: 6M0J). .....	43
<b>Figura 12:</b> Sobreposição do RBD da <i>Spike</i> (em $\epsilon = 78,4$ ) otimizado parcialmente e totalmente e, recorte do sítio de ligação, evidenciando o deslocamento do resíduo Gln493, que na otimização parcial (C na cor rosa) estava com a sua cadeia lateral direcionada para o exterior do bolso (C na cor laranja), enquanto na otimização total (C na cor lilás) ficou posicionada para o interior (C na cor verde). .....	50
<b>Figura 13:</b> A) Ligações de hidrogênio (tracejado amarelo) entre os resíduos Asn501 e Thr500 do RBD (C em laranja) e a Tyr41 do peptídeo SAP1 (C em verde). B) Ligação de hidrogênio intramolecular (tracejado azul) entre o NH da ligação peptídica e a hidroxila da treonina. ....	53
<b>Figura 14:</b> A) Ligações de hidrogênio (tracejado amarelo) entre os resíduos Asn501 e Thr500 do RBD (C em laranja) e a Tyr41 do peptídeo SAP2 (C em verde). B) Ligação de hidrogênio intramolecular (tracejado azul) entre o NH da ligação peptídica e a hidroxila da treonina. ....	53
<b>Figura 15:</b> A) Ligações de hidrogênio (tracejado amarelo) entre os resíduos Asn501 e Thr500 do RBD (C em laranja) e a Tyr41 do peptídeo SAP6 (C em verde). B) Ligação de hidrogênio intramolecular (tracejado azul) entre o NH da ligação peptídica e a hidroxila da treonina. ....	54
<b>Figura 16:</b> Ligações de H (tracejado amarelo) entre os resíduos Tyr449 e Gln498 do RBD (C em laranja) e Asp38 e Gln42 do peptídeo SAP1 (C em verde).....	55
<b>Figura 17:</b> Ligações de H (tracejado amarelo) entre os resíduos Tyr449 e Gln498 do RBD (C em laranja) e Asp38 e Gln42 do peptídeo SAP2 (C em verde).....	55
<b>Figura 18:</b> Ligações de H (tracejado amarelo) entre os resíduos Tyr449 e Gln498 do RBD (C em laranja) e Asp38 e Gln42 do peptídeo SAP6 (C em verde).....	56

<b>Figura 19:</b> Ligações de H (tracejado amarelo) entre os resíduos Gly446 e Gly447 do RBD (C em laranja) e Gln42 do peptídeo SAP1 (C em verde). ....	57
<b>Figura 20:</b> Ligações de H (tracejado amarelo) entre os resíduos Gly446 e Gly447 do RBD (C em laranja) e Gln42 do peptídeo SAP6 (C em verde). ....	57
<b>Figura 21:</b> Exemplos de mecanismo propostos para a formação da ligação covalente entre os resíduos Tyr449 do RBD da <i>Spike</i> e os aminoácidos modificados presentes nos peptídeos (A) SAP <sub>n</sub> -Z1S e (B) SAP <sub>n</sub> -Z1P. Mecanismo similar pode ser proposto para os peptídeos contendo o grupo 1,2,3-triazol SAP <sub>n</sub> -Z2S e SAP <sub>n</sub> -Z2P. ....	65
<b>Figura 22:</b> Distância entre o nitrogênio do anel triazólico e o hidrogênio da hidroxila da Tyr449 (C em rosa), nos complexos RBD-SAP1 modificados com grupo sulfonil-triazol. A) RBD-SAP1-Z1 com 2,8 Å; B) RBD-SAP1-Z2 com 2,8 Å. Obs: apenas os hidrogênios polares são mostrados; e, nos complexos modificados contendo o grupo fosfonil-triazol, esses valores se alteraram para 2,6 Å. ....	66
<b>Figura 23:</b> Distância entre o nitrogênio do anel trizólico e o hidrogênio da hidroxila da Tyr449 (C em rosa), nos complexos RBD-SAP2 modificados com grupo sulfonil-triazol. A) RBD-SAP2-Z1 com 3,3 Å; B) RBD-SAP2-Z2 com 4,5 Å. Obs: apenas os hidrogênios polares são mostrados; e, os mesmos valores foram encontrados nos complexos modificados contendo o grupo fosfonil-triazol. ....	66
<b>Figura 24:</b> Distância entre o nitrogênio do anel trizólico e o hidrogênio da hidroxila da Tyr449 (C em rosa), nos complexos RBD-SAP6 modificados com grupo sulfonil-triazol. A) RBD-SAP6-Z1 com 5,3 Å; B) RBD-SAP6-Z2 com 3,5 Å. Obs: apenas os hidrogênios polares são mostrados; e, os mesmos valores foram encontrados nos complexos modificados contendo o grupo fosfonil-triazol. ....	67

## LISTA DE EQUAÇÕES

<b>Equação 1</b> - Energia para o estiramento de ligações químicas .....	35
<b>Equação 2</b> - Energia de van der Waals .....	36
<b>Equação 3</b> - Energia eletrostática .....	36
<b>Equação 4</b> - Equação de Schrödiger .....	37
<b>Equação 5</b> - Função de onda molecular.....	37
<b>Equação 6</b> - Fórmula da entalpia padrão de formação.....	47
<b>Equação 7</b> - Fórmula da entalpia de interação dos complexos.....	59
<b>Equação 8</b> - Equilíbrio de formação do complexo RBD-SAP .....	59
<b>Equação 9</b> - Fórmula da constante de dissociação .....	59
<b>Equação 10</b> - Fórmula da entalpia de reação dos complexos covalentes .....	69

## LISTA DE ABREVIATURAS, SIGLAS E SÍMBOLOS

ACE2	Enzima Conversora de Angiotensina 2
AM1	Modelo Austin 1
BCoV	Coronavírus bovino
CADD	Planejamento de Fármacos Auxiliado por Computadores
COVID-19	Doença do Coronavírus 19
CoVs	Coronavírus
EPS	Potencial Eletrostático
Ergic	Compartimento Intermediário Retículo Endoplasmático-Golgi
ESPII	Emergência em Saúde Pública de Importância Internacional
FCoV	Coronavírus felino
FDA	Food and Drug Administration
FP	Peptídeo de fusão
HCoVs	Coronavírus humanos
HF	Hartree-Fock
iBV	Vírus da bronquite infecciosa aviária
IC <sub>50</sub>	Concentração requerida para atingir 50% do efeito inibitório máximo
ICTV	Comitê Internacional de Taxonomia de Vírus
IUPAC	International Union of Pure and Applied Chemistry
K <sub>b</sub>	Constante de força
K <sub>d</sub>	Constante de dissociação
LBDD	Planejamento de Fármacos Baseado nos Ligantes
LCAO	Combinação Linear de Orbitais Atômicos
LMO	Orbitais Moleculares Localizados
MERS	Síndrome Respiratória do Oriente Médio
MERS-CoV	Síndrome Respiratória do Oriente Médio do Coronavírus
MHV	Vírus da hepatite do camundongo
MINDO/3	Modified Intermediate Neglect of Differential Overlap
mM	Milimol
MNDO	Modified Neglect of Diatomic Overlap
MOPAC	Molecular Orbital Package
mRNA	RNA mensageiro

NDDO	Neglect of Differential Diatomic Overlap
OM	Orbital Molecular
OMS	Organização Mundial da Saúde
OPAS	Organização Pan-Americana da Saúde
ORFs	Open Reading Frames
PD	Domínio da peptidase
PDB	Protein Data Bank
PEDV	Vírus da diarreia epidêmica suína
PHEV	Vírus da encefalomielite hemaglutinante suína
PLIP	Perfil de Interação Proteína-ligante
PMx	Método Paramétrico
PRCV	Coronavírus respiratório suíno
RBD	Receptor Binding Domain
RM1	Modelo Recife 1
RNA	Ácido ribonucleico
SAPs	Peptídeos derivados da ACE direcionados à <i>Spike</i>
SAR	Relação estrutura-atividade
SARS	Síndrome Respiratória Aguda Grave
SARS-CoV	Síndrome Respiratória Aguda Grave do Coronavírus
SARS-CoV-2	Síndrome Respiratória Aguda Grave do Coronavírus 2
SBDD	Planejamento de Fármacos Baseado na Estrutura do Alvo
SuTEx	Troca Enxofre-triazol
TGEV	Transmissível vírus da gastroenterite
TMPRSS2	Serina Protease Transmembranar do tipo 2
UTI	Unidade de Tratamento Intensivo
UTRs	Regiões não Traduzidas
VOC	Variant of Concern
$\Delta H_f^0$	Entalpia padrão de formação
$\Delta H_{\text{int}}$	Entalpia de interação
$\psi(R,r)$	Função de onda
$\Psi_{Ai}$	Função de onda de orbitais atômicos
$\Psi_M$	Função de onda molecular
$\epsilon$	Constante dielétrica

## SUMÁRIO

<b>1. INTRODUÇÃO.....</b>	15
<b>2. REVISÃO BIBLIOGRÁFICA .....</b>	17
2.1 COVID-19 .....	17
2.1.1 Coronavírus: classificação e histórico.....	18
2.1.2 Morfologia do SARS-CoV-2 e ação viral em células hospedeiras.....	20
2.1.3 Mutações, linhagens, cepas e variantes.....	23
2.2 Química Medicinal: conceito e reconhecimento molecular .....	26
2.2.1 Interações intermoleculares envolvidas no reconhecimento molecular do ligante no sítio receptor.....	26
2.2.2 Peptídeos terapêuticos .....	31
2.2.3 Inibição covalente .....	33
2.3 Métodos e Ferramentas Computacionais .....	34
2.3.1 Métodos clássicos .....	35
2.3.2 Métodos quânticos .....	37
<b>3. OBJETIVOS .....</b>	41
<b>4. MÉTODOS E FERRAMENTAS COMPUTACIONAIS .....</b>	42
4.1 Obtenção dos complexos proteína-peptídeo .....	42
4.2 Cálculos dos complexos nativos com o método semi-empírico .....	43
4.3 Cálculos dos complexos modificados com o método semi-empírico.....	45
4.4 Visualização molecular.....	46
<b>5. RESULTADOS E DISCUSSÃO .....</b>	47
5.1 Otimização das moléculas separadas e em complexo nativo .....	47
5.2 Análise das interações intermoleculares na interface de ligação .....	51
5.3 Entalpia de interação dos complexos nativos e a correlação com a inibição viral .....	59
5.4 Propostas de modificação estrutural nos peptídeos como ligantes covalentes .....	61
5.5 Análise da formação do complexo covalente .....	64
<b>6. CONCLUSÃO.....</b>	70
<b>7. REFERÊNCIAS .....</b>	72
<b>ANEXOS .....</b>	78
A – Relação das ligações de hidrogênio identificadas no complexo nativo RBD-SAP1, no arquivo <i>arc</i> , após a otimização total.....	78
B – Ligações de H intramoleculares envolvendo o resíduo Gln493 do RBD.....	88
C - Ligações de hidrogênio intermoleculares (< -1,0 kcal/mol; $\epsilon = 4,0$ ).....	89
D - Demais interações intermoleculares ( $\epsilon = 4,0$ ) .....	92
E – Ilustração gráfica da interface de interação entre o RBD da <i>Spike</i> do SARS-CoV-2 e o SAP6 .....	94
F- Reatividade x Especificidade das sondas SuTEx e SuFEx .....	95
G - Distância entre o nitrogênio do anel 1,2,4-triazol substituído do peptídeo SAP1 e o hidrogênio da hidroxila da Tyr449 do RBD, nos complexos derivados .....	96

## **1. INTRODUÇÃO**

No final de 2019, o mundo conheceu uma nova doença viral, denominada de COVID-19, identificada pela primeira vez em pacientes que tinham desenvolvido pneumonia sem causa identificável, na cidade de Wuhan, na China (European Centre for Disease Prevention and Control).

Vírus isolados de pacientes do Hospital Jinyintan, em Wuhan, tiveram suas sequências genômicas determinadas e análises por bioinformática indicaram que eles tinham características típicas dos coronavírus. Os coronavírus causam várias doenças respiratórias, algumas das quais se tornaram epidêmicas, como a epidemia de Síndrome Respiratória Aguda Grave (SARS), de 2002-2003, e a epidemia de Síndrome Respiratória do Oriente Médio (MERS), de 2012. O coronavírus causador da COVID-19 recebeu a denominação de SARS-CoV-2 devido à sua alta homologia com o vírus SARS-CoV (WHO-China Joint Mission On Coronavirus Disease 2019 (COVID-19)).

A COVID-19 se espalhou rapidamente pelo mundo e, no início de março de 2020, foi elevada à categoria de pandemia pela Organização Mundial da Saúde (OMS). De acordo com Organização Pan-Americana da Saúde (OPAS), o termo ‘pandemia’ se refere à distribuição geográfica de uma doença e não à sua gravidade. A doença é considerada altamente infecciosa, tendo como principal forma de contágio o contato com a pessoa infectada, que transmite o vírus por meio de gotículas, com cargas virais, provenientes da fala, tosse e espirros, ou quando se toca em uma superfície contaminada e depois nos olhos, nariz ou boca (BORGES *et al.*, 2020).

Até a data atual, o número global de infectados superou os 768,98 milhões e o número de mortes ultrapassou os 6,95 milhões, de acordo com os dados estatísticos do *Our World In Data* (<https://ourworldindata.org/coronavirus>), sendo que eles continuam sendo alterados. Muitos pacientes necessitam de internação em Unidade de Tratamento Intensivo (UTI), devido aos efeitos respiratórios graves da doença, o que sobrecarregou muito, nos anos iniciais da pandemia, o sistema de saúde em muitos países, incluindo o Brasil, onde o número de infectados e mortos se aproxima, respectivamente, de 37,73 milhões e 704,8 mil, segundo o Ministério da Saúde do Brasil.

Existem compostos já autorizados e aprovados pela *Food and Drug Administration* (FDA) para uso emergencial contra a COVID-19. De acordo com a agência, um produto autorizado é aquele que pode ser eficaz no diagnóstico, tratamento ou prevenção de uma doença

ou condição grave, desde que os benefícios sejam superiores aos riscos conhecidos e potenciais. Já um produto aprovado é o medicamento seguro e eficaz, com sua aplicação já indicada em bula. Foram autorizados anticorpos monoclonais direcionados para o SARS-CoV-2 (Bebtelovimabe e Evusheld (tixagevimabe co-embalado com cilgavimabe)); medicamentos antivirais (Paxlovid (nirmatrelvir e ritonavir) e Lagevrio (molnupiravir)); imunomoduladores (Olumiant (baricitinibe)), para pacientes de 2 a menos de 18 anos que necessitem de oxigênio suplementar ou ventilação mecânica invasiva; e Actemra (tocilizumabe)); sedativos (Propofol-Lipuro 1%), para pacientes que estão intubados e necessitam de ventilação mecânica em UTI. Entre os aprovados estão: Veklury (Remdesivir), para usos adulto e pediátrico (com 28 dias de idade ou mais, com pelo menos 3 kg), hospitalizados ou não, com sintomas de leve a moderados e alto risco de progressão; Olumiant (baricitinibe), para adultos hospitalizados que necessitem de oxigênio suplementar e/ou ventilação mecânica invasiva ou não (FDA, 2022).

Além desses compostos, há vacinas monovalentes eficazes disponíveis e entre as que foram aprovadas no Brasil estão: **CoronaVac**, produzida pelo Instituto Butantan em parceria com a farmacêutica chinesa Sinovac, que utiliza a técnica de vírus inativado (morto); **AstraZeneca**, desenvolvida pela farmacêutica de mesmo nome em parceria com a Universidade de Oxford e produzida no Brasil pela Fundação Oswaldo Cruz (Fiocruz), emprega a tecnologia do vetor viral, que utiliza uma versão modificada e segura do vetor com “pedaços” de proteína (antígeno) de um agente infeccioso; **Pfizer**, imunizante da farmacêutica Pfizer em parceria com o laboratório BioNTech, se baseia na técnica do RNA mensageiro (mRNA), que instrui o organismo a produzir as proteínas que são encontradas na superfície do novo coronavírus, e desse modo estimula a resposta imunológica; **Janssen**, do grupo Johnson & Johnson, aplicada em apenas uma dose e também utiliza a tecnologia do vetor viral, com base em um tipo específico de adenovírus modificado geneticamente para não se replicar em humanos. E, atualmente, está sendo disponibilizada para a população a **vacina bivalente** contra a COVID-19, também produzida pela Pfizer, que age da mesma forma que as vacinas monovalentes ao estimular a produção de anticorpos e células de defesa, mas que possui ação contra a cepa original e as subvariantes da Ômicron (INSTITUTO BUTANTAN, 2022; BIOMANGUINHOS/FIOCRUZ, 2022; MINISTÉRIO DA SAÚDE, 2022).

Devido ao surgimento constante de novas cepas, a pesquisa por medicamentos ainda é necessária. Dentro dessa perspectiva, o uso de estratégias racionais para o planejamento de novos fármacos, incluindo o uso de técnicas computacionais, pode trazer uma contribuição efetiva à solução do problema. Esta abordagem é chamada de Planejamento de Fármacos

Auxiliado por Computadores (CADD, do inglês *Computer Assisted Drug Design*). O CADD pode ser empregado desde os estágios de identificação até a otimização estrutural de compostos protótipo. Em comparação com os métodos tradicionais de descoberta de medicamentos, os métodos orientados de planejamento podem trazer uma redução no tempo e no custo envolvidos no processo de desenvolvimento de novos medicamentos.

## 2. REVISÃO BIBLIOGRÁFICA

### 2.1 COVID-19

A cidade de Wuhan, localizada na província de Hubei, na China, vivenciou casos de pneumonia sem causa aparente, no final de 2019. Esse surto se espalhou rapidamente para outras áreas, tendo seu agente causador - um novo coronavírus, descoberto logo em seguida. Inicialmente, o novo coronavírus foi denominado de 2019-nCov e, em seguida, devido à sua alta homologia (~ 80%) com o SARS-CoV, foi renomeado em janeiro de 2020 por cientistas chineses, para SARS-CoV-2 (Síndrome Respiratória Aguda Grave do Coronavírus 2, do inglês *Severe Acute Respiratory Syndrome CoronaVirus 2*). A transmissão zoonótica da doença foi associada ao mercado de frutos do mar e animais vivos da cidade de Wuhan (CAVALCANTE *et al.*, 2020; YUKI, FUJIOGI e KOUTSOGIANNAKI, 2020).

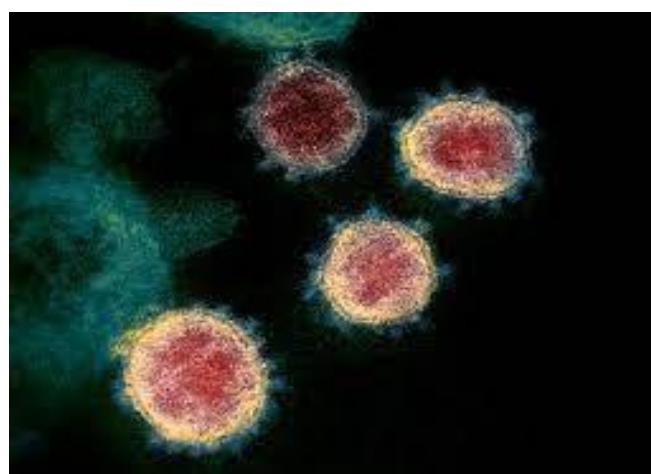
Na China foram registrados, nos primeiros 30 dias, 11.821 casos e 259 óbitos. Ainda em janeiro a doença foi registrada em países da Ásia, Europa e América do Norte, e neste período já era reconhecido que a transmissão entre humanos desempenhava um papel importante no surto subsequente. Assim, a Organização Mundial da Saúde (OMS) declarou, em 30 de janeiro de 2020, Emergência em Saúde Pública de Importância Internacional (ESPII) e em 11 de março de 2020 decretou a pandemia. A doença causada por esse vírus foi chamada de COVID-19, do inglês *Corona Virus Disease 2019* (CAVALCANTE *et al.*, 2020; YUKI, FUJIOGI e KOUTSOGIANNAKI, 2020).

A COVID-19 gerou impactos em todo o mundo, sejam eles relacionados à saúde física e/ou mental, à economia global, à educação, à cultura, à política e ao convívio social. Após o surto inicial na China, o vírus se espalhou pelo mundo e no início de abril de 2020 o número de pacientes com COVID-19 eram mais altos nos seguintes países: EUA, Espanha, Itália, Alemanha, França e China (YUKI, FUJIOGI e KOUTSOGIANNAKI, 2020; CAVALCANTE *et al.*, 2020). No Brasil, a condução da pandemia tornou o país uma ameaça global ao restante

do mundo, tendo em vista o alto número de casos e mortes pela doença e a lentidão na aquisição e distribuição de vacinas. No país, a campanha de vacinação se iniciou somente em 17 de janeiro de 2021, segundo dados da Fiocruz, e contribuiu para a redução no número de casos graves e óbitos em decorrência da doença. Em 05 de maio de 2023, a OMS decretou o fim da ESPII referente à COVID-19. Tal decisão não indica que a doença deixou de ser uma ameaça à saúde pública, visto que sua propagação continua sendo caracterizada como pandemia, mas sim que os países devem fazer a sua transição de emergência para o quadro de doenças infecciosas, conforme destacou Tedros Adhanom, diretor-geral da organização (BIO-MANGUINHOS/FIOCRUZ, 2022; OPAS, 2023).

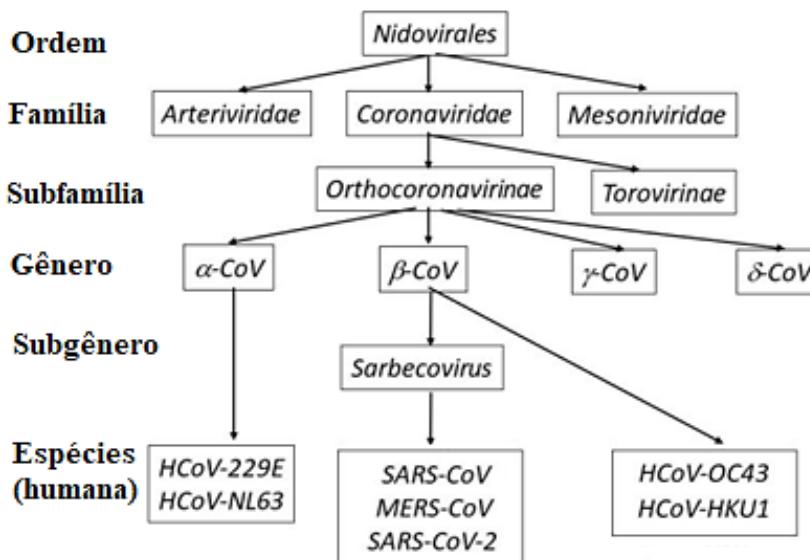
### 2.1.1 Coronavírus: classificação e histórico

Os coronavírus (CoVs), pertencentes a ordem *Nidovirales*, família *Coronaviridae*, na subfamília *Orthocoronavirinae*, causam infecção respiratória em mamíferos e em espécies de aves. Seu nome deriva da palavra em latim *corona*, que significa coroa ou halo, devido aos picos em sua superfície, observáveis em microscópio eletrônico (Figura 1), que leva a um formato semelhante a uma coroa. Os CoVs são envelopados e possuem um genoma de RNA de fita simples e sentido positivo com tamanhos variando de 26 a 32 Kb (quilobases) de comprimento, tendo o maior genoma conhecido para um vírus de RNA (SHARMA *et al.*, 2021; MALIK, 2020).



**Figura 1:** Partículas do vírus SARS-CoV-2, em microfotografia em cores realçadas produzida por microscópio eletrônico (National Geographic, 2022).

Dentro da subfamília, os coronavírus são classificados quanto ao gênero, de acordo com o Grupo de Estudos do Coronavírus do Comitê Internacional de Taxonomia de Vírus (ICTV, do inglês *International Committee on Taxonomy of Viruses*) em: *Alfacoronavirus* ( $\alpha$ -CoV), *Betacoronavirus* ( $\beta$ -CoV), *Gammacoronavirus* ( $\gamma$ -CoV) e *Deltacoronavirus* ( $\delta$ -CoV) (Figura 2). Os gêneros  $\alpha$ -CoV e  $\beta$ -CoV geralmente causam doenças respiratórias em humanos e infectam apenas mamíferos, enquanto  $\gamma$ -CoV e  $\delta$ -CoV ocasionam uma patologia gastrointestinal em animais, sendo encontrados principalmente em aves, apesar de também contagiar os mamíferos. Os SARS-CoV, MERS-CoV e SARS-CoV-2 pertencem ao subgênero *Sarbecovírus* e são responsáveis por causar infecções altamente patogênicas, enquanto os demais  $\alpha$  e  $\beta$ -CoVs causam infecções respiratórias leves (GRÄF, 2020; OLIVEIRA, MATOS e SIQUEIRA, 2020).



**Figura 2:** Classificação dos Coronavírus (adaptado de MALIK, 2020).

A família dos CoVs é conhecida há muito tempo e com o passar dos anos foi aumentando devido ao reconhecimento de novas espécies. O primeiro coronavírus foi detectado na década de 1930, isolado de embriões de galinha, e foi denominado como vírus da bronquite infecciosa aviária (iBV), seguido, na década de 1940, pelo surgimento do vírus da hepatite do camundongo (MHV) e outros coronavírus de mamíferos. Os primeiros coronavírus humanos (HCoVs), HCoV-229E e HCoV-OC43, foram descobertos entre 1966 e 1967. Além destes, antes de 2003, outros vírus com genoma completo foram conhecidos: coronavírus bovino (BCoV), vírus da encefalomielite hemaglutinante suína (PHEV), vírus transmissível da

gastroenterite (TGEV), vírus da diarreia epidêmica suína (PEDV), coronavírus respiratório suíno (PRCV), e coronavírus felino (FCoV) (WOO *et al.*, 2009; V'KOVSKI *et al.*, 2021).

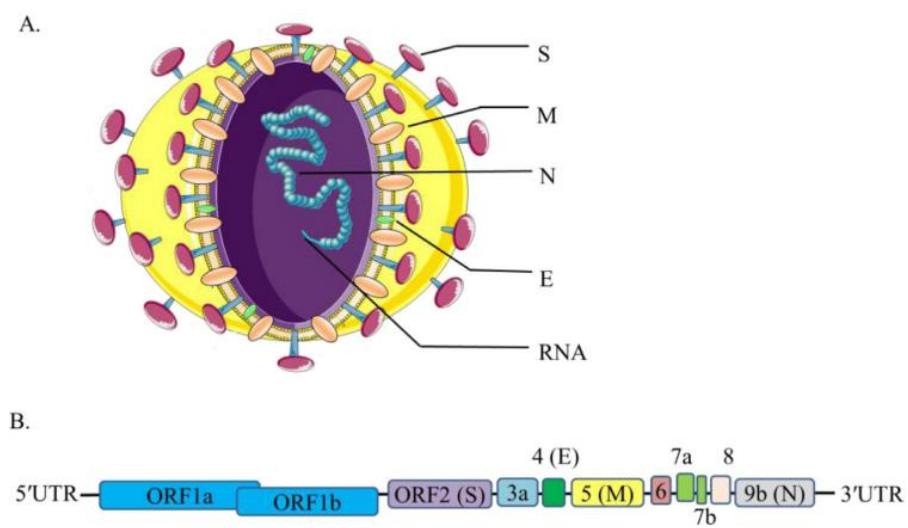
Os CoVs humanos eram conhecidos, até a metade de 2002, por ocasionarem infecções respiratórias mais brandas, com sintomas semelhantes ao resfriado comum. Porém, ao final desse mesmo ano, uma nova espécie de coronavírus se tornou conhecida, provocando um surto, associado à pneumonia, à falência respiratória e a óbitos em Guangdong, na China. O agente causador foi o coronavírus da Síndrome Respiratória Aguda Grave, uma zoonose denominada de SARS-CoV, que causou cerca de 8.000 casos documentados, entre os quais 10% tiveram consequências letais. O surgimento desse vírus foi associado a mutações ocorridas em coronavírus de morcegos, que eram considerados os principais reservatórios dessas espécies. Tendo em vista a transmissão de humano para o humano, medidas de saúde pública foram tomadas, conseguindo conter a disseminação internacional. Entre as medidas estavam: o isolamento social de pacientes infectados e a restrição de viagens. Após a epidemia de SARS, foram identificados os coronavírus HCoV-NL63 e HCoV-HKU1, em consequência do aumento na quantidade de triagem e sequenciamento de vírus. O MERS-CoV - coronavírus causador da Síndrome Respiratória do Oriente Médio, causou mais de 2.500 casos em humanos a partir de 2012, com uma taxa de letalidade de 36%. Ele também teve origem em morcegos e estabeleceu um hospedeiro não humano em camelos dromedários. O coronavírus que causou mais casos e óbitos até hoje foi o SARS-CoV-2, que surgiu e se espalhou rapidamente a partir de 2019 (BORGES *et al.*, 2020; V'KOVSKI *et al.*, 2021; WOO *et al.*, 2009).

### **2.1.2 Morfologia do SARS-CoV-2 e ação viral em células hospedeiras**

O SARS-CoV-2 é um vírus de RNA de fita simples e sentido positivo, envelopado e pertencente ao gênero  $\beta$ -coronavírus. Tem 96% do genoma semelhante ao coronavírus de morcego BatCoV RaTG13, em comparação com SARS-CoV e MERS-CoV, que possuem 78% e 51,8%, respectivamente. Além disso, seu genoma sequenciado tem aproximadamente 30 Kb de comprimento e assim como o SARS-CoV, o novo coronavírus também emergiu a partir das mutações de coronavírus de morcegos (ANDERSEN *et al.*, 2020; ZHOU *et al.*, 2020; GORBALENYA *et al.*, 2020 *apud* BORGES *et al.*, 2020; ZHANG *et al.*, 2021; SHARMA *et al.*, 2021; YAN *et al.*, 2020).

Nos CoVs, o envelope que envolve o nucleocapsídeo (N) é formado por uma bicamada lipídica e nela estão ancoradas as proteínas estruturais do vírus, que são essenciais para a

montagem e infecção, como: a glicoproteína *Spike* (S), a proteína de membrana (M) e a proteína do envelope (E) (Figura 3). O genoma do vírus está organizado na ordem de cap 5' e cauda 3', com duas Regiões não Traduzidas (UTRs) entre as duas extremidades. Os quadros de leituras abertas (ORFs, do inglês *Open Reading Frame*) 1a e 1b são traduzidos em poliproteínas precursoras das proteínas não estruturais, que estão envolvidas nas etapas de transcrição e replicação viral. O restante do material genético é codificado em proteínas estruturais (S, M, E e N) e acessórias, que não possuem função conhecida (LI *et al.*, 2020; HELMY *et al.*, 2020 *apud* BORGES *et al.*, 2020; ALMEIDA *et al.*, 2020).

