

***In silico* analysis of the protein-protein interaction of the SARS-CoV-2 spike protein**

Análise *in silico* da interação proteína-proteína da proteína spike do SARS-CoV-2

Análisis *in silico* de la interacción proteína-proteína de la proteína espiga SARS-CoV-2

Received: 06/06/2024 | Revised: 06/18/2024 | Accepted: 06/19/2024 | Published: 06/21/2024

Andressa Souza Marques¹

ORCID: <https://orcid.org/0000-0003-4332-9079>

Universidade Federal da Bahia, Brasil

E-mail: andressamarques.biotec@gmail.com

Taiane de Macêdo Gondim

ORCID: <https://orcid.org/0000-0002-6802-4661>

Empresa Brasileira de Serviços Hospitalares – EBSEH, Brasil

Universidade Federal da Bahia, Brasil

E-mail: taigondim@hotmail.com

Fulvia Soares Campos de Sousa²

ORCID: <https://orcid.org/0000-0002-2779-5478>

Universidade Federal da Bahia, Brasil

E-mail: fulvia@ufba.br

Antônio Pedro Fróes de Farias³

ORCID: <https://orcid.org/0000-0002-2625-9470>

Universidade Federal da Bahia, Brasil

E-mail: froes_pedro@hotmail.com

Bruno Silva Andrade⁴

ORCID: <https://orcid.org/0000-0002-8031-9454>

Fundação Oswaldo Cruz, Brasil

Universidade Estadual de Santa Cruz, Brasil

Fundação Universidade Federal do Tocantins, Brasil

E-mail: bandrade@uesb.edu.br

Soraya Castro Trindade

ORCID: <https://orcid.org/0000-0001-7125-9114>

Universidade Estadual de Feira de Santana, Brasil

E-mail: soraya.castrotrindade@gmail.com

José Tadeu Raynal da Rocha Filho

ORCID: <https://orcid.org/0000-0002-2771-0235>

Universidade Federal da Bahia, Brasil

E-mail: jtraynal@hotmail.com

Roberto Meyer

ORCID: <https://orcid.org/0000-0002-4727-4805>

Universidade Federal da Bahia, Brasil

E-mail: rmeyer@ufba.br

Abstract

The pandemic caused by the SARS-CoV-2 virus has represented a global challenge with a significant impact on public health since its emergence in 2019. Understanding the interactions between this virus and the human immune system is essential for the development of new strategies more effective therapies and diagnoses. This study aimed to predict epitopes for SARS-CoV-2 T and B cells, as well as evaluate the interaction of the spike protein with other viral proteins, using bioinformatics methods. SARS-CoV-2 protein sequences were collected from UniProt. Epitopes for T cells were predicted *in silico* using specific HLA alleles from the Bahia population. Epitopes for B cells were predicted using the IEDB server with multiple methods based on amino acid features. Protein-protein interaction was analyzed using the STRING database. The result is 10,671 peptides related to various SARS-CoV-2 viral proteins, including the spike, essential for the infection and pathogenesis of COVID-19. In addition to spike, proteins such as ORF3a, ORF7a, and ORF8 showed significant immunogenic potential. Protein-protein interaction analysis revealed that proteases such as TMPRSS2 and TMPRSS11D are crucial for viral entry and are potential therapeutic targets. This study expands the understanding of the molecular interactions of SARS-CoV-2, highlighting new therapeutic targets and clinical complications associated with COVID-19. The results provide valuable insights for the

¹ Programa de Pós-Graduação em Imunologia - Universidade Federal da Bahia, Brasil.

² Laboratório de Imunologia e Biologia Molecular, Instituto de Ciências da Saúde, Universidade Federal da Bahia, Brasil

³ Pesquisador do Laboratório de Imunologia e Biologia Molecular, Instituto de Ciências da Saúde, Universidade Federal da Bahia, Brasil

⁴ Programas de Pós-graduação em Biotecnologia (PPGBIOTEC-UEFS/Fiocruz-Ba), Genética e Biologia Molecular (PPGGBM-UESC), Química (PPGQUI-UESB), e Biotecnologia (PPGBIOTEC-UFT). Brasil.

development of targeted therapeutic strategies and improved diagnostics, contributing to the mitigation of the global pandemic.

