

Universidade Federal do Rio Grande do Sul
Instituto de Química
Bacharel em Química

Alexandre Cardoso Capsi Pires

Análise da viabilidade de peptídeos derivados da hidrólise de caseinato ovino como
inibidores da enzima DPP-IV via docagem molecular

Porto Alegre
2023

Alexandre Cardoso Capsi Pires

**Análise da viabilidade de peptídeos derivados da hidrólise de caseinato ovino
como inibidores da enzima DPP-IV via docagem molecular**

Trabalho de Conclusão de Curso
apresentado como requisito parcial à
obtenção do título de bacharel em Química do
Instituto de Química da Universidade Federal
do Rio Grande do Sul.

Orientador: Paulo Augusto Netz

Co-orientador: Rodrigo dos Santos Fuscaldó

Porto Alegre

2023

FICHA CATALOGRÁFICA

CIP - Catalogação na Publicação

Pires, Alexandre Cardoso Capsi
Análise da viabilidade de peptídeos derivados da
hidrólise de caseinato ovino como inibidores da enzima
DPP-IV via docagem molecular / Alexandre Cardoso Capsi
Pires. -- 2023.
29 f.
Orientador: Paulo Augusto Netz.

Coorientador: Rodrigo dos Santos Fuscaldo.

Trabalho de conclusão de curso (Graduação) --
Universidade Federal do Rio Grande do Sul, Instituto
de Química, Bacharelado em Química, Porto Alegre,
BR-RS, 2023.

1. Docagem Molecular. 2. Drug Design. 3. Química
Computacional. 4. DPP-IV. I. Netz, Paulo Augusto,
orient. II. Fuscaldo, Rodrigo dos Santos, coorient.
III. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

FOLHA DE APROVAÇÃO

Alexandre Cardoso Capsi Pires

**Análise da viabilidade de peptídeos derivados da hidrólise de caseinato ovino
como inibidores da enzima DPP-IV via docagem molecular**

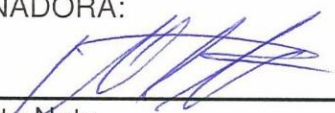
Trabalho de Conclusão de Curso
apresentado como requisito parcial à
obtenção do título de bacharel em Química do
Instituto de Química da Universidade Federal
do Rio Grande do Sul.

Orientador: Paulo Augusto Netz

Co-orientador: Rodrigo Fuscaldo

Aprovada em: Porto Alegre, 5 de abril de 2023.

BANCA EXAMINADORA:



Dr. Paulo Augusto Netz
Instituto de Química

Dr. Hubert Karl Stassen
Instituto de Química

Dra. Daniele Raupp
Instituto de Química

DEDICATÓRIA

Dedico este trabalho de conclusão de curso
à minha avó, Raimunda Gomes Cardoso,
que sonhava em ver um neto formado.

AGRADECIMENTOS

Gostaria de agradecer a todos que fizeram parte desse percurso que foi minha graduação em Química, em especial a meus pais Leandro e Nádia Pires por possibilitar meus estudos durante minha vida inteira e oferecer apoio incondicional e viabilizar minhas escolhas profissionais e que eu me mudasse para estudar em um estado diferente, minha avó Raimunda Gomes, que me deu amor e apoio incondicionais por 23 anos da minha vida e foi um dos motivos para me fazer buscar uma graduação, minhas tias Eliana Cardoso e Simone Capsi, meu professor de química do ensino médio Welff Junior, que inspirou meu interesse em química, meus amigos Guilherme Machado, Daniel Stamm, Flávia Bernardi, Brenda Rolt e Joana Zanchetti pelo apoio e momentos vividos ao meu lado nesses 7 anos de graduação, minha namorada Francielli Menezes, pelo amor, apoio e ajuda durante meus últimos semestres de graduação, meu co-orientador Rodrigo Fuscaldo, com quem tive interações breves, mas muito construtivas, pelas explicações e disponibilidade que me ajudaram muito a ter uma base para este trabalho e meu orientador Paulo Netz por seu tempo e disponibilidade e por ter me ensinado sobre os assuntos relevantes para este trabalho, assim como a ajuda em lidar com o script e os diversos softwares que foram utilizados.

RESUMO

Foram obtidos 57 peptídeos pelo ICTA através da hidrólise de caseinato ovino e foram detectadas atividades antidiabéticas devido à inibição da enzima DPP-IV em 8 desses peptídeos, sendo eles: DIPNPIGSENSGKITMPLW, EPVLGPVRGPFILV, IASAEPTVHSTPTTEAVVN, KNMAIHPR, KPWTQPKTNAIPYVRYL, LYQEPVLGPVRGPFILV, RFVVAPFPEVFR e RHPHPHLSFM. Esses peptídeos foram analisados via docagem molecular com o objetivo de analisar a viabilidade como inibidores da DPP-IV para que os mais promissores sejam sintetizados pelo ICTA. Dois requisitos foram determinados como necessários para que exista a propriedade inibitória, que são a afinidade (em kcal/mol) e a interação com os resíduos-chave (tríade catalítica, S1 e S2). Foram utilizados as ferramentas e programas: Protein Data Bank, UniProt, AlphaFold, CHARMM-GUI, AutoDock 4.2, AutoDock Tools, AutoDock Vina, VinaSplit, OpenBabel, RasMol e VMD. Para fazer a docagem foram obtidos os modelos 3D das estruturas da proteína e dos oito ligantes, esses modelos foram checados/corrigidos, foi feita a preparação dos ligantes e a montagem da grade de potencial, então foi feita a docagem molecular em triplicata através do AutoDock Vina. Foram feitas 24 simulações no total, que foram analisadas nos quesitos afinidade (kcal/mol) e os arquivos resultantes foram convertidos para permitir a sua análise de interação com os resíduos-chave pelo RasMol. As conformações que atingiram a ambos os requisitos para apresentar atividade inibitória são DIPNPIGSENSGKITMPLW, EPVLGPVRGPFILV, KNMAIHPR, RFVVAPFPEVFR e RHPHPHLSFM.

Palavras-chave: DPP-IV; docagem molecular; drug design; química computacional.

ABSTRACT

57 peptides were obtained by ICTA by the hydrolysis of ovine caseinate and antidiabetic activity was detected in eight of those peptides, namely: DIPNPIGSENSGKITMPLW, EPVLGPVRGPFILV, IASAEPTVHSTPTTEAVVN, KNMAIHPR, KPWTQPKTNAIPYVRYL, LYQEPVLGPVRGPFILV, RFVVAPFPEVFR and RHPHPHLSFM. They were analysed via molecular docking to evaluate their viability as DPP-IV inhibitors so that the most promising ones would be synthesized by ICTA. Two requirements were determined as necessary for the inhibitory property existence, they are the affinity (kcal/mol) and the interaction with the key-residues (catalytic triad, S1 and S2). The tools and programs used were: Protein Data Bank, UniProt, AlphaFold, CHARMM-GUI, AutoDock 4.2, AutoDock Tools, AutoDock Vina, VinaSplit, OpenBabel, RasMol and VMD. In order to carry out the molecular docking, the 3D models of the structures of the protein and the eight ligands were obtained, they were checked/corrected, the ligand preparation was made and the grid box was constructed. Then the molecular docking was carried out using AutoDock Vina in triplicate. 24 simulations were made in total, the affinity analysis was made, then the resulting files were converted to allow RasMol to analyse the interaction of the ligand with the key-residues. The conformations that met both of the requirements to allow for inhibitory activities are DIPNPIGSENSGKITMPLW, EPVLGPVRGPFILV, KNMAIHPR, RFVVAPFPEVFR, and RHPHPHLSFM.

