

The Long-COVID Syndrome: Neoantigens as Driving Force for the Onset of Autoimmune Diseases

Rolf Marschalek^{1,*}

¹Goethe-University Frankfurt, Inst. Pharm Biology, Max-von-Laue-Str. 9, 60438 Frankfurt/Main, Germany

*Correspondence should be addressed to Rolf Marschalek, rolf.marschalek@em.uni-frankfurt.de

Received date: February 14, 2025, **Accepted date:** March 08, 2025

Citation: Marschalek R. The Long-COVID Syndrome: Neoantigens as Driving Force for the Onset of Autoimmune Diseases. J Cell Immunol. 2025;7(1):26-31.

Copyright: © 2025 Marschalek R, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

In the fifth year of the pandemic, SARS-CoV-2 variants continue to unveil new insights into particular mechanism of human immunology but also new clues on the interaction of SARS-CoV-2 proteins with host proteins. Nearly all yet known SARS-CoV-2 variants had the ability to cause a post-infection disease termed the Long COVID (LC) syndrome. The LC syndrome encompasses various autoimmune diseases ranging from localized organ damage to systemic effects like the myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS). The complexity of these diverse disease conditions is challenging for physicians. Several studies indicate that SARS-CoV-2 infections may trigger the production of functional autoantibodies (fAAB's), a fact that could well explain the yet known post-infection disease conditions. Therefore, the human target proteome of both SARS-CoV-2 viral proteases, NSP3/M^{Pro} and NSP5/3CL^{Pro}, was investigated. Based on these results, SARS-CoV-2 proteases target ~6,500 human proteins, while well-known common viruses were ranging only between 2 to less than 100 proteins. Thus, the capability of SARS-CoV-2 to generate reactivity against these neoantigens in infected cells is outstanding. Here, the hypothesis is posed that the neoantigen production during a SARS-CoV-2 infection is the major cause for the onset of the LC syndrome, which is currently affecting ~10% of all infected individuals worldwide.

Keywords: Long COVID, SARS-CoV-2, Proteases M^{pro} and 3-CL^{pro}, Neoantigens, Auto antibodies

Introduction

The SARS-CoV-2 pandemic, which began in early 2020, evolved in distinct phases [<https://www.worldometers.info/coronavirus/>]. In 2020, our limited knowledge about this novel virus in combination with severe hospitalized cases unleashed its lethal potential: hyper-inflammation, interferon signaling inhibition, and the emergence of functional autoantibodies causing widespread organ damage, vascular damage in combination with thrombotic events, and neurological symptoms [1]. In December 2020, a mortality rate around ~3% was observed globally. At the end of 2020, a new Alpha variant (B.1.1.7) was first observed in the UK, while the Beta variant (B.1.351) was diagnosed in Africa.

In 2021, the introduction of four vaccines led to widespread inoculation efforts, with e.g. over 13 billion vaccinations

worldwide. The observed vaccine side effects were manageable, such as myocarditis (21 events/million mRNA-based jabs) and increased thrombotic events (37 events/million vector-based jabs) [2]. In 2021 also new variants like the Gamma variant (B.1.1.28) in Brazil, while the Delta variant (B.1.1.7) was emerging in India, and later in Europe.

In 2022, the new Omicron variant (B.1.1.529) emerged from South Africa [3,4], possibly originating from an asymptomatic long-term infection in an immunosuppressed individual [5,6]. With a highly infectious yet less lethal nature, Omicron posed challenges due to immune evasion, leading to a first COVID_19 vaccine adjustment. Despite late arrival, bivalent vaccines targeting original and Omicron strains (BA.1/2, BA.4/5) were introduced in Sept 2022, but only a 2-3% of the population were using the newly arrived vaccines against the Omicron variants [7], and therefore, a significant portion of the

population worldwide became repeatedly infected (estimated between 50-60%).

In 2023, novel side effects of Omicron infections were observed, including prolonged or recurrent illnesses and immunosuppression. A first monovalent vaccine against novel Omicron subvariants was developed (XBB.1.5), though their usage diminished as the pandemic waned. Despite its reduced lethality (~0.46% worldwide), the permanently altering virus displays enhanced infectivity and compromised immune responses in infected individuals [8]. Studies revealed a heightened risk of Long COVID with repeated infections [9], though vaccination showed promise in mitigating this risk, emphasizing the importance of targeted vaccination strategies [10,11]. However, many questions concerning the molecular mechanism for the onset Long COVID are still unanswered.