Keywords: SARS-CoV-2; T cells; B cells; Epitopes; Protein-protein interaction.

Resumo

A pandemia causada pelo vírus SARS-CoV-2 tem representado um desafio global de impacto significativo para a saúde pública desde seu surgimento em 2019. A compreensão das interações entre este vírus e o sistema imunológico humano é essencial para o desenvolvimento de novas estratégias terapêuticas e diagnósticos mais eficazes. Este estudo teve como objetivo realizar a predição de epítomos para células T e B do SARS-CoV-2, bem como avaliar a interação da proteína Spike com outras proteínas virais, utilizando métodos de bioinformática. Foram coletadas sequências de proteínas do SARS-CoV-2 do UniProt. Epítomos para células T foram preditos *in silico* utilizando alelos HLA específicos da população da Bahia. Epítomos para células B foram preditos utilizando o servidor IEDB com múltiplos métodos baseados em características de aminoácidos. A interação proteína-proteína foi analisada utilizando o banco de dados STRING. Como resultado tem-se 10.671 peptídeos relacionados a várias proteínas virais do SARS-CoV-2, incluindo a spike, essenciais para a infecção e patogênese da COVID-19. Além da spike, proteínas como ORF3a, ORF7a, e ORF8 mostraram potencial imunogênico significativo. A análise de interação proteína-proteína revelou que proteases como TMPRSS2 e TMPRSS11D são cruciais para a entrada viral, sendo potenciais alvos terapêuticos. Este estudo amplia o entendimento das interações moleculares do SARS-CoV-2, destacando novos alvos terapêuticos e complicações clínicas associadas à COVID-19. Os resultados fornecem insights valiosos para o desenvolvimento de estratégias terapêuticas direcionadas e aprimoramento dos diagnósticos, contribuindo para a mitigação da pandemia global.

Palavras-chave: SARS-CoV-2; Células T; Células B; Epítomos; Interação proteína-proteína.

Resumen

La pandemia causada por el virus SARS-CoV-2 ha representado un desafío global con un impacto significativo en la salud pública desde su aparición en 2019. Comprender las interacciones entre este virus y el sistema inmunológico humano es esencial para el desarrollo de nuevas estrategias terapéuticas y diagnósticos más eficaces. Este estudio tuvo como objetivo predecir epítomos para las células T y B del SARS-CoV-2, así como evaluar la interacción de la proteína de pico con otras proteínas virales, utilizando métodos bioinformáticos. Métodos: Se recogieron secuencias de proteínas del SARS-CoV-2 de UniProt. Los epítomos de las células T se predijeron *in silico* utilizando alelos HLA específicos de la población de Bahía. Los epítomos de las células B se predijeron utilizando el servidor IEDB con múltiples métodos basados en las características de los aminoácidos. La interacción proteína-proteína se analizó utilizando la base de datos STRING. El resultado son 10.671 péptidos relacionados con diversas proteínas virales del SARS-CoV-2, incluida la espiga, esenciales para la infección y patogénesis del COVID-19. Además del pico, proteínas como ORF3a, ORF7a y ORF8 mostraron un potencial inmunogénico significativo. El análisis de la interacción proteína-proteína reveló que proteasas como TMPRSS2 y TMPRSS11D son cruciales para la entrada viral y son posibles objetivos terapéuticos. Este estudio amplía la comprensión de las interacciones moleculares del SARS-CoV-2, destacando nuevas dianas terapéuticas y complicaciones clínicas asociadas con el COVID-19. Los resultados proporcionan información valiosa para el desarrollo de estrategias terapéuticas específicas y diagnósticos mejorados, contribuyendo a la mitigación de la pandemia mundial.

Palabras clave: SARS-CoV-2; Células T; Células B; Epítomos; Interacción proteína-proteína.

1. Introduction

The SARS-CoV-2 virus was responsible for the global COVID-19 pandemic and, to this day, represents a major challenge for public health (Bar-On et al., 2020; Díaz-Castrillón & Toro-Montoya, 2020; Bussani et al., 2023). Since the first cases reported in 2019, the scientific community has been searching for effective strategies to identify, control and combat the spread of the virus (Srivastana & Saxena, 2020; Kampf et al., 2020; Solanki et al., 2023). In this scenario, understanding the interactions of SARS-CoV-2 and its host's immune system is essential for the development of targeted therapies and more effective diagnostic tests (Díaz-Castrillón & Toro-Montoya, 2020; Eckerle & Meyer, 2020; Majra et al., 2021).