Keywords: DPP-IV; molecular docking; drug design; computational chemistry.

LISTA DE ILUSTRAÇÕES

Figura 1: Modelo 3D da DPP-IV, tríade catalítica, socket S1 e socket S2.....	16
Figura 2: Modelos 3D das estruturas dos peptídeos.....	18
Figura 3: Modelos 3D das estruturas secundárias dos peptídeos.....	18
Figura 4: Script utilizado para fazer a docagem utilizando AutoDock Vina.....	20
Figura 5: Exemplo de análise utilizando RasMol.....	22
Figura 6: Modelos 3D das estruturas das conformações mais promissoras.....	24

LISTA DE ABREVIATURAS E SIGLAS

DPP-IV: Dipeptil Peptidase IV

.pdb: arquivo formato protein data bank

.pdbqt: arquivo formato protein data bank com carga e tipo atômico definidos

.txt: arquivo de texto

S1: socket 1

S2: socket 2

Simbologia de uma letra para resíduos de aminoácidos:

A: Alanina

D: Ácido Aspártico

E: Ácido Glutâmico

F: Fenilalanina

G: Glicina

H: Histidina

I: Isoleucina

K: Lisina

L: Leucina

M: Metionina

N: Asparagina

P: Prolina

Q: Glutamina

R: Arginina

S: Serina

T: Treonina

V: Valina

W: Triptofano

Y: Tirosina

SUMÁRIO

1 INTRODUÇÃO	10
2 REVISÃO BIBLIOGRÁFICA	12
3 OBJETIVOS	15
3.1 OBJETIVOS ESPECÍFICOS.....	15
4 DESENVOLVIMENTO	16
4.1 OBTENÇÃO DOS MODELOS 3D.....	16
4.2 PREPARAÇÃO DOS LIGANTES	18
4.3 DOCAGEM MOLECULAR.....	19
5 ANÁLISE DOS RESULTADOS E DISCUSSÃO.....	21
6 CONCLUSÃO	25
REFERÊNCIAS.....	26

1 INTRODUÇÃO

Com o intuito de desenvolver um novo fármaco utilizando peptídeos, foram analisados no Instituto de Ciência e Tecnologia de Alimentos (ICTA), na UFRGS, 57 peptídeos obtidos através da hidrólise de caseinato ovino e foram detectadas atividades antidiabéticas através da inibição da enzima DPP-IV. Desses 57 peptídeos foram selecionados os 8 mais prováveis de possuírem a propriedade inibidora da DPP-IV para serem analisados via docagem molecular, a fim de identificar os peptídeos bioativos que possuem a propriedade inibidora de DPP-IV mais acentuada para que o laboratório colaborador possa sintetizar os peptídeos mais promissores.

A Dipeptil Peptidase 4 (DPP-IV) é uma enzima responsável pela degradação da incretina, uma substância que estimula a liberação de insulina, por isso, inibidores são de grande interesse e existem numerosas estruturas conhecidas com potencial inibitório aprovadas para o tratamento de diabetes tipo 2.

Os fármacos presentes no mercado para combater a diabetes como insulina, hipoglicemiantes e secretagogos são efetivos, mas possuem alguns efeitos colaterais e não são acessíveis financeiramente. Uma possível solução para isso é a fabricação de fármacos utilizando peptídeos.

Existem muitos benefícios em utilizar peptídeos como fármacos por possuírem baixa toxicidade, além de serem encontrados na natureza, o que poupa etapas no desenho do medicamento.

A docagem molecular já é bastante explorada para drug design, inclusive para a identificação de inibidores de DPP-IV, por ser um método eficiente e barato que simula conformações e utiliza funções de escore para identificar as melhores de acordo com os critérios escolhidos.

O modelo 3D da enzima foi obtido através do Protein Data Bank e os modelos 3D dos peptídeos foram obtidos utilizando o UniProt para descobrir proteínas que possuam a sequência de resíduos que precisamos, os arquivos foram obtidos pelo AlphaFold e os peptídeos foram extraídos das estruturas fornecidas. Os peptídeos foram então otimizados pelo programa CHARMM-GUI para obtermos o modelo 3D necessário para a docagem.

Foram utilizados os programas AutoDock 4.2 e AutoDock Tools para fazer a preparação dos ligantes, o programa AutoDock Vina para fazer a docagem e o programa Vina Split para separar os resultados da docagem em arquivos individuais.

Para analisar a afinidade das conformações, foram lidos os arquivos log de texto que possuem uma tabela com as afinidades das melhores conformações de cada simulação.

Para analisar a interação das conformações com os resíduos-chave, foram utilizados o programa OpenBabel para converter arquivos .pdbqt em .pdb e o programa RasMol para analisar os sítios de ligação das conformações promissoras.

Também foi utilizado o programa Visual Molecular Dynamics (VMD) para gerar as imagens das estruturas 3D.

2 REVISÃO BIBLIOGRÁFICA

Um estudo feito utilizando a FINDRISC (Finnish Diabetes Risk Score) em 2018 resultou em cerca de 22,1% dos 17.580 brasileiros testados estando sob risco de desenvolver a diabetes tipo 2 nos próximos 10 anos.¹ Segundo dados da WHO (World Health Organization), a diabetes tipo 2 representa cerca de 90% dos casos de diabetes no mundo.²

Já são conhecidos vários inibidores de DPP-IV que têm seus efeitos positivos comprovados quanto à degradação da incretina e consequente regulação da glicose no sangue e são aprovados para o tratamento de diabetes tipo 2.³

Atualmente existem no mercado a insulina, hipoglicemiantes e secretagogos para combater a diabetes e os três são efetivos para seu propósito, mas podem possuir alguns efeitos colaterais como retenção de fluidos, problemas digestivos, infecções do trato urinário e problemas cardíacos, portanto, uma alternativa mais acessível e com menores efeitos colaterais seria muito benéfica para a população.⁴

Peptídeos são sequências curtas de aminoácidos contendo de 2 a 50 aminoácidos, enquanto uma proteína é composta de mais de 50 aminoácidos e pode ser dividida em peptídeos. A utilização de peptídeos como fármacos é muito vantajosa em alguns sentidos, como a baixa toxicidade, grande seletividade e baixa acumulação em tecidos, e por serem encontrados na natureza, várias etapas no desenho racional de fármacos são poupadas.⁵

A Dipeptil Peptidase 4 (DPP-IV) é uma enzima responsável pela degradação da incretina (que estimula a liberação de insulina), da alfa-amilase e da lipase pancreática e existem numerosas estruturas com potencial inibitório de sua atividade, por isso, inibidores são de grande interesse na fabricação de medicamentos.⁶

A docagem molecular é um método bastante explorado para drug design e para a identificação dos inibidores da DPP-IV, que simula as interações moleculares para prever as interações ligante-receptor a nível molecular, além da orientação e afinidade dos receptores e ligantes, o que viabiliza seu uso para selecionar potenciais fármacos de maneira mais eficiente e barata. É muito eficaz para múltiplos arranjos intermoleculares e substratos, além de explorar efetivamente a hipersuperfície de potencial com o uso de algoritmos.