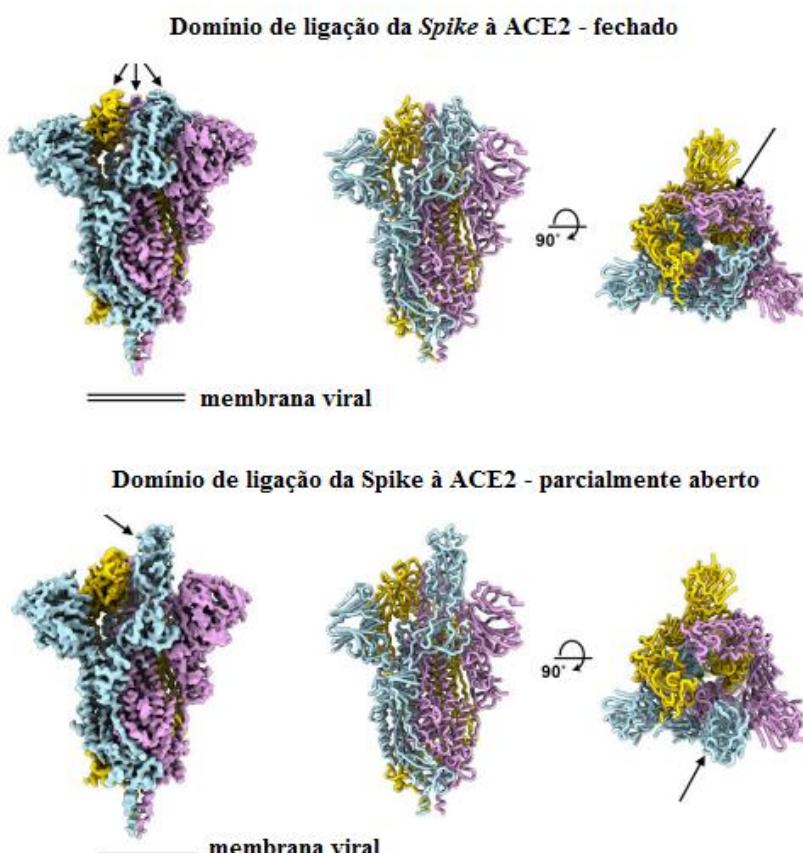


**Figura 3:** Estrutura e genoma do SARS-CoV-2 (LI *et al.*, 2020).

A glicoproteína *Spike* dos coronavírus SARS-CoV, MERS-CoV e SARS-CoV-2 possui entre 1104 e 1273 aminoácidos, sendo dividida em duas regiões: N-terminal (S1) e C-terminal (S2). A primeira é a responsável por determinar a propensão de infecção do vírus numa determinada célula, além da sua transmissibilidade. É nela que está contido o domínio de ligação ao receptor (RBD, do inglês *Receptor Binding Domain*), que possui cerca de 200 resíduos de aminoácidos, consistindo nos seguintes subdomínios: um interno (núcleo) e um externo, sendo que este último contém duas alças externas à superfície, que se ligam com a enzima conversora de angiotensina II (ACE2, do inglês *Angiotensin-Converting Enzyme 2*). O RBD sofre transformações conformacionais, que expõem ou ocultam de maneira transitória os sítios determinantes da ligação ao receptor; esses estados da subunidade são chamados de conformação ‘para cima’ ou estado parcialmente aberto, que corresponde uma posição acessível à ACE2 humana (hACE2), e ‘para baixo’ ou estado fechado, que representa uma posição

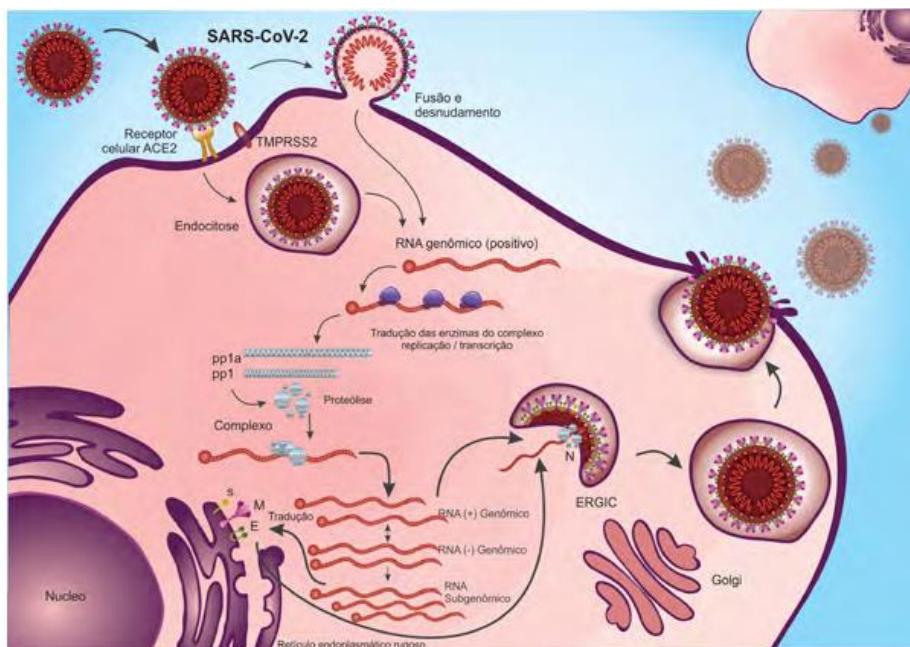
inacessível ao receptor (Figura 4) (LIU *et al.*, 2020; BORGES *et al.*, 2020; ALMEIDA *et al.*, 2020; WALLS *et al.*, 2020).

Assim como o SARS-CoV, o SARS-CoV-2 se liga à ACE2 presente nas células do pulmão, intestino, fígado, coração, endotélio vascular, testículo e rim, por meio da *Spike*. A *Spike* é a glicoproteína responsável pelo reconhecimento do receptor (através da subunidade S1) e fusão da membrana do vírus com a célula do hospedeiro (através da subunidade S2). Nesse processo de reconhecimento viral, a *Spike* é clivada nas subunidades S1 e S2 com o auxílio de proteases celulares. Entre as diversas proteases celulares que podem atuar nas proteínas virais, tanto o SARS-CoV quanto SARS-CoV-2 são afetados pela ação da enzima Serina Protease Transmembranar do tipo 2 (TMPRSS2). As duas espécies de coronavírus podem também entrar na célula por meio da via endocítica, a partir da clivagem por cathepsinas lisossomais. A subunidade S1 se liga ao domínio da peptidase (PD) da ACE2 pelo domínio da ligação ao receptor, desestabilizando o trímero pré-fusão (HOFFMANN *et al.*, 2020; SHANG *et al.*, 2020; ZHANG, *et al.*, 2021, p. 2; YAN *et al.*, 2020; LIMA *et al.*, 2020; LI *et al.*, 2020, p. 2; BORGES *et al.*, 2020).



**Figura 4:** Estruturas obtidas por crio-EM da glicoproteína *Spike* do SARS CoV-2, mostrando as duas conformações adotadas pelo RBD, que permitem ou não o acesso à hACE2 (adaptado de WALLS *et al.*, 2020).

Após a ligação ao receptor, a subunidade S2 é exposta causando a fusão do envelope viral com a membrana celular, por meio do peptídeo de fusão (FP). Em seguida, ocorre o desnudamento e liberação do RNA no citoplasma da célula (Figura 5); esse material genético é imediatamente transcrito e codificado para as poliproteínas pp1a e pp1b, que são então traduzidas e clivadas, formando o complexo replicase-transcriptase e favorecendo a síntese de RNAs genômico e subgenômico virais, onde o primeiro é traduzido em proteínas estruturais no retículo endoplasmático rugoso, seguindo posteriormente para o Compartimento Intermediário Retículo Endoplasmático-Golgi (ERGIC), onde formará os vírions maduros, em conjunto com a proteína N. Após a maturação, os vírus são encaminhados para a via secretória e liberados pelo processo de brotamento, que é quando eles adquirem o envelope por meio da membrana da célula hospedeira. Após a sua saída, eles infectam outras células e estimulam a resposta imunológica do organismo (HOFFMANN *et al.*, 2020; SHANG *et al.*, 2020; TANG *et al.*, 2020; FER, PERLMAN, 2015 *apud* BORGES *et al.*, 2020, p. 11 e 12, cap. 2).



**Figura 5:** Representação da replicação e morfogênese dos coronavírus (BORGES *et al.*, 2020).

### 2.1.3 Mutações, linhagens, cepas e variantes

A probabilidade de o vírus sofrer mutações aumenta com a sua disseminação entre a população, de modo que quanto mais ele se espalha, maior é a sua replicação. A maioria dessas transformações causa pouco ou nenhum impacto na habilidade do vírus em provocar infecções

e doenças. Porém, dependendo da sua localização no material genético, essas mutações podem influenciar em propriedades como transmissão e/ou gravidade da doença (VALVERDE, 2021).

A mutação é um processo natural que está associado à evolução do vírus, principalmente para aqueles que possuem como material genético o ácido ribonucleico (RNA). Ela ocorre durante a replicação viral, devido a falhas na etapa de reparo de erros na síntese da fita de RNA viral. As mutações ocorrem naturalmente na mudança de vetores, como, por exemplo, de um animal para humanos, com a finalidade de impedir a resposta imunológica e aperfeiçoar a interação do vírus com a célula hospedeira. A maior parte das mutações são prejudiciais aos próprios vírus e, assim não causam grandes impactos na evolução da doença, podendo ser benéfica ao hospedeiro (VALVERDE, 2021).

As linhagens apresentam mutações semelhantes e, portanto, são definidas como entidades/organismos com um ancestral em comum. Elas surgem por meio de alterações que são prejudiciais a essas entidades, em sua maioria. A cepa deriva da mutação, quando esta altera pelo menos uma das características fenotípicas do vírus, ou seja, sua capacidade de transmissão, multiplicação, indução dos sintomas no hospedeiro ou estímulo da resposta imunológica diferente de seu ascendente (VALVERDE, 2021).

Os vírus são considerados como organismos simples, com material genético pequeno que contém informações relacionadas às suas características moleculares e biológicas. Variantes são “sequências genéticas virais que diferem em uma ou mais mutações”. Foram detectadas em todo o mundo, até 2021, aproximadamente mil variantes do SARS-CoV-2. Desse total, cerca de 60-100 circulam no Brasil. As variantes de preocupação (ou VOC, do inglês *Variant of Concern*) compõem as linhagens com mais transmissibilidade, “patogenicidade e/ou maior escape dos mecanismos protetores induzidos pelas vacinas”. O monitoramento dessas variantes é muito importante para determinar o espalhamento do vírus pela comunidade (VALVERDE, 2021).

Na Tabela 1 estão relacionadas algumas das variantes de preocupação e suas respectivas mutações da proteína *Spike*, adquiridas pelo SARS-CoV-2. Essas mutações são responsáveis por elevar a transmissibilidade e/ou virulência, podendo reduzir a eficácia da vacina. Atualmente a variante Ômicron é a dominante na maior parte dos países, tendo sido considerada como variante de preocupação pela OMS, em novembro de 2021, por ter um número maior de mutações em comparação com as demais variantes, sendo que a maior parte dessas alterações ocorre na proteína *Spike*. Vale ressaltar que na tabela não estão listadas todas as mutações das variantes (HADJ HASSINE, 2022; GHOSH *et al.*, 2022, INSTITUTO BUTANTAN, 2022).

**Tabela 1:** Variantes de preocupação e mutações de interesse da *Spike* (continua).

Variante	Linhagem	Mutações	Mês e ano da primeira detecção	País da primeira detecção
Alfa	B.1.1.7	N501Y, D614G P681H, E484K Y144-L452R S494P, A570D T716I, S982A, D1118H, K1191N	Setembro 2020	Reino Unido
Beta	B.1.351	K417N, E484K N501Y, D614G D80A, D215G, P384L, E516Q	Outubro 2020	África do Sul
Gama	P.1	K417T, E484K N501Y, D614G L18F, T20N, P26S, D138Y, R190S, H655Y, P681H, T1027I.	Janeiro 2021	Brasil
Delta	B.1.617.2	L452R, D614G P681R, L452Q	Dezembro 2020	Índia
Ômicron	B.1.1.529	A67V, T95I, G142D, Y145D, N211I, L212I, G339D, R346K, S371L, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493R, Q496S, Q498R, N501Y, Y505H, T547K, D614G,	Novembro 2021	África do Sul

---

H655Y, N679K,  
P681H, N764K,  
D796Y, N856K,  
Q954H, N969K,  
L981F

---

Fonte: HADJ HASSINE, 2022 (adaptado); GHOSH *et al.*, 2022; INSTITUTO BUTANTAN, 2022.

## 2.2 Química Medicinal: conceito e reconhecimento molecular

A Química Medicinal, é uma ciência multidisciplinar, que engloba disciplinas como a farmacologia, a bioquímica, a química orgânica sintética e a química computacional, entre outras. De acordo com a União Internacional de Química Pura e Aplicada (IUPAC, do inglês, *International Union of Pure and Applied Chemistry*), a Química Medicinal tem por missão planejar, descobrir, identificar e preparar os compostos biologicamente ativos, bem como, estudar seu metabolismo, interpretar o mecanismo de ação do composto a nível molecular e construir a relação entre a estrutura química do composto e a sua atividade farmacológica (SAR) (VERLI, 2005; LIMA, 2007).

O estudo do reconhecimento molecular na Química Medicinal pode contribuir nas etapas de planejamento e/ou desenvolvimento de novos compostos químicos que modulem um determinado alvo biológico. Esse reconhecimento se baseia na presença de interações intermoleculares específicas entre o ligante e seu biorreceptor, que em geral é uma proteína (FOKOUÉ *et al.*, 2020).

### 2.2.1 Interações intermoleculares envolvidas no reconhecimento molecular do ligante no sítio receptor

Qualitativamente, as interações intermoleculares determinam o grau de afinidade e a especificidade da ligação entre ligante-receptor. Assim, o estabelecimento dessas interações se torna importante para que se tenha o efeito biológico ou farmacológico da molécula bioativa. As interações podem sozinhas estabelecer a ação da molécula bioativa ou serem complementadas por reações com proteína-alvo, que é característico das inibições enzimáticas irreversíveis, nas quais se forma uma ligação covalente entre o inibidor e o sítio catalítico da enzima (BARREIRO e FRAGA, 2015; SANT'ANNA, 2009).

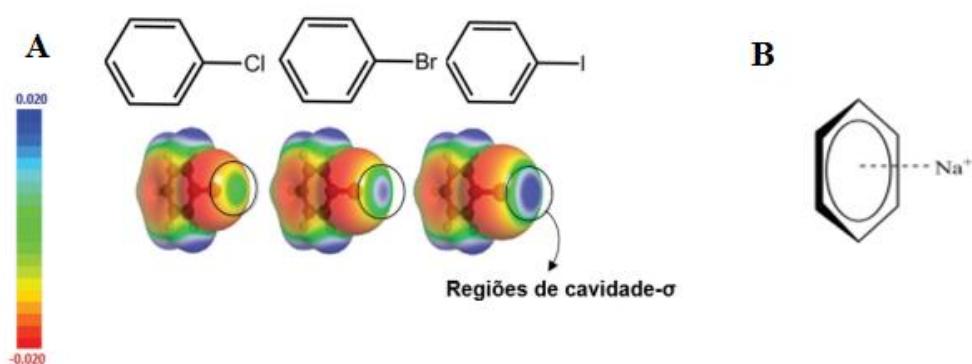
As forças de atração eletrostáticas resultam da atração entre dipolos e/ou íons de cargas opostas, que é dependente da magnitude das cargas e dos dipolos, da distância entre as cargas e/ou dipolos, assim como da constante dielétrica do meio. Alguns aminoácidos, presentes nos biorreceptores, encontram-se ionizados em pH fisiológico e, dessa maneira, podem interagir com grupos carregados positivamente ou negativamente dos fármacos. As interações podem ser entre cargas inteiras ou parciais, como as decorrentes da diferença de eletronegatividade entre átomos ligados covalentemente, formando dipolos permanentes. Nesse caso, podemos ter as chamadas interações **dipolo-dipolo**, que ocorrem entre cargas parciais opostas presentes em dipolos adjacentes, e **íon-dipolo** (BARREIRO e FRAGA, 2015). Deve ser lembrado que dipolos podem surgir mesmo em ligações entre átomos de mesma eletronegatividade, em razão do movimento eletrônico que pode temporariamente concentrar maior densidade eletrônica em um dos átomos da ligação, formando os chamados dipolos instantâneos. Nesse caso, podem surgir as chamadas interações dipolo instantâneo-dipolo induzido, mediadas pelas forças de dispersão de London.

As **ligações de hidrogênio**, são responsáveis por manter, por exemplo, as conformações bioativas de macromoléculas essenciais à vida, tais como nas  $\alpha$  hélices e folhas  $\beta$  das proteínas e entre as bases purínicas e pirimidínicas dos ácidos nucleicos. A ligação de hidrogênio clássica é uma interação não covalente que ocorre entre heteroátomos muito eletronegativos (O, N ou F) e um átomo de H já ligado covalentemente a um desses heteroátomos (X); há uma forte polarização da ligação X-H, que leva a uma carga parcial positiva elevada no H. A energia associada a uma ligação de hidrogênio varia de -2 a -10 kcal/mol (BARREIRO e FRAGA, 2015; FOKOUÉ *et al.*, 2020).

As **ligações de halogênio** são decorrentes de efeitos quânticos da ligação carbono-halogênio (Cl, Br ou I) na densidade eletrônica do átomo de halogênio, formando a chamada cavidade- $\sigma$ , uma região de baixa densidade eletrônica e, consequentemente, com potencial eletrostático positivo no átomo de halogênio, oposta à ligação  $\sigma$  com o carbono (Figura 6A). É por meio dela que os átomos de halogênios interagem, de modo linear, com átomos ou grupos que podem atuar como base de Lewis neutras ou aniônicas (como N, O, S e P) e, dependendo do átomo de halogênio envolvido, sua energia varia entre -1 e -5 kcal/mol (BARREIRO e FRAGA, 2015; FOKOUÉ *et al.*, 2020; ERDELYI, 2012).

A interação **cáton- $\pi$**  é não covalente e ocorre entre um cátion adjacente e um sistema rico em elétrons  $\pi$  (Figura 6B), preferencialmente quando o cátion está localizado sobre o plano do sistema  $\pi$  a uma distância média de 2 a 4 Å. A interação cátion- $\pi$  libera uma energia

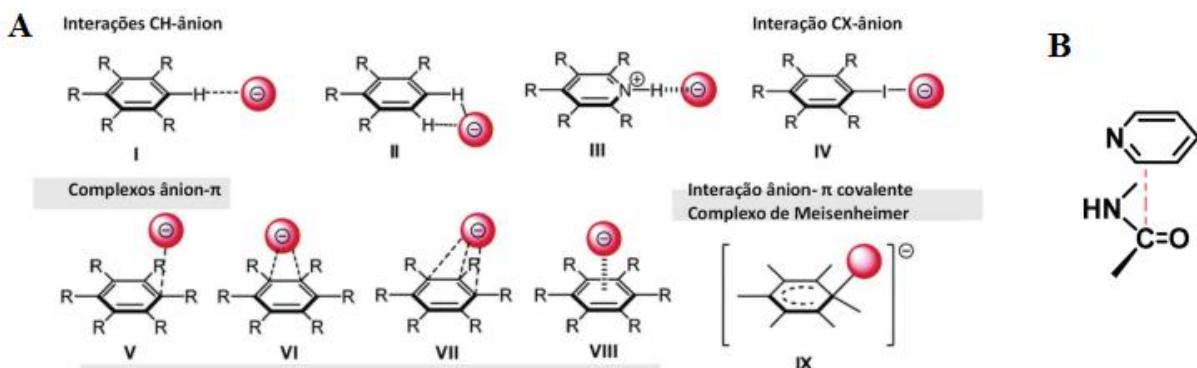
significativa em meio aquoso; experimentalmente foi demonstrado que elas podem estabilizar as energias de interação em -2 a -5 kcal/mol, tornando-as competitivas com ligações de hidrogênio nas interações ligante-receptor e proteína-proteína (BARREIRO e FRAGA, 2015; FOKOUÉ *et al.*, 2020).



**Figura 6:** (A) Exemplos de regiões de cavidade- $\sigma$  em halogênios; (B) exemplo de interação cátion- $\pi$  (adaptado de FOKUE *et al.*, 2020).

Há também a interação **ânion- $\pi$** , que ocorre entre ânions um sistema aromático deficiente em elétrons (Figura 7A), podendo este estar carregado positivamente ou neutro, mas com baixa densidade eletrônica sobre o anel (por ex., anéis triazínicos ou sistemas perfluorados). Dessa maneira, se caracteriza como uma interação iônica ou íon-dipolo, que pode ocorrer com um ângulo entre 0° e 90°. Essa interação dificilmente ocorre em condições biológicas, visto que as cadeias laterais dos aminoácidos repelem as cargas negativas próximas à sua nuvem  $\pi$  (BARREIRO e FRAGA, 2015; FOKOUÉ *et al.*, 2020).

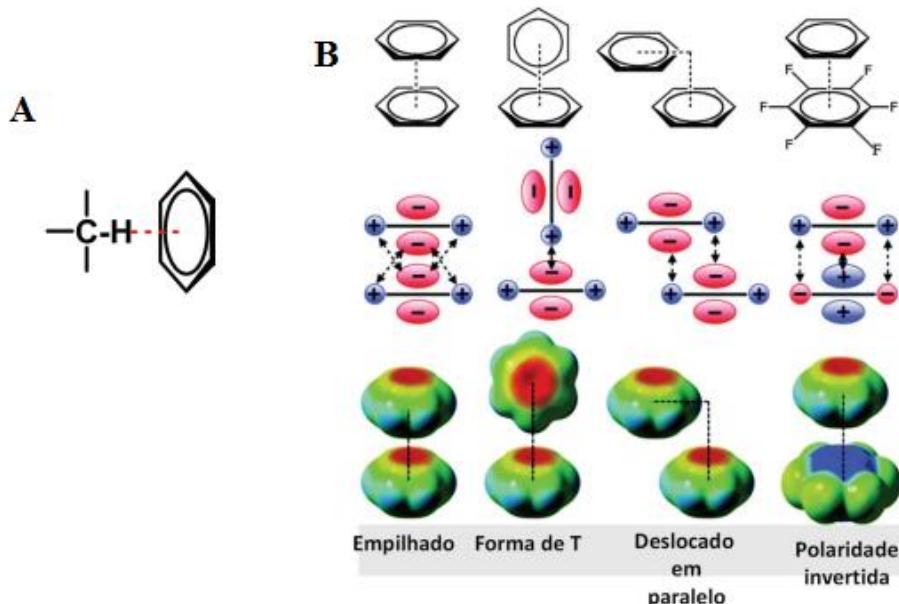
O **empilhamento amida- $\pi$**  se dá pela interação molecular não covalente entre um grupo amida e um sistema rico em elétrons  $\pi$  (Figura 7B), principalmente quando polarizado pela presença de heteroátomo. Essa interação é considerada um tipo de dipolo-dipolo. Por meio de estudos computacionais, com o sistema *N*-metil-acetamida/piridina, foi demonstrado que a energia envolvida nessa interação se torna mais favorável à medida que o ângulo entre os dipolos varia de 0° a 180° e 240°, considerados como melhores valores (FOKOUÉ *et al.*, 2020; BARREIRO e FRAGA, 2015).



**Figura 7:** (A) Possíveis modos de interação entre um ânion e um anel aromático; (B) Interação amida- $\pi$  (adaptado de FOKUE *et al.*, 2020).

Há também a interação **alquila-arila**, que segundo alguns autores é um tipo de ligação de hidrogênio não clássica, que ocorre pela polarização da ligação C-H pela presença de um anel com sistema  $\pi$  (Figura 8A). Ela pode ter uma significativa função na conformação e estabilidade dos complexos formados entre ligante-proteína e proteína-proteína. Normalmente, o C-H está posicionado perpendicularmente em relação ao sistema  $\pi$ , apontando para seu centro numa distância de 3,4 a 4,4 Å. O tipo de fragmento ligado ao CH doador afeta a energia de interação desse grupo com o sistema  $\pi$ , sendo que a interação é mais favorecida por grupos retiradores de elétrons, com energias variando entre -1 a -5 kcal/mol (FOKOUÉ *et al.*, 2020; BARREIRO e FRAGA, 2015).

No **empilhamento  $\pi$** , a interação acontece entre anéis aromáticos e/ou heteroaromáticos (Figura 8B), devido à polarização da densidade eletrônica ou pela atração, em alguns casos, entre o anel aromático e a carga parcial positiva  $\sigma$  de um átomo presente nele. Há quatro modos para esse tipo de interação: empilhado, onde a energia de interação é menos favorável: -1,48 kcal/mol; forma de T, no qual os planos dos anéis podem chegar a 90° e a energia de interação é cerca de -2,50 kcal/mol; deslocado em paralelo, com energia também de aproximadamente -2,50 kcal/mol; e, polaridade invertida, que é favorecida a partir, por exemplo, da fluoração de um dos anéis aromáticos (FOKOUÉ *et al.*, 2020; BARREIRO e FRAGA, 2015).



**Figura 8:** (A) Interação alquila-arila; (B) Exemplos de empilhamento  $\pi$  e suas respectivas orientações (adaptado de FOKUE *et al.*, 2020).

As **interações hidrofóbicas** também são consideradas importantes, porém mal compreendidas, em geral, na Química Medicinal. Elas podem ser definidas como a tendência de hidrocarbonetos (ou grupos lipofílicos) em se agregar em meio aquoso, pois essa aproximação faz com que as moléculas de água, que os circundam, sejam liberadas para o meio e, assim, aumentam a entropia do sistema. Dessa maneira, as moléculas de água que retornam para o solvente são capazes de fazerem mais ligações de hidrogênio do que se estivessem na superfície da estrutura. Em resumo, as interações hidrofóbicas provocam o aumento da entropia, pela liberação de moléculas de água, além da redução da entalpia, devido à formação de novas ligações de hidrogênio (BARREIRO e FRAGA; FOKOUÉ *et al.*, 2020).

As **interações ortogonais**, consideradas como um tipo de dipolo-dipolo ocorre na aproximação ortogonal de dois grupos funcionais dipolares. Elas são encontradas entre os grupos  $C-X\cdots C=O$ ,  $C=O\cdots C=O$  e quando dipolos como  $C\equiv N$ ,  $NO_2$ ,  $S=O$  e espécies nucleofílicas fracas, como compostos contendo as funções éter, sulfeto, álcool e imina interagem entre si ou com uma carbonila. O ângulo envolvido nessas interações varia entre  $70^\circ$  e  $110^\circ$ , numa distância média de 3 a 4 Å. Devido às interações ortogonais, foi observado em dados cristalográficos que o ângulo muitas vezes acaba perdendo a planaridade em até  $5^\circ$  (BARREIRO e FRAGA; FOKOUÉ *et al.*, 2020).

Além dessas interações, na Química Medicinal, as **ligações covalentes** desempenham um papel importante. Elas possuem elevada energia para serem quebradas e são raramente (ou

muito lentamente) desfeitas em condições biológicas, na temperatura comum desse sistema. Logo, complexos fármaco-receptor envolvendo ligações covalentes resultam em inibição enzimática irreversível ou inativação do sítio receptor. A formação desse tipo de ligação pode ocorrer entre fármacos que possuem grupos com caráter eletrofílico acentuado (como ácidos de Lewis) e bionucleófilos orgânicos (BARREIRO e FRAGA; FOKOUE *et al.*, 2020).

### **2.2.2 Peptídeos terapêuticos**

Os peptídeos são moléculas que, devido ao seu tamanho, estrutura e bioatividade, se encontram entre as moléculas pequenas e as proteínas. Eles estão presentes por todo o corpo humano, em células e tecidos e estão, portanto, envolvidos em processos fisiológicos vitais. São compostos por resíduos de aminoácidos ligados, de modo covalente em cadeia, por meio de uma ligação amida substituída denominada de ligação peptídica. Essa ligação ocorre a partir da reação entre um grupo amino (-NH<sub>2</sub>) de um aminoácido com o grupo carboxila (-COOH) de outro aminoácido, através de uma reação de desidratação, tendo em vista que uma molécula de água é liberada (BERG *et al.* 2007). Tanto peptídeos, quanto as proteínas possuem essas características em comum, diferindo apenas no comprimento, onde o peptídeo geralmente tem entre 2 e 50 resíduos aminoácidos e, a proteína é composta por centenas ou até milhares deles (ZHANG, 2020; SANTOS *et al.*, 2022; LI *et al.*, 2018).

O uso terapêutico de peptídeos teve início em 1922, com a insulina, composta por 51 aminoácidos – extraída do pâncreas bovino e suíno – no tratamento de pacientes com diabetes mellitus tipo 1, descoberta que rendeu o prêmio Nobel, em 1923, aos cientistas Frederic G. Banting e John J. R. Macleod, pela inovação. Em 1950, com a possibilidade da elucidação da sequência peptídica e da síntese, tanto a ocitocina sintética quanto a vasopressina foram clinicamente utilizadas. Os peptídeos ganharam espaço no mercado farmacêutico devido aos avanços das técnicas de síntese química e biotecnologia. Entre 1960 e 1970, aprovou-se 5 fármacos dessa classe, sendo que esta quantidade subiu para 29 nas duas décadas seguintes. A partir do ano 2000 até 2019, 52 aquisições foram feitas ao mercado farmacêutico mundial, que estimou movimentar, em 2019, U\$ 1,2 trilhão anualmente, onde 5% desse valor correspondia aos peptídeos terapêuticos. Em 2021, havia cerca de 80 medicamentos à base de peptídeos no mercado global, enquanto 150 estavam em desenvolvimento clínico e, entre 400-600 em estudos pré-clínicos (SANTOS *et al.*, 2020; WANG *et al.*, 2022; MUTTENTHALER *et al.*, 2021; LAU e DUNN, 2018; ZHANG, 2020; SANTOS *et al.*, 2022).

Comparando os peptídeos com as moléculas pequenas, eles apresentam como vantagens a alta especificidade, boa eficácia e segurança e, como desvantagens, a baixa estabilidade, meia-vida curta e menor permeabilidade de membrana. Já em relação às proteínas, possuem como vantagem a permeabilidade de membrana, tendo como desvantagens menor estabilidade, baixa imunogenicidade e meia-vida curta. E, são esses fatores negativos que também afetam a biodisponibilidade oral do peptídeo, devido ao curto tempo de meia-vida plasmática, que é da ordem de minutos para algumas espécies não modificadas, em razão da existência das peptidases, que facilmente hidrolisam as ligações peptídicas quando não há proteção na exposição, além das taxas rápidas de eliminação renal e hepática do organismo humano. Ademais, os peptídeos de baixa lipofilicidade, alta flexibilidade e alta capacidade de ligação de hidrogênio intermolecular apresentam permeabilidade reduzida às membranas celulares e, assim, são impedidos de atingirem os receptores intracelulares e cruzar as barreiras fisiológicas em áreas de absorção, como o trato intestinal delgado. De acordo com LAU e colaboradores, em 2018 os peptídeos em desenvolvimento clínico, visando alvos extracelulares, ultrapassou a porcentagem de 90% (WANG *et al.*, 2022; ZHANG, 2020; LAU e DUNN, 2018; LI *et al.*, 2018).

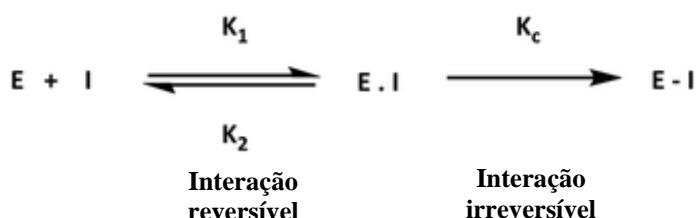
Como consequência da baixa biodisponibilidade oral, as vias de administração subcutânea, intravenosa ou intranasal se tornam fortes candidatas à aplicação no desenvolvimento de fármacos baseados em peptídeos, sendo consideradas como opções menos atraentes para indicação no tratamento de pacientes com doenças crônicas. Contudo já existem métodos que visam melhorar as propriedades farmacológicas dos peptídeos, como exemplo, o aumento em seu tamanho, através da conjugação com outras moléculas, que minimizaria a filtração e a depuração renal, estendendo sua meia-vida. Para aumentar sua resistência proteolítica, três abordagens podem ser utilizadas, em conjunto ou individualmente: 1) aumento do impedimento estérico ao redor dos locais de clivagem, a partir do uso de macromoléculas; 2) inibição da ligação e ação da enzima, por meio da alteração de peptídeos naturais que são determinantes para que os processos de reconhecimento e clivagem enzimática ocorram; 3) aumento da rigidez conformacional, com a incorporação de aminoácidos com cadeias laterais volumosas ou ciclizadas, de modo a inibir que o peptídeo adote a conformação ideal para a clivagem (ZHANG, 2020; LAU e DUNN, 2018; LI *et al.*, 2018).

Novas aplicações de peptídeos terapêuticos estão sendo exploradas. No tocante à COVID-19, foi demonstrado que peptídeos derivados da hACE2, projetados com base nas interfaces de ligação RBD/ACE2 de SARS-CoV-2 e SARS-CoV, são capazes de inibir a

infecção mediada pela glicoproteína *Spike* de ambos os vírus, com valores de IC<sub>50</sub> na faixa milimolar baixa (LARUE *et al.*, 2020).

### 2.2.3 Inibição covalente

Em geral, os inibidores covalentes são compostos destinados a formar ligação covalente com um alvo determinado, que pode ser reversível ou irreversível, dependendo do grupo ogiva (ou seja, aquele que formará a ligação) escolhido para atingir resíduos de aminoácidos específicos, dentre os quais se destacam a cisteína, serina, treonina, tirosina e lisina. A ligação entre o inibidor covalente e seu alvo ocorre em duas etapas: 1) equilíbrio de formação do complexo de interação, que consiste numa etapa reversível e não covalente; 2) formação da ligação covalente, a partir da reação com um resíduo específico do receptor, resultando no estado final essencialmente irreversível (Figura 9), que visa prolongar a resposta terapêutica, devido ao reforço da interação fármaco-receptor (SUTANTO *et al.*, 2020; BERDAN *et al.*, 2021).



**Figura 9:** Mecanismo de interação entre um inibidor covalente e seu alvo (SUTANTO *et al.*, 2020).

Durante o processo de pesquisa e desenvolvimento de fármacos, cerca de 90% deles falham na fase II da etapa dos ensaios clínicos, em consequência da alta toxicidade e baixa eficácia. Diante disso, os inibidores covalentes apresentam as seguintes vantagens: 1) melhor eficiência; 2) menor dose, devido ao baixo valor de IC<sub>50</sub> e a longa interação com o alvo; 3) aumento da adesão de pacientes, como consequência do tópico anterior; e 4) capacidade de ligação em cavidades rasas do receptor. Contudo, alguns possuem como desvantagem a toxicidade elevada, associada ao grupo ogiva, que pode ser minimizada a partir de um maior conhecimento a respeito do mecanismo de ação e da reatividade desse grupo, que possibilitaria a realização de ajustes em suas propriedades (PAUL *et al.*, 2021; SUTANTO *et al.*, 2020; BAUER, 2015).

## 2.3 Métodos e Ferramentas Computacionais

Os métodos e ferramentas computacionais vem desempenhando um importante papel no planejamento e desenvolvimento de fármacos, visto que melhoram a eficiência de custo e tempo. Podendo ser aplicados em todas as etapas do processo, desde a triagem até os estágios pré-clínicos, reduzindo o risco de falha. Esta abordagem é conhecida como Planejamento de Fármacos Auxiliado por Computadores (CADD). As estratégias empregadas para o CADD são divididas em: 1) Planejamento de Fármacos Baseado nos Ligantes (LBDD, do inglês *Ligand Based Drug Design*) e 2) Planejamento de Fármacos Baseado na Estrutura do Alvo (SBDD, do inglês *Structure Based Drug Design*). A primeira estratégia depende do conhecimento prévio sobre os compostos bioativos, incluindo suas estrutura e atividade, para se construir modelos com efeitos biológicos semelhantes. As estruturas podem ser usadas diretamente ou substituídas por descritores variados. O LBDD é geralmente empregado quando a estrutura 3D da proteína alvo não é conhecida. No SBDD é explorado o modo de interação entre ligantes e uma proteína alvo, de forma a se estudar os fatores estruturais importantes para a interação e a partir daí propor novos candidatos ou modificações em ligantes já existentes (SHAKER *et al.*, 2021; SANT'ANNA, 2002).