A fundamental point of virus-host interaction is the identification of specific segments of viral proteins that the human immune system is capable of recognizing. (Zhang et al., 2020; Sohail et al., 2021). The *in silico* identification of epitopes for T and B cells has become a valuable strategy in the search for possible therapeutic targets and the development of increasingly efficient vaccines (Shang et al., 2020; Pan et al., 2021; Sohail et al., 2021; Chakraborty et al., 2021; Mukherjee et al., 2020). Understanding how these epitopes stimulate immune responses can provide us with fundamental insights for developing

immunization strategies (Yarmarkovich et al., 2020; Grifoni et al., 2021; Shomuradova et al., 2020; Devi et al., 2021; Fathollahi et al., 2023; Weingarten-Gabbay et al., 2024).

Furthermore, understanding protein-protein interactions plays a crucial role in the pathogenesis of SARS-CoV-2. The spike (S) protein is extremely relevant due to its function of mediating the entry of the virus into host cells, and understanding how the spike interacts with other proteins of the virus itself and the consequences on host cells is essential to identify new targets therapies and identify the mechanisms used by the virus after infection (Arabi-Jeshvaghani et al., 2023; Navish & Uthayakumar et al., 2023; Kumar et al., 2023; Ozger, 2023; Zhou et al., 2023).

Advances in understanding virus-host interactions and predicting epitopes lead to the development of new therapies and even the improvement of existing ones. Directing these strategies to specific targets can boost the effectiveness of treatments and improve the sensitivity and specificity of tests. In this context, the article aim to predict epitopes for T and B cells, and evaluate the interaction of the spike protein with other SARS-CoV-2 proteins.

This study aims to apply bioinformatics methods to predict specific SARS-CoV-2 epitopes recognized by T and B cells using genomic and proteomic data. Furthermore, we seek to investigate the detailed interaction of the spike protein with other viral proteins, with a focus on identifying potential therapeutic targets and understanding the molecular mechanisms underlying viral pathogenesis.

2. Methods

2.1 Collection of protein sequences

The complete SARS-CoV-2 proteome was obtained from UniProt (ID: 2697049) via the UniProt online platform (<https://www.uniprot.org/>). The proteome includes all proteins encoded by the virus genome, essential for viral replication and interaction.

2.2 Allele selection

Human Leukocyte Antigen (HLA) alleles were selected using the Allele Frequencies database (<http://www.allelefrequencies.net/>), which contains genetic data from global populations. High frequency alleles were chosen in the State of Bahia (HLA-A*02:01, HLA-A*02:06, HLA-B*15:01, HLA-B*35:03, HLA-B*44:02, HLA-B*44:03), aiming to represent local genetic diversity.

2.3 *In silico* epitope mapping for T cells

The SARS-CoV-2 proteome was subjected to *in silico* epitope mapping for T cells using the IEDB platform (Immune Epitope Database; <http://tools.iedb.org/main/tcell/>). This tool identifies amino acid sequences with the potential to bind to class I Major Histocompatibility Complexes (MHC). Each viral protein was evaluated for its affinity with HLA alleles selected from the Bahia population. Epitopes with 100% affinity were identified and manually filtered to ensure accurate results.

2.4 *In silico* epitope mapping for B cells

The proteome was submitted to mapping on the IEDB platform (<http://tools.iedb.org/main>). The server has a collection of methods to predict linear B cell epitopes based on amino acid characteristics. The methods chosen were the BepiPred-2.0 server predicts B cell epitopes from the input protein sequence, amino acid residues above the default value (0.5) are predicted with part of an epitope (Jespersen et al., 2017), the evaluation of β -sheets for the prediction of antibody epitopes (Pellequer et al., 1993; Chou & Fasman, 1978), the surface accessibility method developed based on the surface accessibility

scale of an antibody (Emini et al., 1985), flexibility scale based on the mobility of protein segments (Karplus & Schulz, 1985), antigenicity scale – semi-empirical method – which uses physicochemical properties of amino acid residues and the frequency of occurrence in segmental epitopes already known experimentally (Kolaskar & Tongaonkar, 1990), and hydrophilicity prediction – construction of a hydrophilic scale based on peptide retention times during high-performance liquid chromatography (HPLC) (Parker & Hodges, 1986).