A docagem simula a melhor conformação de acordo com os parâmetros escolhidos e a partir dela diz a afinidade, orientação e modo de interação entre ligante

e receptor, que pode ocorrer pelo modelo chave-fechadura ou pelo modelo de encaixe induzido. O modelo chave-fechadura ocorre quando o receptor é rígido e o ligante se adequa a ele para achar a conformação que faz o encaixe correto, o que demonstra a importância da compatibilidade geométrica. O modelo de encaixe induzido ocorre quando tanto o receptor quanto os ligantes são flexíveis o suficiente para mudar suas conformações para fazer o encaixe correto, procurando a estrutura mais estável e com energia livre mínima com o uso de um algoritmo.⁷

São utilizadas funções de escore para avaliar qual a melhor opção de acordo com os parâmetros escolhidos e listar por afinidade em kcal/mol, para então serem analisados para conferir se estão interagindo suficientemente com os resíduos-chave, sendo esses a tríade catalítica composta por serina (na posição 630), ácido aspártico (na posição 708) e um tautômero prototrópico da histidina com o nitrogênio protonado na posição delta (na posição 740),⁸ o socket 1 (S1), composto por tirosinas (nas posições 631, 662 e 666), valinas (nas posições 656 e 711) e triptofano (na posição 659), e o socket 2 (S2), composto por argininas (nas posições 125 e 358), glutaminas (nas posições 203 e 206), fenilalanina (na posição 357) e serina (na posição 209). Essas interações são necessárias para inibir a enzima.^{2,4} O objetivo é encontrar conformações com uma afinidade de -7,5 kcal/mol ou menos que atendam aos requisitos que dão a propriedade inibidora da DPP-IV.

Foram utilizados os programas Protein Data Bank, UniProt, AlphaFold, CHARMM-GUI para a preparação dos modelos 3D dos ligantes.

O Protein Data Bank é um banco de dados 3D estruturais de macromoléculas biológicas determinados experimentalmente, que atualmente é o banco de dados mais relevante para esse tipo de dados.⁹

O UniProt é um banco de dados de sequências proteicas que são revisados por especialistas para determinar o nível de evidência experimental ou teórica para cada proteína e que permite a busca de sequências de peptídeos para encontrar uma proteína que contenha essa sequência.¹⁰

O AlphaFold Protein Structure Database é um banco de dados de predições de estruturas proteicas 3D feito com inteligência artificial, permitindo a busca de proteínas que não têm seus modelos disponibilizados em outros bancos de dados acessíveis.¹¹

O CHARMM-GUI é um programa acadêmico de pesquisa usado para mecânica e dinâmica de macromoléculas, além da manipulação de dados, permitindo a correção de arquivos .pdb criados ou não formatados corretamente.¹²

Foram utilizados os programas AutoDock, AutoDock Tools, AutoDock Vina e VinaSplit para a docagem molecular.

O AutoDock e AutoDock Tools são dois programas que devem ser usados em conjunto para fazer uma docagem molecular, o AutoDock faz uso de algoritmos de busca de conformações que podem ser genéticos ou lamarckianos, cálculos semiempíricos de energia livre e permite a docagem de receptores rígidos com ligantes rígidos ou flexíveis.¹³ O AutoDock Tools é utilizado para facilitar o processo de entrada de modelos 3D, a protonação, cálculo de cargas e a especificação das ligações rotacionais para o Auto Dock.

O AutoDock Vina é um programa de docagem molecular que utiliza algoritmos de busca de conformações genéticos, cálculos semiempíricos de energia livre e permite a docagem de receptores rígidos com ligantes rígidos ou flexíveis. Além disso, não possui a limitação de graus de liberdade que o AutoDock possui (o AutoDock é recomendado para graus de liberdade torcionais próximos de 10 e possui um limite de 32), assim permitindo a docagem de ligantes maiores e mais complexos, que é o caso desse experimento.¹⁴

O VinaSplit é um programa que separa o arquivo de saída do AutoDock Vina que contém as nove melhores conformações em nove arquivos com uma conformação cada.

Para a análise dos resultados, foram utilizados OpenBabel, RasMol e Visual Molecular Dynamics.

O OpenBabel é uma ferramenta de conversão de estruturas químicas em diferentes formatos utilizados para dados químicos que também permite a pesquisa ou análise de dados em formatos comumente utilizados na química computacional.¹⁵

O RasMol é um programa de visualização de estruturas 3D que permite a demonstração de diferentes maneiras de cadeias ou moléculas específicas para análise com uma interface simples.

O Visual Molecular Dynamics (VMD) é um programa de visualização e análise de estruturas 3D que possui variados tipos de renderização de imagem e métodos de coloração e também suporta uma alta resolução.

3 OBJETIVOS

Objetivo geral: Estudar mediante métodos computacionais as interações entre a proteína DPP-IV e oito peptídeos com possível ação inibitória, visando a produção de fármacos de combate à diabetes.

3.1 OBJETIVOS ESPECÍFICOS

Obter a estrutura 3D da proteína DPP-IV utilizando banco de dados e modelar a estrutura 3D secundária de oito peptídeos obtidos pela hidrólise de caseinato ovino utilizando bancos de dados e ferramentas de inteligência artificial.

Preparar receptor e ligantes para o estudo via docagem molecular, além da preparação do arquivo de configuração para a docagem.

Realizar a docagem em triplicata dos oito ligantes na proteína DPP-IV.

Analisar detalhadamente os resultados de afinidade e de interação com os resíduos-chave da proteína dos peptídeos mais promissores para a posterior síntese.

