

ESTUDOS *IN SILICO* E *IN VITRO* PARA BUSCA DE POTENCIAIS FÁRMACOS CONTRA O SARS-CoV-2 (COVID-19)¹

William Jefferson Silva Sena², Bruno Silva Andrade³

RESUMO

Uma base de dados de peptídeos dispõe aos seus usuários um conjunto de informações bem organizadas e definidas, que servem como fonte para o que eles necessitarão. Por conta de como a mesma é tratada, isso impõe algumas restrições quanto ao contexto em que ela é situada. Para expandir esse contexto e diminuir essa limitação, buscamos mesclar diversas bases de dados de peptídeos, mantendo as informações que elas possuem em comum, além de acrescentar aquelas destinadas aos nossos propósitos. Como resultado disso, obtivemos uma base de dados com mais de 280 mil peptídeos, sendo que mais da metade deles possuem atividade imunomoduladora, a qual foi identificada como a mais presente entre esses peptídeos, e utilizada como fonte para o desenvolvimento de uma *api*.

PALAVRAS-CHAVE: peptídeos, banco de dados, scraping, *api*.

IN SILICO AND *IN VITRO* STUDIES TO SEARCH FOR POTENTIAL DRUGS AGAINST SARS-CoV-2 (COVID-19)¹

ABSTRACT

A peptide database delivers to their users a set of well-organized and defined information, that act as a source for what they need. Because of how it is treated, this impose some limitations regarding the context in which it is located. For expanding this context and decrease this limitation, we seek to merge several peptide databases, keeping their information in common, in addition to adding those intended to our purposes. As a result of this, we obtained a database with more than 280 thousand peptides, with more than half of them to having immunomodulatory activity, which was identified as the most present among these peptides and used as source to *api* development.

KEYWORDS: peptides, database, scraping, *api*.

INTRODUÇÃO

¹ Programa Institucional de Bolsas de Iniciação em Desenvolvimento Tecnológico e Inovação (Pibiti).

² Graduando em Sistemas de Informação pela Universidade Estadual do Sudoeste da Bahia - Campus Jequié. Bolsista de Iniciação Científica do Laboratório de Bioinformática e Química Computacional (LBQC/UESB). Rua José Moreira Sobrinho, s/n, Jequiezinho, 45200000 - Jequié, BA - Brasil.

³ Professor Titular (Medicina) - Universidade Estadual do Sudoeste da Bahia, Jequié - BA
Doutor em Biotecnologia (UEFS/FIOCRUZ) - Modelagem Molecular e Desenho de Fármacos/Química Medicinal. Coordenador do Laboratório de Bioinformática e Química Computacional - LBQC/UESB. Rua José Moreira Sobrinho, s/n, Jequiezinho, 45200000 - Jequié, BA - Brasil.

Um banco de dados é uma coleção organizada de informações estruturadas, normalmente armazenadas eletronicamente em um sistema de computador [4]. Essa coleção, geralmente, representa uma parte do mundo real, que precisa ser expresso computacionalmente. A partir disso, o contexto definido para esse projeto foram os peptídeos - biomoléculas que contém de dois a dezenas de resíduos de aminoácidos unidos entre si através de ligações peptídicas [3].

Atualmente, existem diversas bases de dados com enfoque em peptídeos, cada uma com suas especificações diante do contexto em que serão utilizadas. Alguns exemplos são: a *FermFooDb* [2], que armazena peptídeos de alimentos fermentados e a *CancerPPD* [8], que engloba peptídeos com atividades anticâncer e afins.

Apesar de existirem bases de dados que mesclam as informações de diversas outras bases, elas ainda possuem as limitações do contexto em que são utilizadas, como foi dito anteriormente. Por conta disso, construímos um banco de dados que armazena essas informações, sem se prender a esse ponto, unindo o que for obtido de diversas fontes. E, por fim, disponibilizá-lo em forma de *api* online.

MATERIAIS E MÉTODOS

A seleção dos peptídeos foi feita a partir de bases de dados de peptídeos públicas, são elas: APD3 [10], CancerPPD [8], DBAASP [6], FermFooDb [2], MAHMI [1], PepBDB [11], THPdb [9]. Para extrair os dados de cada base utilizamos *web scraping* com *Python* e *Selenium* – o qual consistiu na automação do navegador para captura dos dados, que por sua vez foram salvos em formato *csv*.

Além dos dados referentes as sequências de peptídeos, também foram obtidas as suas estruturas 3D em formato *pdb*. Com exceção das bases de dados que já forneciam essas estruturas, utilizamos dois *softwares* para obter as restantes: o *software* de busca, modelagem e *docking* de peptídeos – *BioPep* [5]; e o programa de inteligência artificial para predição de estruturas de proteínas/peptídeos – *AlphaFold* [7].