In order to explain the LC syndrome, immunology offers several explanations that may seem straightforward, while treating physicians face a plethora of puzzling symptoms. Long COVID may manifest with a vascular system damage [12,13], involves nearly all organs [14], and displays clear signs of a dysregulated immune system. Patients exhibit systemic inflammatory responses driven by a misled innate immune system, particularly the complement system [15]. Quite worrying was the observed loss of CD8⁺ T cells and erratic humoral immune reactions [8,16]. Neurocognitive issues may stem from inhibitory effects of specific SARS-CoV-2 proteins on the interferon signaling pathway [17], the surge in autoantibody production [18], and reactivation of latent Epstein-Barr viruses [19].

Where do all these autoantibodies originate? The first hint came from a study on SARS-CoV-2-infected hamsters in 2020 but was later also observed for COVID-19 patients: 90% of all antibodies post-infection target the body's own proteins rather than viral proteins. This isn't a coincidence. The SARS-CoV-2 virus carries two protease genes in its genome, NSP3 and NSP5, encoding the two viral proteases M^{Pro} and 3CL^{Pro}, respectively [20]. Both proteases cleave the viral polyprotein into 16 individual proteins, crucial for viral replication. Their "consensus cleavage sites" (CCS), "L-GG•A/K" and "LQ•A/S", facilitate this process [21]. Consequently, cells expressing these viral proteases may not only cleave the viral polyprotein but also all host proteins that display these CCS in infected cells.

Material and Methods

Motif search

Using a bioinformatics tool (Motif Search), the consensus cleavage sequences (CCS) of these proteases were scanned against the human proteome (~72,387 proteins in human KEGG database; [24,25]). Motif Search is available from the website

<https://www.genome.jp/tools/motif/MOTIF2.html>. By using the following parameters, the target proteome of all proteases listed in (Figure 1A) allowed us to determine precisely how many human proteins are potentially hydrolyzed by the 2 SARS-CoV-2 proteases. These data have been exported from the website and used for further analysis. For further details, please see the Supplementary data file.

Results

The neoantigenic potential of SARS-CoV-2

To address this issue, the CCS of known proteases from various viruses (e.g., VZV, HSV1/2, HSV6/7, CMV, EBV) were used to compare it with the CCS of both proteases encoded by SARS-CoV-2. In addition, known proteases from rhinoviruses, flaviviruses, hepaciviruses, and parasitic protists like Plasmodia, Trypanosoma, and Leishmania were investigated. This provided an initial insight into the potential immunological threat posed by the two SARS-CoV-2 proteases. The gathered CCS are detailed in the Table shown in Figure 1A, with amino acid residues depicted as single letters and identified target proteins within the human proteome (KEGG database; ~72,000 proteins) and listed as complete target proteome in the Supplementary Excel files.

This analysis unveils an intriguing contrast across various viral and human parasites. Latent viruses, typically well-controlled by the immune system in healthy individuals, exhibit a very limited number of cellular proteins that are potentially hydrolyzed by corresponding proteases. For instance, EBV targets only 2 human proteins, HSV6/7 only 3, and Herpesviridae 8 proteins (by the "release protease") or 50 proteins (by the "maturation protease"), while CMV targets 68 human proteins. The risk of dangerous neoantigen formation in these infections appears low, suggesting that "mimotopes" (viral protein fragments resembling human proteins) may play a much more significant role for autoimmune disease development, as seen e.g. in EBV-associated autoimmune diseases (like e.g. systemic lupus erythematosus, Sjögren's syndrome or symptoms associated with myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS). By contrast, the findings for SARS-CoV-2 are striking: NSP3/M^{Pro} protease targets 602 human proteins, while the main protease NSP5/3CL^{Pro} recognizes 6,439 human proteins (Figure 1B). This extensive targeting of human proteins likely contributes to the development of individual autoimmune diseases following SARS-CoV-2 infection.