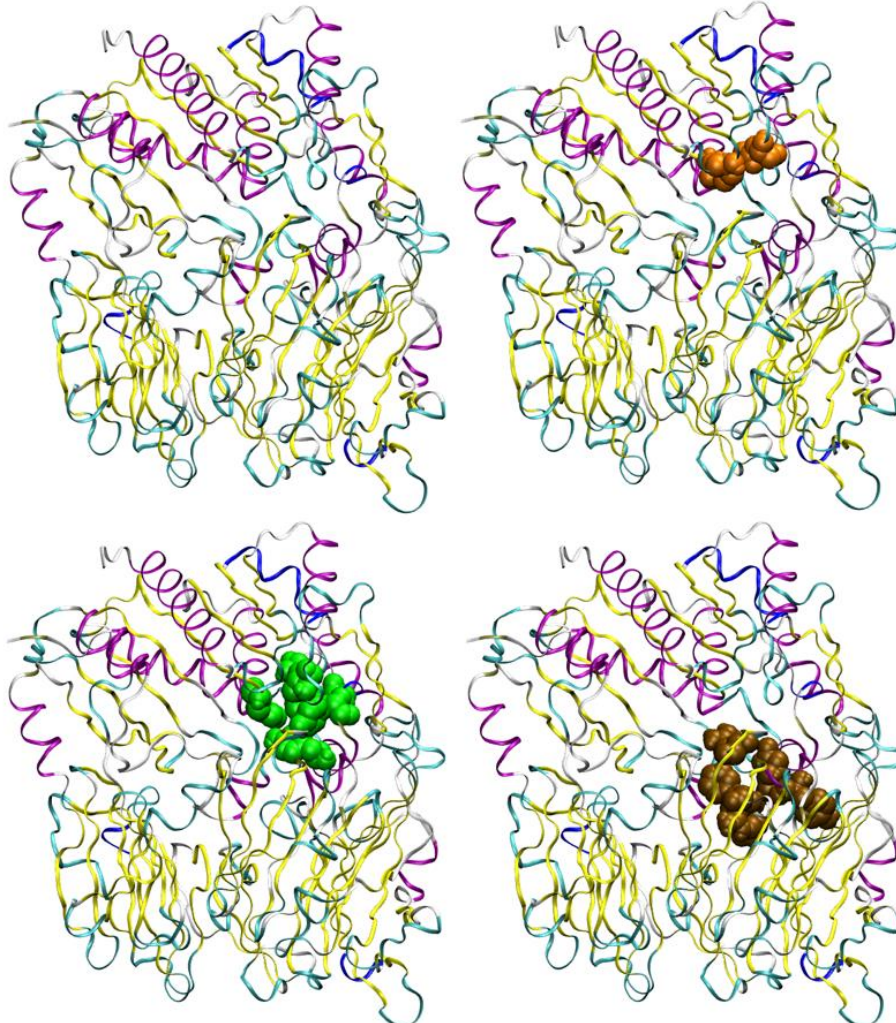
4 DESENVOLVIMENTO

4.1 OBTENÇÃO DOS MODELOS 3D

A proteína receptora é a DPP-IV e os oito peptídeos selecionados pelo ICTA foram os seguintes: DIPNPIGSENSGKITMPLW, EPVLGPVRGPFILV, IASAEPTVHSTPTTEAVVN, KNMAIHPR, KPWTQPKTNAIPYVRYL, LYQEPVLGPVRGPFILV, RFVVAPFPEVFR e RHPHPLSFM, todos são oligopeptídeos e estão representados pela simbologia de uma letra para resíduos de aminoácidos.

Primeiramente foram obtidos os modelos 3D dos receptores e ligantes. O modelo da enzima foi obtido no formato .pdb pela base de dados Protein Data Bank, com o código de identificação 4A5S, uma estrutura cristalina de DPP-IV humana em complexo com um inibidor. Na Figura 1 está o modelo 3D da DPP-IV, além da tríade catalítica, socket S1 e socket S2.

Figura 1: Da esquerda para a direita e de cima para baixo estão representados o modelo 3D da DPP-IV, a tríade catalítica, o socket S1 e o socket S2. Imagem gerada pelo VMD.



Os modelos dos peptídeos foram obtidos utilizando primeiramente a ferramenta de pesquisa UniProt para identificar quatro proteínas que possuem os oito peptídeos em sua estrutura, todas pertencentes a ovelhas, sendo elas a alfa-S1-caseína, de código de identificação P04653 e gene CSN1S1 contendo RFVVAPFPEVFR e DIPNPIGSENSGKITMPLW em sua estrutura (nas posições 37-48 e 196-214, respectivamente), a alfa-S2-caseína, de código de identificação P04654 e gene CSN1S2 contendo KNMAIHPR e KPWTQPKTNAIPYVRYL em sua estrutura (nas posições 40-47 e 207-223, respectivamente), a beta-caseína, de código de identificação P11839 e gene CSN2 contendo EPVLGPVVRGPFILV e LYQEPVLGPVVRGPFILV em sua estrutura (nas posições 205-222 e 208-222, respectivamente), e a kappa-caseína, de código de identificação P02669 e gene CSN3 contendo RHPHPHLSFM e IASAEPTVHSTPTTEAVVN em sua estrutura (nas posições 118-127 e 146-164, respectivamente). A alfa-S1-caseína, a beta-caseína e a kappa-caseína têm sua existência comprovada experimentalmente, enquanto a alfa-S2-caseína tem a existência de uma transcrição dessa proteína indicada pela expressão do gene.

Os modelos das quatro proteínas foram obtidos no formato .pdb através do AlphaFold, utilizando seus códigos de identificação.

Após isso foi feita a extração dos peptídeos do modelo da proteína, para isso os arquivos .pdb foram convertidos em arquivos .txt , que foram editados para manter apenas as moléculas nas posições de interesse, assim gerando oito arquivos de texto contendo os modelos 3D dos peptídeos, que foram então convertidos para o formato .pdb novamente.

Todos os modelos foram analisados utilizando o programa CHARMM-GUI, no modo PDB Reader & Manipulator, onde foi feito o upload dos arquivos .pdb com a opção Check/Correct PDB Format e não foi selecionada nenhuma manipulação, e então foram obtidos os 9 modelos 3D necessários (receptor e oito ligantes) para fazer a docagem molecular. Na Figura 2 estão representados os modelos 3D dos oito ligantes. Na Figura 3 estão representadas as estruturas secundárias dos modelos 3D dos oito ligantes.

Figura 2: Da esquerda para a direita e de cima para baixo estão representados os modelos 3D das estruturas dos oito ligantes: DIPNPIGSENSGKITMPLW, EPVLGPVRGPFILV, IASAPTVHSTPTTEAVVN, KNMAIHPR, KPWTQPKTNAIPYVRYL, LYQEPVLGPVRGPFILV, RFVVAPFPEVFR e RHPHPHLSFM. Imagem gerada pelo VMD.