Cabe ressaltar que a molécula bioativa não é investigada somente através da aplicação de programas de computador. A descoberta de um fármaco, devido à sua complexidade, abrange o trabalho de uma equipe multidisciplinar, que pode usar um conjunto de métodos computacionais como meio de otimizar o processo de pesquisa e desenvolvimento dos compostos com atividade biológica, numa constante troca de informações e decisões com grupos envolvidos com a síntese e a avaliação das atividades dessas moléculas (SANT'ANNA, 2009).

Qualquer que seja o método de modelagem molecular, ele depende em algum momento da geração de estruturas moleculares com qualidade comparável à de métodos experimentais. Para isso, métodos baseados em duas diferentes aproximações físicas são predominantemente aplicados: mecânica clássica, que incluem os métodos da mecânica molecular e dinâmica molecular; e mecânica quântica, que inclui os métodos *ab initio* e semi-empíricos. De modo que a escolha entre eles é dependente das propriedades a serem avaliadas, da precisão desejada e da capacidade computacional disponível para a realização dos cálculos (SANT'ANNA, 2009).

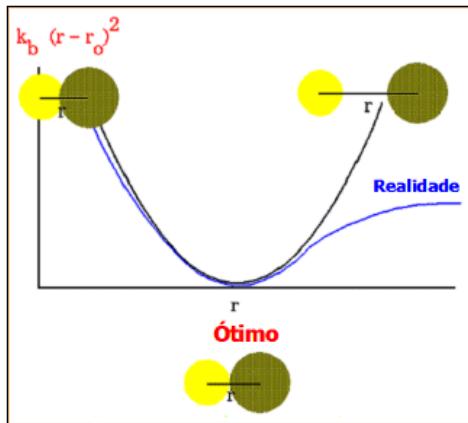
## 2.3.1 Métodos clássicos

### 2.3.1.1 Mecânica molecular

Nesse método, as moléculas não são representadas por núcleos e elétrons, mas como um conjunto de esferas unidas por molas com constantes de força características. A energia é calculada a partir do campo de força, que é um conjunto de funções de energia associadas a essa representação. Essas funções de energia são relacionadas a posições de equilíbrio do sistema (descritas através de parâmetros como distâncias de ligação, ângulos de ligação, ângulos diedros, distâncias de van der Waals, ligações de hidrogênio, interações eletrostáticas etc.) e determinam penalidades energéticas para o afastamento dos parâmetros de seus valores “ideais”. A fim de reduzir essas penalidades (ou seja, tornar a energia a um mínimo), usa-se um algoritmo matemático, num processo conhecido como minimização de energia. Em geral, as constantes de força necessárias são avaliadas por meio de dados espectroscópicos ou por cálculos quânticos. A principal vantagem da mecânica molecular é a rapidez na avaliação de sistemas moleculares complexos. Esse modelo é justificável porque os parâmetros associados a conjuntos de átomos, como distâncias e ângulos de ligação, permanecem razoavelmente constantes entre estruturas diferentes, desde que o tipo e a hibridação dos átomos envolvidos sejam os mesmos (BARREIRO *et al.*, 1997; SANT’ANNA, 2015; SANT’ANNA, 2009).

Um exemplo de uma das funções do campo de força é o termo de energia para o estiramento de ligações químicas. Uma ligação química tem o comportamento identificado pela curva da realidade (Figura 10), na qual quando os átomos estão infinitamente distantes, a energia tende à zero; enquanto à medida que são aproximados a energia é reduzida até se atingir a distância ótima, que corresponde ao mínimo da energia. Se os átomos são aproximados além desse ponto, a energia volta a aumentar. A curva representada pela equação 1 têm em comum com a curva da realidade a região de mínimo de energia; nessa função, “ $r$ ” é a distância variável; “ $r_0$ ” é o comprimento ótimo da ligação; “ $K_b$ ” é a constante de força, que informa a rigidez da ligação, de maneira que quanto maior for o seu valor, mais rígida será a ligação e, consequentemente, o estiramento da ligação irá requerer mais energia (SANT’ANNA, 2009).

$$E = \sum K_b (r - r_0)^2 \quad (1)$$



**Figura 10:** Energia em função da distância de ligação descrita pelo modelo da mecânica molecular em comparação com o comportamento real (SANT'ANNA, 2009).

Atualmente, os campos de força incluem termos para interações atômicas à distância, que consideram efeitos entre os átomos separados por mais de três ligações. Esses termos são conhecidos como “termos não ligados” e são de dois tipos: de van der Waals e eletrostático (ou de Coulomb). O termo de van der Waals é comumente calculado pela soma dos parâmetros de repulsão e de atração de Lennard-Jones (equação 2), na qual A e B são constantes e “ $r_{ij}$ ” é a distância entre os átomos i e j. Nesta equação,  $A/r_{ij}^{12}$  representa as repulsões a curta distância, enquanto  $-B/r_{ij}^6$  descreve as atrações, geradas pelas forças de dispersão de London. A partir da aplicação da lei de Coulomb, entre pares de átomos com cargas parciais ou reais ( $q_i$  e  $q_j$ ) separados por certa distância ( $r_{ij}$ ), é obtido o termo eletrostático (equação 3), onde  $D_q$  é a constante dielétrica efetiva, que é dependente do meio, sendo em estudos de química medicinal o uso mais comum da constante dielétrica da água ( $\epsilon = 78,4$ ) (SANT'ANNA, 2015).

$$E = \frac{A}{r_{ij}^{12}} - \frac{B}{r_{ij}^6} \quad (2)$$

$$E = \frac{q_i q_j}{D_q r_{ij}} \quad (3)$$

### 2.3.2 Métodos quânticos

Um dos fatores limitantes nas aproximações clássicas é que os elétrons não são considerados. Assim, nas aplicações que precisem considerá-los, é necessário o uso de métodos quanto-mecânicos, tendo em vista que, nestes casos, a estrutura molecular é vista como um conjunto de núcleos e elétrons. O desenvolvimento desses métodos teve como principal complicador o comportamento dual partícula-onda do elétron. Para incluir esse comportamento no cálculo da energia  $E$  do sistema, Schrödinger propôs a equação 4, na qual  $\psi(R,r)$  é a função de onda, que depende das coordenadas de todos os núcleos ( $R$ ) e elétrons ( $r$ ), e  $H$  é o chamado operador Halmitoniano (a soma de operadores para as energias cinética e potencial) (SANT'ANNA, 2015; SANT'ANNA, 2009).

$$E \psi(R,r) = H \psi(R,r) \quad (4)$$

Por ser bastante complicada, a resolução da equação 4 passa por algumas aproximações para que possa ser utilizada, na prática, em sistemas de interesse. Uma delas é chamada aproximação de Born-Oppenheimer: os núcleos são mantidos fixos em certas posições do espaço (coordenadas  $R$  fixas) para que a equação seja resolvida apenas para os elétrons submetidos ao potencial produzido pelos núcleos “congelados”. Para que se obtenha a energia total, a energia de repulsão entre os núcleos é calculada e somada com a energia eletrônica calculada pela resolução da equação de Schrödinger. Na aproximação LCAO (Combinação Linear de Orbitais Atômicos; do inglês, *Linear Combination of Atomic Orbitals*), a função de onda molecular  $\Psi_M$  é substituída por um somatório de funções de onda atômicas (equação 5), na qual  $\Psi_{Ai}$  e  $C_i$  são, respectivamente, as funções de onda dos orbitais atômicos e os coeficientes que permitem a atribuição de “pesos” na participação de cada orbital atômico em cada orbital molecular (SANT'ANNA, 2015).

$$\Psi_M = \sum C_i \Psi_{Ai} \quad (5)$$

Como não são conhecidas soluções analíticas da equação de Schrödinger para sistemas com mais do que um elétron, a resolução para átomos ou moléculas multieletrônicos é alcançada por meio de métodos numéricos, como o chamado método Hartree-Fock, dentro da aproximação de Born-Oppenheimer. No método Hartree-Fock a função de onda multieletrônica é obtida pela solução numérica de um conjunto de funções de onda de um elétron (SANT'ANNA, 2015).

### **2.3.2.1 Métodos *ab-initio***

Nos métodos *ab initio* não é feita nenhuma aproximação adicional na resolução da equação de Schrödinger. Além disso, os orbitais atômicos necessários são obtidos pelos chamados conjuntos de base, que são um grupo de funções matemáticas adequadas para representar cada orbital atômico (SANT'ANNA, 2015; SANT'ANNA, 2009; SANT'ANNA, 2002).

O conjunto de base mínimo inclui somente um número de funções de onda necessárias para a acomodação de todos os elétrons de um átomo, como por exemplo, 1s para os átomos de H e He, as funções 1s e 2s para Li e Be, e as funções 1s, 2s, 2px, 2py e 2pz para os demais elementos do 2º período. Também são incluídas funções do tipo 2p para descrever melhor os átomos de Li e Be (SANT'ANNA, 2015).

Os conjuntos de base de valência dividida permitem uma maior flexibilidade na descrição dos elétrons, visto que a representação deles é dividida em dois ou mais componentes. Os conjuntos de base de polarização permitem pequenos deslocamentos dos elétrons em relação às ligações que conectam os núcleos, a partir da introdução de funções de número quântico secundário maior (funções do tipo p para os átomos de H e funções do tipo d para os demais). Os conjuntos de base contendo funções difusas representam melhor sistemas aniônicos e estados excitados e são obtidos pela adição de orbitais s e p muito difusos para melhorar a descrição de pares de elétrons de alta energia (SANT'ANNA, 2015; SANT'ANNA, 2009).

### **2.3.2.2 Métodos semi-empíricos**

São cálculos de orbital molecular (OM), que consistem na introdução de parâmetros empíricos ou previamente calculados, e que envolvem apenas os elétrons de valência na resolução da equação de Schrödinger, já que os elétrons de camada interna são tratados como

partes de um cerne (“caroço”), que contém o núcleo. Além disso, uma parte das integrais referentes à resolução da equação de Schrödinger são negligenciadas e substituídas por parâmetros obtidos a partir de dados experimentais, de modo a se ter uma diminuição do tempo de cálculo (SANT’ANNA, 2015; SANT’ANNA, 2009).

A partir dos trabalhos de Dewar e colaboradores, nas décadas de 1970 e 1980, uma série de programas foi desenvolvida tendo como objetivo de tornar os cálculos semi-empíricos de orbitais moleculares acessíveis aos não especialistas. Além de fornecer também informações relativas às estruturas químicas, com precisão e custo razoável de tempo de cálculo, o que só foi alcançado parcialmente (SANT’ANNA, 2015; SANT’ANNA, 2009).

Os métodos MINDO/3 (*Modified Intermediate Neglect of Differential Overlap*) e MNDO (*Modified Neglect of Diatomic Overlap*) foram as primeiras versões propostas pelos autores, mas os métodos semi-empíricos AM1 (*Austin Model 1*) e o PM3 (*Parametric Method 3*) – este desenvolvido por J. J. P. Stewart, tiveram resultados com qualidade para serem aceitos. Ambos possuem aproximações semelhantes, diferindo nas parametrizações utilizadas (SANT’ANNA, 2015; SANT’ANNA, 2009).

Os métodos semi-empíricos são aplicados de modo limitado a sistemas contendo apenas os elementos para os quais tiveram os parâmetros desenvolvidos. No caso do método AM1, estes elementos são: H, Na, K, Rb, Zn, Hg, B, Al, C, Si, Ge, Sn, N, P, O, S, F, Cl, Br e I; o método PM3 não contêm parâmetros para B, mas é parametrizado para outros elementos que não estão contidos no método AM1, como: Be, Mg, Cd, Ga, In, Tl, As, Sb, Bi, Se e Te (SANT’ANNA, 2015; SANT’ANNA, 2009).

Novas parametrizações, com desempenhos superiores aos métodos semi-empíricos anteriores, foram propostas nos últimos anos: RM1 (*Recife Model 1*) - desenvolvido na Universidade Federal de Pernambuco juntamente com Stewart; e PM6 (*Parametric Method 6*) e PM7 (*Parametric Method 7*), versões mais recentes do método PM3, que dispõem de parâmetros para 70 elementos da tabela periódica (SANT’ANNA, 2015; SANT’ANNA, 2009).

O método de orbitais semi-empírico PM7 é um dos mais recentes dessa categoria e apresenta algumas melhorias em comparação com seus antecessores. Duas mudanças foram feitas no conjunto de aproximações: uma modificação para melhorar a descrição de interações não covalentes e dois pequenos erros no formalismo NDDO foram corrigidos. Com as alterações realizadas, o PM7 apresenta uma precisão significativa, com uma redução dos erros nas geometrias em aproximadamente 30% em relação ao seu antecessor, PM6 (STEWART, 2013; ROCHA *et al.*, 2019).

Analizando o erro médio para os cálculos de entalpia dos métodos frente a dois conjuntos de referência de moléculas (Tabela 2), verifica-se que o desempenho do PM7 é muito bom, visto que seu erro encontra-se abaixo de 1,0 kcal/mol, semelhante aos métodos derivados do PM6 e, performance superior a métodos computacionalmente mais custosos, como o funcional de densidade B3LYP. Além disso, em virtude da sua rapidez nos cálculos, o PM7 (como os demais métodos semi-empíricos) pode ser aplicado a sistemas contendo um grande número de átomos, incluindo proteínas e seus complexos (STEWART, 2013).

**Tabela 2:** Erros médios para os conjuntos de dados S22 e S66 (kcal/mol).

Referência	PM7	PM6-DH2	PM6-DH+	PM6	B3LYP
S22	0,74	0,38	0,58	3,38	2,75
S66	0,78	0,66	0,64	2,68	2,29

Fonte: Adaptado de STEWART, 2013.

### **3. OBJETIVOS**

#### **3.1 Objetivo geral**

A partir dos resultados de LARUE e colaboradores (2021), demonstrando a capacidade de peptídeos derivados da hACE2 em inibir a infecção viral pelo SARS-CoV-2, o objetivo da pesquisa foi estudar, com o método semi-empírico de orbitais moleculares PM7, complexos entre o RBD completo da proteína *Spike* do SARS-CoV-2 e os peptídeos derivados da hACE2 para se entender o processo de interação peptídeo-proteína à nível atômico, obtendo-se subsídios para a proposta de novos peptídeos, candidatos à ligantes da proteína, capazes de se conectar covalentemente com o RBD e, assim, elevar a eficiência da interação e da inibição.

#### **3.2 Objetivos específicos**

- Otimizar as estruturas dos complexos nativos, formados entre o RBD da proteína *Spike* do SARS-CoV-2 e os peptídeos derivados da hACE2, bem como, do alvo e ligantes separados;
- Analisar a entalpia de formação dos peptídeos e RBD separados e suas interações intramoleculares;
- Obter as informações estruturais e energéticas sobre a interação intermolecular nos complexos nativos;
- Planejar as modificações estruturais nos peptídeos derivados da hACE2 de forma a permitir a formação da ligação covalente com o RBD da proteína *Spike* do SARS-CoV-2 e, com isso, propor candidatos a inibidores mais eficientes da infecção pelo vírus;
- Verificar as informações estruturais e energéticas sobre a interação intermolecular nos complexos, após a modificação nos peptídeos;
- Simular a reação de obtenção dos complexos covalentes com os peptídeos modificados e avaliar as suas respectivas entalpias de reação.

## 4. MÉTODOS E FERRAMENTAS COMPUTACIONAIS

### 4.1 Obtenção dos complexos proteína-peptídeo

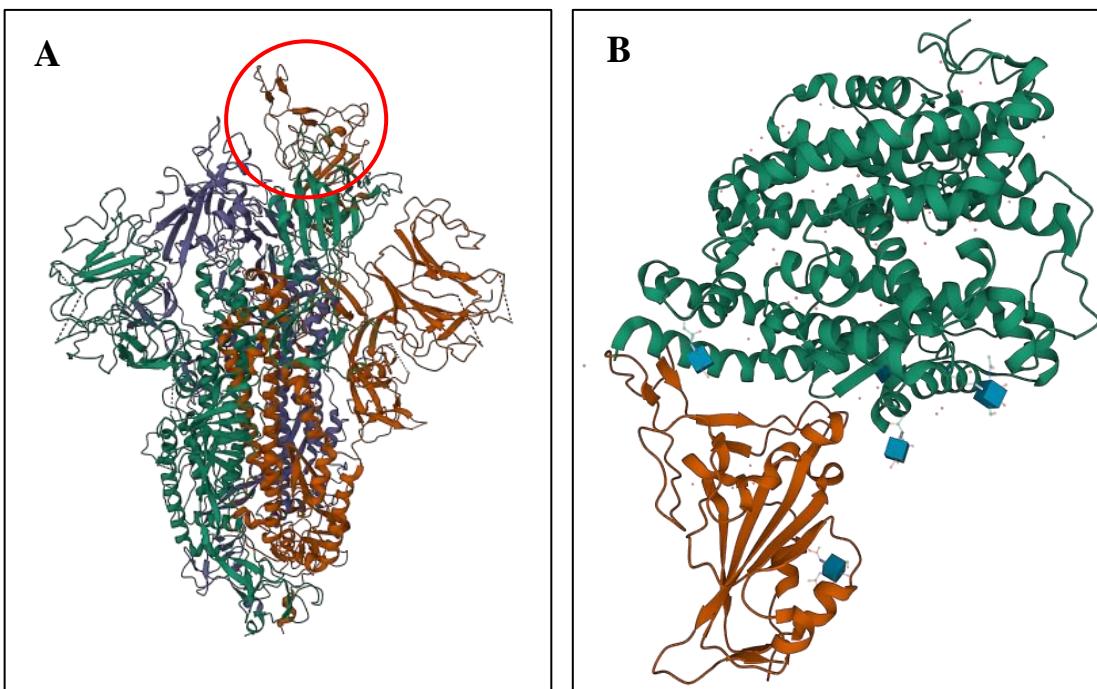
Os seis peptídeos derivados da hACE2, denominados de SAP1 a SAP6, foram selecionados com base no estudo de LARUE *et al.* (2021) sobre a inibição da infecção pelo SARS-CoV-2 *in vitro* (Tabela 3). Entre eles, quatro são sequência de aminoácidos contidos na  $\alpha$ -hélice 1, um da  $\alpha$ -hélice 3 e um da  $\alpha$ -hélice 11. Apenas os peptídeos SAP1, SAP2 e SAP6 apresentam em sua sequência o motivo **EDLFYQ** (em vermelho, na tabela), proposto como essencial para a interação e inibição da infecção. Esses peptídeos foram projetados para terem pelo menos três pontos de interação com a proteína (destaque em negrito, na tabela) e comprimento de 6 a 30 resíduos, para que fossem adequados à síntese (LARUE *et al.*, 2021).

**Tabela 3:** Peptídeos avaliados *in vitro* como inibidores da infecção viral pelo SARS-CoV-2.

Peptídeo	Sequência de aminoácidos	Localização na hACE2
SAP1	27- <b>TFLDKFNHEAEDLFYQ</b> -42	$\alpha$ -hélice 1
SAP2	37- <b>EDLFYQSSL</b> -45	$\alpha$ -hélice 1
SAP3	79- <b>LAQMYPL</b> -85	$\alpha$ -hélice 3
SAP4	352- <b>GKGDFRIL</b> -359	$\alpha$ -hélice 11
SAP5	24- <b>QAKTFLDKFNHEA</b> -36	$\alpha$ -hélice 1
SAP6	37- <b>EDLFYQ</b> -42	$\alpha$ -hélice 1

Fonte: Adaptado de LARUE *et al.*, 2021.

Na figura 11A tem-se a representação da estrutura cristalográfica, adquirida no *Protein Data Bank* (<https://www.rcsb.org/>), da glicoproteína *Spike* nativa do SARS-CoV-2 (PDB ID: 7FB1) no estado aberto, onde pode-se ver o RBD (destacado pelo círculo vermelho), localizado na subunidade S1 da proteína. Nesse estudo, utilizou-se a estrutura cristalográfica do RBD da *Spike* do SARS-CoV-2 ligado à hACE2 (PDB ID: 6M0J, Figura 11B), com resolução de 2,45 Å (LAN *et al.*, 2020). No programa Pymol 0.99rc6 (Delano Scientific LLC) foram selecionados o RBD completo (cadeia E) e a sequência de cada peptídeo de interesse, extraído da hACE2 (cadeia A). O complexo proteína-peptídeo gerado dessa forma, contendo mais de 3000 átomos, foram exportados em formato *pdb*.



**Figura 11:** Estrutura cristalográfica **A)** da glicoproteína *Spike* nativa do SARS-CoV-2, no estado aberto, com RBD destacado pelo círculo vermelho (PDB ID: 7FB1); **B)** do RBD da *Spike* nativa do SARS-CoV-2 (cor laranja) ligado com a hACE2 (cor verde); em azul estão os quatro *N*-acetil- $\beta$ -glucosamina glicanos (NAG) ligados a Asn90, Asn322 e Asn546 da hACE2 e a Asn343 do RBD; as esferas vermelhas representam as moléculas de água (PDB ID: 6M0J).

#### 4.2 Cálculos dos complexos nativos com o método semi-empírico

Os complexos obtidos tiveram a energia minimizada com o método semi-empírico PM7 (STEWART, 2013), disponível no pacote computacional MOPAC 2016 (<http://openmopac.net/>), versões 18.117L (Linux) e 21.319W (Windows), um programa semiempírico de química quântica baseado na aproximação NDDO (*Neglect of Differential Diatomic Overlap*) de Dewar e Thiel (STEWART, 1990).

O tratamento envolveu as seguintes etapas:

- **Etapa 1:** Adição de hidrogênios e definição do estado de ionização dos resíduos através da adição das palavras-chave “ADD-H SITE=(IONIZE) PDBOUT” na linha de comando do arquivo de entrada, onde a palavra-chave PDBOUT indica o formato de saída *pdb* para a estrutura otimizada;
- **Etapa 2:** Cálculo da carga total do sistema através da palavra-chave “CHARGES”;
- **Etapa 3:** Otimização inicial dos hidrogênios, aplicando o modelo de contínuo COSMO para incluir o efeito do solvente, com o uso das palavras-chave “MOZYME CHARGE=n PDBOUT EPS=n.nn NOOPT OPT-H EF”, onde “MOZYME” substitui o método SCF

padrão por um método de orbitais moleculares localizados (LMO), que permite que sistemas com muitos átomos sejam calculados mais rapidamente; “EPS” define a constante dielétrica do solvente para o valor indicado pelo usuário (foi utilizado  $\epsilon = 4,0$  que é o valor comumente empregado para o interior de proteínas e  $\epsilon = 78,4$ , para simulações em meio aquoso); “CHARGE=n” define a carga total do sistema; e EF ativa a rotina *Eigenvector Following* em substituição ao método de otimização de geometria padrão para sistemas com mais de 100 variáveis, o otimizador L-BFGS. O otimizador EF usa mais memória e é mais lento do que o L-BFGS, mas a convergência em geral ocorre mais suavemente;

- **Etapa 4:** Otimização parcial preliminar do sistema, que é uma técnica usada para otimizar as posições dos átomos em uma molécula enquanto mantém alguns átomos ou grupos de átomos fixos em suas posições originais. O procedimento foi aplicado à proteína isolada e aos complexos, fazendo-se a liberação das coordenadas atômicas dos peptídeos e a fixação das coordenadas atômicas da proteína. Essa etapa foi realizada a partir da alteração das coordenadas cartesianas de fixas (*flag +0*) para livres (*flag +1*) dos átomos diferentes de hidrogênio nos SAPs, tanto separados quanto no complexo, mantendo-se fixa (*flag +0*) a geometria do RBD; foram usadas as palavras-chave “MOZYME CHARGE=n PDBOUT EPS=n.nn EF MMOK”, onde “MMOK” indica a correção da mecânica molecular para as ligações peptídicas;
- **Etapa 5:** Otimização total do sistema, que consistiu na liberação das coordenadas atômicas de todos os átomos tanto da proteína isolada quanto em complexo. A liberdade de movimento dos átomos permite que uma geometria de menor energia e, portanto, mais estável seja encontrada. Nessa etapa ocorreu a alteração de todas as coordenadas cartesianas de fixas para livre da proteína, a partir do uso das palavras-chave “MOZYME CHARGE=n PDBOUT EPS=n.nn EF MMOK OPT”. Devido à grande quantidade de átomos, na casa dos milhares, em sistemas com problemas de convergência utilizou-se a palavra-chave “GNORM=10”, que sinaliza o término da otimização da geometria quando o gradiente (a primeira derivada da energia em relação às coordenadas atômicas) for menor do que 10;
- **Etapa 6:** retirada da matriz do peptídeo do arquivo contendo o complexo proteína-ligante, seguida da submissão dos cálculos de carga da molécula (etapa 2) e otimização total (etapa 5), em EPS=n.nn;
- **Etapa 7:** Realizada após a etapa de otimização do sistema, consistiu na determinação das ligações de hidrogênio intermoleculares e intramoleculares com energia liberada igual ou

maior do que 1,0 kcal/mol com o acréscimo das palavras-chave “1SCF DISP(1.0)”, onde “1SCF” determina que seja calculado apenas um ciclo do procedimento SCF. A relação das interações foi obtida por meio de arquivo de saída (*arc*), em formato de texto, no qual as linhas com cadeias “A” e “E” indicam interações intermoleculares (ANEXO A, destacado em negrito), enquanto as linhas somente com “A” ou “E” representam as interações intramoleculares;

- **Etapa 8:** Consistiu na identificação das demais interações intermoleculares, bem como das suas respectivas distâncias, pelo servidor PLIP (do inglês *Protein-ligand Interaction Profiler*) (ADASME, 2021). Antes da submissão ao servidor, no arquivo *pdb* gerado na etapa 5, foi alterada a sigla “ATOM” para “HETATM”, na matriz referente ao peptídeo, para identificá-lo como ligante.

#### 4.3 Cálculos dos complexos modificados com o método semi-empírico

Os complexos que apresentaram os melhores resultados de entalpia de interação, tiveram seus peptídeos submetidos a modificações estruturais no resíduo Asp38, para permitir a formação da ligação covalente com a Tyr449 do RBD da proteína (vide Resultados para maior detalhamento). As alterações foram realizadas por edição dos arquivos de formato *pdb*, gerados na etapa de otimização parcial.

O tratamento envolveu as seguintes etapas:

- **Etapa 1:** A fim de obter-se a entalpia de formação do peptídeo modificado, foi extraída a sua matriz do arquivo contendo o complexo não-covalente e antes da troca do átomo de hidrogênio. Em seguida, fez-se o cálculo da carga total utilizando “CHARGES” e a otimização total das coordenadas atômicas, a partir das seguintes palavras-chave “MOZYME CHARGE=n PDBOUT EPS=n.nn OPT MMOK”, onde o valor de “EPS” foi definido pelas constantes dielétricas utilizadas anteriormente;
- **Etapa 2:** Realizada antes da troca de hidrogênio entre a Tyr449 do RBD e o nitrogênio do anel triazólico, do resíduo Asp38 modificado do peptídeo. Essa etapa permitiu a estabilidade do intermediário da reação, tendo em vista que manteve o hidrogênio capturado ligado ao nitrogênio do grupo de saída. Consistiu na otimização parcial do complexo, onde foram liberadas as coordenadas atômicas tanto do peptídeo quanto do resíduo Tyr449 do RBD, com o uso das palavras-chave “MOZYME CHARGE=n PDBOUT EPS=n.nn MMOK T=5D xyz”. Os outros átomos dos resíduos da proteína permaneceram fixos. A cópia do

arquivo *arc* gerado foi submetida a otimização total e os valores obtidos foram utilizados nos cálculos de entalpia de interação;

- **Etapa 3:** Após a etapa de otimização anterior, fez-se a determinação das ligações de hidrogênio intermoleculares e intramoleculares com energia liberada igual ou maior do que 1,0 kcal/mol;
- **Etapa 4:** Consistiu na identificação das demais interações intermoleculares pelo PLIP;
- **Etapa 5:** Executada após a formação da ligação covalente. Ela forneceu o cálculo da entalpia de formação do complexo covalente otimizado totalmente. Utilizou-se as seguintes palavras-chave: “MOZYME CHARGE=n PDBOUT EPS=n.nn OPT MMOK T=5D GNORM=10 xyz”. Neste caso, “GNORM=10” foi diretamente empregado para que o cálculo fosse executado com mais rapidez, visto que as coordenadas atômicas estavam totalmente livres e isso demandaria um maior tempo computacional;
- **Etapa 6:** Do resultado gerado na etapa anterior, fez-se a separação do complexo covalente e o grupo de saída em arquivos diferentes, submetendo-os a um novo cálculo de entalpia, em ambos os meios dielétricos (EPS).

#### 4.4 Visualização molecular

Através do programa Pymol 0.99rc6 (Delano Scientific LLC) foi possível visualizar a estrutura molecular dos complexos, bem como, as ligações de hidrogênio envolvidas na interação entre a proteína *Spike* e os peptídeos derivados da hACE2, além das principais mudanças conformacionais ocorridas.

## 5. RESULTADOS E DISCUSSÃO

### 5.1 Otimização das moléculas separadas e em complexo nativo

A entalpia padrão de formação ( $\Delta H_f^\theta$ ) de um composto químico é a variação da energia na reação de formação deste composto a partir dos elementos que o compõem, na pressão de 1 atm e na temperatura de 25 °C (298 K) (ATKINS *et al.*, 2012; BROWN, 2005). No programa MOPAC,  $\Delta H_f^\theta$  é definida como:

$$\Delta H_f^\theta = E_{elect} + E_{nuc} - E_{isol} + E_{atom} + E_{hbd} \quad (6)$$

onde  $E_{elect}$  é a energia eletrônica,  $E_{nuc}$  é a energia de repulsão nuclear,  $-E_{isol}$  é a energia necessária para extrair todos os elétrons de valência de todos os átomos do sistema,  $E_{atom}$  é a energia de atomização de todos os átomos do sistema, e  $E_{hbd}$  é a soma das energias das ligações de hidrogênio e das forças de dispersão (que são impressas com o uso da palavra-chave “DISP”).

Quando EPS=n está presente, a energia devida à estabilização do sistema pelo solvente está localizada na linha "DIELECTRIC ENERGY" no arquivo de saída do MOPAC. A desestabilização do solvente devido ao sistema não é impressa. As energias por solvatação estão nesse caso presentes nos termos  $E_{elect}$  e  $E_{nuc}$ , ou seja, os termos de solvatação estão incluídos nessas duas grandezas, não podendo ser separados delas.

Considerando, que inicialmente o peptídeo estará no meio aquoso (e mais polar) para só depois se aproximar do RBD, que é uma região de menor polaridade, realizou-se os cálculos de  $\Delta H_f^\theta$  e energia de ligação de hidrogênio com as duas constantes dielétricas. A constante dielétrica da água ( $\epsilon = 78,4$ ) é usada na representação do meio aquoso em cálculos de contínuo. A constante dielétrica igual a 4,0 é frequentemente utilizada em representações de contínuo para simular o efeito da matriz proteica em regiões no interior da proteína, um valor considerado adequado para representar a polarização eletrônica e pequenas flutuações da cadeia polipeptídica, embora outros valores, como  $\epsilon = 8, 10, 11$  e  $20$ , também sejam propostos na literatura (LI *et al.*, 2013).

No arquivo *arc* gerado após a determinação das interações, são descritos: a) os resíduos que interagem por ligações de hidrogênio e suas respectivas distâncias e energias (iguais ou

mais favoráveis do que -1,0 kcal/mol); b) a energia total das ligações de hidrogênio, que considera também os valores menos favoráveis do que -1,0 kcal/mol.

Na tabela 4, verifica-se que a  $\Delta H_f^\theta$  dos peptídeos na região do RBD, ou seja, quando imersos em uma matriz essencialmente proteica ( $\epsilon = 4,0$ ), é menos favorável em relação aos respectivos valores em meio aquoso ( $\epsilon = 78,4$ ). Provavelmente, isso ocorre porque as cargas das cadeias laterais de seus aminoácidos são desestabilizadas em um meio de baixa polaridade, como esperado. Contudo, nesse mesmo meio dielétrico, a energia total das ligações de hidrogênio para cada SAP é mais favorável, com exceção do peptídeo SAP3, que se manteve constante.