This analysis is substantiated by a recently published study estimating approximately 4,460 human proteins as targets of SARS-CoV-2 proteases by using bioinformatic algorithms [25]. Conversely, rhinoviruses target around half of the human proteome (11,577 proteins), and flaviridae, hepaciviruses, and parasitic protists potentially target nearly the entire human proteome (15,133 to 20,516 proteins for the former, and 7,198

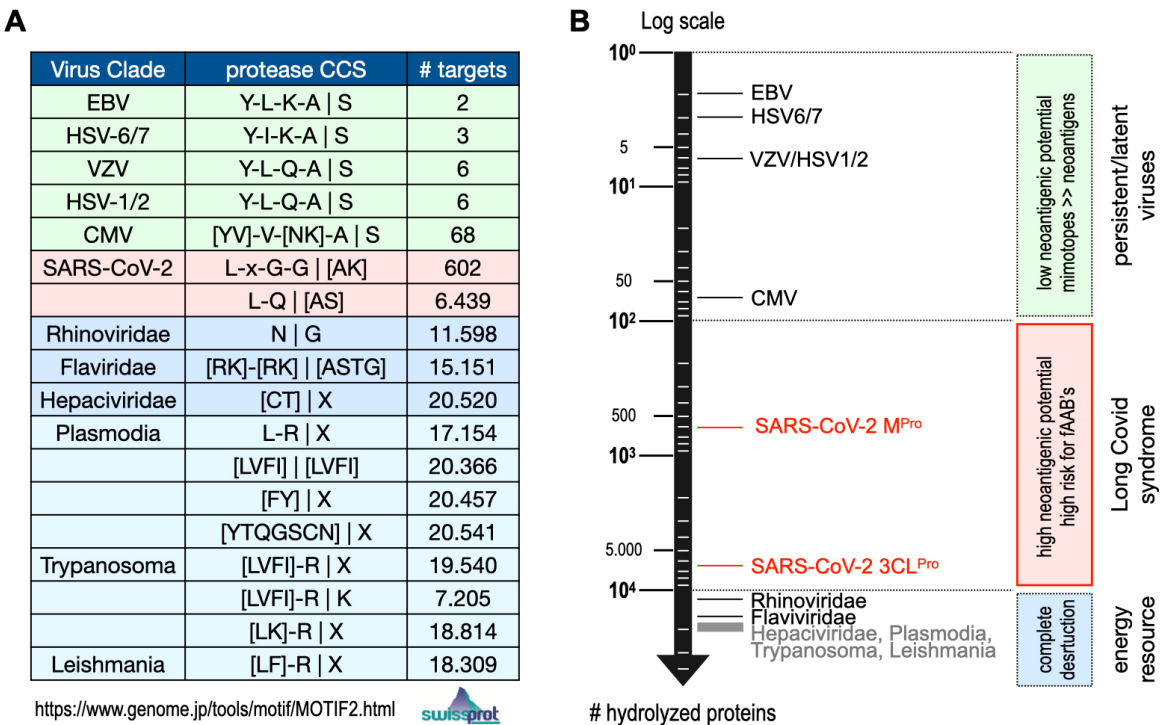


Figure 1. The neoantigen potential of proteases from various virus clades and parasitic protists. **A.** The CCS for different viruses and parasites are presented in the single-letter amino acid code, with hydrolysis occurring between the 2 amino acids separated by a vertical bar. If multiple amino acids were possible at one position, they were shown in square brackets. **B.** Target proteins of different proteases deriving either from different viruses or parasitic pathogens were categorized into 3 groups. Latent viruses with few target proteins (<100), and SARS-CoV-2 displaying a large number of target proteins (602 and 6,439). Known proteases of rhinoviruses, flaviviruses, and hepaciviruses target many proteins, ranging from 11,598 to 20,520. Proteases of parasitic pathogens target between 7,205 and 20,541 proteins of the human proteome.

to 20,453 for the latter), which likely serves as energy resource in their respective parasitic life cycles. Thus, a complete destruction of the proteome of infected cells likely diminishes the immunological significance of these proteolytic products, apart from periodical fever episodes upon the lytic disruption of these infected cells (**Figure 1B**).