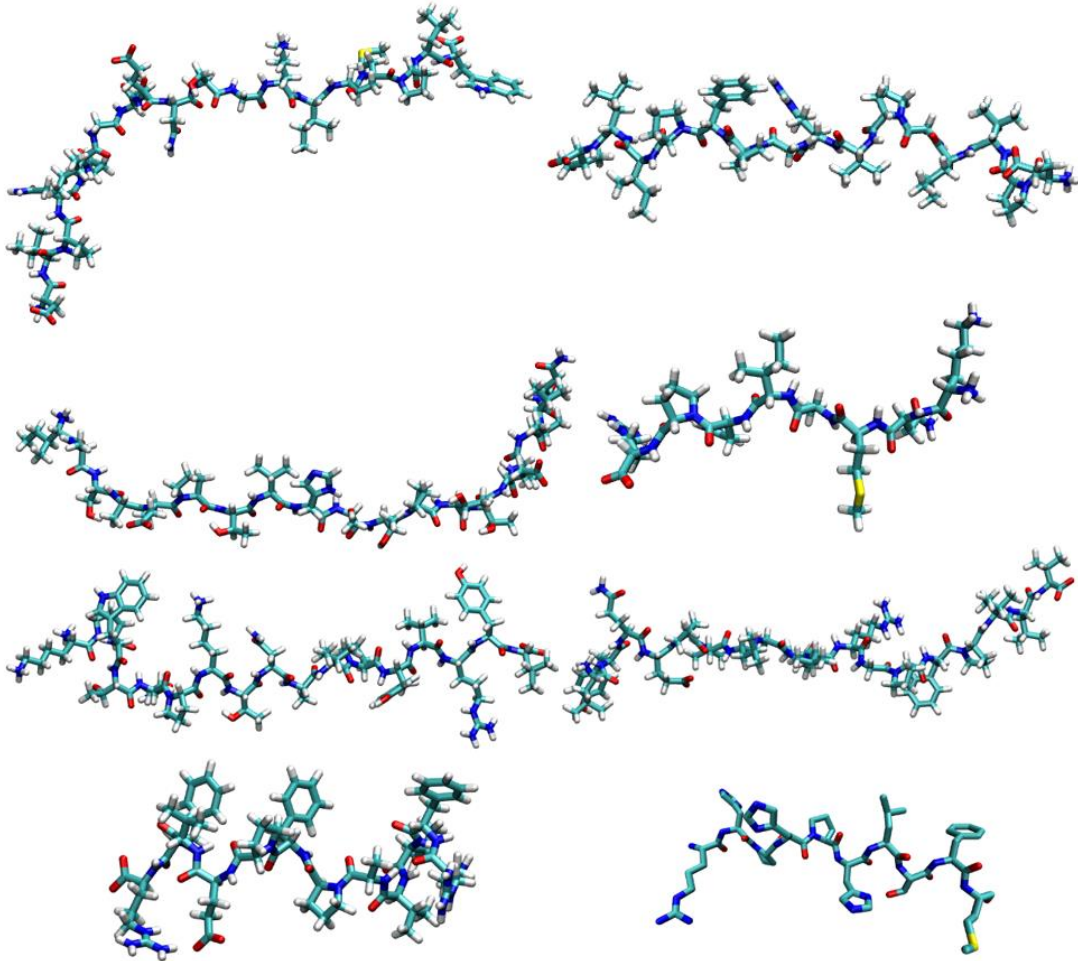
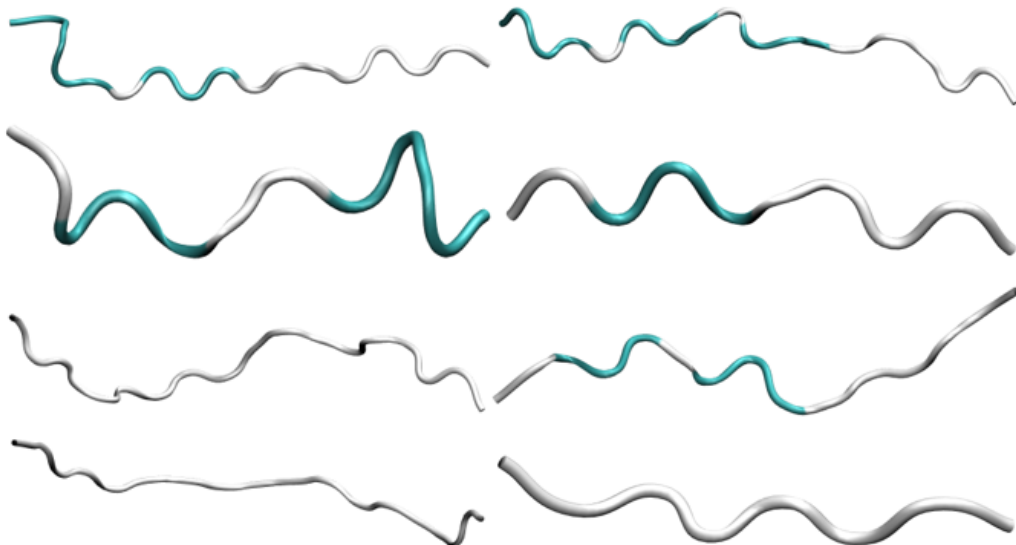


Figura 3: Da esquerda para a direita e de cima para baixo estão representados os modelos 3D, com as estruturas secundárias em azul, das estruturas dos oito ligantes: DIPNPIGSENSGKITMPLW, EPVLGPVRGPFILV, IASAPTVHSTPTTEAVVN, KNMAIHPR, KPWTQPKTNAIPYVRYL, LYQEPVLGPVRGPFILV, RFVVAPFPEVFR e RHPHPHLSFM. Imagem gerada pelo VMD.



4.2 PREPARAÇÃO DOS LIGANTES

O próximo passo é a preparação dos ligantes, que foi feita utilizando os programas AutoDock 4.2 e AutoDock Tools, primeiro fazendo a leitura do modelo 3D do receptor, editando a estrutura para adicionar os hidrogênios que não estavam presentes e gerando um novo arquivo .pdb.

Em seguida foi feita a leitura do receptor e o AutoDock Tools fez a otimização dos parâmetros, deletando a água e utilizando uma constante para simular um meio contínuo de água e deletando cadeias e ligantes extras que não serão importantes para a simulação, também são geradas as grades de potencial ao redor do receptor, assim gerando um novo arquivo .pdbqt. Isso foi feito para cada um dos ligantes, resultando em oito arquivos .pdbqt.

Além disso, também foi criado um arquivo vina.conf com as configurações da grade de potenciais para futuro uso no AutoDock Vina, contendo as dimensões configuradas em 63x63x63 (com base nas dimensões utilizadas no AutoDock, que são 96x96x96, com uma resolução de 0,375 angstrom, esse valor foi escolhido para manter a caixa do mesmo tamanho), o ponto central da grade, nas coordenadas 13,532 no eixo X, 26,462 no eixo Y e 55,847 no eixo Z, e o nível de exaustão 10, um parâmetro interno para insistência do programa em determinada conformação.

4.3 DOCAGEM MOLECULAR

Para a docagem molecular foi utilizado o programa AutoDock Vina, executado em Linux nos computadores do Grupo de Química Teórica (GQT) da UFRGS.

Foram utilizando os arquivos vina.conf, ligantes e receptor para iniciar a docagem. As simulações foram feitas em triplicatas em 100 passos para as conformações, selecionando as nove conformações com melhor afinidade em cada uma das simulações feitas por cada ligante.

Para isso foi feito um script para fazer as vinte e quatro simulações (três para cada ligante) em sequência, que está demonstrado na Figura 4.