**Tabela 4:** Parâmetros termodinâmicos dos peptídeos calculados com o método PM7.

Peptídeo	nº resíduos de aminoácidos	$\epsilon = 78,4$		$\epsilon = 4,0$	
		$\Delta H_f^\theta$ (kcal/mol) <sup>1</sup>	Evolução de H (kcal/mol) <sup>2</sup>	$\Delta H_f^\theta$ (kcal/mol) <sup>1</sup>	Evolução de H (kcal/mol) <sup>2</sup>
SAP1	16	-1829,3	-62,61	-1681,84	-70,00
SAP2	9	-1138,0	-26,17	-1052,58	-28,80
SAP3	7	-564,0	-7,64	-501,14	-7,40
SAP4	8	-532,9	-13,00	-461,66	-15,83
SAP5	13	-1167,2	-50,98	-1068,13	-54,83
SAP6	6	-849,2	-11,47	-760,65	-13,61

<sup>1</sup> Calculado diretamente pelo PM7;

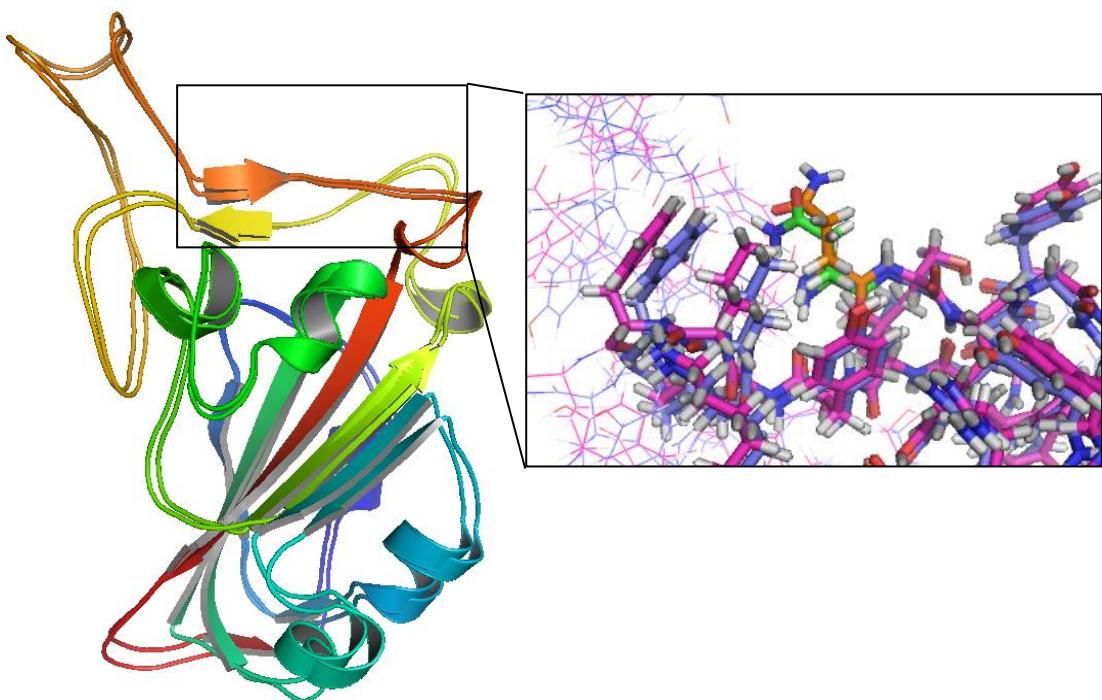
<sup>2</sup> Energia total das ligações de hidrogênio.

É possível verificar que os SAPs 1, 5 e 2, que contém nesta ordem mais aminoácidos em sua estrutura primária, estabelecem um maior número de ligações de hidrogênio intramoleculares, com base nas suas respectivas energias relacionadas a essas interações. A partir disso, pode-se dizer que esses peptídeos possivelmente preservam melhor sua estrutura secundária em  $\alpha$ -hélice, sendo conformacionalmente mais estáveis em meio aquoso e no dielétrico de baixa polaridade ( $\epsilon = 4,0$ ), considerando que este tipo de interação em conjunto com outras, como as hidrofóbicas e iônicas, são as responsáveis por manter o arranjo tridimensional de peptídeos e proteínas (FITTOLANI *et al.*, 2020).

A entalpia padrão de formação do RBD da *Spike* foi avaliada em dois tipos de otimização, em meio aquoso ( $\epsilon = 78,4$ ): com as coordenadas atômicas dos átomos diferentes de H fixas e livres, obtendo-se uma redução de 1.247,83 kcal/mol no valor da  $\Delta H_f^\rho$  nas condições de liberação total das coordenadas atômicas. Isso mostra que, mesmo considerando as limitações da representação do solvente como um contínuo, a conformação cristalográfica difere consideravelmente daquela em água, já que com as coordenadas atômicas totalmente livres, a otimização da geometria molecular pode avançar mais na direção do mínimo de energia, permitindo, por exemplo, a formação de mais ligações de hidrogênio intramoleculares. De fato, a energia total das ligações de hidrogênio passou de -606,16 kcal/mol na otimização parcial para -901,75 kcal/mol na otimização total.

No sistema totalmente otimizado em meio aquoso, ocorreram movimentos nas cadeias laterais dos resíduos de aminoácidos do RBD, gerando interações entalpicamente mais estáveis, refletindo na redução da energia total do sistema. O deslocamento mais proeminente foi observado no resíduo Gln493, que na otimização parcial (Figura 12) estava com seu grupo amida da cadeia lateral voltado para o exterior da cavidade da proteína, enquanto na otimização total esse grupo ficou posicionado para o interior do bolso, formando um número maior de ligações de hidrogênio intramoleculares (ANEXO B). Isso se manteve quando a constante dielétrica menor foi aplicada, tendo em geral, seus valores de distância e energia reduzidos. Esses resultados indicam a importância de manter todas as coordenadas livres durante a otimização para se obter resultados de melhor qualidade para sistemas extensos como os estudados nesse projeto.

Visto que o RBD é na verdade parte integrante da *Spike* (vide figura 11), uma representação de contínuo mais adequada para descrever o ambiente ao seu redor provavelmente envolveria um valor de  $\epsilon$  mais baixo, próximo daquele usado para descrever o interior de proteínas. Assim, a fim de se usar condições que poderiam representar melhor aquelas do sítio de interação real, que está apenas parcialmente em contato com a água, realizou-se a otimização total do RBD com  $\epsilon = 4,0$ , obtendo-se  $\Delta H_f^\rho$  de -16.238,32 kcal/mol e energia total de ligações de hidrogênio intramoleculares igual a -910,05 kcal/mol. Assim, no restante do trabalho, tanto para os complexos com peptídeos nativos quanto para os modificados e durante o estudo da formação da ligação covalente, consideraremos o RBD inserido em um contínuo com constante dielétrica igual a 4,0.



**Figura 12:** Sobreposição do RBD da *Spike* (em  $\epsilon = 78,4$ ) otimizado parcialmente e totalmente e, recorte do sítio de ligação, evidenciando o deslocamento do resíduo Gln493, que na otimização parcial (C na cor rosa) estava com a sua cadeia lateral direcionada para o exterior do bolso (C na cor laranja), enquanto na otimização total (C na cor lilás) ficou posicionada para o interior (C na cor verde).

Nos complexos nativos, observa-se que a energia total das forças de dispersão é mais favorável do que das ligações de hidrogênio e, assim, pode ser considerada como a interação mais importante para a afinidade proteína-peptídeo e a estabilidade do complexo (Tabela 5). O RBD-SAP1 se destaca por ter mais interações intermoleculares, seguido pelos complexos RBD-SAP5 e RBD-SAP2, como era esperado, visto que seus peptídeos possuem um número maior de resíduos de aminoácidos, respectivamente. Com o propósito de comparar o comportamento das interações num meio mais polar, realizou-se também os cálculos em meio aquoso ( $\epsilon = 78,4$ ) e observou-se a mesma tendência em relação às forças de dispersão.

**Tabela 5:** Energias de interação nos complexos nativos calculados pelo método PM7 (continua).

Complexos	Otimização total ( $\epsilon = 4,0$ )		Otimização total ( $\epsilon = 78,4$ )	
	Evolução de H <sup>1</sup>	Evolução de dispersão <sup>2</sup>	Evolução de H <sup>1</sup>	Evolução de dispersão <sup>2</sup>
	(kcal/mol)	(kcal/mol)	(kcal/mol)	(kcal/mol)
RBD-SAP1	-95,56	-159,51	-63,64	-144,57
RBD-SAP2	-21,76	-51,18	-22,30	-50,71

RBD-SAP3	-11,33	-28,64	-19,05	-31,54
RBD-SAP4	-21,50	-29,86	-0,77	-11,17
RBD-SAP5	-36,31	-74,30	-14,81	-62,81
RBD-SAP6	-14,45	-20,69	13,07	-0,84

<sup>1</sup> Energia total, calculada pela diferença de energia entre produto e reagentes;

<sup>2</sup> Calculado pela diferença de energia entre produto e reagentes.

## 5.2 Análise das interações intermoleculares na interface de ligação

A relação completa das interações intermoleculares identificadas pelo MOPAC e pelo PLIP, em  $\epsilon = 4,0$ , encontram-se nos ANEXOS C e D, respectivamente. Observa-se que os SAPs 3, 4 e 5 interagem diferentemente com a proteína, já os peptídeos 1, 2 e 6 apresentam resíduos semelhantes na interface de interação com o RBD da *Spike* (Tabela 6), pertencentes ao motivo estrutural **EDLFYQ**, que foi proposto como uma sequência essencial para a inibição da infecção (LARUE *et al.*, 2021). Seguiremos nossa análise aqui com esses três peptídeos somente.

**Tabela 6:** Resíduos semelhantes na interface de interação proteína-peptídeo.

Resíduos do RBD	Resíduos do peptídeo			Tipo de interação
	SAP1	SAP2	SAP6	
Tyr 505	Glu 37	Glu 37	-	hidrofóbica
Asn 501	Tyr 41	Tyr 41	Tyr 41	ligação de hidrogênio
Thr 500				
	Tyr 41	Tyr 41	Tyr 41	hidrofóbica
Gln 498	Asp 38	Asp 38	Asp 38	ligação de hidrogênio
	Gln 42	-	Gln 42	
Tyr 449	Asp 38	Asp 38	Asp 38	ligação de hidrogênio
	Gln 42	Gln 42	Gln 42	
Gly447	Gln 42	Gln 42	Gln 42	ligação de hidrogênio
Gly446				

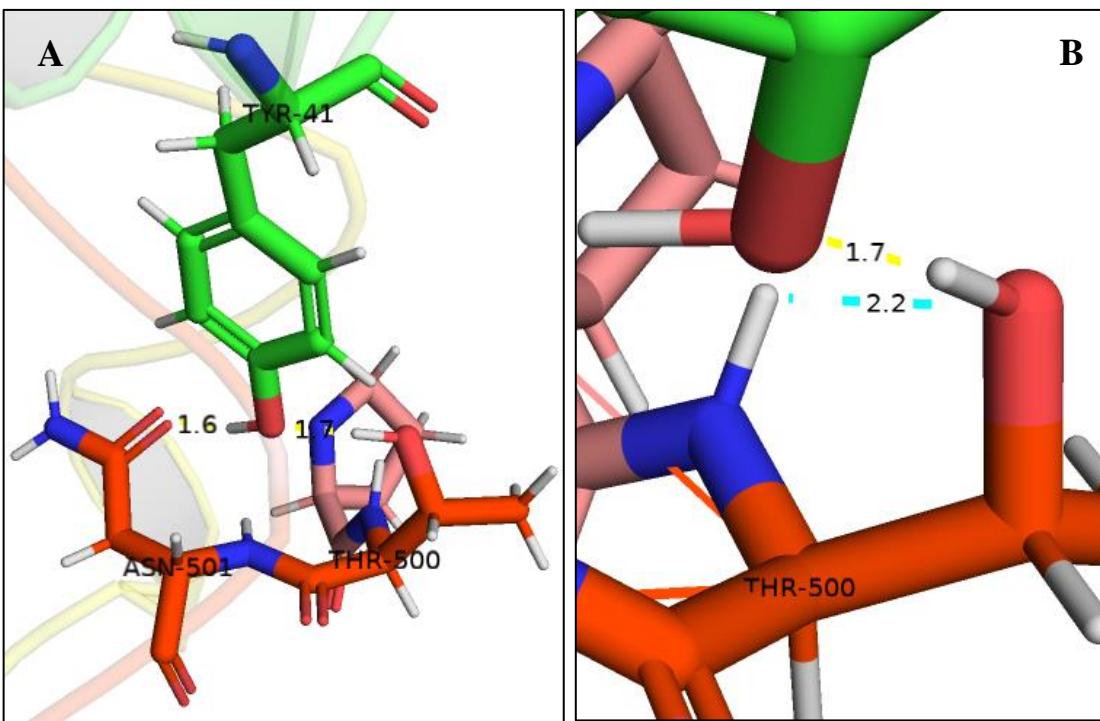
Esses resíduos são também destacados na estrutura cristalográfica resolvida por LAN *et al.* (2020) (ANEXO E), na qual os grupos carboxilato de Glu37 e Asp38, amida de Gln42 e a hidroxila de Tyr41 de SAP6 formam ligações de hidrogênio com Tyr505, Tyr449 e Thr500 do RBD, respectivamente. A otimização total dos complexos permitiu que mais resíduos do RBD interagissem com os peptídeos, por ligação de hidrogênio (Gly446, Gly447 e Gln498) e interações hidrofóbicas (Tyr505 e Gln498).

Como esperado, de todos os peptídeos, o SAP1 foi o que obteve mais pontos de interação com a proteína, daqueles descritos na tabela 3 (em negrito), nos quais todos formaram ligações de hidrogênio com o alvo e, entre eles, 8 de 10 fizeram ponte salina ou interação hidrofóbica com o RBD. Além disso, a interação entre o Asp38 dos SAPs e a Tyr449 do RBD, se destacou como a de maior contribuição para a afinidade entre a proteína e o peptídeo, devido à energia mais favorável frente às demais ligações de hidrogênio intermoleculares encontradas.

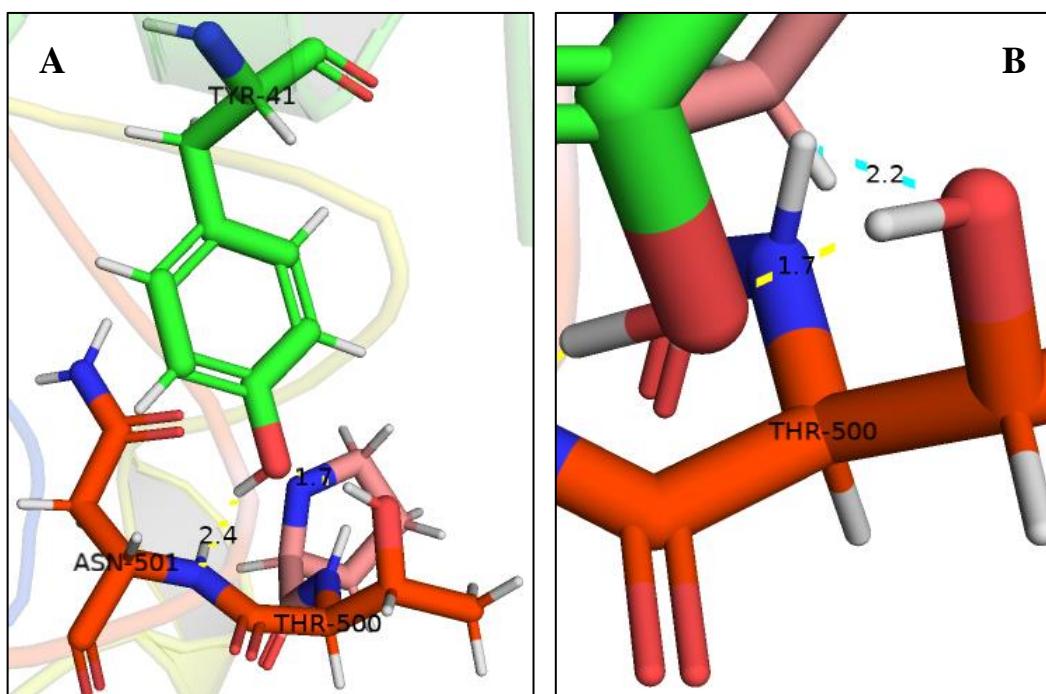
O movimento da cadeia lateral da Gln493 do RBD, relatado no tópico 6.2, não é observado no complexo com SAP1, visto que esse resíduo interage com a Lys31 do peptídeo após a otimização total, com energia mais favorável que -2 kcal/mol (ANEXO C). Tanto no SAP2 quanto no SAP6 é observado o movimento da cadeia lateral da glutamina, já que não há interações entre o resíduo e os peptídeos. Ela ocorre somente no complexo com SAP1 devido ao seu maior tamanho, em comparação aos demais SAPs.

Nos três complexos, foram encontradas interações entre o resíduo Tyr41 dos peptídeos e os resíduos Thr500 e Asn501 do RBD (Tabela 7; Figuras 13 a 15). Nos SAPs 1 e 6 (Tabela 7; Figuras 13 e 15), ambas as ligações de hidrogênio apresentaram distâncias semelhantes, variando somente na energia, na qual a interação entre a tirosina e a treonina se mostrou mais forte. A presença da interação intramolecular entre o NH da ligação peptídica e o oxigênio da hidroxila da treonina, a uma distância de 2,2 Å, pode contribuir para o aumento da atração entre o doador e acceptor da ligação de hidrogênio entre a Tyr41 e a Thr500, tornando-a mais forte, visto que a intensidade da carga parcial positiva do hidrogênio ligado ao oxigênio se torna maior.

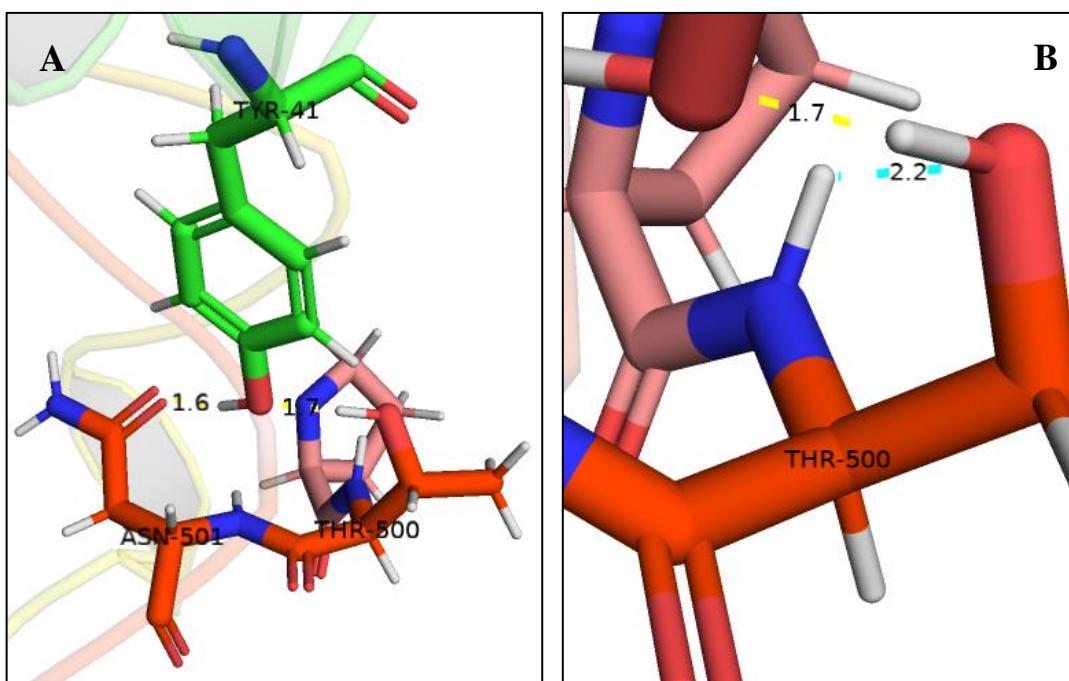
Na interação da tirosina com a asparagina, o SAP2 (Tabela 7; Figura 14) forma ligação de hidrogênio mais fraca, porque diferente dos outros peptídeos, em que o acceptor é o oxigênio da carbonila da cadeia lateral da Asn501, nele o acceptor é o nitrogênio da ligação peptídica do mesmo resíduo. Possivelmente na busca pelo mínimo de energia na otimização, este tenha sido o primeiro ponto encontrado, de maneira que para a formação da interação com o oxigênio, que é considerada mais estável, houvesse uma barreira energética e, assim, o procedimento de minimização de energia pode não ter convergido para esse mínimo.



**Figura 13:** **A)** Ligações de hidrogênio (tracejado amarelo) entre os resíduos Asn501 e Thr500 do RBD (C em laranja) e a Tyr41 do peptídeo **SAP1** (C em verde). **B)** Ligação de hidrogênio intramolecular (tracejado azul) entre o NH da ligação peptídica e a hidroxila da treonina.



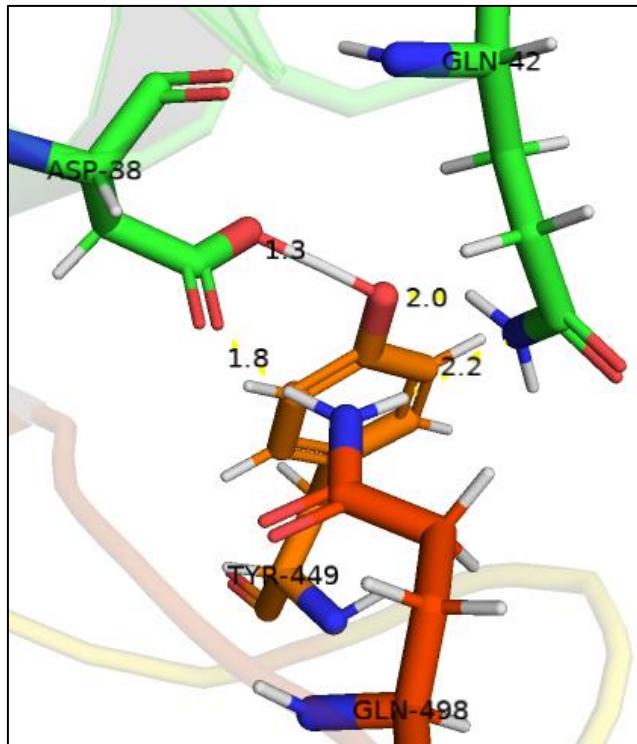
**Figura 14:** **A)** Ligações de hidrogênio (tracejado amarelo) entre os resíduos Asn501 e Thr500 do RBD (C em laranja) e a Tyr41 do peptídeo **SAP2** (C em verde). **B)** Ligação de hidrogênio intramolecular (tracejado azul) entre o NH da ligação peptídica e a hidroxila da treonina.



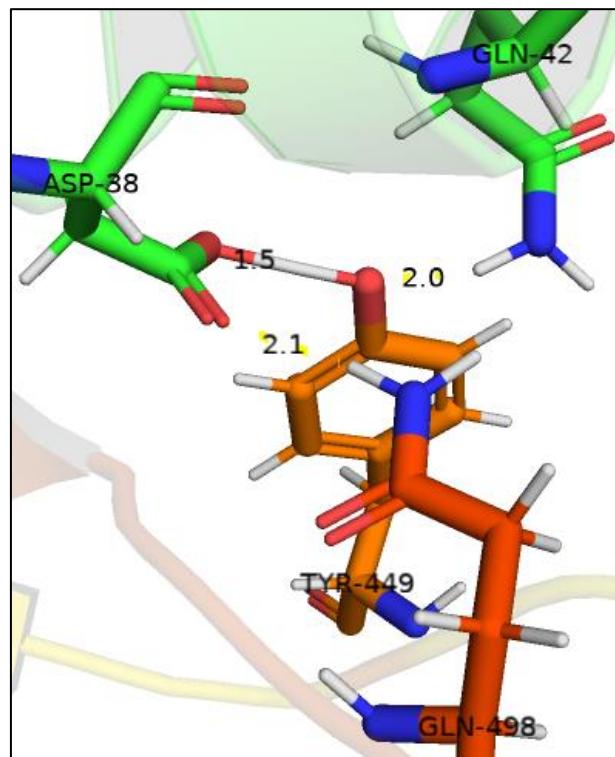
**Figura 15:** A) Ligações de hidrogênio (tracejado amarelo) entre os resíduos Asn501 e Thr500 do RBD (C em laranja) e a Tyr41 do peptídeo SAP6 (C em verde). B) Ligação de hidrogênio intramolecular (tracejado azul) entre o NH da ligação peptídica e a hidroxila da treonina.

Na ligação de hidrogênio entre a Tyr449 do RBD e os resíduos Asp38 e Gln42 dos peptídeos (Tabela 7; Figuras 16 a 18), a interação com o aspartato ocorre em menor distância e é mais favorável, podendo ser verificado pelo aspecto de ligação covalente na representação adotada pelo Pymol, visto que o grupo aceptor carboxilato tem carga negativa, o que reforça a interação com o doador de ligação de hidrogênio. Já o grupo doador amida da glutamina, por ser neutro, interage mais fracamente com a tirosina aceptor da proteína.

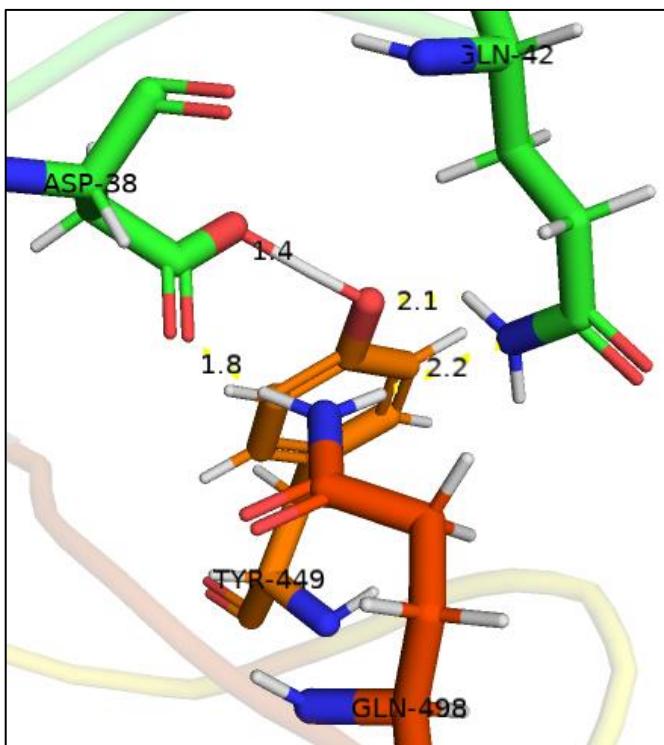
Na interação entre o Asp38 e a Gln42 dos peptídeos com a Gln498 da proteína (Tabela 7; Figuras 16 a 18), apenas o SAP2 (Tabela 7; Figura 17) não apresentou ligação de hidrogênio entre a sua glutamina e o resíduo do RBD. Além disso, observou-se que a interação da proteína com o aspartato também se mostrou mais intensa. Entretanto, ao comparar as duas interações do aspartato com a proteína, Asp38-Tyr449 e Asp38-Gln498, é possível constatar que a segunda apresenta valores mais altos tanto em distância quanto em energia. Possivelmente, isso se deve ao fato do grupo amida de Gln498, doador da ligação de hidrogênio, estar envolvido em uma segunda doação de ligação de hidrogênio com o grupo amida do resíduo Gln42, aceptor.



**Figura 16:** Ligações de H (tracejado amarelo) entre os resíduos Tyr449 e Gln498 do RBD (C em laranja) e Asp38 e Gln42 do peptídeo **SAP1** (C em verde).

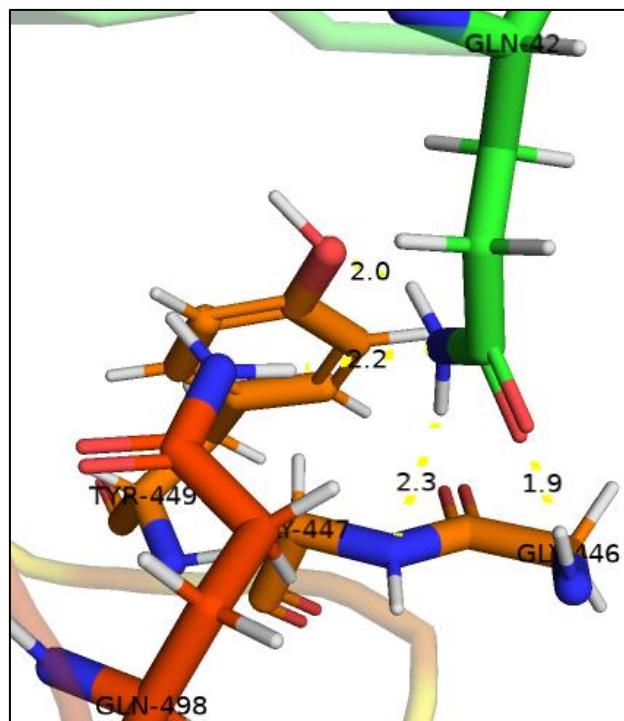


**Figura 17:** Ligações de H (tracejado amarelo) entre os resíduos Tyr449 e Gln498 do RBD (C em laranja) e Asp38 e Gln42 do peptídeo **SAP2** (C em verde).

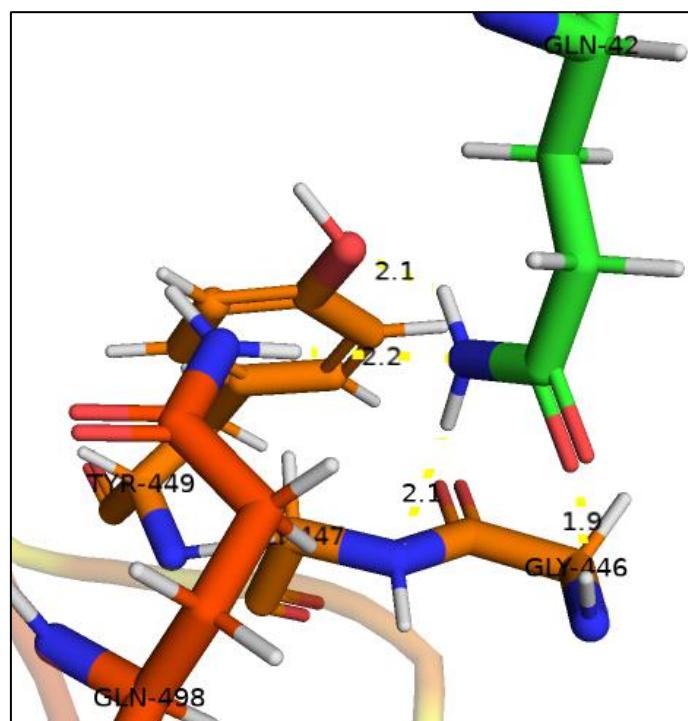


**Figura 18:** Ligações de H (tracejado amarelo) entre os resíduos Tyr449 e Gln498 do RBD (C em laranja) e Asp38 e Gln42 do peptídeo **SAP6** (C em verde).

Apenas os SAPs 1 e 6 estabeleceram conexão com os resíduos Gly446 e Gly447 do RBD (Tabela 7; Figuras 19 e 20). Verifica-se que a variação nas distâncias, entre as interações no complexo, foi mínima, sendo: 0,4 Å para RBD-SAP1 e 0,2 Å para RBD-SAP6. Todavia, a ligação de hidrogênio mais curta do RBD-SAP6, formada entre Gln42-Gly446, tem energia menos negativa. Em ambos os complexos, a glutamina dos peptídeos estabelece interações com outros resíduos próximos da proteína, onde a ocorrida com a Gln498 é entalpicamente mais estável. Acredita-se que no cálculo da interação entre a glutamina com a Gly446, essa energia possa ter sido considerada e, devido ao fato do complexo com o SAP6 ter tido um maior valor para essa ligação de hidrogênio, consequentemente contribuiu para o aumento energético na conexão Gln42-Gly446.



**Figura 19:** Ligações de H (tracejado amarelo) entre os resíduos Gly446 e Gly447 do RBD (C em laranja) e Gln42 do peptídeo **SAP1** (C em verde).



**Figura 20:** Ligações de H (tracejado amarelo) entre os resíduos Gly446 e Gly447 do RBD (C em laranja) e Gln42 do peptídeo **SAP6** (C em verde).

Foram identificadas as seguintes interações hidrofóbicas entre o primeiro e o penúltimo resíduo da sequência dos peptídeos - **EDLFYQ**: Glu37-Tyr505, nos complexos com os SAPs 1 e 2, com a distância variando na faixa de 3,46 a 3,53 Å; e Tyr41-Gln498, nos três sistemas, com um intervalo entre 3,43 e 3,90 Å (ANEXO D). Não há parâmetros que definam um comprimento ótimo para avaliar a intensidade dessa interação, em forte ou fraca. Todavia, ela é muito importante para o reconhecimento molecular, porque aumenta a entropia do sistema através da liberação de moléculas de água do sítio ativo para o meio solvente. Além disso, melhora a entalpia, visto que auxilia na formação das ligações de hidrogênio entre as moléculas de água liberadas (BARREIRO e FRAGA, 2015; FOKOUÉ *et al.*, 2020).

**Tabela 7:** Ligações de H semelhantes na interface de interação, calculadas pelo método PM7 (continua).

Complexo	Interação		Ligaçāo de H <sup>1</sup>	
	Resíduos do RBD	Resíduos do SAP	D-H (Å)	Energia (kcal/mol)
RBD-SAP1	Asn501 Thr500	Tyr41	1,70	-4,30
			1,69	-4,87
	Tyr449	Asp38	1,30	-4,85
		Gln42	2,00	-1,77
	Gln498	Asp38	1,82	-2,59
		Gln42	2,20	-2,84
	Gly447 Gly446	Gln42	2,26	-2,52
			1,86	-2,76
RBD-SAP2	Asn501 Thr500	Tyr41	2,42	-1,81
			1,70	-4,65
	Tyr449	Asp38	1,50	-4,98
		Gln42	2,04	-1,18
	Gln498	Asp38	2,12	-2,13
RBD-SAP6	Asn501 Thr500	Tyr41	1,76	-2,46
			1,69	-4,85
	Tyr449	Asp38	1,45	-4,86
		Gln42	2,13	-1,42

Gln498	Asp38	1,80	-2,30
	Gln42	2,25	-2,65
Gly447		2,14	-2,97
Gly446	Gln42	1,93	-2,49

<sup>1</sup>Mais estável do que -1,0 kcal/mol.