2.5 Analysis of protein-protein interactivity

To investigate interactions between SARS-CoV-2 viral proteins, the STRING tool (Search Tool for the Retrieval of Interacting Genes/Proteins; <https://string-db.org/>) was used. The STRING network provides a visualization of the molecular interactions between different proteins, based on experimental and predictive data from diverse sources. High confidence parameters (minimum score of 0.700) were used to guarantee the accuracy and relevance of the interactions identified.

3. Results and Discussion

The T and B cell prediction results in 10,671 peptides (these epitopes were selected through filtering in an Excel spreadsheet.) that are related to different protein fragments of SARS-CoV-2 involved in both the infection process and the clinical picture of COVID-19. In our study we observed peptides from the entire SARS-CoV-2 genome. In addition to protein S, other proteins drew attention because they have affinity for human HLA. Evaluating the function of these proteins where these sequences are found can bring insights for new research. The Table 1 below shows the proteins in which these peptides are contained and the functions they play in viral infection.

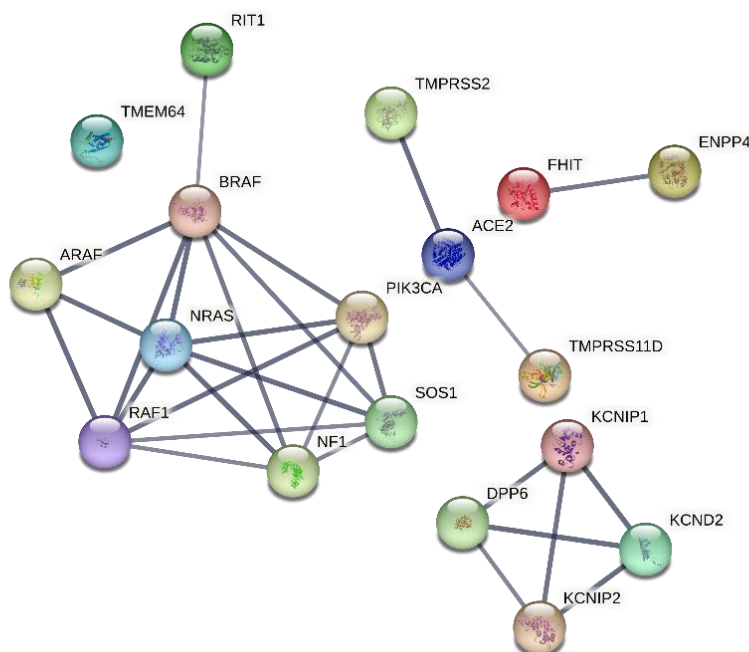
Table1 - Function of proteins containing epitopes.

Protein	Function
Replicase Polyprotein 1a and Replicase Polyprotein 1ab	Essential components in the virus replication process. They are translated into several enzymes that have essential functions in the replication of viral genetic material.
Spike	Responsible for virus entry into host cells through interaction with the ACE2 receptor on human cells.
ORF3a, ORF6, ORF7a, ORF7b, ORF8, ORF9b, ORF9c, ORF10	They refer to open reading frames. Specific sequences of RNA that can be translated into proteins. Each sequence encodes a specific SARS-CoV-2 protein, with functions that may involve modulating the immune response and regulating gene expression.
Putative ORF3d protein	Function not fully described.
Nucleoprotein	It acts in the formation of the protein capsule that surrounds the viral genetic material. Its role is fundamental in maintaining the integrity of the viral genome.

Source: Authors himself.

This information and analysis allows us to identify potential targets for therapeutic interventions along with information on protein-protein interactions. In our study, we searched for information on how the S protein interacts with other proteins in order to provide some possible answers on how the pathology and infectious process takes place.

Figure 1 - Graphical representation of the interaction networks of the S protein with other SARS-CoV-2 proteins. Direct and indirect interactions of Spike with functional associations are observed.