Figura 4: Script utilizado para fazer a docagem.

```
#!/bin/csh -f
foreach l (DIPNPIGSENSGKITMPLW EPVLGPVVRGPFPILV IASAEPVHSTPTTEAWN KNMAIHR KPWTQPKTNAIPYVRYL LYQEPVLGPVVRGPFPILV RFVWAPFPEVFR RHPHPLSFM)
  mkdir result_DPP4_fit_${l}
  foreach r (DPP4)
    foreach t (1 2 3)
      echo Processing receptor $r ligand $l replica $t
      /home/alexandre/work --config vina.conf --cpu 1 --receptor $r.pdbqt --ligand $l.pdbqt --out ${r}_${l}_out_${t}.pdbqt --log ${r}_${l}log_${t}.txt
      /home/alexandre/work --input ${r}_${l}_out_${t}.pdbqt --ligand ${r}_${l}_${t}_
      mv ${r}_${l}* result_DPP4_fit_${l}
    end
  end
end
cd ./result_DPP4_fit_${l}
cat ../DPP4.pdbqt > complex_${r}_${l}.pdbqt
foreach t (1 2 3)
  cat ${r}_${l}_${t}_1.pdbqt >> complex_${r}_${l}.pdbqt
end
cd ..
end
exit
```

O script consiste nos comandos: `foreach` (implementa um ciclo em que as variáveis são obtidas por uma ou mais listas), `mkdir` (criar diretório), `echo` (apresentar uma linha de texto ou string), `mv` (mover e renomear arquivos ou diretórios), `cd` (acessar diretório), `cat` (concatenar: lista, combina e escreve arquivos no formato padrão), `end` (termina um ciclo) e `exit` (fecha o terminal e finaliza todos os processos sendo executados no terminal), além dos arquivos discutidos anteriormente com seus endereços e os diretórios relevantes.

O script foi executado com o comando `nohup` (`nohangups`: faz o processo ocorrer no plano de fundo, para que o terminal possa ser utilizado enquanto o processo é executado) para iniciar a docagem molecular. Durante o processo foi utilizado o comando `top` (apresenta todos os processos sendo executados no sistema) para conferir se as simulações estão sendo executadas corretamente. As simulações demandaram cerca de 100 horas no total.

Para cada ligante, foram obtidos 34 arquivos resultantes, que incluem a estrutura do complexo formado, um arquivo contendo as nove conformações resultantes um para cada simulação da triplicata (3 arquivos), nove arquivos contendo as conformações de melhor afinidade para cada uma das três simulações feitas por ligante (27 arquivos), e um arquivo log em formato de texto contendo os resultados para cada simulação da triplicata (3 arquivos).

5 ANÁLISE DOS RESULTADOS E DISCUSSÃO

Primeiramente foram checados os 24 arquivos log que mostram as melhores conformações de cada simulação de acordo com a afinidade. A melhor conformação de cada simulação está representada na Tabela 1.

Tabela 1: Resultados das afinidades para a melhor conformação de cada simulação feita na docagem. Em negrito estão os resultados mais promissores.

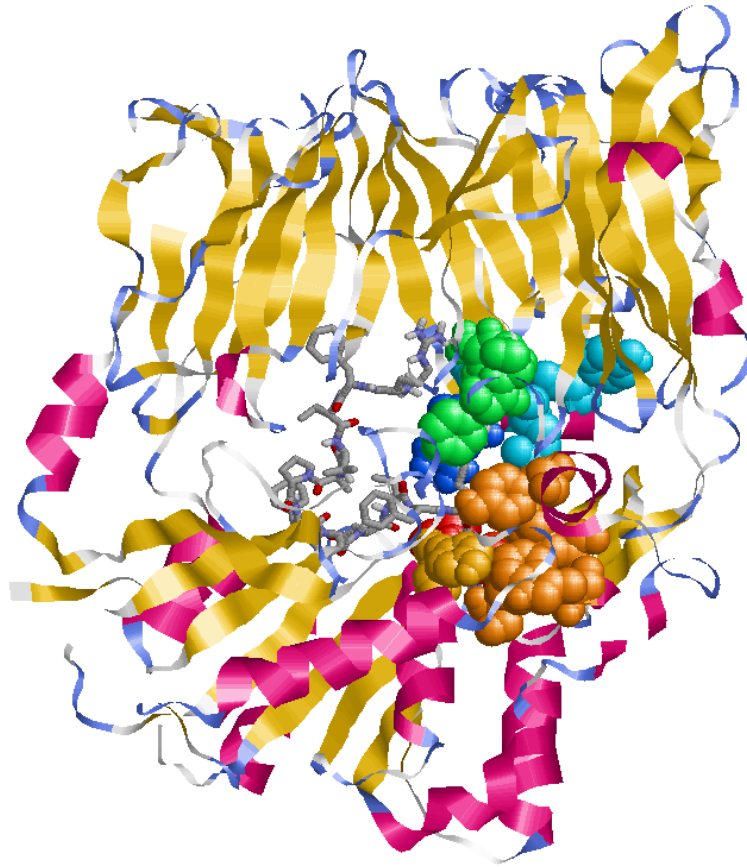
Ligante	Afinidade da primeira simulação	Afinidade da segunda simulação	Afinidade da terceira simulação
DIPNPIGSENSGKITMPLW	-7,7 kcal/mol	-7,7 kcal/mol	-7,7 kcal/mol
EPVLGPVRGPFILV	-8,9 kcal/mol	-8,7 kcal/mol	-8,5 kcal/mol
IASAEPTVHSTPTTEAVVN	-7,4 kcal/mol	-7,0 kcal/mol	-7,6 kcal/mol
KNMAIHR	-7,6 kcal/mol	-7,6 kcal/mol	-7,6 kcal/mol
KPWTQPKTNAIPYVRYL	-7,8 kcal/mol	-7,7 kcal/mol	-7,9 kcal/mol
LYQEPVLGPVRGPFILV	-7,9 kcal/mol	-7,9 kcal/mol	-8,9 kcal/mol
RFVVAPFPEVFR	-9,2 kcal/mol	-7,9 kcal/mol	-7,8 kcal/mol
RHPHPHLSFM	-8,6 kcal/mol	-6,9 kcal/mol	-7,3 kcal/mol

As conformações que mais se destacaram foram as três do EPVLGPVRGPFILV, a terceira do LYQEPVLGPVRGPFILV e as primeiras do RFVVAPFPEVFR e RHPHPHLSFM.

Após isso, foram convertidos os arquivos .pdbqt da conformação de melhor afinidade de cada simulação para .pdb e concatenados com o arquivo .pdb da proteína, para formar um arquivo do complexo para cada ligante de cada simulação individualmente.

As 24 conformações foram então analisadas com o programa RasMol, onde foi selecionada a cadeia da proteína e as opções Display>Ribbons e Colours>Structure. Em seguida foi selecionada a cadeia do ligante e as opções Display>Sticks e Colours>CPK. Após isso, foram selecionados os resíduos-chave nas posições 630, 708 e 740 para a tríade catalítica, 631, 656, 659, 662, 666 e 711 para o S1 e 125, 203, 206, 209, 357 e 358 para o S2 e as opções Display>Spacefill e Colours>Group para todos. Na Figura 5 está um exemplo de análise utilizando o RasMol.

Figura 5: Exemplo de análise de interação da melhor conformação da primeira simulação da RFVVAPFPEVFR utilizando o RasMol.



Os resultados das análises estão dispostos na Tabela 2.