### 5.3 Entalpia de interação dos complexos nativos e a correlação com a inibição viral

O cálculo da entalpia de interação ( $\Delta H_{int}$ ) foi realizado pela diferença entre os valores de entalpia padrão de formação ( $\Delta H_f^0$ ) dos complexos (RBD-SAP) e dos componentes separados (RBD e SAP) (equação 7). Para melhor representar o que ocorre no sistema real, consideramos o peptídeo em fase aquosa ( $\varepsilon = 78,4$ ); como o RBD e o respectivo complexo RBD-SAP estão inseridos na matriz proteica, as estruturas foram otimizadas com  $\varepsilon = 4$ , conforme o equilíbrio de formação do complexo expresso pela equação 8. A constante de dissociação  $K_d$ , que mede a propensão do peptídeo em se dissociar do complexo, é definida pela equação 9. Logo, o peptídeo com menor  $K_d$  tem uma afinidade maior com o RBD e uma tendência menor em se dissociar do complexo (VALLE, 2008).

$$\Delta H_{int} = (\Delta H_f^0(\text{complexo RBD-SAP})_{(ptm)} - [(\Delta H_f^0(\text{RBD})_{(ptm)} + (\Delta H_f^0(\text{SAP}))_{(aq)}]) \quad (7)$$



$$K_d = \frac{[\text{RBD}][\text{SAP}]}{[\text{RBD.SAP}]} \quad (9)$$

LARUE *et al.* determinaram inicialmente o valor de IC<sub>50</sub>, que identificou quais seriam os melhores candidatos, *in vitro*, a inibidores da infecção pelo SARS-CoV-2 e, a partir desse dado foi definida experimentalmente a constante de dissociação para cada complexo. Contudo, os SAPs 3, 4 e 5 tiveram valores muito altos de IC<sub>50</sub> (> 7,5 mM), de modo que se determinou  $K_d$  somente para os demais SAPs. Como o nosso estudo envolve apenas o RBD da *Spike* e os SAPs, a comparação mais adequada entre os dados teóricos e os experimentais é realizada entre

os valores de  $K_d$  e a entalpia de interação, já que esses parâmetros envolvem diretamente a afinidade entre a *Spike* do SARS-CoV-2 e os peptídeos. Deve-se destacar, no entanto, que a comparação mais adequada seria com a variação da energia livre de Gibbs resultante da interação, visto que esta é diretamente associada com as constantes de equilíbrio. No entanto, dado o número elevado de átomos dos sistemas modelados e as grandes exigências computacionais para o cálculo da entropia por métodos quânticos, neste estudo avaliamos somente a entalpia de interação. Nesse caso, uma boa correlação seria esperada somente se o fenômeno de interação fosse dominado pela parcela entálpica da energia livre de Gibbs.

De fato, verificou-se que os valores calculados da  $\Delta H_{int}$  revelaram a mesma tendência que a constante de dissociação nos ensaios de *binding* no RBD da *Spike* do SARS-CoV-2 (Tabela 8). Isto é, quanto mais favorável é o valor de  $\Delta H_{int}$ , menor é o valor de  $K_d$ , demonstrando uma maior afinidade e estabilidade do complexo. Dessa maneira, o SAP1 se destaca como melhor ligante do RBD, por ter a melhor correlação desses parâmetros, seguido pelos peptídeos SAP6 e SAP2, respectivamente. A comparação com os dados de IC<sub>50</sub> não é adequada nesse caso, porque se refere à inibição de infecção pelo SARS-CoV-2, um fenômeno bem mais complexo do que a simples interação SAP-RBD.

**Tabela 8:** Parâmetros calculados pelo método semi-empírico PM7 ( $\varepsilon = 4,0$ ) para os complexos e dados obtidos *in vitro*<sup>1</sup> para os peptídeos.

Complexos	Otimização total				
	$\Delta H_{int}$ (kcal/mol)	Evolução de H <sup>a</sup> (kcal/mol)	Eforças de dispersão <sup>b</sup> (kcal/mol)	$K_d$ (mM) <sup>1</sup>	IC <sub>50</sub> (mM) <sup>1</sup>
RBD-SAP1	-205,37	-95,56	-159,51	$0,53 \pm 0,01$	$2,39 \pm 0,20$
RBD-SAP2	-51,38	-21,76	-51,18	$10,7 \pm 4,2$	$3,72 \pm 0,37$
RBD-SAP3	-1,95	-11,33	-28,64	-	$> 7,5$
RBD-SAP4	11,50	-21,50	-29,86	-	$> 7,5$
RBD-SAP5	-75,64	-36,31	-74,30	-	$> 7,5$
RBD-SAP6	-54,61	-14,45	-20,69	$1,36 \pm 0,14$	$1,90 \pm 0,14$

<sup>1</sup>LARUE *et al.* (2021);

<sup>a</sup>Energia total, calculada pela diferença de energia entre produto e reagentes;

<sup>b</sup>Calculado pela diferença de energia entre produto e reagentes.

Comparando as interações identificadas pelo PM7, nota-se que, em geral, as forças de dispersão têm maior contribuição na manutenção da interação entre a proteína e o peptídeo. Todavia, na entalpia de interação outros fatores são importantes como, por exemplo, as mudanças conformacionais resultantes do processo de interação entre o peptídeo e o RBD, alcançáveis pela liberação total das coordenadas atômicas do sistema durante a busca pelo mínimo de energia.

Análises teóricas já haviam proposto o uso de peptídeos derivados da hACE2 como candidatos para interferir na infecção viral pelo SARS-CoV-2 (HAN e KRAL, 2020). O estudo de LARUE *et al.* (2021) forneceu a prova de conceito e demonstrou a viabilidade na inibição, utilizando peptídeos baseados no próprio receptor da *Spike*, agindo na interface de interação. Por outro lado, os dados indicam que a atividade biológica alcançada não é muito elevada, já que os valores de IC<sub>50</sub> estão na faixa milimolar (mM) baixa, assim como os valores de K<sub>d</sub>.

#### **5.4 Propostas de modificação estrutural nos peptídeos como ligantes covalentes**

Inibidores covalentes têm sido explorados no planejamento de medicamentos, principalmente por causa de seu tempo de residência prolongado e seu potencial para alvos onde moléculas de ligação reversível não tiveram sucesso (HILLEBRAND e GEHRINGER, 2022). Esse tipo de ligante pode alcançar elevada potência em consequência da formação de ligações covalentes com resíduos na proteína-alvo. Assim, a incorporação de um grupo capaz de formar ligação covalente com o alvo, conhecido como grupo reativo ou grupo ogiva, pode ser uma estratégia interessante para melhorar a potência dos peptídeos derivados da hACE2. Por outro lado, devido à sua reatividade, uma deficiência frequentemente apontada nos inibidores covalentes é que eles em geral têm baixa seletividade, um problema que deverá ser minimizado pela boa complementaridade entre a estrutura dos SAPs e o RBD da *Spike*, visto que a sequência de aminoácidos dos peptídeos é a mesma do receptor.

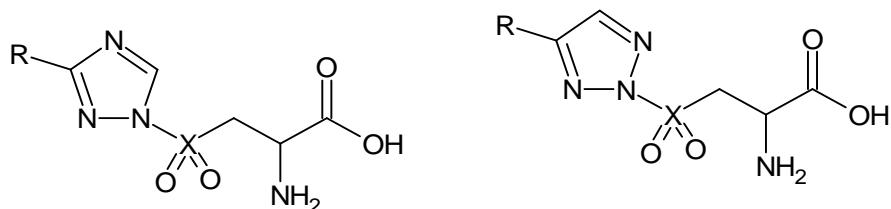
Alguns resíduos nucleofílicos podem ser explorados em alvos potenciais para ligantes covalentes, como a cisteína, serina, treonina, tirosina e lisina. De fato, como descrito anteriormente, há alguns resíduos de tirosina no RBD da *Spike* na região de contato com os peptídeos SAPs 1, 2 e 6 e uma forte interação presente em todos os complexos envolveu uma ligação de hidrogênio entre os resíduos Tyr449 do RBD e Asp38 dos peptídeos.

A química da troca enxofre-triazol (SuTEx, do inglês *Sulfur-Triazole Exchange*), foi apresentada por HAHM e colaboradores (2020) como uma plataforma ajustável para o

desenvolvimento de ligantes covalentes, mostrando que modificações no grupo de saída 1,2,4-triazol podem fornecer ligantes com quimiosseletividade aumentada em ~5 vezes para tirosinas em relação a outros aminoácidos nucleofílicos.

Baseado nessa proposta, projetou-se a modificação na cadeia lateral do resíduo Asp38 dos peptídeos para torná-lo um grupo ogiva do tipo SuTEx para reagir covalentemente com a Tyr449 do RBD. Dessa maneira, os SAPs 1, 2 e 6, que apresentaram os melhores resultados de interação com a proteína e que possuem resíduos de tirosina na região de contato, foram submetidos a essa alteração. Assim, fez-se a troca do carboxilato do aspartato inicialmente pelos grupos 1,2,4-triazol-sulfonila (aqui denominado de Z1) e 1,2,3-triazol-sulfonila (denominado de Z2) (Tabela 9). Por sua semelhança, também foram propostos peptídeos modificados contendo o grupo ogiva fosfonil-triazol, a fim de avaliar o efeito da troca do átomo de enxofre pelo de fósforo na reatividade do grupo ogiva.

**Tabela 9:** SAPs modificados contendo a sequência **EZLFYQ**, em que **Z** é um dos aminoácidos modificados abaixo, contendo grupos sulfonil-triazol ( $X = S$ ) ou fosfonil-triazol ( $X = P$ ), propostos neste trabalho para ligação covalente com o RBD.



SAP modificado	Anel heterocíclico	R
SAPn-Z1(S ou P)	1,2,4-triazol	H
SAPn-Z2(S ou P)	1,2,3-triazol	H
SAPn-Z3(S ou P)	1,2,4-triazol	C <sub>6</sub> H <sub>5</sub>
SAPn-Z4(S ou P)	1,2,4-triazol	4-FC <sub>6</sub> H <sub>4</sub>
SAPn-Z5(S ou P)	1,2,4-triazol	4-OCH <sub>3</sub> C <sub>6</sub> H <sub>4</sub>

$n = 1, 2$  ou  $6$ .

Antes da formação de um aduto covalente entre os peptídeos modificados e o RBD, deve haver um equilíbrio de formação de um complexo não-covalente. A partir dos dados de entalpia de formação, calculou-se por meio da equação 7 a entalpia de interação, inicialmente somente para os complexos com peptídeos modificados SAPnZ1S e SAPnZ2S ( $n = 1, 2$  ou  $6$ ), não-covalentes e totalmente otimizados. Observou-se que os valores de  $\Delta H_{int}$  (Tabela 10) foram menos favoráveis em relação aos dos complexos com peptídeos não modificados SAP1, SAP2

e SAP6 (Tabela 8). Além disso, as energias tanto das ligações de hidrogênio quanto das forças de dispersão variaram, se tornando mais ou menos favoráveis. Assim como o peptídeo não modificado SAP1, os peptídeos SAP1-Z1S e SAP1-Z2S podem ser considerados como os melhores candidatos a ligante do RBD do SARS-CoV-2 de sua série, em razão dos melhores valores de  $\Delta H_{int}$ .

**Tabela 10:** Parâmetros calculados pelo método semi-empírico PM7 ( $\varepsilon = 4,0$ ), para os complexos de interação com peptídeos modificados com grupos sulfonil-1,2,4-triazol (Z1S) e com grupo sulfonil-1,2,3-triazol (Z2S).

Complexo modificado	$\Delta H_{int}$ (kcal/mol)	Evolução de H <sup>a</sup> (kcal/mol)	Eforças de dispersão <sup>b</sup> (kcal/mol)
RBD-SAP1-Z1S	-101,46	-47,94	-170,11
RBD-SAP1-Z2S	-91,12	-40,19	-160,36
RBD-SAP2-Z1S	-13,86	-6,46	-91,88
RBD-SAP2-Z2S	-26,43	-7,97	-93,49
RBD-SAP6-Z1S	-23,03	-19,92	-85,56
RBD-SAP6-Z2S	-23,28	-23,03	-85,87

<sup>a</sup> Energia total, calculada pela diferença de energia entre produtos e reagentes;

<sup>b</sup> Calculado pela diferença de energia entre produtos e reagentes.

Na tabela 11, são apresentados os valores correspondentes de entalpia de interação para os complexos com peptídeos modificados SAP<sub>n</sub>Z1P e SAP<sub>n</sub>Z2P ( $n = 1, 2$  ou  $6$ ), não-covalentes e totalmente otimizados. Pode-se observar que 3 peptídeos modificados, SAP2-Z1P, SAP6-Z1P e SAP6-Z2P, apresentaram valores da  $\Delta H_{int}$  mais favoráveis do que seus respectivos contendo o grupo sulfonil-triazol, embora ainda não tão favoráveis em relação aos dos complexos com peptídeos não modificados (Tabela 8).

**Tabela 11:** Parâmetros calculados pelo método semi-empírico PM7 ( $\varepsilon = 4,0$ ), para os complexos de interação com peptídeos modificados com grupos ogiva fosfonil-1,2,4-triazol (Z1P) e com grupo fosfonil-1,2,3-triazol (Z2P) (continua).

Complexo modificado	$\Delta H_{int}$ (kcal/mol)	Evolução de H <sup>a</sup> (kcal/mol)	Eforças de dispersão <sup>b</sup> (kcal/mol)
RBD-SAP1-Z1P	-65,55	-43,89	-159,11
RBD-SAP1-Z2P	-75,21	-49,13	-163,47

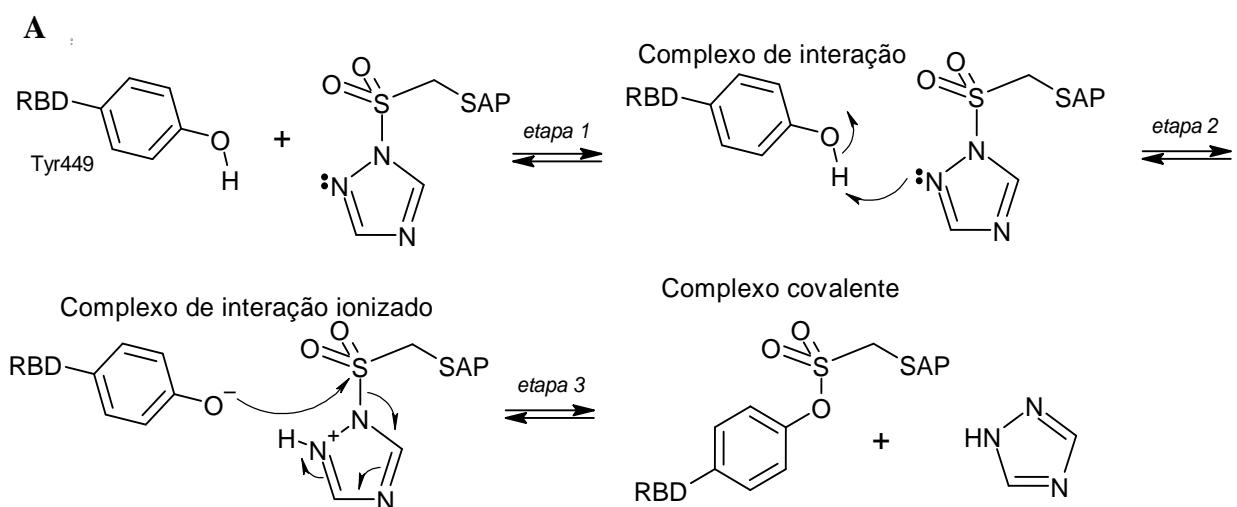
RBD-SAP2-Z1P	-22,53	-6,31	-91,22
RBD-SAP2-Z2P	-20,84	5,75	-86,14
RBD-SAP6-Z1P	-30,78	-18,31	-82,34
RBD-SAP6-Z2P	-32,83	-23,70	-85,77

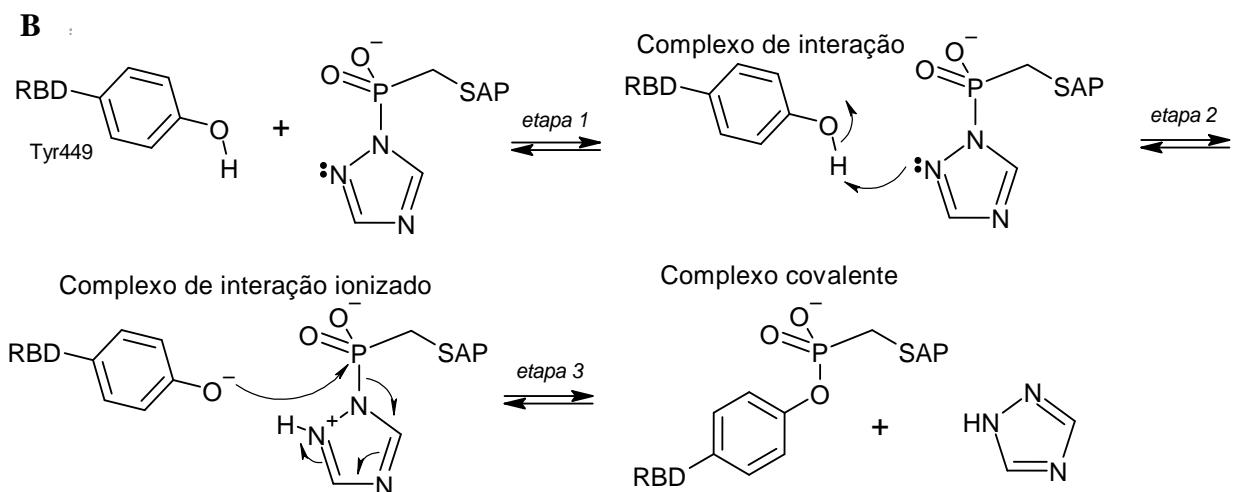
<sup>a</sup> Energia total, calculada pela diferença de energia entre produto e reagentes;

<sup>b</sup> Calculado pela diferença de energia entre produto e reagentes.

## 5.5 Análise da formação do complexo covalente

Após o equilíbrio de formação do complexo não-covalente (Figura 21, etapa 1), o passo posterior é a formação do complexo covalente. Avaliamos a formação para os peptídeos modificados contendo o grupo 1,2,4-triazol (Z1S e Z1P) ou 1,2,3-triazol (Z2S e Z2P). A proposta é que o par de elétrons livres do átomo de nitrogênio do anel triazólico capture o hidrogênio da Tyr449, que está mais próxima, sendo que a carga positiva resultante no anel triazólico protonado será estabilizada por ressonância (Figura 21, etapa 2). O oxigênio desprotonado e agora carregado negativamente, um nucleófilo potente, faz um ataque nucleofílico no enxofre ou fósforo do resíduo modificado do peptídeo, rompendo a ligação S-N ou P-N e formando, em seguida, um complexo covalente, com a liberação do grupo de saída 1,2,4-triazol ou 1,2,3-triazol (Figura 21, etapa 3).

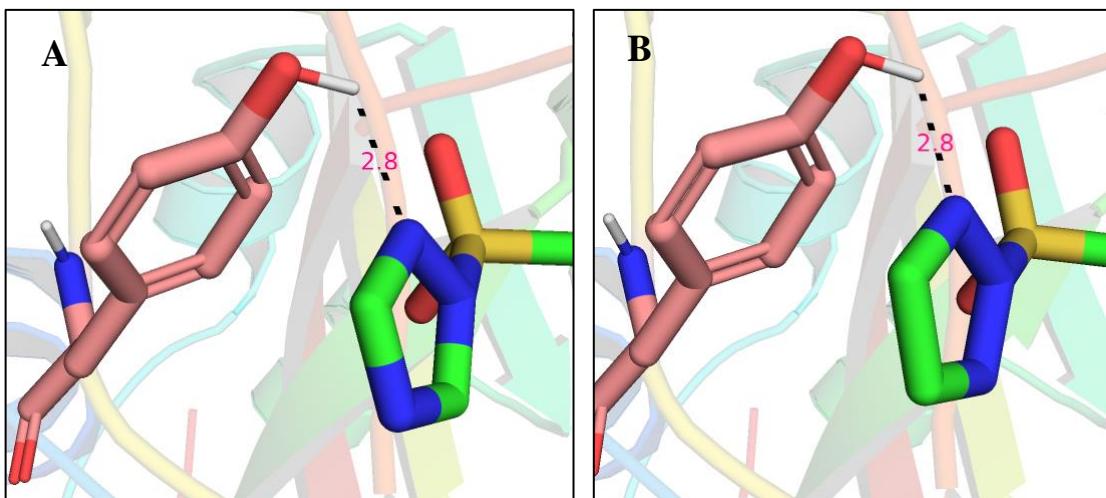




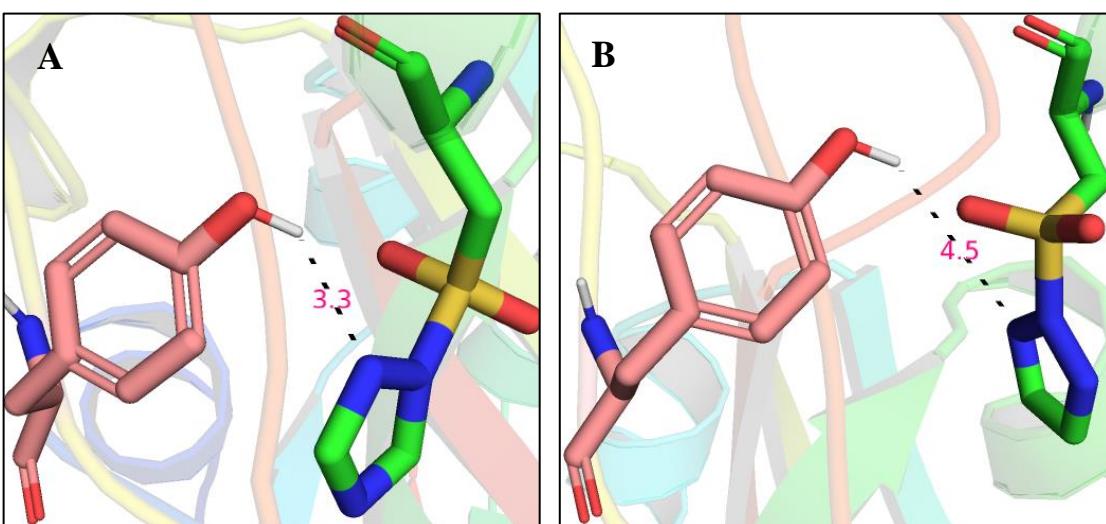
**Figura 21:** Exemplos de mecanismo propostos para a formação da ligação covalente entre os resíduos Tyr449 do RBD da Spike e os aminoácidos modificados presentes nos peptídeos (A) SAP<sub>n</sub>-Z1S e (B) SAP<sub>n</sub>-Z1P. Mecanismo similar pode ser proposto para os peptídeos contendo o grupo 1,2,3-triazol SAP<sub>n</sub>-Z2S e SAP<sub>n</sub>-Z2P.

A viabilidade da etapa 1 do mecanismo de reação (figura 21), ou seja, a captura do hidrogênio da hidroxila da Tyr449, foi avaliada por meio da mensuração da distância entre o nitrogênio do anel triazólico e o hidrogênio da tirosina. Esse é um parâmetro importante, pois por meio dele foi possível ter a noção se a desprotonação seria viável, além de possibilitar a dedução de em quais complexos isso aconteceria mais facilmente. Baseado nos valores fornecidos pelo Pymol, pode-se dizer que os derivados do SAP1, contendo o átomo de enxofre ou fósforo no resíduo modificado Z38, está bem posicionado em relação à tirosina, visto que a distância entre os átomos que reagem é de 2,8 Å (Figura 22) e, assim, pode-se afirmar que nos complexos com esse peptídeo a reação provavelmente se dará mais rapidamente.

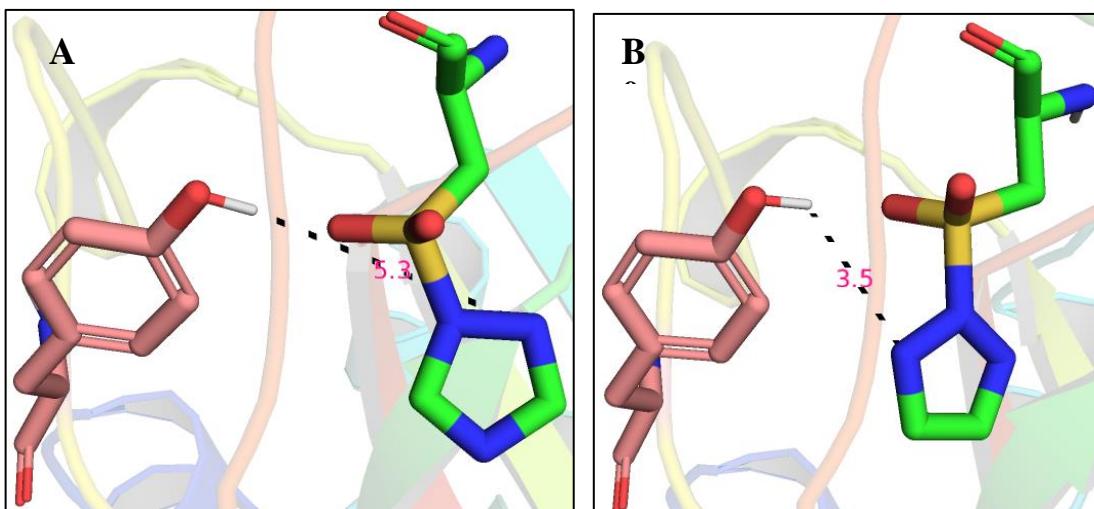
Isso vale também para os complexos com o SAP2 (na modificação Z1; Figura 23A) e SAP6 (na modificação Z2; Figura 24A), onde os peptídeos encontram-se a uma distância de ~ 3,4 Å da Tyr449. Em contrapartida, os peptídeos nos complexos RBD-SAP2-Z2 e RBD-SAP6-Z1 encontram-se mais distantes da tirosina, 4,5 Å (Figuras 23B) e 5,3 Å (24A), respectivamente. Nestes casos, a reação deverá ser mais lenta. Essencialmente os mesmos valores foram encontrados nos complexos contendo o átomo de fósforo no grupo ogiva.



**Figura 22:** Distância entre o nitrogênio do anel triazólico e o hidrogênio da hidroxila da Tyr449 (C em rosa), nos complexos **RBD-SAP1 modificados com grupo sulfonil-triazol**. **A)** RBD-SAP1-Z1 com 2,8 Å; **B)** RBD-SAP1-Z2 com 2,8 Å. Obs: apenas os hidrogênios polares são mostrados; e, nos complexos modificados contendo o grupo fosfonil-triazol, esses valores se alteraram para 2,6 Å.



**Figura 23:** Distância entre o nitrogênio do anel trizólico e o hidrogênio da hidroxila da Tyr449 (C em rosa), nos complexos **RBD-SAP2 modificados com grupo sulfonil-triazol**. **A)** RBD-SAP2-Z1 com 3,3 Å; **B)** RBD-SAP2-Z2 com 4,5 Å. Obs: apenas os hidrogênios polares são mostrados; e, os mesmos valores foram encontrados nos complexos modificados contendo o grupo fosfonil-triazol.



**Figura 24:** Distância entre o nitrogênio do anel trizólico e o hidrogênio da hidroxila da Tyr449 (C em rosa), nos complexos **RBD-SAP6 modificados com grupo sulfonil-triazol**. **A)** RBD-SAP6-Z1 com 5,3 Å; **B)** RBD-SAP6-Z2 com 3,5 Å. Obs: apenas os hidrogênios polares são mostrados; e, os mesmos valores foram encontrados nos complexos modificados contendo o grupo fosfonil-triazol.

No estudo da metodologia SuTEx (HAHM *et al.*, 2020), verificou-se que estruturas contendo o grupo sulfonil-1,2,3-triazol possuem menor seletividade para tirosina frente a lisina ( $\text{Tyr/Lys} = 2,3$ ; ANEXO F), em comparação com aquelas contendo o grupo sulfonil-1,2,4-triazol ( $\text{Tyr/Lys} = 2,8$ ; ANEXO F). Diante dessa característica, os autores realizaram a síntese e a análise de uma série de análogos de 1,2,4-triazol, com diferentes alterações na posição R do anel triazólico (Tabela 9), com a finalidade de elevar essa especificidade. Observou-se que a quimiosseletividade foi melhorada com a adição do grupo fenila ( $\text{Tyr/Lys} = 3,5$ ; ANEXO F) ao anel 1,2,4-triazol e que a substituição no fenil-triazol modificou ainda mais esse comportamento. De maneira que a adição do substituinte flúor na posição *para* do anel, resultou numa especificidade ligeiramente reduzida, em relação aos outros derivados, além de uma proximidade na reatividade com o 1,2,3-triazol. Por outro lado, a adição do grupo metóxi na posição *para* da fenila demonstrou ser altamente seletiva para tirosina ( $\text{Tyr/Lys} = \sim 5$ ; ANEXO F). Assim, a partir desses resultados, foram realizadas *in silico* as substituições no anel heterocíclico 1,2,4-triazol, denominadas de Z3, Z4 e Z5 (tabela 9).

No planejamento dos complexos com os peptídeos derivados foram considerados os parâmetros calculados pelo MOPAC, que somado à distância entre o nitrogênio do anel triazólico e o hidrogênio da cadeia lateral da Tyr449, definiram o peptídeo SAP1 como melhor candidato à ligante do RBD da *Spike* do SARS-CoV-2. Deste modo, para esses novos complexos, foi determinado inicialmente a  $\Delta\text{H}_{\text{int}}$  (Tabela 12), seguindo o mesmo procedimento

utilizado para os peptídeos modificados SAP1-Z1(S ou P) e SAP1-Z2(S ou P), aplicando a equação 7 no cálculo desse parâmetro. Pode-se observar que a introdução da fenila melhorou o valor de  $\Delta H_{int}$ , comparado com RBD-SAP1-Z1S e RBD-SAP1-Z1P. Além disso, na mensuração da distância entre o N triazólico e o H da tirosina, obteve-se o valor de 2,4 Å em todos os sistemas (ANEXO G).

**Tabela 12:**  $\Delta H_{int}$  e energias de interação calculadas pelo método semi-empírico PM7 ( $\epsilon = 4,0$ ), para os complexos derivados.

Complexo derivado	Substituição <sup>1</sup>	$\Delta H_{int}$ (kcal/mol)	Evolução de H <sup>a</sup> (kcal/mol)	Eforças de dispersão <sup>b</sup> (kcal/mol)
RBD-SAP1-Z3S	C <sub>6</sub> H <sub>5</sub>	-111,23	-56,23	-175,96
RBD-SAP1-Z4S	4-FC <sub>6</sub> H <sub>4</sub>	-102,22	-46,80	-169,44
RBD-SAP1-Z5S	4-OCH <sub>3</sub> C <sub>6</sub> H <sub>4</sub>	-98,80	-48,14	-166,09
RBD-SAP1-Z3P	C <sub>6</sub> H <sub>5</sub>	-118,97	-53,34	-163,31
RBD-SAP1-Z4P	4-FC <sub>6</sub> H <sub>4</sub>	-64,77	-36,78	-151,40
RBD-SAP1-Z5P	4-OCH <sub>3</sub> C <sub>6</sub> H <sub>4</sub>	-51,69	-34,27	-144,59

<sup>1</sup>no anel 1,2,4-triazol;

<sup>a</sup>Energia total, calculada pela diferença de energia entre produto e reagentes;

<sup>b</sup>Calculado pela diferença de energia entre produto e reagentes.

Diante dos bons resultados de  $\Delta H_{int}$ , realizou-se o ataque nucleofílico (etapa 3) nos complexos não covalentes RBD-SAP1-Z<sub>n</sub>, onde *n* corresponde a 1,2,3,4 ou 5 (tabela 9). Para simular a reação nas condições mais próximas do real, nos cálculos foi assumido que os peptídeos modificados estão inicialmente em fase aquosa ( $\epsilon = 78,4$ ), enquanto o RBD está no meio de baixa polaridade ( $\epsilon = 4$ ). Conforme o ligante se aproxima do sítio ativo de seu alvo, a solvatação ao redor de ambos diminui, aproximando os resíduos que irão interagir. Os complexos covalentes resultantes, RBD-SAP1-Z<sub>n</sub>(S ou P)cov (onde, Z<sub>n</sub>: modificação 1,2,3,4 ou 5; cov: covalente), situam-se no meio dielétrico de menor constante ( $\epsilon = 4$ ), liberando finalmente o grupo de saída (GS) para o meio aquoso ( $\epsilon = 78,4$ ) na etapa 3 do mecanismo proposto.