Infections with SARS-CoV-2 lead to a critical scenario within infected cells: firstly, infected cells suffer from functional impairments of cellular processes due to the known actions of virus-encoded proteins and their specific interaction with host proteins [26]. However, functional impairment also occurs due to the loss of essential proteins which are hydrolyzed by the two viral proteases. When the “digested content” of cells is presented as neoantigens on the cell surface of infected cells, it will potentially be recognized as “foreign” by CD8⁺ T cells and also presented to CD4⁺ T cells by professional antigen-presenting cells. Thus, an erroneous immune response is likely triggered, leading to organ damaging cytotoxic T-cells as well as autoantibody formation. Also, the fact that persistent SARS-CoV-2 viruses have been identified in Long COVID patients suggests a fundamental immunological challenge

over extended periods, contributing to ongoing health issues [27,28].

The digested target proteome is listed in the accompanying **Supplementary data files**. The overlap between both target proteomes (602 vs. 6439) is rather small and comprises only few proteins, of which one is the IL-12Rβ2 chain. IL-12Rβ1 and IL-12Rβ2 are forming the IL-12R, and IL12Rβ2 and gp130 form the IL-35 receptor, both signaling either to STAT4 (via JAK2/TYK) or to STAT4 and STAT1 (via JAK1/JAK2). STAT4 and JAK2 are both targets of NSP5, while STAT1 is not. But the kinase TYK2 is target of both, NSP3 and NSP5, respectively. Thus, the important IL-12 signaling is compromised in infected cells. IL-12 signaling represents the initial signal deriving from innate immune cells to trigger an appropriate systemic interferon response (type I and type II). This is just a simple example to demonstrate the importance of analyzing the target proteome of both proteases of SARS-CoV-2 in more detail. **Figure 2A** displays the scenario for the IL12 signaling pathway, but summarizes also some examples of other affected receptors, signal transducers and targeted transcription factors (**Figure 2B**).

Discussion

Based on these data, the LC syndrome is a clear autoimmune disorder that exhibits with various individual outcomes, most likely because of the different ability to present neoantigens on individual MHC's. This makes it nearly impossible for treating physicians to find a standard operating procedure for the treatment of affected patients. What other options might be available for Long COVID patients? With the current knowledge, there are limited choices. Surprisingly, vaccination against the Omicron variant may not only prevent severe disease or death but also secondary conditions like LC syndrome [10,11], making them the most cost-effective preventive measure. Alternatively, an oral protease inhibitor like Paxlovid or other potent protease inhibitors designed against NSP3 or NSP5, like e.g. the preclinical drug ML2006a4 developed at the Stanford University (CA, USA), can partially mitigate immune-damaging reactions post-infection. Another potential treatment stems from an ongoing clinical study with the aptamer BC007 from the company "Berlin Cures". This small DNA aptamer blocks functional autoantibodies (fAABs) targeting G protein-coupled receptors (GPCRs), which are implicated in various autoimmune diseases and Long COVID [29]. By binding directly to the CDR loops of these autoantibodies, BC007 inhibits their agonistic effects on many human GPCRs, which regulate critical organ functions. The presence of fAABs can profoundly affect physiology and lead

to diverse symptoms, including heart problems, high blood pressure, diabetes type II, dementia, and chronic fatigue syndrome. The outcomes of these studies will determine if BC007 offers relief or even a cure for these conditions, provided the existing damage is not irreversible.

In conclusion, preventing SARS-CoV-2 Omicron infections may help to prevent also the development of the LC syndrome. Several studies have shown that vaccinations after infections may lower the risk of developing such autoimmune diseases by roughly 50%. In case of an infection, the inhibition of the main protease NSP5 by small molecules will definitively help to lower also the risk for developing autoimmune diseases.