Com todos esses dados e arquivos organizados, desenvolvemos uma *api* com *Python*, *Django Rest Framework*, banco de dados *PostgreSQL* e *Docker*. Por meio dessa *api* é possível realizar buscas por sequências homólogas (*BLASTp*) e efetuar o *download* dos arquivos *pdb* e *fasta* (arquivo de texto com a sequência do peptídeo).

RESULTADOS E DISCUSSÃO

A partir da captura de dados das bases citadas, foi possível obter mais de 320 mil peptídeos, que posteriormente passaram por uma filtragem para capturar apenas as sequências válidas, resultando em um total de 286.913 peptídeos, assim como descrito na tabela 1.

TABELA 1: Quantidade de peptídeos obtidos por base de dados.

Bases de dados	Quantidade de peptídeos
APD3	2.810
CancerPPD	4.290
DBAASP	14.576
FermFooDb	2.156
MAHMI	265.918
PepBDB	9.122
ThPdb	209
Total	286.913

Fonte: autoral.

Dentre esses peptídeos, também foi efetuada uma análise das atividades biológicas mais presentes nos mesmos (tabela 2) e, como é possível perceber na tabela 1, a base de dados que mais contribuiu para isso foi a MAHMI ^[1], responsável por mais de 90% deles. Sendo a atividade imunomodulatória a mais presente, utilizamos apenas aqueles que a possuíam, levando em consideração que um único peptídeo pode apresentar diversas atividades.

TABELA 2: Atividade biológicas mais presentes entre os peptídeos obtidos.

Atividades	Bases de dados				Total
	APD3	CancerPPD	FermFooDb	MAHMI	
Imunomodulatória	0	0	14	173935	173949
Anti-inflamatória	34	0	1	71256	71291
Migração/Adesão	0	0	0	14875	14875
Citotóxica	0	0	0	5852	5852
Anti-Gram+	2105	0	0	0	2105
Anti-Gram-	1942	0	0	0	1942
Antifúngica	803	251	0	0	1054

Fonte: autoral.

CONCLUSÕES/CONSIDERAÇÕES

A criação dessa api mostrou-se trabalhosa, por conta de envolver ferramentas e tecnologias que não tínhamos um contato prévio, como o *Docker* e o *Django Rest Framework*. Em questão de funcionalidades ainda temos muito o que implementar, mas o que temos até o momento já alcançou nossas expectativas.

Atualmente, essa *api* está disponível online e é utilizada no site *PIABA (Peptides With Immune and Antimicrobial Activity-based Associations Database)*, com acesso através do - <http://44.217.73.143>.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] BLANCO-MÍGUEZ, A.; GUTIÉRREZ-JÁCOME, A.; FDEZ-RIVEROLA, F.; LOURENÇO, A.; SÁNCHEZ, B. MAHMI database: a comprehensive MetaHit-based resource for the study of the mechanism of action of the human microbiota. **Oxford**, 2017. Disponível em: <http://doi.org/10.1093/database/baw157>.
- [2] CHAUDHARY, A. et al. FermFoodb: A database of bioactive peptides derived from fermented foods. **Heliyon** 7(4): e06668, 2021.
- [3] MACHADO, A. et al. Sínteses química e enzimática de peptídeos: princípios básicos e aplicações. **Química nova**, v. 27, p. 781-789, 2004.
- [4] ORACLE. What is a Database. **Oracle**, 2023. Disponível em: <https://www.oracle.com/br/database/what-is-database/>.
- [5] PALMEIRA, L. S.; SENA, W. J. S. **BioPep**. Zenodo. Disponível em: <https://doi.org/10.5281/zenodo.5781778>.
- [6] PIRTSKHALAVA, M. et al. DBAASP v3: database of antimicrobial/cytotoxic activity and structure of peptides as a resource for development of new therapeutics. **Nucleic Acids Research**, v. 49, i. D1, p. D288–D297, 2021. Disponível em: <https://doi.org/10.1093/nar/gkaa991>.
- [7] TUNYASUVUNAKOOL, K.; ADLER, J.; WU, Z. *et al.* Highly accurate protein structure prediction for the human proteome. **Nature**, p. 590–596 (2021). Disponível em: <https://doi.org/10.1038/s41586-021-03828-1>.
- [8] TYAGI et al. CancerPPD: a database of anticancer peptides and proteins. **Nucleic Acids Research**, v. 43: D837-43, 2015.
- [9] UESMANI, S. et al. THPdb: Database of FDA-approved peptide and protein therapeutics. **PLoS ONE**, 12(7) e0181748, 2017.
- [10] WANG, G.; LI, X.; WANG, Z. APD3: the antimicrobial peptide database as a tool for research and education. **Nucleic Acids Research**, v. 44, D1087-D1093, 2016.
- [11] WEN, Z.; HE, J.; TAO, H.; HUANG, S. PepBDB: a comprehensive structural database of biological peptide-protein interactions. **Bioinformatics**, 2018. Disponível em: <http://doi.org/10.1093/bioinformatics/bty579>.