A partir dos dados de  $\Delta H_f^0$  dos reagentes e produtos, calculou-se a entalpia de reação nos sistemas, contendo tanto o átomo de enxofre quanto o de fósforo no grupo ogiva, por meio da equação 10. Verificou-se que todas as reações são exotérmicas (Tabela 13), confirmando a

afinidade proteína-ligante mesmo após as alterações estruturais nos peptídeos. Além disso, os complexos contendo o átomo de enxofre no grupo ogiva são mais estáveis entalpicamente, ou seja, tem a maior tendência de serem formados e persistirem no meio em que foram estudados, possivelmente devido às interações mais fortes e específicas presentes na interface RBD-SAP.

$$\Delta H_r = [(\Delta H_f(RBD-SAPnZcov)_{(pm)} + (\Delta H_f(GS)_{(aq)})] - [(\Delta H_f(RBD)_{(ptn)} + (\Delta H_f(SAPnZ)_{(aq)})] \quad (10)$$

**Tabela 13:**  $\Delta H_r$  calculada pelo método semi-empírico PM7 ( $\epsilon = 4,0$ ), para os complexos covalente.

Com átomo de enxofre		Com átomo de fósforo	
Complexo covalente	$\Delta H_r$ (kcal/mol)	Complexo covalente	$\Delta H_r$ (kcal/mol)
RBD-SAP1-Z1Scov	-128,73	RBD-SAP1-Z1Pcov	-100,93
RBD-SAP1-Z2Scov	-139,14	RBD-SAP1-Z2Pcov	-109,89
RBD-SAP1-Z3Scov	-115,80	RBD-SAP1-Z3Pcov	-34,74
RBD-SAP1-Z4Scov	-111,70	RBD-SAP1-Z4Pcov	-36,15
RBD-SAP1-Z5Scov	-124,20	RBD-SAP1-Z5Pcov	-20,12

Nos dois sistemas covalentes, tanto o complexo RBD-SAP1-Z2Scov quanto RBD-SAP1-Z2Pcov, que possuem o resíduo Z38 do peptídeo modificado pelo 1,2,3-triazol, apresentaram uma entalpia mais favorável. Sendo considerados como os SAPs de maior reatividade prevista nas séries de complexos propostas. Embora a questão da seletividade seja muito importante quando pensamos em futuros fármacos, deve ser lembrado que a sugestão de se basear em peptídeos construídos a partir do próprio alvo da *Spike*, a hACE2, deve conferir elevada seletividade aos peptídeos. Nesse sentido, entendemos que a quimiosseletividade do grupo ogiva seja menos importante, visto que é a estrutura do peptídeo que irá conferir a especificidade necessária para que esse grupo só venha a reagir com o resíduo alvo na *Spike*, a Tyr449.

Adicionalmente, mesmo que a formação do complexo de interação do SAP1 modificado tenha sido entalpicamente mais favorável do que a formação dos complexos covalentes do SAP1, deve ser lembrado que complexos covalentes são em geral mais estáveis do que os complexos não-covalentes, sendo essencialmente irreversíveis ou de reversibilidade longa, o que deve se refletir em um efeito mais prolongado de ação dos derivados propostos.

## 6. CONCLUSÃO

Nossos estudos indicam que todos os peptídeos derivados da hACE2 possuem estabilidade tanto em meio aquoso quanto na região mais próxima à cavidade da proteína, que é o local de baixa polaridade, devido às ligações de hidrogênio identificadas, que preservam as suas conformações nativas.

A otimização total da geometria dos peptídeos derivados da hACE2 e do RBD da *Spike* – separadas e em complexo, não só conferiu uma estabilidade conformacional a eles, mas permitiu que mais resíduos interagissem, ampliando a afinidade entre eles.

A sequência **EDLFYQ** é de fato importante para a interação proteína-peptídeo, tendo em vista que os complexos com os SAPs 1, 6 e 2, que a possuem em sua estrutura primária, apresentaram os melhores valores de entalpia de interação, respectivamente, com a efetiva participação de resíduos desse motivo estrutural em interações com o RBD.

Os resultados teóricos de  $\Delta H_{int}$  ajudaram a esclarecer os experimentais, pois tem a mesma tendência que a constante de dissociação ( $K_d$ ) de cada complexo, de modo que os SAPs 1, 6 e 2, nesta ordem, que possuem maiores afinidades com o alvo, de fato estabelecem interações com resíduos específicos do RBD da *Spike*.

Dentre essas interações, a mais relevante envolveu os resíduos Asp38 dos peptídeos e a Tyr449 do RBD. Isso permitiu o planejamento de modificações estruturais nos peptídeos baseadas na química SuTEx, usada para criar ligantes covalentes seletivos para tirosina, com a introdução de um resíduo modificado na posição 38.

Os novos peptídeos apresentaram, em geral, valores de  $\Delta H_{int}$  favoráveis com o RBD, embora não tão favoráveis quanto seus respectivos nativos. Apesar do desfavorecimento na entalpia de interação dos complexos modificados, os peptídeos SAP1-Z1(S ou P) e SAP1-Z2(S ou P) tiveram excelentes resultados, visto que possuem uma sequência maior de aminoácidos que interagem com o RBD, com mais resíduos contendo cadeias laterais polares, apolares e carregadas ou não, em comparação com os SAPs 2 e 6 modificados.

Adicionalmente, nossos resultados mostram a viabilidade da proposta, visto que complexos de interação puderam ser convertidos em complexos ligados covalentemente com valores de entalpias de reação favoráveis em todos os casos avaliados, sendo os sistemas com o átomo de enxofre no grupo ogiva mais estáveis em relação àqueles contendo o átomo de fósforo. Somada à seletividade pelo RBD da *Spike* do SARS-CoV-2, a capacidade de formação de ligações

covalentes estáveis confere aos peptídeos modificados aqui propostos características potenciais de inibição da infecção viral com ação duradoura.

É importante destacar que a nossa proposta de inibição covalente combina tanto a seletividade quanto a reatividade, isso porque utiliza um ligante com a mesma sequência peptídica do receptor natural do RBD da *Spike* do SARS-CoV-2, contendo um grupo ogiva com reatividade comprovada na literatura.

Além disso, ficou confirmado que o método de orbitais semi-empírico PM7 pode ser aplicado de forma prática em cálculos de otimização estrutural de sistemas contendo um grande número de átomos, incluindo proteínas e seus complexos, devido a rapidez e precisão nos cálculos de entalpias de formação. Somado ao uso da aproximação do contínuo para incluir o efeito do meio externo no sistema molecular, descrições realistas e práticas de sistemas de interesse em Química Medicinal são acessíveis aos grupos de pesquisa. Por ser um método quântico, sua aplicação é particularmente útil em sistemas em que ocorram formação de ligações covalentes, como é o caso de inibidores enzimáticos covalentes, exemplificados neste estudo.

## 7. REFERÊNCIAS

ADASME *et al.* PLIP 2021: expanding the scope of the protein-ligand interaction profiler to DNA and RNA. NAR 2021

ALMEIDA, J. O. de *et al.* COVID-19: Fisiopatologia e alvos para intervenção terapêutica. **Rev. Virtual Quim.** ISSN, v. 12, n. 6, p. 1464-1497, 2020.

ANDERSEN, Kristian G. *et al.* The proximal origin of SARS-CoV-2. **Nature medicine**, v. 26, n. 4, p. 450-452, 2020.

ATKINS, Peter; JONES, Loretta & LAVERMAN, Leroy. **Princípios de Química:- Questionando a Vida Moderna e o Meio Ambiente**. 5<sup>a</sup> edição. Porto Alegre: Bookman Editora, 2012. 252 p.

BARREIRO, Eliezer J. & FRAGA, Carlos Alberto Manssour. **Química Medicinal: as bases moleculares da ação dos fármacos**. 3<sup>a</sup> edição. Porto Alegre: Artmed, 2015. cap 1.

BARREIRO, Eliezer J. & FRAGA, Carlos Alberto Manssour. **Química Medicinal: as bases moleculares da ação dos fármacos**. 3<sup>a</sup> edição. Porto Alegre: Artmed, 2015. 566 p.

BARREIRO, Eliezer J. *et al.* Modelagem molecular: uma ferramenta para o planejamento racional de fármacos em química medicinal. **Química nova**, v. 20, p. 300-310, 1997.

BAUER, Renato A. Covalent inhibitors in drug discovery: from accidental discoveries to avoided liabilities and designed therapies. **Drug discovery today**, v. 20, n. 9, p. 1061-1073, 2015.

BERDAN, Viktoriya Y.; KLAUSER, Paul C. & WANG, Lei. Covalent peptides and proteins for therapeutics. **Bioorganic & Medicinal Chemistry**, v. 29, p. 115896, 2021.

BERG, Jeremy M.; TYMOCZKO, John L. & STRYER, Lubert. **Biochemistry** (Loose-Leaf). 5<sup>a</sup> edição. Macmillan, 2007.

BIO-MANGUINHOS/FIOCRUZ. **Vacinação contra a Covid-19 no Brasil completa um ano.** 18 de janeiro de 2022. Disponível em: <<https://portal.fiocruz.br/noticia/vacinacao-contra-covid-19-no-brasil-completa-um-ano>> Acesso em: 04 jun 2023.

BIO-MANGUINHOS/FIOCRUZ. **Vacinas de Vetores Virais.** 04 de março de 2022. Disponível em: <<https://www.bio.fiocruz.br/index.php/br/perguntas-frequentes/69-perguntas-frequentes/perguntas-frequentes-vacinas/2824-vacinas-de-vetores-virais-2#:~:text=As%20vacinas%20de%20vetores%20virais,sistema%20de%20entrega%20de%20informa%C3%A7%C3%A3o%20%9D.>> Acesso em: 27 out 2022.

BORGES, A. A.; SUZUKAWA, A. A.; ZANLUCA, C.; SANTOS & C. N. D. dos. SARS-CoV-2: origem, estrutura, morfogênese e transmissão. In: BARRAL-NETTO, M.; BARRETO, M. L.; PINTO JUNIOR, E. P.; ARAGÃO, E. (org.). **Construção de conhecimento no curso da pandemia de COVID-19: aspectos biomédicos, clínico-assistenciais, epidemiológicos e**

**sociais.** Salvador: Edufba, 2020. v. 1. Disponível em: <<https://books.scielo.org/id/hg5rg/pdf/barral-9786556302447.pdf>> Acesso em: 20 mai 2022.

BROWN, Theodore Lawrence. **Química, a ciência central.** 9<sup>a</sup> edição. São Paulo: Pearson Prentice Hall, 2005. 148 p.

CAVALCANTE, João Roberto *et al.* COVID-19 no Brasil: evolução da epidemia até a semana epidemiológica 20 de 2020. **Epidemiologia e Serviços de Saúde**, v. 29, 2020.

ERDELYI, Mate. Halogen bonding in solution. **Chemical Society Reviews**, v. 41, n. 9, p. 3547-3557, 2012.

European Centre for Disease Prevention and Control. **Latest evidence on COVID-19.** Disponível em: <<https://www.ecdc.europa.eu/en/covid-19/latest-evidence>> Acesso em: 5 mai 2020.

FITTOLANI, Giulio; SEEBERGER, Peter H.; DELBIANCO, Martina. Helical polysaccharides. **Peptide Science**, v. 112, n. 1, p. e24124, 2020.

FOKOUÉ, Harold H. *et al.* Há algo novo no reconhecimento molecular aplicado à Química Medicinal? **Química Nova**, v. 43, p. 78-89, 2020.

FOOD AND DRUG ADMINISTRATION (Estados Unidos). Coronavirus (COVID-19) | Drugs. FDA, 08 de novembro de 2022. Disponível em: <<https://www.fda.gov/drugs/emergency-preparedness-drugs/coronavirus-covid-19-drugs>> Acesso em: 27 out 2022.

Fotos de microscópio eletrônico mostram partículas do coronavírus. **National Geographic Brasil.** Disponível em: <Fotos de microscópio eletrônico mostram partículas do coronavírus | National Geographic ([nationalgeographicbrasil.com](http://nationalgeographicbrasil.com))> Acesso em: 03 out 2022.

GHOSH, Nimisha; NANDI, Suman & SAHA, Indrajit. A review on evolution of emerging SARS-CoV-2 variants based on spike glycoprotein. **International Immunopharmacology**, p. 108565, 2022.

GRÄF, T. Diversidade dos coronavírus, origem e evolução do SARS-COV-2. In: BARRAL-NETTO, M.; BARRETO, M. L.; PINTO JUNIOR, E. P.; ARAGÃO, E. (org.). **Construção de conhecimento no curso da pandemia de COVID-19: aspectos biomédicos, clínico-assistenciais, epidemiológicos e sociais.** Salvador: Edufba, 2020. v. 1. Disponível em: <<https://books.scielo.org/id/hg5rg/pdf/barral-9786556302447.pdf>> Acesso em: 20 mai 2022.

HADJ HASSINE, Ikbel. Covid-19 vaccines and variants of concern: A review. **Reviews in medical virology**, v. 32, n. 4, p. e 2313, 2022.

HAHM, H. S.; TOROITICH, E. K.; BORNE, A. L.; BRULET, J. W.; LIBBY, A. H.; YUAN, K.; WARE, T.B.; MCCLOUD, R. L.; CIANCONE, A.M. & HSU, K. L. Global targeting of functional tyrosines using sulfur-triazole exchange chemistry. **Nat. Chem. Biol.** 2020;16, 150-159. doi: 10.1038/s41589-019-0404-5. Epub 2019 Nov 25. PMID: 31768034; PMCID: PMC6982592.

HAN, Y. e KRAL, P. (2020) Computational Design of ACE2-Based Peptide Inhibitors of SARS-CoV-2. *ACS Nano* 14, 5143–5147, 2020.

HILLEBRAND, Laura & GEHRINGER, Matthias. Never Gonna Give You Up—Current Developments in Covalent Protein Kinase Inhibitors. *CHIMIA*, v. 76, n. 5, p. 435-435, 2022.

HOFFMANN, Markus *et al.* SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *cell*, v. 181, n. 2, p. 271-280. e8, 2020.

INSTITUTO BUTANTAN (Brasil). **Quais são as diferenças entre as vacinas contra Covid-19 que estão sendo aplicadas no Brasil?** COVID – Tira dúvida. São Paulo, 2022. Disponível em: <<https://butantan.gov.br/covid/butantan-tira-duvida/tira-duvida-noticias/quais-sao-as-diferencias-entre-as-vacinas-contra-covid-19-que-estao-sendo-aplicadas-no-brasil>> Acesso em: 27 out 2022.

INSTITUTO BUTANTAN (Brasil). **Seis fatos sobre a ômicron, a variante mais transmissível da Covid-19.** Notícias. São Paulo, 2022. Disponível em: <<https://butantan.gov.br/noticias/seis-fatos-sobre-a-omicron-a-variante-mais-transmissivel-da-covid-19>> Acesso em: 27 out 2022.

LAN, Jun *et al.* Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*, v. 581, n. 7807, p. 215-220, 2020.

LARUE, Ross C. *et al.* Rationally designed ACE2-derived peptides inhibit SARS-CoV-2. *Bioconjugate Chemistry*, v. 32, n. 1, p. 215-223, 2021.

LAU, Jolene L. & DUNN, Michael K. Therapeutic peptides: Historical perspectives, current development trends, and future directions. *Bioorganic & medicinal chemistry*, v. 26, n. 10, p. 2700-2707, 2018.

LI, Heng *et al.* Coronavirus disease 2019 (COVID-19): current status and future perspectives. *International journal of antimicrobial agents*, v. 55, n. 5, p. 105-951, 2020.

LI, Lin *et al.* On the dielectric “constant” of proteins: smooth dielectric function for macromolecular modeling and its implementation in DelPhi. *Journal of chemical theory and computation*, v. 9, n. 4, p. 2126-2136, 2013.

LI, Yaohao *et al.* Methods for engineering therapeutic peptides. *Chinese Chemical Letters*, v. 29, n. 7, p. 1074-1078, 2018.

LIMA, Lídia M. Química medicinal moderna: desafios e contribuição brasileira. *Química Nova*, v. 30, p. 1456-1468, 2007.

LIMA, Luana Nepomuceno Gondim Costa; DE SOUSA, Maisa Silva; LIMA, Karla Valéria Batista. As descobertas genômicas do SARS-CoV-2 e suas implicações na pandemia de COVID-19. *Journal of Health & Biological Sciences*, v. 8, n. 1, p. 1-9, 2020.

LIU, Zhixin *et al.* Composition and divergence of coronavirus spike proteins and host ACE2 receptors predict potential intermediate hosts of SARS-CoV-2. *Journal of medical virology*, v. 92, n. 6, p. 595-601, 2020.

MALIK, Yasmin A. Properties of coronavirus and SARS-CoV-2. **The Malaysian journal of pathology**, v. 42, n. 1, p. 3-11, 2020.

MANICA, James. **Anestesiologia**. Editora: Artmed, 2018, p. 693.

MINISTÉRIO DA SAÚDE (Brasil). **Vacinas bivalentes e monovalentes são igualmente eficazes e protegem contra a Covid-19**. Notícias. Brasil, 2022. Disponível em: <<https://www.gov.br/saude/pt-br/assuntos/noticias/2023/fevereiro/vacinas-bivalentes-e-monovalentes-sao-igualmente-eficazes-e-protegem-contra-a-covid-19-saiba-mais>> Acesso em: 22 mar 2023.

MUTTENTHALER, M., King; G. F., Adams, D. J. & Alewood, P. F. Trends in peptide drug discovery. **Nature reviews Drug discovery**, v. 20, n. 4, p. 309-325, 2021.

OLIVEIRA, Marilia Santini de; MATOS, Aline da Rocha & SIQUEIRA, Marilda Agudo Mendonça Teixeira de. Conhecendo o Sars-CoV-2 e a COVID-19. **Observatório Covid-19 Fiocruz**, 2020. Disponível em: <[https://www.arca.fiocruz.br/bitstream/handle/icict/49920/cap5\\_conhecendo\\_sars-Cov-2\\_covid\\_19.pdf;jsessionid=B4AF5DAE1F1C2209C4FFD31D693256A2?sequence=2](https://www.arca.fiocruz.br/bitstream/handle/icict/49920/cap5_conhecendo_sars-Cov-2_covid_19.pdf;jsessionid=B4AF5DAE1F1C2209C4FFD31D693256A2?sequence=2)> Acesso em: 17 jun 2023.

Organização Pan-Americana da Saúde (OPAS). **Histórico da pandemia de COVID-19**. Disponível em: <<https://www.paho.org/pt/covid19/historico-da-pandemia-covid-19>> Acesso em: 10 mai 2022.

Organização Pan-Americana da Saúde (OPAS). **OMS declara fim da Emergência de Saúde Pública de Importância Internacional referente à COVID-19**. Disponível em: <<https://www.paho.org/pt/noticias/5-5-2023-oms-declara-fim-da-emergencia-saude-publica-importancia-internacional-referente>> Acesso em: 04 jun 2023.

PAUL, Debleena *et al.* Artificial intelligence in drug discovery and development. **Drug discovery today**, v. 26, n. 1, p. 80, 2021.

ROCHA, Sheisi Fonseca Leite da Silva *et al.* **Desenvolvimento de um modelo empírico de predição da seletividade e da atividade de inibidores da Shp2 utilizando o método semi-empírico PM7**. 2019. Tese (Doutorado) – Curso de Pós-graduação em Química, Universidade Federal Rural do Rio de Janeiro, Seropédica, 2019.

SANT'ANNA, Carlos Maurício R. Métodos de Modelagem Molecular para Estudo e Planejamento de Compostos Bioativos: uma introdução. **Revista Virtual de Química**, p. 49-57, 2009.

SANT'ANNA, Carlos Maurício R. Uma Introdução à Modelagem Molecular aplicada à Química Medicinal. In BARREIRO, Eliezer J. & FRAGA, Carlos Alberto Manssour. **Química Medicinal: as bases moleculares da ação dos fármacos**. 3<sup>a</sup> edição. Porto Alegre: Artmed, 2015. cap 5.

SANT'ANNA, Carlos Mauricio R. Glossário de termos usados no planejamento de fármacos (recomendações da IUPAC para 1997). **Química Nova**, v. 25, p. 505-512, 2002.

SANTOS, Karina B. *et al.* Highly flexible ligand docking: benchmarking of the DockThor program on the LEADS-PEP protein-peptide data set. **Journal of Chemical Information and Modeling**, v. 60, n. 2, p. 667-683, 2020.

SANTOS, Paulo Pitasse *et al.* **Planejamento, síntese e avaliação da atividade biológica de peptídeos e conjugados peptídeo-fármacos, com atividade antibacteriana e antiparasitária.** Tese (Doutorado em Química) – Programa de Pós-graduação em Química, Universidade Federal Rural do Rio de Janeiro, Seropédica. 2022. Disponível em: <<https://tede.ufrrj.br/handle/jspui/5642>> Acesso em: 28 out 2022.

SHAKER, Bilal *et al.* In silico methods and tools for drug discovery. **Computers in Biology and Medicine**, v. 137, p. 104851, 2021.

SHANG, Jian *et al.* Cell entry mechanisms of SARS-CoV-2. **Proceedings of the National Academy of Sciences**, v. 117, n. 21, p. 11727-11734, 2020.

SHARMA, Anshika; AHMAD FAROUK, Isra & LAL, Sunil Kumar. COVID-19: A review on the novel coronavirus disease evolution, transmission, detection, control and prevention. **Viruses**, v. 13, n. 2, p. 202, 2021.

STEWART, James JP. MOPAC: a semiempirical molecular orbital program. **Journal of computer-aided molecular design**, v. 4, n. 1, p. 1-103, 1990.

STEWART, James JP. Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters. **Journal of molecular modeling**, v. 19, n. 1, p. 1-32, 2013.

SUTANTO, Fandi; KONSTANTINIDOU, Markella & DÖMLING, Alexander. Covalent inhibitors: a rational approach to drug discovery. **RSC medicinal chemistry**, v. 11, n. 8, p. 876-884, 2020.

TANG, Tiffany *et al.* Coronavirus membrane fusion mechanism offers a potential target for antiviral development. **Antiviral research**, v. 178, p. 104792, 2020.

V'KOVSKI, Philip *et al.* Coronavirus biology and replication: implications for SARS-CoV-2. **Nature Reviews Microbiology**, v. 19, n. 3, p. 155-170, 2021.

VALLE, Sandra Costa. **Receptor insulin-like e transporte de nutrientes energéticos em Neohelice granulata: sazonalidade.** Tese (Doutorado em Ciências) – Programa de Pós-graduação em Ciências Biológicas e Fisiologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, p. 58. 2008. Disponível em: <https://www.lume.ufrgs.br/handle/10183/14913> Acesso em: 23 out 2022.

VALVERDE, Ricardo. O que são mutações, linhagens, cepas e variantes? **Agência Fiocruz de Notícias**, 2021. Disponível em: <<https://portal.fiocruz.br/noticia/o-que-sao-mutacoes-linhagens-cepas-e-variantes-1>> Acesso em: 03 out 2022.

VERLI, Hugo & BARREIRO, Eliezer J. Um paradigma da química medicinal: a flexibilidade dos ligantes e receptores. **Química nova**, v. 28, p. 95-102, 2005.

WALLS, Alexandra C. *et al.* Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. **Cell**, v. 181, n. 2, p. 281-292, 2020.

WANG, Lei; WANG, Nanxi; ZHANG, Wenping; CHENG, Xurui; YAN, Zhibin; SHAO, Gang; WANG, Xi; WANG, Rui & Fu, Caiyun. Therapeutic peptides: Current applications and future directions. **Signal Transduction and Targeted Therapy**, v. 7, n. 1, p. 48, 2022.

**WHO. World Health Organization.** China Joint Mission On Coronavirus Disease 2019. Disponível em: <[https://www.who.int/emergencies/diseases/novel-coronavirus-2019?adgroupsurvey={adgroupsurvey}&gclid=Cj0KCQjwy5maBhDdARIsAMxrkw1ZqxhUS\\_8maKp1PK7NRHnxFo4waes2bW7601GOpmv-XCI-LiniOAAAsvfEALw\\_wcB](https://www.who.int/emergencies/diseases/novel-coronavirus-2019?adgroupsurvey={adgroupsurvey}&gclid=Cj0KCQjwy5maBhDdARIsAMxrkw1ZqxhUS_8maKp1PK7NRHnxFo4waes2bW7601GOpmv-XCI-LiniOAAAsvfEALw_wcB)> Acesso em: 22 mai 2020.

WOO, Patrick CY *et al.* Coronavirus diversity, phylogeny and interspecies jumping. **Experimental Biology and medicine**, v. 234, n. 10, p. 1117-1127, 2009.

YAN, Renhong *et al.* Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. **Science**, v. 367, n. 6485, p. 1444-1448, 2020.

YUKI, Koichi; FUJIOGI, Miho & KOUTSOGIANNAKI, Sophia. COVID-19 pathophysiology: A review. **Clinical immunology**, v. 215, p. 108427, 2020.

ZHANG, Qianqian *et al.* Molecular mechanism of interaction between SARS-CoV-2 and host cells and interventional therapy. **Signal Transduction and Targeted Therapy**, v. 6, n. 1, p. 1-19, 2021.

ZHANG, Yajie *et al.* Just how prevalent are peptide therapeutic products? A critical review. **International Journal of Pharmaceutics**, v. 587, p. 119491, 2020.

ZHOU, Peng *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. **nature**, v. 579, n. 7798, p. 270-273, 2020.

## ANEXOS

### A – Relação das ligações de hidrogênio identificadas no complexo nativo RBD-SAP1, no arquivo *arc*, após a otimização total.

SUMMARY OF PM7 CALCULATION, Site No: 22345

MOPAC2016 (Version: 18.117L)  
Wed May 17 09:19:04 2023

Empirical Formula: C1078 H1583 N277 O317 S8 = 3263 atoms

MOZYME PDBOUT EPS=4.0 1SCF DISP(1.0)

1SCF WAS SPECIFIED, SO BFGS WAS NOT USED  
SCF FIELD WAS ACHIEVED

HEAT OF FORMATION = -18207.98023 KCAL/MOL = -76182.18927 KJ/MOL

TOTAL ENERGY = -6931069.70971 KCAL/MOL  
ENERGY OF ATOMS = 6916928.14541 KCAL/MOL  
SUM = -14141.56430 KCAL/MOL  
DISPERSION ENERGY = -3032.24532 KCAL/MOL  
H-BOND ENERGY = -1068.21653 KCAL/MOL  
MM CORR. FOR -CO-NH- = 34.04592 KCAL/MOL  
SUM = -18207.98023 KCAL/MOL

No. OF HYDROGEN BONDS = 282 (H-bond Energy < -1.0 Kcal/mol)

List of hydrogen bonds found						
No.	Donor	R(D-H)	Hydrogen		Acceptor	H-bond energy
1	"ATOM 3120 OE2 GLU E 516"	1.469	"ATOM 1266 HH TYR E 396"	"ATOM 1257 OH TYR E 396"		-4.93
	Kcal/mol					
2	"ATOM 1405 OE1 GLU E 406"	1.508	"ATOM 2800 HH TYR E 495"	"ATOM 2791 OH TYR E 495"		-4.88
	Kcal/mol					
3	"ATOM 243 OH TYR A 41"	1.688	"ATOM 2869 HG1 THR E 500"	"ATOM 2864 OG1 THR E 500"		-4.87
	Kcal/mol					
4	"ATOM 1284 OD2 ASP E 398"	1.665	"ATOM 1676 HH TYR E 423"	"ATOM 1667 OH TYR E 423"		-4.85
	Kcal/mol					
5	"ATOM 2048 OH TYR E 449"	1.158	"ATOM 2057 HH TYR E 449"	"ATOM 188 OD2 ASP A 38"		-4.85 Kcal/mol
6	"ATOM 766 OD1 ASP E 364"	1.620	"ATOM 803 HG SER E 366"	"ATOM 798 OG SER E 366"		-4.79
	Kcal/mol					
7	"ATOM 931 OG SER E 375"	1.692	"ATOM 2989 HH TYR E 508"	"ATOM 2980 OH TYR E 508"		-4.65
	Kcal/mol					
8	"ATOM 1904 O ASN E 439"	1.704	"ATOM 1970 HG SER E 443"	"ATOM 1965 OG SER E 443"		-4.63
	Kcal/mol					
9	"ATOM 487 O PHE E 347"	1.689	"ATOM 2092 HH TYR E 451"	"ATOM 2083 OH TYR E 451"		-4.46
	Kcal/mol					
10	"ATOM 2123 OH TYR E 453"	1.699	"ATOM 2937 HH TYR E 505"	"ATOM 2928 OH TYR E 505"		-4.46
	Kcal/mol					
11	"ATOM 1205 O THR E 393"	1.717	"ATOM 3226 HG1 THR E 523"	"ATOM 3221 OG1 THR E 523"		-4.36
	Kcal/mol					
12	"ATOM 2387 OD1 ASP E 467"	1.682	"ATOM 2422 HG SER E 469"	"ATOM 2417 OG SER E 469"		-4.36
	Kcal/mol					

13 "ATOM 2630 O GLY E 485" 1.670 "ATOM 2698 HH TYR E 489" "ATOM 2689 OH TYR E 489" -4.34  
 Kcal/mol  
 14 "ATOM 127 O HIS A 34" 1.682 "ATOM 2132 HH TYR E 453" "ATOM 2123 OH TYR E 453" -4.33  
 Kcal/mol  
 15 "ATOM 2879 OD1 ASN E 501" 1.695 "ATOM 252 HH TYR A 41" "ATOM 243 OH TYR A 41" -4.30  
 Kcal/mol  
 16 "ATOM 1077 OG1 THR E 385" 1.716 "ATOM 1057 HG SER E 383" "ATOM 1052 OG SER E 383" -4.27  
 Kcal/mol  
 17 "ATOM 1955 OD2 ASP E 442" 1.706 "ATOM 1900 HG SER E 438" "ATOM 1895 OG SER E 438" -4.10  
 Kcal/mol  
 18 "ATOM 2199 O ARG E 457" 1.692 "ATOM 2491 HH TYR E 473" "ATOM 2482 OH TYR E 473" -4.08  
 Kcal/mol  
 19 "ATOM 63 OD2 ASP A 30" 1.410 "ATOM 1578 2HZ LYS E 417" "ATOM 1566 NZ LYS E 417" -3.56  
 Kcal/mol  
 20 "ATOM 642 NZ LYS E 356" 1.179 "ATOM 654 2HZ LYS E 356" "ATOM 379 OE2 GLU E 340" -3.53  
 Kcal/mol  
 21 "ATOM 3195 O PRO E 521" 1.750 "ATOM 1212 HG1 THR E 393" "ATOM 1207 OG1 THR E 393" -3.42  
 Kcal/mol  
 22 "ATOM 900 OG SER E 373" 1.787 "ATOM 884 HG SER E 371" "ATOM 879 OG SER E 371" -3.38  
 Kcal/mol  
 23 "ATOM 1648 OD1 ASN E 422" 1.660 "ATOM 2144 H ARG E 454" "ATOM 2133 N ARG E 454" -3.19  
 Kcal/mol  
 24 "ATOM 2388 OD2 ASP E 467" 1.614 "ATOM 2215 HE ARG E 457" "ATOM 2203 NE ARG E 457" -3.18  
 Kcal/mol  
 25 "ATOM 1284 OD2 ASP E 398" 1.592 "ATOM 633 2HH2 ARG E 355" "ATOM 620 NH2 ARG E 355" -3.17  
 Kcal/mol  
 26 "ATOM 2619 OE1 GLU E 484" 1.630 "ATOM 87 1HZ LYS A 31" "ATOM 76 NZ LYS A 31" -3.14  
 Kcal/mol  
 27 "ATOM 1148 OD2 ASP E 389" 1.397 "ATOM 1107 3HZ LYS E 386" "ATOM 1094 NZ LYS E 386" -3.11  
 Kcal/mol  
 28 "ATOM 2619 OE1 GLU E 484" 1.580 "ATOM 2710 H PHE E 490" "ATOM 2699 N PHE E 490" -3.08  
 Kcal/mol  
 29 "ATOM 1283 OD1 ASP E 398" 1.675 "ATOM 629 HE ARG E 355" "ATOM 617 NE ARG E 355" -2.99  
 Kcal/mol  
 30 "ATOM 2804 O GLY E 496" 1.755 "ATOM 2885 1HD2 ASN E 501" "ATOM 2880 ND2 ASN E 501" -2.97  
 Kcal/mol  
 31 "ATOM 149 OE2 GLU A 35" 1.585 "ATOM 88 2HZ LYS A 31" "ATOM 76 NZ LYS A 31" -2.97  
 Kcal/mol  
 32 "ATOM 1954 OD1 ASP E 442" 1.767 "ATOM 3013 2HH2 ARG E 509" "ATOM 3000 NH2 ARG E 509" -2.88  
 Kcal/mol  
 33 "ATOM 704 OG SER E 359" 1.751 "ATOM 675 HE ARG E 357" "ATOM 663 NE ARG E 357" -2.87  
 Kcal/mol  
 34 "ATOM 2660 OD1 ASN E 487" 1.681 "ATOM 10 3H THR A 27" "ATOM 1 N THR A 27" -2.84  
 Kcal/mol  
 35 "ATOM 261 NE2 GLN A 42" 2.202 "ATOM 2843 1HE2 GLN E 498" "ATOM 2836 NE2 GLN E 498" -2.84  
 Kcal/mol  
 36 "ATOM 1955 OD2 ASP E 442" 1.616 "ATOM 3009 HE ARG E 509" "ATOM 2997 NE ARG E 509" -2.80  
 Kcal/mol  
 37 "ATOM 1133 OD1 ASN E 388" 1.677 "ATOM 784 H TYR E 365" "ATOM 772 N TYR E 365" -2.77  
 Kcal/mol  
 38 "ATOM 260 OE1 GLN A 42" 1.864 "ATOM 2013 H GLY E 446" "ATOM 2009 N GLY E 446" -2.76  
 Kcal/mol  
 39 "ATOM 1893 O SER E 438" 1.734 "ATOM 1956 H ASP E 442" "ATOM 1948 N ASP E 442" -2.74  
 Kcal/mol  
 40 "ATOM 3033 O VAL E 511" 1.890 "ATOM 1831 H ILE E 434" "ATOM 1823 N ILE E 434" -2.68  
 Kcal/mol  
 41 "ATOM 2387 OD1 ASP E 467" 1.865 "ATOM 2418 H SER E 469" "ATOM 2412 N SER E 469" -2.65  
 Kcal/mol  
 42 "ATOM 1219 O ASN E 394" 1.950 "ATOM 3121 H GLU E 516" "ATOM 3112 N GLU E 516" -2.65  
 Kcal/mol  
 43 "ATOM 1222 OD1 ASN E 394" 1.729 "ATOM 709 HG SER E 359" "ATOM 704 OG SER E 359" -2.64  
 Kcal/mol  
 44 "ATOM 148 OE1 GLU A 35" 1.838 "ATOM 2775 H SER E 494" "ATOM 2769 N SER E 494" -2.64  
 Kcal/mol  
 45 "ATOM 2681 O TYR E 489" 1.655 "ATOM 2483 H TYR E 473" "ATOM 2471 N TYR E 473" -2.64  
 Kcal/mol