Besides the above-mentioned precautions before (variant-specific Omicron vaccines) and during infection with SARS-CoV-19 (inhibitors of proteases to block virus replication), there are only very few other possibilities. One potential treatment is the use of a small DNA aptamer named BC007 which somehow neutralizes fAAB's directed against G-protein coupled receptors (GPCR's). There are currently two studies in Germany (Berlin Cure's and the Charité in Berlin, Germany) which are trying to demonstrate the effectiveness of this small DNA molecule in Long COVID patients. A previous study has already demonstrated that BC007 is effective in several autoimmune diseases [29]. Apart from these studies there are currently only very few other options, e.g. to wash out these

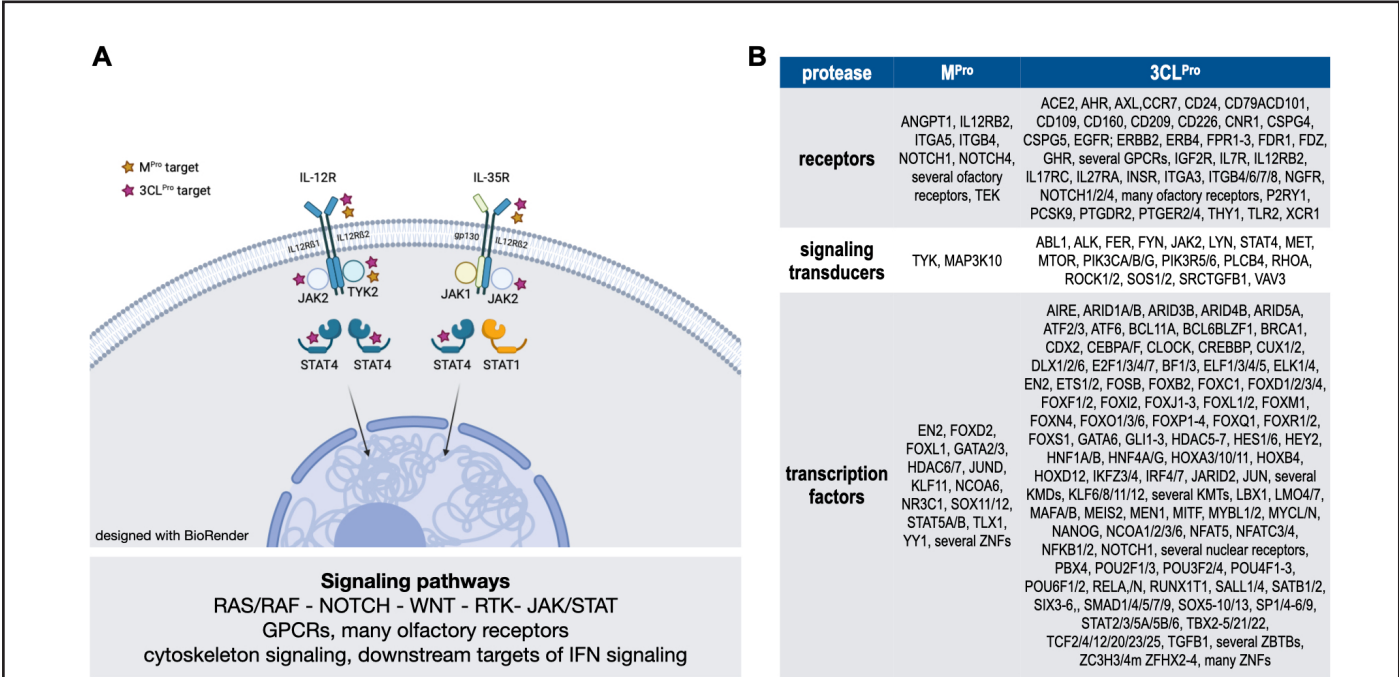


Figure 2. Exemplary effect of SARS-CoV-2 proteases M^{Pro} and 3CL^{Pro}. **A.** The IL-12 and IL-35 signaling pathways are classical pathways via the associated kinases JAK1, JAK2, and TYK2, involving the STAT factors STAT4 and STAT1. Except JAK1 and STAT1 and gp130, all involved molecules are targeted by the two SARS-CoV-2 proteases M^{Pro} and 3CL^{Pro}. **B.** Table summarizing in part the identified receptors, signal transducers and transcription factors for the two proteases M^{Pro} and 3CL^{Pro}, respectively. A complete list of targeted proteins is available in the Supplementary data files "NSP3_602targets.xlsx" and NSP5_6439targets.xlsx", respectively.

fAAB's from the blood (therapeutic apheresis), however, as the B-cells producing these fAAB's are still in place, it is only a question of time when the Long-COVID disease phenotype will come back.