Source: Authors himself.

Spike (S) is capable of interacting with cellular proteases and can also access these cells through two distinct pathways. The S2 subunit is cleaved by Transmembrane Serine Proteases (TTSPs), in a first way, it contains the transmembrane protease associated with the plasma membrane, cell surface serine 2 (TMPRSS2) and promotes direct fusion between the viral envelope and the cell membrane through ACE2. In a second pathway, virions are directed to an endosome

and then the S2 subunit is cleaved by the lysosomal protease cathepsin B and L, responsible for the process of modeling the cellular matrix and production of peptide neurotransmitters (Bugge et al., 2009; Hook et al., 2012; Turk et al., 2000).

TMPRSS2 is a well-characterized human TTSP and is classified into three types that are classified according to the portion of the transmembrane domain, thus type I has a carboxy-terminal transmembrane domain; type II has an amino-terminal transmembrane domain; and type III has a glycosyl-phosphatidylinositol domain, a membrane anchoring portion (Tharapell et al., 2020; Kishimoto et al., 2021). Type II has 20 proteases that are divided between the families of hepsins/transmembrane proteases/serine (TMPRSS2), human airway trypsin-like protease (HAT)/differential expression in squamous cell carcinoma (DESC), matriptase and Corin (Tharapell et al., 2020; Kishimoto et al., 2021; Szabo & Bugge, 2007). Our *in silico* analyzes suggest that a direct interaction occurs between TMPRSS11D – belonging to the HAT/DESC family – ACE2 and TMPRSS2, as shown in Figure 1.

TMPRSS11D is co-expressed with ACE2 in bronchial epithelial cells and pneumocytes, together with TMPRSS2 which cleaves the S protein, dividing it into S1 and S2, increasing the virus's entry performance into the host cell (Shulla et al., 2011; Bittmann et al., 2020; Kishimoto et al., 2021). TMPRSS2 and TMPRSS11D emerge as promising targets for the development of therapeutic targets, once their activity is inhibited, effective strategies for preventing viral infection and the development of targeted and effective therapies can be devised..

In contrast, our analysis suggests that proteins such as KCND2 (voltage-gated potassium channel) and TMEM64 (transmembrane protein 64) act in distinct functions and are not directly involved in viral entry. KCND2 is part of voltage-regulated potassium channels (VGPCs), mutations in this gene are related to arrhythmias and sudden cardiac death, symptoms frequently reported after SARS-CoV-2 infection (Dehghani-Samani et al., 2019; Sasha et al., 2022; Huseynov et al., 2023). TMEM64 is a regulator of Ca²⁺ signaling pathways mediated by the receptor activator of NF-KB ligand (RANKL) involved in the regulation of osteoclast differentiation, this suggests that the virus may affect this regulation and cause arthralgias in patients with COVID-19 (Kim et al., 2013; Disser et al., 2020). These proteins play important roles in specific and distinct cellular processes, however, understanding these nuances in protein interactions is essential for the development of therapeutic strategies.

ORF3a (Ap3A) is an accessory protein located in late endosomes and plays its role in viral exit through lysosomal trafficking, inhibiting autophagic activity by blocking the fusion between autophagosomes and amphisomes with lysosomes (Miao et al., 2021). ENPP4 and FHIT, in our prediction, act concomitantly, promoting the hydrolysis of extracellular AP3A to produce AMP and ADP. ENPP4 promotes platelet aggregation in human plasma, favoring thrombus formation. FHIT, when mediated by viruses, increases the release of reactive oxygen species (ROS) and induces cell apoptosis (Albright et al., 2012; Albright et al., 2014; Zavalhia et al., 2018; Trapasso et al., 2008; Borza et al., 2022).