46 "ATOM 1779 O THR E 430" 1.897 "ATOM 3103 H PHE E 515" "ATOM 3092 N PHE E 515" -2.61  
 Kcal/mol  
 47 "ATOM 1292 O SER E 399" 1.816 "ATOM 604 H ASN E 354" "ATOM 596 N ASN E 354" -2.60  
 Kcal/mol  
 48 "ATOM 144 O GLU A 35" 1.920 "ATOM 201 H LEU A 39" "ATOM 193 N LEU A 39" -2.59  
 Kcal/mol  
 49 "ATOM 187 OD1 ASP A 38" 1.821 "ATOM 2844 2HE2 GLN E 498" "ATOM 2836 NE2 GLN E 498" -2.59  
 Kcal/mol  
 50 "ATOM 954 O PHE E 377" 1.730 "ATOM 991 2HZ LYS E 378" "ATOM 979 NZ LYS E 378" -2.55  
 Kcal/mol  
 51 "ATOM 940 O THR E 376" 1.809 "ATOM 1847 H ALA E 435" "ATOM 1842 N ALA E 435" -2.54  
 Kcal/mol  
 52 "ATOM 2565 N CYS E 480" 2.383 "ATOM 2588 2HD2 ASN E 481" "ATOM 2582 ND2 ASN E 481" -2.53  
 Kcal/mol  
 53 "ATOM 2115 O TYR E 453" 1.689 "ATOM 2761 H GLN E 493" "ATOM 2752 N GLN E 493" -2.53  
 Kcal/mol  
 54 "ATOM 2752 N GLN E 493" 2.133 "ATOM 2767 1HE2 GLN E 493" "ATOM 2760 NE2 GLN E 493" -2.53  
 Kcal/mol  
 55 "ATOM 1845 O ALA E 435" 1.950 "ATOM 944 H THR E 376" "ATOM 937 N THR E 376" -2.52  
 Kcal/mol  
 56 "ATOM 2016 N GLY E 447" 2.257 "ATOM 269 1HE2 GLN A 42" "ATOM 261 NE2 GLN A 42" -2.52  
 Kcal/mol  
 57 "ATOM 1800 O CYS E 432" 1.782 "ATOM 3070 H LEU E 513" "ATOM 3062 N LEU E 513" -2.50  
 Kcal/mol  
 58 "ATOM 2160 O LEU E 455" 1.840 "ATOM 1579 3HZ LYS E 417" "ATOM 1566 NZ LYS E 417" -2.48  
 Kcal/mol  
 59 "ATOM 2064 OD1 ASN E 450" 1.933 "ATOM 479 HE ARG E 346" "ATOM 467 NE ARG E 346" -2.48  
 Kcal/mol  
 60 "ATOM 1222 OD1 ASN E 394" 1.725 "ATOM 677 2HH1 ARG E 357" "ATOM 665 NH1 ARG E 357" -2.45  
 Kcal/mol  
 61 "ATOM 3115 O GLU E 516" 1.916 "ATOM 1209 H THR E 393" "ATOM 1202 N THR E 393" -2.44  
 Kcal/mol  
 62 "ATOM 2350 OE2 GLU E 465" 1.834 "ATOM 2218 1HH2 ARG E 457" "ATOM 2206 NH2 ARG E 457" -2.39  
 Kcal/mol  
 63 "ATOM 1810 O VAL E 433" 1.802 "ATOM 980 H LYS E 378" "ATOM 971 N LYS E 378" -2.38  
 Kcal/mol  
 64 "ATOM 2811 O PHE E 497" 1.854 "ATOM 1980 H LYS E 444" "ATOM 1971 N LYS E 444" -2.38  
 Kcal/mol  
 65 "ATOM 1050 O SER E 383" 2.049 "ATOM 1116 H LEU E 387" "ATOM 1108 N LEU E 387" -2.37  
 Kcal/mol  
 66 "ATOM 347 O PHE E 338" 1.900 "ATOM 413 H PHE E 342" "ATOM 402 N PHE E 342" -2.33  
 Kcal/mol  
 67 "ATOM 2179 O PHE E 456" 1.859 "ATOM 2154 2HH1 ARG E 454" "ATOM 2142 NH1 ARG E 454" -2.33  
 Kcal/mol  
 68 "ATOM 2415 O SER E 469" 1.800 "ATOM 2155 1HH2 ARG E 454" "ATOM 2143 NH2 ARG E 454" -2.32  
 Kcal/mol  
 69 "ATOM 637 O LYS E 356" 1.858 "ATOM 1272 H ALA E 397" "ATOM 1267 N ALA E 397" -2.32  
 Kcal/mol  
 70 "ATOM 1393 OD2 ASP E 405" 1.860 "ATOM 1378 2HH2 ARG E 403" "ATOM 1365 NH2 ARG E 403" -2.32  
 Kcal/mol  
 71 "ATOM 2722 O PRO E 491" 2.100 "ATOM 2187 H PHE E 456" "ATOM 2176 N PHE E 456" -2.31  
 Kcal/mol  
 72 "ATOM 1303 O PHE E 400" 1.986 "ATOM 3021 H VAL E 510" "ATOM 3014 N VAL E 510" -2.31  
 Kcal/mol  
 73 "ATOM 1233 O VAL E 395" 1.846 "ATOM 688 H ILE E 358" "ATOM 680 N ILE E 358" -2.30  
 Kcal/mol  
 74 "ATOM 3120 OE2 GLU E 516" 1.772 "ATOM 1229 2HD2 ASN E 394" "ATOM 1223 ND2 ASN E 394" -2.29  
 Kcal/mol  
 75 "ATOM 1432 O ARG E 408" 1.951 "ATOM 1536 2HE2 GLN E 414" "ATOM 1528 NE2 GLN E 414" -2.28  
 Kcal/mol  
 76 "ATOM 1185 O PHE E 392" 2.037 "ATOM 3237 H VAL E 524" "ATOM 3230 N VAL E 524" -2.28  
 Kcal/mol  
 77 "ATOM 71 O LYS A 31" 1.795 "ATOM 150 H GLU A 35" "ATOM 141 N GLU A 35" -2.27  
 Kcal/mol  
 78 "ATOM 2026 O ASN E 448" 2.020 "ATOM 2084 H TYR E 451" "ATOM 2072 N TYR E 451" -2.23  
 Kcal/mol

79 "ATOM 1921 OD1 ASN E 440" 2.041 "ATOM 1871 HE1 TRP E 436" "ATOM 1859 NE1 TRP E 436" -2.21  
 Kcal/mol  
 80 "ATOM 1052 OG SER E 383" 1.970 "ATOM 1079 H THR E 385" "ATOM 1072 N THR E 385" -2.20  
 Kcal/mol  
 81 "ATOM 2657 O ASN E 487" 1.942 "ATOM 2514 H ALA E 475" "ATOM 2509 N ALA E 475" -2.19  
 Kcal/mol  
 82 "ATOM 132 NE2 HIS A 34" 2.174 "ATOM 1375 1HH1 ARG E 403" "ATOM 1364 NH1 ARG E 403" -2.18  
 Kcal/mol  
 83 "ATOM 713 O ASN E 360" 2.115 "ATOM 295 H ASN E 334" "ATOM 287 N ASN E 334" -2.17  
 Kcal/mol  
 84 "ATOM 2755 O GLN E 493" 1.868 "ATOM 2124 H TYR E 453" "ATOM 2112 N TYR E 453" -2.16  
 Kcal/mol  
 85 "ATOM 599 O ASN E 354" 1.873 "ATOM 1295 H SER E 399" "ATOM 1289 N SER E 399" -2.16  
 Kcal/mol  
 86 "ATOM 2072 N TYR E 451" 2.224 "ATOM 2070 1HD2 ASN E 450" "ATOM 2065 ND2 ASN E 450" -2.16  
 Kcal/mol  
 87 "ATOM 2519 N GLY E 476" 2.415 "ATOM 2667 2HD2 ASN E 487" "ATOM 2661 ND2 ASN E 487" -2.15  
 Kcal/mol  
 88 "ATOM 3000 NH2 ARG E 509" 2.346 "ATOM 495 H PHE E 347" "ATOM 484 N PHE E 347" -2.14  
 Kcal/mol  
 89 "ATOM 1583 O ILE E 418" 1.978 "ATOM 1650 H ASN E 422" "ATOM 1642 N ASN E 422" -2.13  
 Kcal/mol  
 90 "ATOM 796 O SER E 366" 1.888 "ATOM 873 2HD2 ASN E 370" "ATOM 867 ND2 ASN E 370" -2.12  
 Kcal/mol  
 91 "ATOM 124 N HIS A 34" 2.673 "ATOM 175 H GLU A 37" "ATOM 166 N GLU A 37" -2.11  
 Kcal/mol  
 92 "ATOM 3017 O VAL E 510" 2.057 "ATOM 1311 H PHE E 400" "ATOM 1300 N PHE E 400" -2.11  
 Kcal/mol  
 93 "ATOM 3233 O VAL E 524" 1.806 "ATOM 1193 H PHE E 392" "ATOM 1182 N PHE E 392" -2.10  
 Kcal/mol  
 94 "ATOM 575 O TRP E 353" 1.738 "ATOM 2376 HE ARG E 466" "ATOM 2364 NE ARG E 466" -2.10  
 Kcal/mol  
 95 "ATOM 2256 O ASN E 460" 1.894 "ATOM 2217 2HH1 ARG E 457" "ATOM 2205 NH1 ARG E 457" -2.09  
 Kcal/mol  
 96 "ATOM 1680 O LYS E 424" 1.748 "ATOM 2333 H PHE E 464" "ATOM 2322 N PHE E 464" -2.09  
 Kcal/mol  
 97 "ATOM 3249 O CYS E 525" 1.873 "ATOM 755 H ALA E 363" "ATOM 750 N ALA E 363" -2.09  
 Kcal/mol  
 98 "ATOM 763 O ASP E 364" 2.046 "ATOM 828 H LEU E 368" "ATOM 820 N LEU E 368" -2.08  
 Kcal/mol  
 99 "ATOM 683 O ILE E 358" 2.065 "ATOM 1237 H VAL E 395" "ATOM 1230 N VAL E 395" -2.07  
 Kcal/mol  
 100 "ATOM 1392 OD1 ASP E 405" 1.806 "ATOM 1449 1HH1 ARG E 408" "ATOM 1438 NH1 ARG E 408" -2.07  
 Kcal/mol  
 101 "ATOM 2913 O GLY E 504" 2.019 "ATOM 1374 HE ARG E 403" "ATOM 1362 NE ARG E 403" -2.06  
 Kcal/mol  
 102 "ATOM 2029 OD1 ASN E 448" 1.994 "ATOM 1992 3HZ LYS E 444" "ATOM 1979 NZ LYS E 444" -2.05  
 Kcal/mol  
 103 "ATOM 2913 O GLY E 504" 1.942 "ATOM 1394 H ASP E 405" "ATOM 1386 N ASP E 405" -2.05  
 Kcal/mol  
 104 "ATOM 449 O THR E 345" 2.075 "ATOM 481 2HH1 ARG E 346" "ATOM 469 NH1 ARG E 346" -2.05  
 Kcal/mol  
 105 "ATOM 2075 O TYR E 451" 1.971 "ATOM 2792 H TYR E 495" "ATOM 2780 N TYR E 495" -2.05  
 Kcal/mol  
 106 "ATOM 1793 O GLY E 431" 1.855 "ATOM 1015 H TYR E 380" "ATOM 1003 N TYR E 380" -2.04  
 Kcal/mol  
 107 "ATOM 2384 O ASP E 467" 1.933 "ATOM 2152 HE ARG E 454" "ATOM 2140 NE ARG E 454" -2.03  
 Kcal/mol  
 108 "ATOM 2349 OE1 GLU E 465" 1.951 "ATOM 2295 H LYS E 462" "ATOM 2286 N LYS E 462" -2.03  
 Kcal/mol  
 109 "ATOM 62 OD1 ASP A 30" 1.958 "ATOM 138 HD1 HIS A 34" "ATOM 130 ND1 HIS A 34" -2.02  
 Kcal/mol  
 110 "ATOM 304 O LEU E 335" 1.945 "ATOM 299 1HD2 ASN E 334" "ATOM 294 ND2 ASN E 334" -2.02  
 Kcal/mol  
 111 "ATOM 519 OG SER E 349" 2.025 "ATOM 2101 H LEU E 452" "ATOM 2093 N LEU E 452" -2.02  
 Kcal/mol

112 "ATOM 1882 OD1 ASN E 437" 2.174 "ATOM 1909 H ASN E 439" "ATOM 1901 N ASN E 439" -2.01  
 Kcal/mol  
 113 "ATOM 974 O LYS E 378" 1.796 "ATOM 1814 H VAL E 433" "ATOM 1807 N VAL E 433" -2.01  
 Kcal/mol  
 114 "ATOM 3049 O VAL E 512" 1.963 "ATOM 1285 H ASP E 398" "ATOM 1277 N ASP E 398" -2.00  
 Kcal/mol  
 115 "ATOM 2702 O PHE E 490" 1.938 "ATOM 2768 2HE2 GLN E 493" "ATOM 2760 NE2 GLN E 493" -1.99  
 Kcal/mol  
 116 "ATOM 1249 O TYR E 396" 1.927 "ATOM 3087 H SER E 514" "ATOM 3081 N SER E 514" -1.99  
 Kcal/mol  
 117 "ATOM 1616 OD2 ASP E 420" 1.868 "ATOM 1555 H GLY E 416" "ATOM 1551 N GLY E 416" -1.97  
 Kcal/mol  
 118 "ATOM 184 O ASP A 38" 1.964 "ATOM 263 H GLN A 42" "ATOM 253 N GLN A 42" -1.97  
 Kcal/mol  
 119 "ATOM 2972 O TYR E 508" 2.132 "ATOM 1344 H ILE E 402" "ATOM 1336 N ILE E 402" -1.96  
 Kcal/mol  
 120 "ATOM 2522 O GLY E 476" 2.268 "ATOM 2507 1HE2 GLN E 474" "ATOM 2500 NE2 GLN E 474" -1.95  
 Kcal/mol  
 121 "ATOM 3084 O SER E 514" 1.870 "ATOM 1258 H TYR E 396" "ATOM 1246 N TYR E 396" -1.94  
 Kcal/mol  
 122 "ATOM 2811 O PHE E 497" 1.950 "ATOM 2031 H ASN E 448" "ATOM 2023 N ASN E 448" -1.94  
 Kcal/mol  
 123 "ATOM 1523 O GLN E 414" 1.794 "ATOM 1697 2HZ LYS E 424" "ATOM 1685 NZ LYS E 424" -1.92  
 Kcal/mol  
 124 "ATOM 1460 OE1 GLN E 409" 2.292 "ATOM 1604 H ALA E 419" "ATOM 1599 N ALA E 419" -1.91  
 Kcal/mol  
 125 "ATOM 1280 O ASP E 398" 1.859 "ATOM 3053 H VAL E 512" "ATOM 3046 N VAL E 512" -1.90  
 Kcal/mol  
 126 "ATOM 2345 O GLU E 465" 1.879 "ATOM 591 HE1 TRP E 353" "ATOM 579 NE1 TRP E 353" -1.90  
 Kcal/mol  
 127 "ATOM 374 O GLU E 340" 1.946 "ATOM 441 H ALA E 344" "ATOM 436 N ALA E 344" -1.90  
 Kcal/mol  
 128 "ATOM 2256 O ASN E 460" 2.023 "ATOM 2252 HG SER E 459" "ATOM 2247 OG SER E 459" -1.89  
 Kcal/mol  
 129 "ATOM 1974 O LYS E 444" 1.960 "ATOM 2020 H GLY E 447" "ATOM 2016 N GLY E 447" -1.89  
 Kcal/mol  
 130 "ATOM 1932 O LEU E 441" 2.131 "ATOM 482 1HH2 ARG E 346" "ATOM 470 NH2 ARG E 346" -1.88  
 Kcal/mol  
 131 "ATOM 2136 O ARG E 454" 2.054 "ATOM 1655 2HD2 ASN E 422" "ATOM 1649 ND2 ASN E 422" -1.88  
 Kcal/mol  
 132 "ATOM 727 O CYS E 361" 2.072 "ATOM 3252 H CYS E 525" "ATOM 3246 N CYS E 525" -1.88  
 Kcal/mol  
 133 "ATOM 1460 OE1 GLN E 409" 1.912 "ATOM 1588 H ILE E 418" "ATOM 1580 N ILE E 418" -1.87  
 Kcal/mol  
 134 "ATOM 1583 O ILE E 418" 1.862 "ATOM 1668 H TYR E 423" "ATOM 1656 N TYR E 423" -1.86  
 Kcal/mol  
 135 "ATOM 1554 O GLY E 416" 1.826 "ATOM 1617 H ASP E 420" "ATOM 1609 N ASP E 420" -1.86  
 Kcal/mol  
 136 "ATOM 1561 O LYS E 417" 1.877 "ATOM 1654 1HD2 ASN E 422" "ATOM 1649 ND2 ASN E 422" -1.84  
 Kcal/mol  
 137 "ATOM 2980 OH TYR E 508" 2.267 "ATOM 1889 2HD2 ASN E 437" "ATOM 1883 ND2 ASN E 437" -1.84  
 Kcal/mol  
 138 "ATOM 1086 N LYS E 386" 2.619 "ATOM 1135 H ASN E 388" "ATOM 1127 N ASN E 388" -1.83  
 Kcal/mol  
 139 "ATOM 2689 OH TYR E 489" 1.896 "ATOM 2662 H ASN E 487" "ATOM 2654 N ASN E 487" -1.82  
 Kcal/mol  
 140 "ATOM 1075 O THR E 385" 2.054 "ATOM 1149 H ASP E 389" "ATOM 1141 N ASP E 389" -1.82  
 Kcal/mol  
 141 "ATOM 20 O PHE A 28" 2.044 "ATOM 101 H PHE A 32" "ATOM 90 N PHE A 32" -1.81  
 Kcal/mol  
 142 "ATOM 290 O ASN E 334" 1.970 "ATOM 741 H VAL E 362" "ATOM 734 N VAL E 362" -1.81  
 Kcal/mol  
 143 "ATOM 1750 OD1 ASP E 428" 1.735 "ATOM 1752 H ASP E 428" "ATOM 1744 N ASP E 428" -1.79  
 Kcal/mol  
 144 "ATOM 40 O LEU A 29" 2.045 "ATOM 118 H ASN A 33" "ATOM 110 N ASN A 33" -1.79  
 Kcal/mol

145 "ATOM 2349 OE1 GLU E 465" 1.863 "ATOM 2216 1HH1 ARG E 457" "ATOM 2205 NH1 ARG E 457"-1.79  
 Kcal/mol  
 146 "ATOM 2873 N ASN E 501" 2.934 "ATOM 1914 2HD2 ASN E 439" "ATOM 1908 ND2 ASN E 439" -1.79  
 Kcal/mol  
 147 "ATOM 2061 O ASN E 450" 2.013 "ATOM 524 HG SER E 349" "ATOM 519 OG SER E 349" -1.77  
 Kcal/mol  
 148 "ATOM 1473 O ILE E 410" 2.075 "ATOM 1023 HH TYR E 380" "ATOM 1014 OH TYR E 380" -1.77  
 Kcal/mol  
 149 "ATOM 2048 OH TYR E 449" 1.997 "ATOM 270 2HE2 GLN A 42" "ATOM 261 NE2 GLN A 42" -1.77  
 Kcal/mol  
 150 "ATOM 1205 O THR E 393" 2.184 "ATOM 3223 H THR E 523" "ATOM 3216 N THR E 523" -1.76  
 Kcal/mol  
 151 "ATOM 1908 ND2 ASN E 439" 2.475 "ATOM 2953 1HE2 GLN E 506" "ATOM 2946 NE2 GLN E 506" -1.76  
 Kcal/mol  
 152 "ATOM 367 O GLY E 339" 1.977 "ATOM 435 2HD2 ASN E 343" "ATOM 429 ND2 ASN E 343" -1.75  
 Kcal/mol  
 153 "ATOM 3219 O THR E 523" 1.891 "ATOM 730 H CYS E 361" "ATOM 724 N CYS E 361" -1.74  
 Kcal/mol  
 154 "ATOM 2495 O GLN E 474" 2.267 "ATOM 2240 2HZ LYS E 458" "ATOM 2228 NZ LYS E 458" -1.74  
 Kcal/mol  
 155 "ATOM 1405 OE1 GLU E 406" 1.842 "ATOM 1366 H ARG E 403" "ATOM 1355 N ARG E 403" -1.73  
 Kcal/mol  
 156 "ATOM 1003 N TYR E 380" 3.056 "ATOM 1038 H VAL E 382" "ATOM 1031 N VAL E 382" -1.73  
 Kcal/mol  
 157 "ATOM 2671 O CYS E 488" 1.974 "ATOM 2631 H GLY E 485" "ATOM 2627 N GLY E 485" -1.72  
 Kcal/mol  
 158 "ATOM 71 O LYS A 31" 2.161 "ATOM 161 H ALA A 36" "ATOM 156 N ALA A 36" -1.72  
 Kcal/mol  
 159 "ATOM 1134 ND2 ASN E 388" 2.489 "ATOM 799 H SER E 366" "ATOM 793 N SER E 366" -1.72  
 Kcal/mol  
 160 "ATOM 2958 O PRO E 507" 1.881 "ATOM 1896 H SER E 438" "ATOM 1890 N SER E 438" -1.71  
 Kcal/mol  
 161 "ATOM 2228 NZ LYS E 458" 1.994 "ATOM 2239 1HZ LYS E 458" "ATOM 2445 OE2 GLU E 471" -1.71  
 Kcal/mol  
 162 "ATOM 4 O THR A 27" 2.226 "ATOM 77 H LYS A 31" "ATOM 68 N LYS A 31" -1.70 Kcal/mol  
 163 "ATOM 2928 OH TYR E 505" 2.272 "ATOM 2805 H GLY E 496" "ATOM 2801 N GLY E 496" -1.69  
 Kcal/mol  
 164 "ATOM 544 O TYR E 351" 2.217 "ATOM 2389 H ASP E 467" "ATOM 2381 N ASP E 467" -1.68  
 Kcal/mol  
 165 "ATOM 2831 O GLN E 498" 1.981 "ATOM 2881 H ASN E 501" "ATOM 2873 N ASN E 501" -1.67  
 Kcal/mol  
 166 "ATOM 1826 O ILE E 434" 1.958 "ATOM 3037 H VAL E 511" "ATOM 3030 N VAL E 511" -1.67  
 Kcal/mol  
 167 "ATOM 1294 OG SER E 399" 2.233 "ATOM 509 H ALA E 348" "ATOM 504 N ALA E 348" -1.65  
 Kcal/mol  
 168 "ATOM 2542 OG1 THR E 478" 2.352 "ATOM 2523 H GLY E 476" "ATOM 2519 N GLY E 476" -1.65  
 Kcal/mol  
 169 "ATOM 470 NH2 ARG E 346" 2.152 "ATOM 2035 1HD2 ASN E 448" "ATOM 2030 ND2 ASN E 448" -1.63  
 Kcal/mol  
 170 "ATOM 565 O ALA E 352" 2.207 "ATOM 608 1HD2 ASN E 354" "ATOM 603 ND2 ASN E 354" -1.62  
 Kcal/mol  
 171 "ATOM 3065 O LEU E 513" 1.986 "ATOM 1803 H CYS E 432" "ATOM 1797 N CYS E 432" -1.61  
 Kcal/mol  
 172 "ATOM 1855 O TRP E 436" 1.967 "ATOM 3001 H ARG E 509" "ATOM 2990 N ARG E 509" -1.61  
 Kcal/mol  
 173 "ATOM 1406 OE2 GLU E 406" 1.958 "ATOM 1376 2HH1 ARG E 403" "ATOM 1364 NH1 ARG E 403" -1.60  
 Kcal/mol  
 174 "ATOM 2289 O LYS E 462" 1.946 "ATOM 2351 H GLU E 465" "ATOM 2342 N GLU E 465" -1.60  
 Kcal/mol  
 175 "ATOM 2030 ND2 ASN E 448" 2.413 "ATOM 2071 2HD2 ASN E 450" "ATOM 2065 ND2 ASN E 450" -1.60  
 Kcal/mol  
 176 "ATOM 2540 O THR E 478" 1.925 "ATOM 2547 HG1 THR E 478" "ATOM 2542 OG1 THR E 478" -1.60  
 Kcal/mol  
 177 "ATOM 2660 OD1 ASN E 487" 2.306 "ATOM 28 H PHE A 28" "ATOM 17 N PHE A 28" -1.60  
 Kcal/mol  
 178 "ATOM 159 O ALA A 36" 1.864 "ATOM 223 H PHE A 40" "ATOM 212 N PHE A 40" -1.59  
 Kcal/mol

179 "ATOM 449 O THR E 345" 1.835 "ATOM 3012 1HH2 ARG E 509" "ATOM 3000 NH2 ARG E 509" -1.59  
 Kcal/mol  
 180 "ATOM 2879 OD1 ASN E 501" 2.087 "ATOM 2837 H GLN E 498" "ATOM 2828 N GLN E 498" -1.57  
 Kcal/mol  
 181 "ATOM 2474 O TYR E 473" 2.151 "ATOM 2690 H TYR E 489" "ATOM 2678 N TYR E 489" -1.57  
 Kcal/mol  
 182 "ATOM 2993 O ARG E 509" 2.074 "ATOM 1866 H TRP E 436" "ATOM 1852 N TRP E 436" -1.56  
 Kcal/mol  
 183 "ATOM 507 O ALA E 348" 2.173 "ATOM 1327 H VAL E 401" "ATOM 1320 N VAL E 401" -1.55  
 Kcal/mol  
 184 "ATOM 1389 O ASP E 405" 2.190 "ATOM 1468 1HE2 GLN E 409" "ATOM 1461 NE2 GLN E 409" -1.55  
 Kcal/mol  
 185 "ATOM 1336 N ILE E 402" 2.916 "ATOM 2981 H TYR E 508" "ATOM 2969 N TYR E 508" -1.54  
 Kcal/mol  
 186 "ATOM 572 N TRP E 353" 2.783 "ATOM 2376 HE ARG E 466" "ATOM 2364 NE ARG E 466" -1.53  
 Kcal/mol  
 187 "ATOM 659 O ARG E 357" 2.234 "ATOM 655 3HZ LYS E 356" "ATOM 642 NZ LYS E 356" -1.52  
 Kcal/mol  
 188 "ATOM 1738 OD1 ASP E 427" 1.805 "ATOM 1740 H ASP E 427" "ATOM 1732 N ASP E 427" -1.52  
 Kcal/mol  
 189 "ATOM 3115 O GLU E 516" 2.413 "ATOM 1224 H ASN E 394" "ATOM 1216 N ASN E 394" -1.52  
 Kcal/mol  
 190 "ATOM 333 O PRO E 337" 2.416 "ATOM 393 H VAL E 341" "ATOM 386 N VAL E 341" -1.51  
 Kcal/mol  
 191 "ATOM 1845 O ALA E 435" 1.939 "ATOM 932 H SER E 375" "ATOM 926 N SER E 375" -1.51  
 Kcal/mol  
 192 "ATOM 1615 OD1 ASP E 420" 1.806 "ATOM 2265 1HD2 ASN E 460" "ATOM 2260 ND2 ASN E 460" -1.51  
 Kcal/mol  
 193 "ATOM 1392 OD1 ASP E 405" 1.872 "ATOM 1451 1HH2 ARG E 408" "ATOM 1439 NH2 ARG E 408" -1.49  
 Kcal/mol  
 194 "ATOM 1006 O TYR E 380" 2.258 "ATOM 1794 H GLY E 431" "ATOM 1790 N GLY E 431" -1.45  
 Kcal/mol  
 195 "ATOM 460 N ARG E 346" 3.017 "ATOM 3012 1HH2 ARG E 509" "ATOM 3000 NH2 ARG E 509" -1.45  
 Kcal/mol  
 196 "ATOM 1609 N ASP E 420" 3.259 "ATOM 1698 3HZ LYS E 424" "ATOM 1685 NZ LYS E 424" -1.44  
 Kcal/mol  
 197 "ATOM 2426 O THR E 470" 1.951 "ATOM 2433 HG1 THR E 470" "ATOM 2428 OG1 THR E 470" -1.43  
 Kcal/mol  
 198 "ATOM 1358 O ARG E 403" 2.061 "ATOM 1420 H VAL E 407" "ATOM 1413 N VAL E 407" -1.42  
 Kcal/mol  
 199 "ATOM 2350 OE2 GLU E 465" 1.592 "ATOM 2305 1HZ LYS E 462" "ATOM 2294 NZ LYS E 462" -1.42  
 Kcal/mol  
 200 "ATOM 2783 O TYR E 495" 1.897 "ATOM 2819 H PHE E 497" "ATOM 2808 N PHE E 497" -1.41  
 Kcal/mol  
 201 "ATOM 2029 OD1 ASN E 448" 2.061 "ATOM 483 2HH2 ARG E 346" "ATOM 470 NH2 ARG E 346" -1.41  
 Kcal/mol  
 202 "ATOM 2938 N GLN E 506" 3.357 "ATOM 1383 H GLY E 404" "ATOM 1379 N GLY E 404" -1.41  
 Kcal/mol  
 203 "ATOM 839 N TYR E 369" 3.027 "ATOM 880 H SER E 371" "ATOM 874 N SER E 371" -1.40  
 Kcal/mol  
 204 "ATOM 2037 N TYR E 449" 2.975 "ATOM 1990 1HZ LYS E 444" "ATOM 1979 NZ LYS E 444" -1.40  
 Kcal/mol  
 205 "ATOM 386 N VAL E 341" 3.111 "ATOM 430 H ASN E 343" "ATOM 422 N ASN E 343" -1.40  
 Kcal/mol  
 206 "ATOM 842 O TYR E 369" 2.122 "ATOM 890 H ALA E 372" "ATOM 885 N ALA E 372" -1.39  
 Kcal/mol  
 207 "ATOM 2393 N ILE E 468" 3.171 "ATOM 2156 2HH2 ARG E 454" "ATOM 2143 NH2 ARG E 454" -1.37  
 Kcal/mol  
 208 "ATOM 766 OD1 ASP E 364" 2.072 "ATOM 799 H SER E 366" "ATOM 793 N SER E 366" -1.36  
 Kcal/mol  
 209 "ATOM 439 O ALA E 344" 1.881 "ATOM 3010 1HH1 ARG E 509" "ATOM 2999 NH1 ARG E 509" -1.35  
 Kcal/mol  
 210 "ATOM 1859 NE1 TRP E 436" 2.552 "ATOM 1928 2HD2 ASN E 440" "ATOM 1922 ND2 ASN E 440" -1.35  
 Kcal/mol  
 211 "ATOM 775 O TYR E 365" 2.068 "ATOM 851 H TYR E 369" "ATOM 839 N TYR E 369" -1.34  
 Kcal/mol