Tabela 2: Resultados das interações com sítios ativos relevantes para a melhor conformação de cada simulação. Em negrito estão os resultados mais promissores.

Ligante	Interações com sítios ativos relevantes na primeira simulação	Interações com sítios ativos relevantes na segunda simulação	Interações com sítios ativos relevantes na terceira simulação
DIPNPIGSENSGKITMPLW	Triade, S1, S2	Triade, S1, S2	Triade, S1, S2
EPVLPVVRGPFPIILV	Triade, S1, S2	Triade, S1, S2	S2
IASAEPTVHSTPTTEAVVN	S2	S2	S2
KNMAIHPR	Triade, S1, S2	Triade, S1, S2	Triade, S1, S2
KPWTQPKTNAIPYVRYL	S2	S2	S2
LYQEPVLPVVRGPFPIILV	S2	S2	S2
RFVVAPFPEVFR	Triade, S1, S2	Nenhuma	Triade, S1, S2
RHPHPLSFM	Triade, S1, S2	Triade, S1, S2	S2

As interações com os três resíduos-chave são essenciais para existir o caráter inibitório da proteína, os peptídeos que apresentaram pelo menos uma conformação com essas interações foram: DIPNPIGSENSGKITMPLW, EPVLGPVRGPFILV, KNMAIHPR, RFVVAPFPEVFR e RHPHPLSFM, com destaque para DIPNPIGSENSGKITMPLW e KNMAIHPR que tiveram as três conformações atendendo aos requisitos.

Todas as conformações dos peptídeos IASAEPTVHSTPTTEAVVN, KPWTQPKTNAIPYVRYL e LYQEPVLGPVRGPFILV apresentaram interações relevantes apenas com o S2, assim como a terceira simulação dos peptídeos EPVLGPVRGPFILV e RHPHPLSFM. A segunda simulação do RFVVAPFPEVFR não demonstrou interações com nenhum sítio ativo relevante.

Comparando as duas tabelas, podemos selecionar os peptídeos com melhores afinidades que atendem aos requisitos estéricos, sendo esses EPVLGPVRGPFILV, RFVVAPFPEVFR e RHPHPLSFM, seguidos de DIPNPIGSENSGKITMPLW e KNMAIHPR.

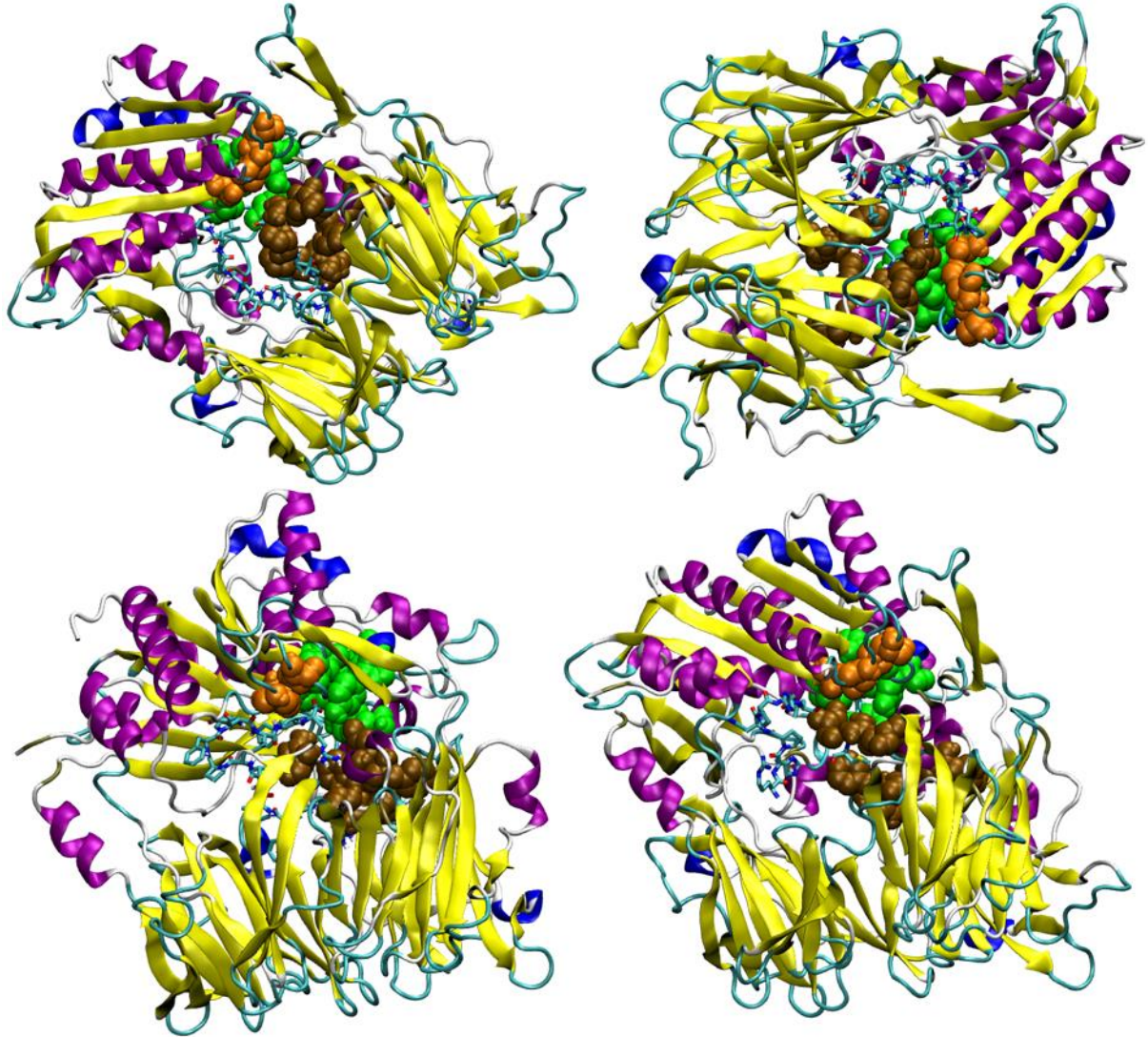
EPVLGPVRGPFILV apresentou conformações com ótimas afinidades e duas delas com boas interações, RFVVAPFPEVFR teve a melhor conformação no geral na primeira simulação de sua triplicata e RHPHPLSFM teve uma conformação com ótima afinidade e interação.

DIPNPIGSENSGKITMPLW e KNMAIHPR apresentaram bons resultados de afinidade e interações em todas as suas simulações, mas com uma afinidade um pouco menor que os três citados anteriormente.

LYQEPVLGPVRGPFILV apresentou uma conformação com ótima afinidade, mas sem interações com os sítios ativos relevantes, não cumprindo um dos requisitos necessários.

Foi utilizado o VMD para gerar imagens das estruturas dos complexos das conformações mais promissoras, que cumpriram ambos os requisitos de afinidade e interação com os resíduos chave e estão representados na Figura 6.