The ORF8 protein plays a role in modulating the host's immune response and the action of RIT1 under this protein suggests the triggering of a signaling cascade directing broad physiological cellular responses through the Ras/mitogen-activated protein kinase (MAPK) signaling pathway. RIT1 is part of the family of Ras guanosine triphosphate hydrolases (GTPases) and has a regulatory function in the proliferation, survival and differentiation of neuronal cells. Mutations in the signal transduction pathway result in disorders called RASopathies (Rauen, 2013; Van et al., 2020). RIT1 mRNS is widely expressed in the lungs, esophagus, blood and spleen (Lonsdale et al., 2013; Fang et al., 2016). Mutations in RIT1 have already been reported in other pathologies. Among the mutations are cardiac anomalies and blood diseases, this may suggest that SARS-CoV-2 can cause changes in RIT1 mRNA in host cells since respiratory complications, cardiac manifestations and vascular manifestations have been reported. (Aly et al., 2020; Han et al., 2022; Mehandru & Merad, 2022; Shukla et al., 2022; Ruggiero et al., 2022; Bowe et al., 2023; Sciaudone et al., 2023). Although the impacts of RIT1 mutations are evident in other pathologies, it is necessary to understand its mechanism at the molecular level in SARS-CoV-2 infections.

4. Conclusion

The present study investigated the prediction of T and B cells in relation to 10,671 SARS-CoV-2 peptides, highlighting the interaction of viral proteins with human HLA. We found that, in addition to the Spike (S) protein, other proteins, such as TMPRSS2 and TMPRSS11D, are crucial for virus entry into host cells and emerge as promising targets for antiviral therapies. We also identified proteins, such as KCND2 and TMEM64, which, although not directly involved in viral entry, are associated with clinical complications of COVID-19. This study provides valuable insights into virus protein-protein interactions, proposing new targets for therapeutic interventions and expanding understanding of the molecular mechanisms of infection which must be tested in bench and in vivo for confirmation.

References

- Arabi-Jeshvaghani, F., Javadi-Zarnaghi, F., & Ganjalikhany, M. R. (2023). Analysis of critical protein–protein interactions of SARS-CoV-2 capping and proofreading molecular machineries towards designing dual target inhibitory peptides. *Scientific Reports*, 13(1), 350.
- Bar-On, Y M, Flamholz, A., Phillips, R., & Milo, R. (2020). SARS-CoV-2 (COVID-19) by the numbers. *eLife*, 9, e57309.
- Bittmann, S., Weissenstein, A., Moschüring-Alieva, E., Bittmann, L., Luchter, E., & Villalon, G. (2020). The role of TMPRSS2 and TMPRSS2-inhibitors in cell entry mechanism of COVID-19. *Journal of Regenerative Biology and Medicine*, 2(3), 1-3.
- Borza, R., Salgado-Polo, F., Moolenaar, W H., & Perrakis, A. (2022). Structure and function of the ecto-nucleotide pyrophosphatase/phosphodiesterase (ENPP) family: Tidying up diversity. *The Journal of Biological Chemistry*, 298(2), 101526.
- Bowe, B., Xie, Y., & Al-Aly, Z. (2023). Postacute sequelae of COVID-19 at 2 years. *Nature Medicine*, 29, 2347–2357.