212 "ATOM 2760 NE2 GLN E 493" 3.115 "ATOM 89 3HZ LYS A 31" "ATOM 76 NZ LYS A 31" -1.34  
 Kcal/mol  
 213 "ATOM 446 N THR E 345" 3.276 "ATOM 3010 1HH1 ARG E 509" "ATOM 2999 NH1 ARG E 509" -1.33  
 Kcal/mol  
 214 "ATOM 2938 N GLN E 506" 2.801 "ATOM 2886 2HD2 ASN E 501" "ATOM 2880 ND2 ASN E 501" -1.33  
 Kcal/mol  
 215 "ATOM 1277 N ASP E 398" 2.654 "ATOM 3053 H VAL E 512" "ATOM 3046 N VAL E 512" -1.33  
 Kcal/mol  
 216 "ATOM 59 O ASP A 30" 2.221 "ATOM 134 H HIS A 34" "ATOM 124 N HIS A 34" -1.33 Kcal/mol  
 217 "ATOM 2941 O GLN E 506" 2.000 "ATOM 1383 H GLY E 404" "ATOM 1379 N GLY E 404" -1.32  
 Kcal/mol  
 218 "ATOM 1456 O GLN E 409" 1.995 "ATOM 1494 H ALA E 411" "ATOM 1489 N ALA E 411" -1.32  
 Kcal/mol  
 219 "ATOM 596 N ASN E 354" 3.337 "ATOM 2380 2HH2 ARG E 466" "ATOM 2367 NH2 ARG E 466" -1.32  
 Kcal/mol  
 220 "ATOM 3221 OG1 THR E 523" 2.037 "ATOM 718 H ASN E 360" "ATOM 710 N ASN E 360" -1.31  
 Kcal/mol  
 221 "ATOM 1527 OE1 GLN E 414" 1.928 "ATOM 1544 H THR E 415" "ATOM 1537 N THR E 415" -1.31  
 Kcal/mol  
 222 "ATOM 323 O CYS E 336" 2.897 "ATOM 368 H GLY E 339" "ATOM 364 N GLY E 339" -1.29  
 Kcal/mol  
 223 "ATOM 1492 O ALA E 411" 2.180 "ATOM 1529 H GLN E 414" "ATOM 1520 N GLN E 414" -1.29  
 Kcal/mol  
 224 "ATOM 525 N VAL E 350" 3.193 "ATOM 567 H ALA E 352" "ATOM 562 N ALA E 352" -1.28  
 Kcal/mol  
 225 "ATOM 1339 O ILE E 402" 1.856 "ATOM 2981 H TYR E 508" "ATOM 2969 N TYR E 508" -1.27  
 Kcal/mol  
 226 "ATOM 1389 O ASP E 405" 2.105 "ATOM 1450 2HH1 ARG E 408" "ATOM 1438 NH1 ARG E 408" -1.27  
 Kcal/mol  
 227 "ATOM 1602 O ALA E 419" 2.096 "ATOM 1698 3HZ LYS E 424" "ATOM 1685 NZ LYS E 424" -1.26  
 Kcal/mol  
 228 "ATOM 469 NH1 ARG E 346" 3.613 "ATOM 2035 1HD2 ASN E 448" "ATOM 2030 ND2 ASN E 448" -1.25  
 Kcal/mol  
 229 "ATOM 3081 N SER E 514" 2.816 "ATOM 1258 H TYR E 396" "ATOM 1246 N TYR E 396" -1.23  
 Kcal/mol  
 230 "ATOM 724 N CYS E 361" 2.985 "ATOM 278 1H THR E 333" "ATOM 271 N THR E 333" -1.22  
 Kcal/mol  
 231 "ATOM 2759 OE1 GLN E 493" 1.947 "ATOM 89 3HZ LYS A 31" "ATOM 76 NZ LYS A 31" -1.22  
 Kcal/mol  
 232 "ATOM 1382 O GLY E 404" 2.539 "ATOM 1440 H ARG E 408" "ATOM 1429 N ARG E 408" -1.21  
 Kcal/mol  
 233 "ATOM 1993 N VAL E 445" 3.468 "ATOM 2020 H GLY E 447" "ATOM 2016 N GLY E 447" -1.21  
 Kcal/mol  
 234 "ATOM 796 O SER E 366" 1.823 "ATOM 868 H ASN E 370" "ATOM 860 N ASN E 370" -1.20  
 Kcal/mol  
 235 "ATOM 3149 O LEU E 518" 2.005 "ATOM 3187 H ALA E 520" "ATOM 3182 N ALA E 520" -1.19  
 Kcal/mol  
 236 "ATOM 1520 N GLN E 414" 3.775 "ATOM 1697 2HZ LYS E 424" "ATOM 1685 NZ LYS E 424" -1.19  
 Kcal/mol  
 237 "ATOM 1879 O ASN E 437" 3.234 "ATOM 1923 H ASN E 440" "ATOM 1915 N ASN E 440" -1.18  
 Kcal/mol  
 238 "ATOM 1127 N ASN E 388" 3.239 "ATOM 1161 H LEU E 390" "ATOM 1153 N LEU E 390" -1.18  
 Kcal/mol  
 239 "ATOM 1580 N ILE E 418" 3.557 "ATOM 1633 H TYR E 421" "ATOM 1621 N TYR E 421" -1.17  
 Kcal/mol  
 240 "ATOM 3062 N LEU E 513" 2.907 "ATOM 1803 H CYS E 432" "ATOM 1797 N CYS E 432" -1.17  
 Kcal/mol  
 241 "ATOM 2890 O GLY E 502" 2.055 "ATOM 2947 H GLN E 506" "ATOM 2938 N GLN E 506" -1.17  
 Kcal/mol  
 242 "ATOM 1156 O LEU E 390" 2.359 "ATOM 3260 H GLY E 526" "ATOM 3256 N GLY E 526" -1.17  
 Kcal/mol  
 243 "ATOM 737 O VAL E 362" 1.915 "ATOM 326 H CYS E 336" "ATOM 320 N CYS E 336" -1.17  
 Kcal/mol  
 244 "ATOM 2945 OE1 GLN E 506" 1.908 "ATOM 1888 1HD2 ASN E 437" "ATOM 1883 ND2 ASN E 437" -1.16  
 Kcal/mol  
 245 "ATOM 1153 N LEU E 390" 3.828 "ATOM 3260 H GLY E 526" "ATOM 3256 N GLY E 526" -1.16  
 Kcal/mol

246 "ATOM 2019 O GLY E 447" 2.241 "ATOM 1990 1HZ LYS E 444" "ATOM 1979 NZ LYS E 444" -1.16  
 Kcal/mol  
 247 "ATOM 2415 O SER E 469" 1.950 "ATOM 2446 H GLU E 471" "ATOM 2437 N GLU E 471" -1.15  
 Kcal/mol  
 248 "ATOM 2890 O GLY E 502" 2.045 "ATOM 2929 H TYR E 505" "ATOM 2917 N TYR E 505" -1.15  
 Kcal/mol  
 249 "ATOM 333 O PRO E 337" 1.935 "ATOM 380 H GLU E 340" "ATOM 371 N GLU E 340" -1.14  
 Kcal/mol  
 250 "ATOM 2492 N GLN E 474" 2.965 "ATOM 2240 2HZ LYS E 458" "ATOM 2228 NZ LYS E 458" -1.14  
 Kcal/mol  
 251 "ATOM 470 NH2 ARG E 346" 3.194 "ATOM 2071 2HD2 ASN E 450" "ATOM 2065 ND2 ASN E 450" -1.13  
 Kcal/mol  
 252 "ATOM 40 O LEU A 29" 1.951 "ATOM 123 2HD2 ASN A 33" "ATOM 117 ND2 ASN A 33" -1.12  
 Kcal/mol  
 253 "ATOM 793 N SER E 366" 3.536 "ATOM 851 H TYR E 369" "ATOM 839 N TYR E 369" -1.12  
 Kcal/mol  
 254 "ATOM 2880 ND2 ASN E 501" 2.602 "ATOM 2837 H GLN E 498" "ATOM 2828 N GLN E 498" -1.12  
 Kcal/mol  
 255 "ATOM 1413 N VAL E 407" 3.247 "ATOM 1462 H GLN E 409" "ATOM 1453 N GLN E 409" -1.12  
 Kcal/mol  
 256 "ATOM 181 N ASP A 38" 3.423 "ATOM 223 H PHE A 40" "ATOM 212 N PHE A 40" -1.10  
 Kcal/mol  
 257 "ATOM 90 N PHE A 32" 3.483 "ATOM 134 H HIS A 34" "ATOM 124 N HIS A 34" -1.09 Kcal/mol  
 258 "ATOM 2259 OD1 ASN E 460" 1.961 "ATOM 2261 H ASN E 460" "ATOM 2253 N ASN E 460" -1.09  
 Kcal/mol  
 259 "ATOM 2096 O LEU E 452" 2.986 "ATOM 553 H TYR E 351" "ATOM 541 N TYR E 351" -1.09  
 Kcal/mol  
 260 "ATOM 1951 O ASP E 442" 1.932 "ATOM 2036 2HD2 ASN E 448" "ATOM 2030 ND2 ASN E 448" -1.09  
 Kcal/mol  
 261 "ATOM 181 N ASP A 38" 3.651 "ATOM 244 H TYR A 41" "ATOM 232 N TYR A 41" -1.09  
 Kcal/mol  
 262 "ATOM 575 O TRP E 353" 1.924 "ATOM 2380 2HH2 ARG E 466" "ATOM 2367 NH2 ARG E 466" -1.09  
 Kcal/mol  
 263 "ATOM 1948 N ASP E 442" 3.631 "ATOM 480 1HH1 ARG E 346" "ATOM 469 NH1 ARG E 346" -1.09  
 Kcal/mol  
 264 "ATOM 517 O SER E 349" 2.036 "ATOM 567 H ALA E 352" "ATOM 562 N ALA E 352" -1.07  
 Kcal/mol  
 265 "ATOM 17 N PHE A 28" 3.265 "ATOM 2662 H ASN E 487" "ATOM 2654 N ASN E 487" -1.07  
 Kcal/mol  
 266 "ATOM 2499 OE1 GLN E 474" 2.067 "ATOM 2571 H CYS E 480" "ATOM 2565 N CYS E 480" -1.06  
 Kcal/mol  
 267 "ATOM 971 N LYS E 378" 3.154 "ATOM 991 2HZ LYS E 378" "ATOM 979 NZ LYS E 378" -1.05  
 Kcal/mol  
 268 "ATOM 56 N ASP A 30" 3.160 "ATOM 123 2HD2 ASN A 33" "ATOM 117 ND2 ASN A 33" -1.04  
 Kcal/mol  
 269 "ATOM 3230 N VAL E 524" 3.489 "ATOM 718 H ASN E 360" "ATOM 710 N ASN E 360" -1.04  
 Kcal/mol  
 270 "ATOM 2342 N GLU E 465" 3.677 "ATOM 1686 H LYS E 424" "ATOM 1677 N LYS E 424" -1.04  
 Kcal/mol  
 271 "ATOM 1718 N PRO E 426" 3.691 "ATOM 1517 H GLY E 413" "ATOM 1513 N GLY E 413" -1.03  
 Kcal/mol  
 272 "ATOM 1072 N THR E 385" 3.695 "ATOM 1135 H ASN E 388" "ATOM 1127 N ASN E 388" -1.03  
 Kcal/mol  
 273 "ATOM 1702 O LEU E 425" 2.045 "ATOM 1517 H GLY E 413" "ATOM 1513 N GLY E 413" -1.03  
 Kcal/mol  
 274 "ATOM 1612 O ASP E 420" 1.918 "ATOM 2275 H LEU E 461" "ATOM 2267 N LEU E 461" -1.02  
 Kcal/mol  
 275 "ATOM 1732 N ASP E 427" 3.212 "ATOM 1767 H PHE E 429" "ATOM 1756 N PHE E 429" -1.02  
 Kcal/mol  
 276 "ATOM 2345 O GLU E 465" 3.036 "ATOM 1686 H LYS E 424" "ATOM 1677 N LYS E 424" -1.02  
 Kcal/mol  
 277 "ATOM 2752 N GLN E 493" 3.305 "ATOM 2165 H LEU E 455" "ATOM 2157 N LEU E 455" -1.02  
 Kcal/mol  
 278 "ATOM 1277 N ASP E 398" 3.570 "ATOM 643 H LYS E 356" "ATOM 634 N LYS E 356" -1.02  
 Kcal/mol  
 279 "ATOM 2522 O GLY E 476" 1.984 "ATOM 2544 H THR E 478" "ATOM 2537 N THR E 478" -1.01  
 Kcal/mol

280 "ATOM 2910 N GLY E 504" 3.485 "ATOM 2947 H GLN E 506" "ATOM 2938 N GLN E 506" -1.01  
Kcal/mol  
281 "ATOM 364 N GLY E 339" 3.364 "ATOM 393 H VAL E 341" "ATOM 386 N VAL E 341" -1.01  
Kcal/mol  
282 "ATOM 1979 NZ LYS E 444" 4.154 "ATOM 483 2HH2 ARG E 346" "ATOM 470 NH2 ARG E 346" -1.01  
Kcal/mol  
TOTAL ENERGY = -300559.87483 EV  
ELECTRONIC ENERGY = -35885471.06237 EV  
CORE-CORE REPULSION = 35585205.01181 EV  
DIELECTRIC ENERGY = -35.52439 EV  
DIPOLE = 222.89026 DEBYE  
NO. OF FILLED LEVELS = 4615  
MOLECULAR WEIGHT = 23751.5094  
COSMO AREA = 7467.71 SQUARE ANGSTROMS  
COSMO VOLUME = 26839.16 CUBIC ANGSTROMS

#### MOLECULAR DIMENSIONS (Angstroms)

Atom	Atom	Distance
H 2562	H 1105	60.81960
H 2871	H 3180	51.30024
H 52	H 284	33.26231

SCF CALCULATIONS = 1  
WALL-CLOCK TIME = 7 MINUTES AND 18.895 SECONDS  
COMPUTATION TIME = 10 MINUTES AND 25.195 SECONDS

**B – Ligações de H intramoleculares envolvendo o resíduo Gln493 do RBD.**

Meio dielétrico	Otimização	Doador	R <sub>(D-H)</sub>	Hidrogênio	Aceptor	Energia (kcal/mol)
			Å			
$\epsilon = 78,4$	Parcial	O Gln E 493	1,979	H Tyr E 453	N Tyr E 453	-2,26
		N Gln E 493	2,367	1HE2 Gln E 493	NE2 Gln E 493	-2,07
	Total	O Gln E 493	1,956	H Tyr E 453	N Tyr E 453	-1,75
		O Phe E 490	1,956	2HE2 Gln E 493	NE2 Gln E 493	-1,14
$\epsilon = 4,0$	Total	N Gln E 493	2,350	1HE2 Gln E 493	NE2 Gln E 493	-2,20
		O Gln E 493	1,941	H Tyr E 453	N Tyr E 453	-1,83
		O Phe E 490	1,952	2HE2 Gln E 493	NE2 Gln E 493	-1,12
		O Tyr E 453	1,843	H Gln E 493	N Gln E 493	-1,98

## C - Ligações de hidrogênio intermoleculares (< -1,0 kcal/mol; ε = 4,0)

- Complexos não modificados e totalmente otimizados:

	<b>Doador</b>	<b>R<sub>(D-H)</sub> (Å)</b>	<b>Hidrogênio</b>	<b>Aceptor</b>	<b>Energia (kcal/mol)</b>
RBD-SAP1	OH Tyr E 449	1,300	HH Tyr E 449	OD2 Asp A 38	-4,85
	OH Tyr A 41	1,688	HG1 Thr E 500	OG1 Thr E 500	-4,87
	OD1 Asn E 501	1,695	HH Tyr A 41	OH Tyr A 41	-4,30
	OD2 Asp A 30	1,410	2HZ Lys E 417	NZ Lys E 417	-3,56
	O His A 34	1,682	HH Tyr E 453	OH Tyr E 453	-4,33
	OE1 Glu E 484	1,630	1HZ Lys A 31	NZ Lys A 31	-3,14
	OE1 Glu A 35	1,838	H Ser E 494	N Ser E 494	-2,64
	OD1 Asn E 487	1,681	3H Thr A 27	N Thr A 27	-2,84
	NE2 Gln A 42	2,202	1HE2 Gln E 498	NE2 Gln E 498	-2,84
	OH Tyr E 449	1,997	2HE2 Gln A 42	NE2 Gln A 42	-1,77
	OD1 Asn E 487	2,306	H Phe A 28	N Phe A 28	-1,60
	NE2 His A 34	2,174	1HH1 Arg E 403	NH1 Arg E 403	-2,18
	NE2 Gln E 493	3,115	3HZ Lys A 31	NZ Lys A 31	-1,34
	N Phe A 28	3,265	H Asn E 487	N Asn E 487	-1,07
	OD1 Asp A 38	1,821	2HE2 Gln E 498	NE2 Gln E 498	-2,59
	OE1 Gln A 42	1,864	H Gly E 446	N Gly E 446	-2,76
	N Gly E 447	2,257	1HE2 Gln A 42	NE2 Gln A 42	-2,52
	OE1 Gln E 493	1,947	3HZ Lys A 31	NZ Lys A 31	-1,22

	<b>Doador</b>	<b>R<sub>(D-H)</sub> (Å)</b>	<b>Hidrogênio</b>	<b>Aceptor</b>	<b>Energia (kcal/mol)</b>
RBD-SAP2	OD2 Asp A 38	1,498	HH Tyr E 449	OH Tyr E 449	-4,98
	OH Tyr A 41	1,700	HG1 Thr E 500	OG1 Thr E 500	-4,65
	OH Tyr E 449	2,044	2HE2 Gln A 42	NE2 Gln A 42	-1,18
	OD1 Asp A 38	2,124	1HE2 Gln E 498	NE2 Gln E 498	-2,13
	OH Tyr E 505	2,313	H Asp A 38	N Asp A 38	-1,99
	N Asn E 501	2,416	HH Tyr A 41	OH Tyr A 41	-1,81
	O411 Leu A 45	1,455	H Gly E 446	N Gly E 446	-3,43

RBD-SAP3	Doador	R <sub>(D-H)</sub> (Å)	Hidrogênio	Aceptor	Energia (kcal/mol)
	OH Tyr E 489	1,727	HH Tyr A 83	OH Tyr A 83	-4,34
	N Ala A 80	2,452	HH Tyr E 489	OH Tyr E 489	-1,95
RBD-SAP4	Doador	R <sub>(D-H)</sub> (Å)	Hidrogênio	Aceptor	Energia (kcal/mol)
	OD2 Asp A 355	1,695	HG1 Thr E 500	OG1 Thr E 500	-3,97
	O Lys A 353	2,272	H Gly E 502	N Gly E 502	-2,00
	OD1 Asn E 501	1,802	1HZ Lys A 353	NZ Lys A 353	-1,93
	NZ Lys A 353	3,395	H Gln E 498	N Gln E 498	-1,13
	OE1 Gln E 498	1,725	2HZ Lys A 353	NZ Lys A 353	-1,42
	ND2 Asn E 501	2,749	1HZ Lys A 353	NZ Lys A 353	-1,32
RBD-SAP5	Doador	R <sub>(D-H)</sub> (Å)	Hidrogênio	Aceptor	Energia (kcal/mol)
	OD2 Asp A 30	1,428	2HZ Lys E 417	NZ Lys E 417	-3,35
	OE1 Glu E 484	1,719	1HZ Lys A 31	NZ Lys A 31	-2,51
	NE2 Gln A 24	2,313	H Ser E 477	N Ser E 477	-2,59
	OE1 Gln A 24	1,897	1HD2 Asn E 487	ND2 Asn E 487	-2,11
	O Ala E 475	1,758	3H Gln A 24	N Gln A 24	-2,06
	NE2 Gln E 493	3,037	3HZ Lys A 31	NZ Lys A 31	-1,34
	O His A 34	1,714	HH Tyr E 453	OH Tyr E 453	-3,99
	OE1 Glu A 35	1,803	H Ser E 494	N Ser E 494	-2,68
	OG Ser E 477	2,139	2HE2 Gln A 24	NE2 Gln A 24	-1,77
RBD-SAP6	Doador	R <sub>(D-H)</sub> (Å)	Hidrogênio	Aceptor	Energia (kcal/mol)
	OH Tyr A 41	1,690	HG1 Thr E 500	OG1 Thr E 500	-4,85
	OD1 Asn E 501	1,755	HH Tyr A 41	OH Tyr A 41	-2,46
	OD1 Asp A 38	1,799	2HE2 Gln E 498	NE2 Gln E 498	-2,30
	OH Tyr E 449	2,125	2HE2 Gln A 42	NE2 Gln A 42	-1,42
	NE2 Gln A 42	2,248	1HE2 Gln E 498	NE2 Gln E 498	-2,65
	OD2 Asp A 38	1,447	HH Tyr E 449	OH Tyr E 449	-4,86
	N Gly E 447	2,135	1HE2 Gln A 42	NE2 Gln A 42	-2,97
	OE1 Gln A 42	1,925	H Gly E 446	N Gly E 446	-2,49

OE1 Glu A 37	1,794	1HH1 Arg E 403	NH1 Arg E 403	-2,01
N Asn E 501	2,432	HH Tyr A 41	OH Tyr A 41	-1,19
OE2 Glu A 37	1,955	2HH1 Arg E 403	NH1 Arg E 403	-1,16

## D - Demais interações intermoleculares ( $\epsilon = 4,0$ )

### - Complexos não modificados e totalmente otimizados:

	<b>Resíduo RBD</b>	<b>Resíduo do peptídeo</b>	<b>A-E (Å)</b>	<b>Tipo de interação</b>
RBD-SAP1	Lys E 417	Asp A 30	2,84	ponte salina
	Phe E 456		3,77	
	Gln E 493	Glu A 35	3,61	hidrofóbica
	Tyr E 505	Glu A 37	3,49	
	Tyr E 453	His A 34	3,86	
	Leu E 455		3,32	
	Tyr E 489	Lys A 31	3,32	
	Phe E 486	Phe A 28	3,43	
	Phe E 456		3,11	
	Tyr E 473	Thr A 27	3,88	
	Tyr E 489		3,59	
	Gln E 498	Tyr A 41	3,75	
RBD-SAP2				
		Asp A 38	3,92	hidrofóbica
	Tyr E 505	Glu A 37	3,46	
			3,53	
	Val E 445	Leu A 45	3,89	
RBD-SAP3		Leu A 45	3,97	hidrofóbica
	Gln E 498	Tyr A 41	3,43	
			3,53	

	<b>Resíduo RBD</b>	<b>Resíduo do peptídeo</b>	<b>A-E (Å)</b>	<b>Tipo de interação</b>
RBD-SAP4	Tyr E 505	Lys A 353	3,63 3,36	hidrofóbica
RBD-SAP5	Lys E 417	Asp A 30	2,91	ponte salina
	Phe E 456		3,69	
	Phe E 486	Ala A 25	3,66	
	Gln E 493	Glu A 35	3,72	hidrofóbica
	Leu E 455	His A 34	3,48	
	Tyr E 489	Lys A 31	3,39	
		Phe A 28	3,79	
	Phe E 456		3,38	
	Tyr E 473	Thr A 27	3,71	
	Tyr E 489		3,58	
RBD-SAP6	Arg E 403	Glu A 37	3,70	ponte salina
	Gln E 498	Tyr A 41	3,90	hidrofóbica

**E – Ilustração gráfica da interface de interação entre o RBD da *Spike* do SARS-CoV-2 e o SAP6**

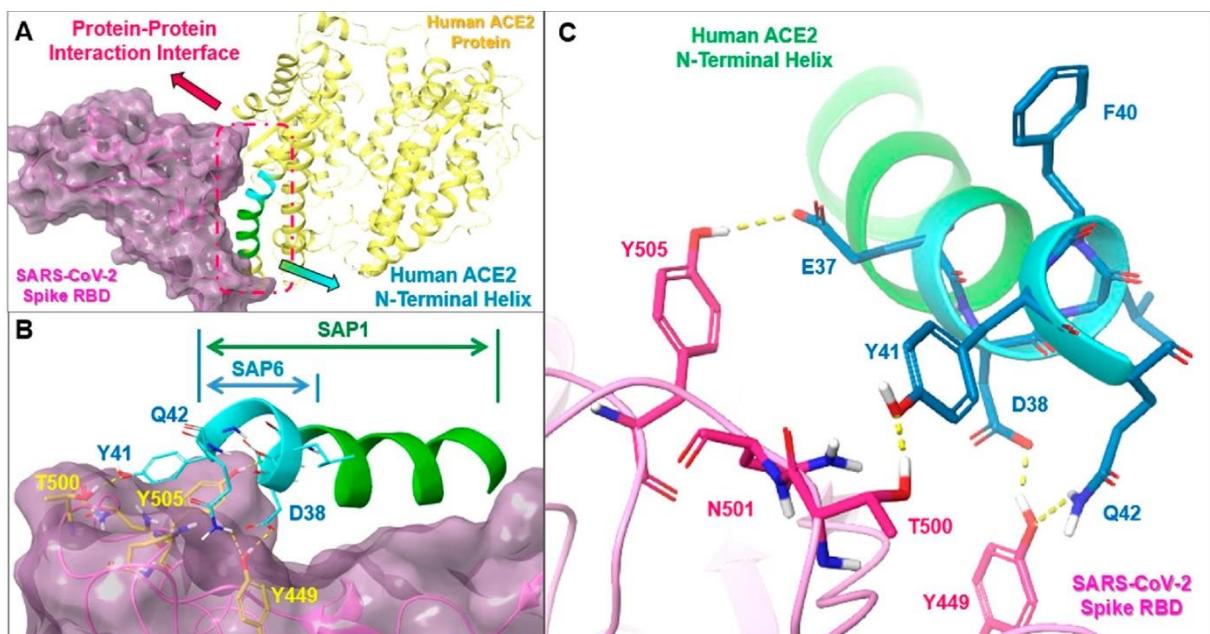
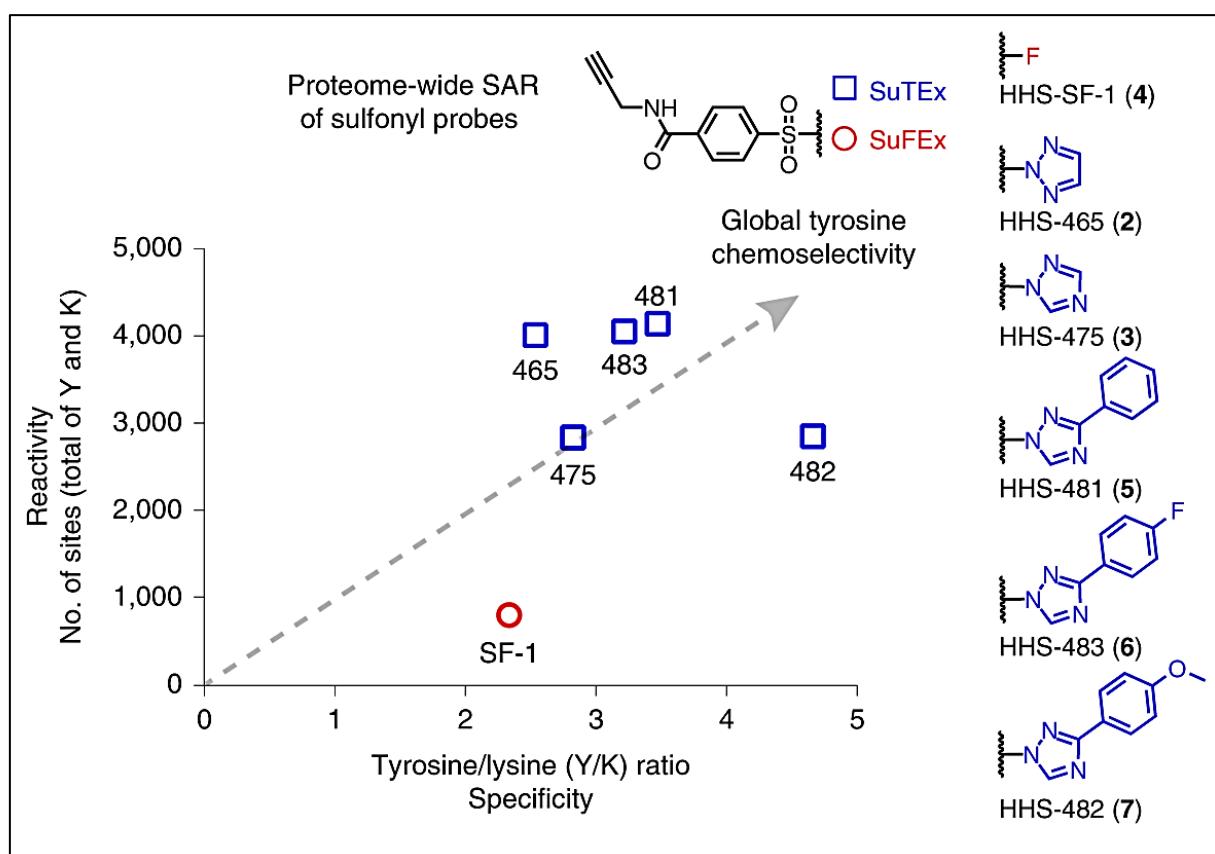


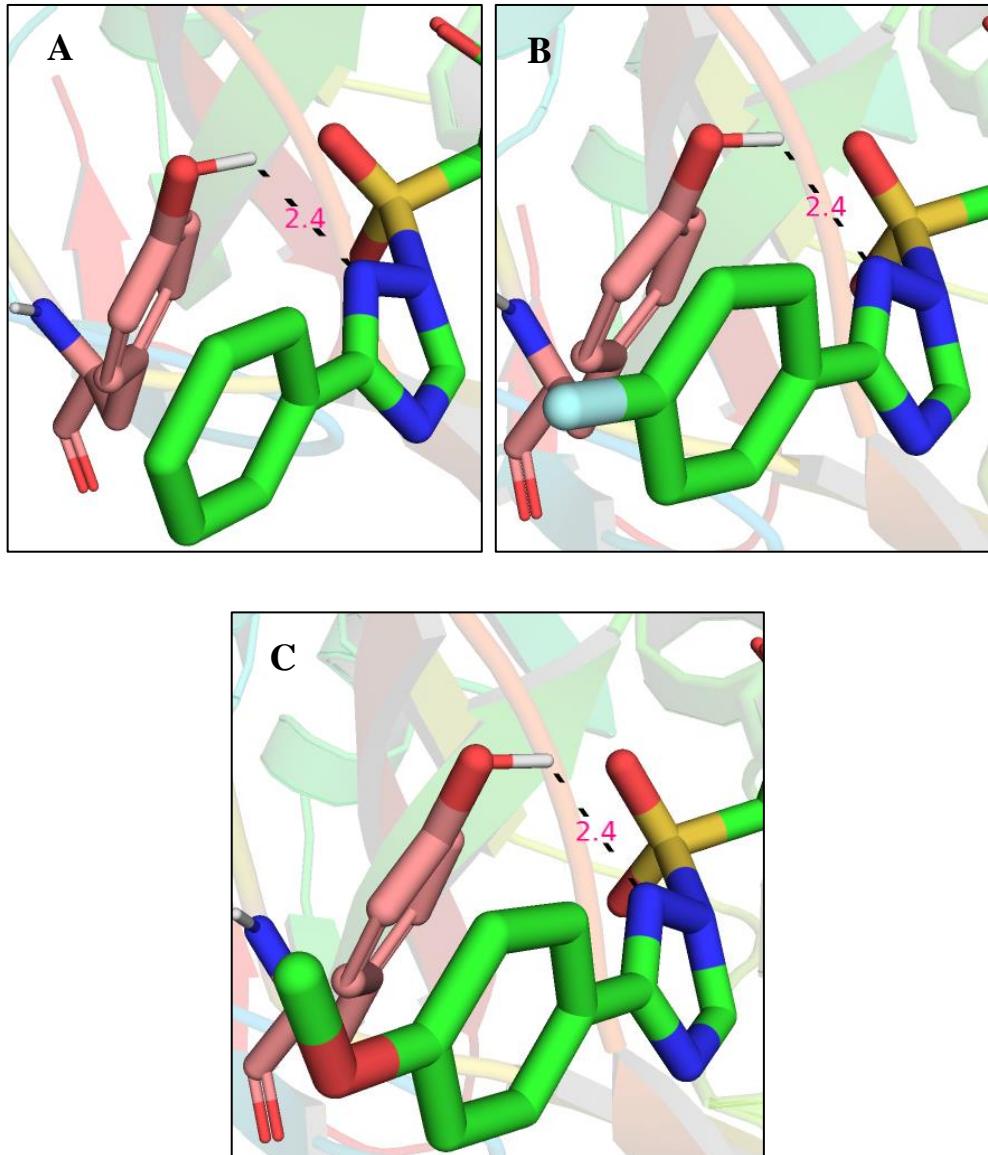
Ilustração gráfica da interface de interação a *Spike* do SARS-CoV-2 e o SAP6. **A)** Visão geral do RBD da *Spike* do SARS-CoV-2 e do modo de interação com a hACE2. A hélice N-terminal do hACE2 está localizada na interface central. **B)** Localização relativa de SAP6 (azul claro) e SAP1 (verde e azul claro). **C)** Rede de interação de *H-bond* entre SAP6 e o RBD da *Spike* do SARS-CoV-2. A Tyr41, Gln42, Asp38 e Glu37 do peptídeo SAP6 estão envolvidos nas interações da ligação H com Thr500, Tyr449 e Tyr505 do RBD da *Spike* do SARS-CoV-2. Estrutura cristalográfica correspondente: Código PDB: 6M0J. <http://www.rcsb.org/structure/6M0J> (LAN *et al.*, 2020).

## F- Reatividade x Especificidade das sondas SuTEx e SuFEx



Reatividade global [número total de locais de tirosina (Y) e lisina (K)] e especificidade (razão Y/K) de locais marcados com sonda de experimentos proteômicos por Cromatografia Líquida Acoplada à Espectrometria de Massas (LC-MS) (HAHM *et al.*, 2020).

**G - Distância entre o nitrogênio do anel 1,2,4-triazol substituído do peptídeo SAP1 e o hidrogênio da hidroxila da Tyr449 do RBD, nos complexos derivados**



Distância entre o nitrogênio do anel trizólico substituído e o hidrogênio da hidroxila da Tyr449 (C em rosa), nos **complexos derivados do SAP1 modificado com grupo sulfonil-triazol**. A) RBD-SAP1-Z3 com 2,4 Å; B) RBD-SAP1-Z4 com 2,4 Å; C) RBD-SAP1-Z5 com 2,4 Å. Obs: apenas os hidrogênios polares são mostrados; e, os mesmos valores foram encontrados nos complexos modificados contendo o grupo fosfonil-triazol.