Figura 6: Da esquerda para a direita e de cima para baixo estão os modelos 3D das estruturas dos complexos das conformações mais promissoras: EPVLGPVRGPFILV 1, EPVLGPVRGPFILV 2, RFVVAPFPEVFR 1 e RHPHPLSFM 1. Imagem gerada pelo VMD.



6 CONCLUSÃO

A docagem molecular utilizando o AutoDock Vina se demonstrou um método eficiente e relativamente rápido para a simulação de ligante-receptor, houve uma variância nas afinidades e interações com resíduos-chave em alguns peptídeos na triplicata devido ao uso de algoritmo genético para a busca de conformações.

O RasMol foi muito efetivo na discriminação e análise do ligante e dos sítios ativos de maneira simples e rápida.

Foram encontrados alguns problemas na execução do AutoDock 4.2, AutoDock Tools e AutoDock Vina nas suas versões para Windows, mas não apresentaram problema algum em sua execução em Linux.

Os peptídeos EPVLGPVRGPFILV, RFVVAPFPEVFR e RHPHPLSFM apresentaram ótimas conformações que sugerem seu caráter inibitório da enzima DPP-IV, assim como DIPNPIGSENSGKITMPLW e KNMAIHPR que tiveram conformações com afinidades satisfatórias e boas interações com os sítios ativos, que também sugerem um caráter inibitório. Os peptídeos IASAEPTVHSTPTTEAVVN, KPWTQPKTNAIPYVRYL e LYQEPVLGPVRGPFILV não obtiveram nenhuma conformação que satisfaça o requisito das interações com os resíduos-chave, por esse motivo não é provável que sejam inibidores da enzima em questão.

Dito isso, os peptídeos que devem ser sintetizados por possuírem grandes possibilidades de serem inibidores da enzima DPP-IV são: EPVLGPVRGPFILV, RFVVAPFPEVFR, RHPHPLSFM, DIPNPIGSENSGKITMPLW e KNMAIHPR.

REFERÊNCIAS

- [1] Correr, C. J., Coura-Vital, W., Frade, J. C. Q. P., Nascimento, R. C. R. M., Nascimento, L. G., Pinheiro, E. B., Ferreira, W. M., Reis, J. S., Melo, K. F. S., Pontarolo, R., Lenzi, M. S. A., Almeida, J. v., Pedrosa, H. C., & João, W. S. J. (2020). Prevalence of people at risk of developing type 2 diabetes mellitus and the involvement of community pharmacies in a national screening campaign: A pioneer action in Brazil. *Diabetology and Metabolic Syndrome*, 12(1). <https://doi.org/10.1186/s13098-020-00593-5>
- [2] Kim, B.-R., Kim, H., Choi, I., Kim, J.-B., Jin, C., & Han, A.-R. (2018). DPP-IV Inhibitory Potentials of Flavonol Glycosides Isolated from the Seeds of *Lens culinaris*: In Vitro and Molecular Docking Analyses. *Molecules*, 23(8), 1998. <https://doi.org/10.3390/molecules23081998>
- [3] Juillerat-Jeanneret, L. (2014). Dipeptidyl Peptidase IV and Its Inhibitors: Therapeutics for Type 2 Diabetes and What Else? *Journal of Medicinal Chemistry*, 57(6), 2197–2212. <https://doi.org/10.1021/jm400658e>
- [4] Pantaleão, S. Q., Maltarollo, V. G., Araujo, S. C., Gertrudes, J. C., & Honorio, K. M. (2015). Molecular docking studies and 2D analyses of DPP-4 inhibitors as candidates in the treatment of diabetes. *Molecular BioSystems*, 11(11), 3188–3193. <https://doi.org/10.1039/c5mb00493d>
- [5] Wang, L., Wang, N., Zhang, W., Cheng, X., Yan, Z., Shao, G., Wang, X., Wang, R., & Fu, C. (2022). Therapeutic peptides: current applications and future directions. In *Signal Transduction and Targeted Therapy* (Vol. 7, Issue 1). Springer Nature. <https://doi.org/10.1038/s41392-022-00904-4>
- [6] Mudgil, P., Kamal, H., Yuen, G. C., & Maqsood, S. (2018). Characterization and identification of novel antidiabetic and anti-obesity peptides from camel milk protein hydrolysates. *Food Chemistry*, 259, 46–54. <https://doi.org/10.1016/j.foodchem.2018.03.082>
- [7] Morrison, J. L., Breitling, R., Higham, D. J., & Gilbert, D. R. (2006). A lock-and-key model for protein-protein interactions. *Bioinformatics*, 22(16), 2012–2019. <https://doi.org/10.1093/bioinformatics/btl338>
- [8] Kim, M. O., Nichols, S. E., Wang, Y., & McCammon, J. A. (2013). Effects of histidine protonation and rotameric states on virtual screening of *M. tuberculosis* RmlC. *Journal of Computer-Aided Molecular Design*, 27(3), 235–246. <https://doi.org/10.1007/s10822-013-9643-9>
- [9] Burley, S. K., Berman, H. M., Bhikadiya, C., Bi, C., Chen, L., di Costanzo, L., Christie, C., Duarte, J. M., Dutta, S., Feng, Z., Ghosh, S., Goodsell, D. S., Green, R. K., Guranovic, V., Guzenko, D., Hudson, B. P., Liang, Y., Lowe, R., Peisach, E., ... Ioannidis, Y. E. (2019). Protein Data Bank: The single global archive for 3D macromolecular structure data. *Nucleic Acids Research*, 47(D1), D520–D528. <https://doi.org/10.1093/nar/gky949>
- [10] Bateman, A., Martin, M. J., O'Donovan, C., Magrane, M., Alpi, E., Antunes, R., Bely, B., Bingley, M., Bonilla, C., Britto, R., Bursteinas, B., Bye-AJee, H., Cowley, A., da Silva, A., de Giorgi, M., Dogan, T., Fazzini, F., Castro, L. G., Figueira, L., ... Zhang, J. (2017). UniProt: The universal protein knowledgebase. *Nucleic Acids Research*, 45(D1), D158–D169. <https://doi.org/10.1093/nar/gkw1099>

- [11] Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., Zidek, A., Green, T., Tunyasuvunakool, K., Petersen, S., Jumper, J., Clancy, E., Green, R., Vora, A., Lutfi, M., ... Velankar, S. (2022). AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research*, 50(D1), D439–D444. <https://doi.org/10.1093/nar/gkab1061>
- [12] Jo, S., Kim, T., Iyer, V. G., & Im, W. (2008). CHARMM-GUI: A web-based graphical user interface for CHARMM. *Journal of Computational Chemistry*, 29(11), 1859–1865. <https://doi.org/10.1002/jcc.20945>
- [13] Morris, G. M., Ruth, H., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). Software news and updates AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry*, 30(16), 2785–2791. <https://doi.org/10.1002/jcc.21256>
- [14] Trott, O., & Olson, A. J. (2009). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*. <https://doi.org/10.1002/jcc.21334>
- [15] O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., & Hutchison, G. R. (2011). Open Babel: An Open chemical toolbox. *Journal of Cheminformatics*, 3(10). <https://doi.org/10.1186/1758-2946-3-33>