- Bugge, T H, Antalis, T M, & Wu, Q. (2009). Type II transmembrane serine proteases. *Journal of Biological Chemistry*, 284(35), 23177-23181.
- Bussani, R., Zentilin, L., Correa, R., Colliva, A., Silvestri, F., Zacchigna, S., & Giacca, M. (2023). Persistent SARS-CoV-2 infection in patients seemingly recovered from COVID-19. *The Journal of Pathology*, 259(3), 254-263.
- Chakraborty, C., Sharma, A R., Bhattacharya, M., Sharma, G., & Lee, S S. (2021). Immunoinformatics approach for the identification and characterization of T cell and B cell epitopes towards the peptide-based vaccine against SARS-CoV-2. *Archives of Medical Research*, 52(4), 362-370.
- Chou, P. Y., & Fasman, G. D. (1978). Prediction of the secondary structure of proteins from their amino acid sequence. *Advances in Enzymology and Related Areas of Molecular Biology*, 47, 45–148.
- Díaz-Castrillón, F. J., & Toro-Montoya, A I. (2020). SARS-CoV-2/COVID-19: el virus, la enfermedad y la pandemia. *Medicina & Laboratorio*, 24(3), 183-205.
- Devi, Y D, Goswami, H B, Konwar, S., Doley, C., Dolley, A., Devi, A., & Namsa, N D. (2021). Immunoinformatics mapping of potential epitopes in SARS-CoV-2 structural proteins. *PLOS ONE*, 16(11), e0258645.
- Disser, N P, De Micheli, A J, Schonk, M M, Konnaris, M A, Piacentini, A N, Edon, D L, Toresdahl, B G, Rodeo, S A, Casey, EK, & Mendias, CL. (2020). Musculoskeletal consequences of COVID-19. *Journal of Bone and Joint Surgery*, 102(14), 1197-1204.
- Eckerle, I., & Meyer, B. (2020). SARS-CoV-2 seroprevalence in COVID-19 hotspots. *The Lancet*, 396(10250), 514-515.
- Emini, E A, Hughes, J V, Perlow, D S, & Boger, J. (1985). Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *Journal of Virology*, 55(3), 836–839.
- Fang, Z., Marshall, C B, Yin, J C, Mazhab-Jafari, M T, Gasmi-Seabrook, G M, Smith, M J, Nishikawa, T., Xu, Y., Neel, B G, & Ikura, M. (2016). Biochemical classification of disease-associated mutants of RAS-like protein expressed in many tissues (RIT1). *Journal of Biological Chemistry*, 291(30), 15641-15652.
- Fathollahi, M., Motamedi, H., Hossainpour, H., Abiri, R., Shahlaei, M., Moradi, S., & Alvandi, A. (2023). Designing a novel multi-epitopes pan-vaccine against SARS-CoV-2 and seasonal influenza: In silico and immunoinformatics approach. *Journal of Biomolecular Structure and Dynamics*, 1-24.
- Grifoni, A., Sidney, J., Vita, R., Peters, B., Crotty, S., Weiskopf, D., & Sette, A. (2021). SARS-CoV-2 human T cell epitopes: Adaptive immune response against COVID-19. *Cell Host & Microbe*, 29(7), 1076-1092.
- Han, Q., Zheng, B., Daines, L., & Sheikh, A. (2022). Long-term sequelae of COVID-19: A systematic review and meta-analysis of one-year follow-up studies on post-COVID symptoms. *Pathogens*, 11(2), 269.
- Hook, V., Funkelstein, L., Wegrzyn, J., Bark, S., Kindy, M., & Hook, G. (2012). Cysteine cathepsins in the secretory vesicle produce active peptides: Cathepsin L generates peptide neurotransmitters and cathepsin B produces beta-amyloid of Alzheimer's disease. *Biochimica et Biophysica Acta*, 1824(1), 89-104