Observation studies with single individuals that were additionally vaccinated after Long COVID symptoms have already developed, showed a partial response in about 50% of the tested people. However, pre-existing co-morbidities may limit such a strategy to use vaccination (after prior infection) as a therapy.

Therefore, the best way to avoid Long-COVID are disease-preventive strategies, which means that people have to vaccinate every year with novel vaccines directed against current omicron strains. This can be done either with the available mRNA vaccines or classical vaccines which are also on the market.

In conclusion, this pandemic is not yet over, as we still have the evolution of novel omicron strains (currently 4 per year), that are still infecting people with the consequence of dying patients or the development of individual autoimmune diseases. Thus, future vaccination strategies are the way of choice and may help to prevent the individual health damage in our societies and to reduce the costs for health systems and economics.

CRediT Authorship Contribution Statement

Rolf Marschalek: Writing – review & editing, Validation, Supervision, Project administration, Funding acquisition, Formal analysis, Conceptualization.

Declaration of Competing Interest

The author declares to have no competing of interest.

Acknowledgements

This work has been in part funded by a grant from the Wilhelm Sander Stiftung (2022.070.1) and a grant from the Deutsche Jose Carreras Leukemia Stiftung (DJCLS 07 R/2022).

Supplementary Materials

A Supplementary data file and 2 Excel files to this article can be found online.

References

1. Davis HE, McCorkell L, Vogel JM, Topol EJ. Long COVID: major findings, mechanisms and recommendations. *Nat Rev Microbiol.* 2023 Mar;21(3):133-46.
2. PEI vaccine security report 23.12.2021: <https://www.pei.de/SharedDocs/Downloads/DE/newsroom/dossiers/>

[sicherheitsberichte/sicherheitsbericht-27-12-20-bis-30-11-21.pdf?blob=publicationFile&v=9](https://www.sicherheitsberichte.de/sicherheitsbericht-27-12-20-bis-30-11-21.pdf?blob=publicationFile&v=9)

3. Viana R, Moyo S, Amoako DG, Tegally H, Scheepers C, Althaus CL, et al. Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. *Nature.* 2022 Mar;603(7902):679-86.
4. Sonnleitner ST, Prelog M, Sonnleitner S, Hinterbichler E, Halbfurter H, Kopecky DBC, et al. Cumulative SARS-CoV-2 mutations and corresponding changes in immunity in an immunocompromised patient indicate viral evolution within the host. *Nat Commun.* 2022 May 10;13(1):2560.
5. Scherer EM, Babiker A, Adelman MW, Allman B, Key A, Kleinhenz JM, et al. SARS-CoV-2 Evolution and Immune Escape in Immunocompromised Patients. *N Engl J Med.* 2022 Jun 23;386(25):2436-8.
6. Guilbaud R, Franco Yusti AM, Leducq V, Zafilaza K, Bridier-Nahmias A, Todesco E, et al. Higher Levels of SARS-CoV-2 Genetic Variation in Immunocompromised Patients: A Retrospective Case-Control Study. *J Infect Dis.* 2024 Apr 12;229(4):1041-9.
7. RKI dashboard: <https://impfdashboard.de/en/>
8. Yin K, Peluso MJ, Luo X, Thomas R, Shin MG, Neidلمان J, et al. Long COVID manifests with T cell dysregulation, inflammation and an uncoordinated adaptive immune response to SARS-CoV-2. *Nat Immunol.* 2024 Feb;25(2):218-25.
9. Bowe B, Xie Y, Al-Aly Z. Acute and postacute sequelae associated with SARS-CoV-2 reinfection. *Nat Med.* 2022 Nov;28(11):2398-405.
10. Lundberg-Morris L, Leach S, Xu Y, Martikainen J, Santosa A, Gisslén M, et al. Covid-19 vaccine effectiveness against post-covid-19 condition among 589722 individuals in Sweden: population-based cohort study. *BMJ.* 2023 Nov 22;383: e076990.
11. Byambasuren O, Stehlik P, Clark J, Alcorn K, Glasziou P. Effect of covid-19 vaccination on long covid: systematic review. *BMJ Med.* 2023 Feb 1;2(1): e000385.
12. Smadja DM, Mentzer SJ, Fontenay M, Laffan MA, Ackermann M, Helms J, et al. COVID-19 is a systemic vascular hemopathy: insight for mechanistic and clinical aspects. *Angiogenesis.* 2021 Nov;24(4):755-88.
13. Evans PC, Rainger GE, Mason JC, Guzik TJ, Osto E, Stamataki Z, et al. Endothelial dysfunction in COVID-19: a position paper of the ESC Working Group for Atherosclerosis and Vascular Biology, and the ESC Council of Basic Cardiovascular Science. *Cardiovasc Res.* 2020 Dec 1;116(14):2177-84.
14. Gan R, Rosoman NP, Henshaw DJE, Noble EP, Georgius P, Sommerfeld N. COVID-19 as a viral functional ACE2 deficiency disorder with ACE2 related multi-organ disease. *Med Hypotheses.* 2020 Nov; 144:110024.
15. Cervia-Hasler C, Brüningk SC, Hoch T, Fan B, Muzio G, Thompson RC, et al. Persistent complement dysregulation with signs of thromboinflammation in active Long Covid. *Science.* 2024 Jan 19;383(6680):eadg7942.