- Huseynov, A., Akin, I., Duerschmied, D., & Scharf, R E. (2023). Cardiac arrhythmias and COVID-19: A scientific statement from the German Cardiac Society (DGK) Working Group on arrhythmias. *Heart*, 109(5), 369-376.
- Jespersen, M C, Peters, B., Nielsen, M., & Marcatili, P. (2017). BepiPred-2.0: Improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Research*, 45(W1), W24–W29.
- Kampf, G., Brüggemann, Y., Kaba, H E J, Steinmann, J., Pfaender, S., Scheithauer, S., & Steinmann, E. (2020). Potential sources, modes of transmission and effectiveness of prevention measures against SARS-CoV-2. *Journal of Hospital Infection*, 106(4), 678-697.
- Karplus, P. A., & Schulz, G. E. (1985). Predição da flexibilidade de cadeia em proteínas - Uma ferramenta para a seleção de antígenos peptídicos. *Naturwissenschaften*, 72, 212-213.
- Karplus, P A., & Schulz, G E. (1985). Prediction of chain flexibility in proteins: A tool for the selection of peptide antigens. *Naturwissenschaften*, 72(4), 212-213.
- Kishimoto, M., Uemura, K., Sanaki, T., Sato, A., Hall, W W, Kariwa, H., Orba, Y., Sawa, H., & Sasaki, M. (2021). Tmprss11d and Tmprss13 activate the SARS-CoV-2 spike protein. *Viruses*, 13(3), 384.
- Kolaskar, A S, & Tongaonkar, P. C. (1990). A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Letters*, 276(1-2), 172–174.
- Kumar, P., Kumar, A., Garg, N., & Giri, R. (2023). An insight into SARS-CoV-2 membrane protein interaction with spike, envelope, and nucleocapsid proteins. *Journal of Biomolecular Structure and Dynamics*, 41(3), 1062-1071.
- Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., & Moore, H. F. (2013). The genotype-tissue expression (GTEx) project. *Nature Genetics*, 45(6), 580-585.
- Majra, D., Benson, J., Pitts, J., & Stebbing, J. (2021). SARS-CoV-2 (COVID-19) superspreader events. *Journal of Infection*, 82(1), 36-40.
- Mehandru, S., & Merad, M. (2022). Pathological sequelae of long-haul COVID. *Nature Immunology*, 23, 194–202.
- Miao, G., Zhao, H., Li, Y., ... & Zhong, Q. (2021). ORF3a of the COVID-19 SARS-CoV-2 virus blocks the HOPS complex-mediated assembly of the SNARE complex required for the formation of autolysosomes. *Developmental Cell*, 56(4), 427-442.e5.
- Mukherjee, S., Tworowski, D., Detroja, R., & Frenkel-Morgenstern, M. (2020). Immunoinformatics and structural analysis for identification of immunodominant epitopes in SARS-CoV-2 as potential vaccine targets. *Vaccines*, 8(2), 290.
- Navish, A A, & Uthayakumar, R. (2023). An exploration on the topologies of SARS-CoV-2/human protein-protein interaction network. *Journal of Biomolecular Structure and Dynamics*, 41(13), 6313-6325.
- Ozger, Z B. (2023). A robust protein language model for SARS-CoV-2 protein–protein interaction network prediction. *Artificial Intelligence in Medicine*, 142, 102574.
- Parker, J M, Guo, D., & Hodges, R S. (1986). New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: Correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. *Biochemistry*, 25(19), 5425–5432.
- Pellequer, J L, Westhof, E., & Van Regenmortel, M H. (1993). Correlation between the location of antigenic sites and the prediction of turns in proteins. *Immunology Letters*, 36(1), 83–99.
- Saha, S A, Russo, A M, Chung, M K, Deering, T F, Lakkireddy, D., & Gopinathannair, R. (2022). COVID-19 and cardiac arrhythmias: A contemporary review. *Current Treatment Options in Cardiovascular Medicine*, 24(6), 87-107.
- Shomuradova, A S, Vagida, M S, Sheetikov, S A, Zornikova, K V, Kiryukhin, D., Titov, A., & Efimov, G A. (2020). SARS-CoV-2 epitopes are recognized by a public and diverse repertoire of human T cell receptors. *Immunity*, 53(6), 1245-1257.
- Shulla, A., Heald-Sargent, T., Subramanya, G., Zhao, J., Perlman, S., & Gallagher, T. (2011). A transmembrane serine protease is linked to the severe acute respiratory syndrome coronavirus receptor and activates virus entry. *Journal of Virology*, 85(2), 873-882.
- Szabo, R., & Bugge, T H. (2008). Type II transmembrane serine proteases in development and disease. *The International Journal of Biochemistry & Cell Biology*, 40(6-7), 1297-1316.
- Tharappel, A M, Samrat, S K, Li, Z., & Li, H. (2020). Targeting crucial host factors of SARS-CoV-2. *ACS Infectious Diseases*, 6(11), 2844-2865.
- Turk, B., Turk, D., & Turk, V. (2000). Lysosomal cysteine proteases: More than scavengers. *Biochimica et Biophysica Acta*, 1477(1-2), 98–111.
- Weingarten-Gabbay, S., Chen, D Y, Sarkizova, S., Taylor, H B, Gentili, M., Hernandez, G M, & Sabeti, P C. (2024). The HLA-II immunopeptidome of SARS-CoV-2. *Cell Reports*, 43(1).
- Yarmarkovich, M., Warrington, J. M., Farrel, A., & Maris, J. M. (2020). Identification of SARS-CoV-2 vaccine epitopes predicted to induce long-term population-scale immunity. *Cell Reports Medicine*, 1(3).
- Zhou, Y., Liu, Y., Gupta, S., Paramo, M I, Hou, Y., Mao, C., & Yu, H. (2023). A comprehensive SARS-CoV-2–human protein–protein interactome reveals COVID-19 pathobiology and potential host therapeutic targets. *Nature Biotechnology*, 41(1), 128-139.