16. Diao B, Wang C, Tan Y, Chen X, Liu Y, Ning L, et al. Reduction and Functional Exhaustion of T Cells in Patients with Coronavirus Disease 2019 (COVID-19). *Front Immunol*. 2020 May 1; 11:827.
17. Lei X, Dong X, Ma R, Wang W, Xiao X, Tian Z, et al. Activation and evasion of type I interferon responses by SARS-CoV-2. *Nat Commun*. 2020 Jul 30;11(1):3810.
18. Dobrowolska K, Zarębska-Michaluk D, Poniedziałek B, Jaroszewicz J, Flisiak R, Rzymiski P. Overview of autoantibodies in COVID-19 convalescents. *J Med Virol*. 2023 Jun;95(6): e28864.
19. Altmann DM, Whettlock EM, Liu S, Arachchillage DJ, Boyton RJ. The immunology of long COVID. *Nat Rev Immunol*. 2023 Oct;23(10):618-34.
20. Kreye J, Reincke SM, Kornau HC, Sánchez-Sendin E, Corman VM, Liu H, et al. A Therapeutic Non-self-reactive SARS-CoV-2 Antibody Protects from Lung Pathology in a COVID-19 Hamster Model. *Cell*. 2020 Nov 12;183(4):1058-69.e19.
21. Ullrich S, Nitsche C. SARS-CoV-2 Papain-Like Protease: Structure, Function and Inhibition. *Chembiochem*. 2022 Oct 6;23(19): e202200327.
22. Lee J, Kenward C, Worrall LJ, Vuckovic M, Gentile F, Ton AT, et al. X-ray crystallographic characterization of the SARS-CoV-2 main protease polyprotein cleavage sites essential for viral processing and maturation. *Nat Commun*. 2022 Sep 3;13(1):5196.
23. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res*. 2000;28:27-30.
24. Kanehisa M, Furumichi M, Sato Y, Matsuura Y, Ishiguro-Watanabe M. KEGG: biological systems database as a model of the real world. *Nucleic Acids Res*. 2025 Jan 6;53(D1): D672-7.
25. Prescott L. SARS-CoV-2 3CLpro whole human proteome cleavage prediction and enrichment/depletion analysis. *Comput Biol Chem*. 2022 Jun; 98:107671.
26. Gordon DE, Jang GM, Bouhaddou M, Xu J, Obernier K, White KM, et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature*. 2020 Jul;583(7816):459-68.
27. Chen B, Julg B, Mohandas S, Bradfute SB. RECOVER Mechanistic Pathways Task Force. Viral persistence, reactivation, and mechanisms of long COVID. *Elife*. 2023 May 4;12: e86015.
28. Proal AD, VanElzakker MB, Aleman S, Bach K, Boribong BP, Buggert M, et al. SARS-CoV-2 reservoir in post-acute sequelae of COVID-19 (PASC). *Nat Immunol*. 2023 Oct;24(10):1616-27.
29. Becker NP, Goettel P, Mueller J, Wallukat G, Schimke I. Functional autoantibody diseases: Basics and treatment related to cardiomyopathies. *Front Biosci (Landmark Ed)*. 2019 Jan 1;24(1):48-95.