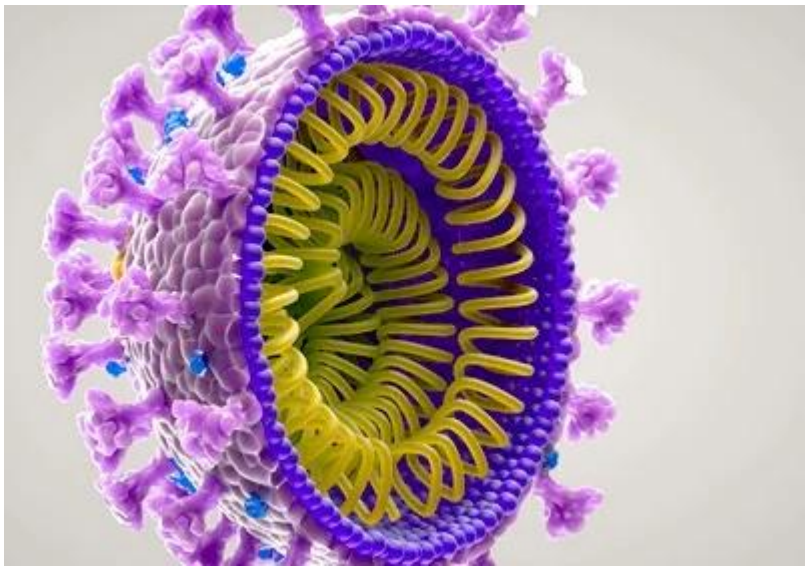
Molecular similarity can be defined as the theoretical explanation for sequence similarities (mainly in antigens) between two (or more) organisms. In this study, the author from India employed an integrative approach to identify antigenic sites in the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spike (S) receptor binding domain (RBD) and to evaluate the molecular similarities of antigenic sites predicted in the SARS-CoV-2 S RBD with proteins/antigens from other different organisms.

The author emphasized that computational biology approaches for epitope prediction, identification, and analysis are well-developed and have been proven highly successful in predicting and identifying both weak and strong antibody epitopes.

SARS-CoV-2 is an enveloped, positive-sense, single-stranded RNA virus. Its genome encodes four structural proteins, namely the spike (S), envelope (E), nucleocapsid (N), and membrane (M) protein. The S protein appears to be a major pathogenic factor that contributes to the unique pathogenesis of SARS-CoV-2. The S protein is a glycosylated homotrimer with each monomer composed of subunits S1 and S2, separated by host cell proteases. The S1 domain comprises the N-terminal domain (NTD), the receptor binding domain (RBD) with a receptor binding motif (RBM), and two C-terminal domains. The RBD in the S1 subunit recognizes human angiotensin-converting enzyme 2 receptor (ACE2) and is responsible for attachment to host cells.



About the study

The author used an integrated approach that included identifying potential antigenic sites and antigenic determinants in SARS-CoV-2 RBD based on its primary sequence and 3-

dimensional structure.

Results

The SARS-CoV-2 spike RBD was highly antigenic with nine potential surface antigenic sites. Seven of nine antigenic sites in the SARS-CoV-2 S protein RBD showed molecular similarities with 54 antigenic determinants found in fifteen pathogenic bacteria, parasites, and viruses, namely in twelve pathogenic bacterial species (*Mycobacterium tuberculosis, Mycobacterium leprae, Bacillus anthracis, Borrelia burgdorferi, Clostridium perfringens, Clostridium tetani, Helicobacter Pylori, Listeria monocytogenes, Staphylococcus aureus, Streptococcus pyogenes, Vibrio cholerae* and *Yersinia pestis*), two malarial parasites (*Plasmodium falciparum* and *Plasmodium knowlesi*) and influenza virus A.

Table from the original article of Dakal TC. List of antigenic proteins from different pathogenic microorganisms and virus showing sequence similarity with predicted antigenic sites' sequence in SARS-CoV-2 spike RBD.

S.No	UniProt ID	Protein Name	Protein Description	Organisms
1	O33084	ESXB_MYCLE	ESAT-6-like protein EsxB	Mycobacterium leprae (strain TN)
2	P01556	CHTB_VIBCH	Cholera enterotoxin subunit B	Vibrio cholerae serotype O1 (strain ATCC 39315/El Tor Inaba N16961)
3	P01558	ELTB_CLOPF	Heat-labile enterotoxin B chain	Clostridium perfringens
4	P02893	CSP_PLAFA	Circumsporozoite protein	Plasmodium falciparum
5	P02894	CSP_PLAKH	Circumsporozoite protein	Plasmodium knowlesi (strain H)
6	P03438	HEMA_I000X	Hemagglutinin	Influenza A virus (strain A/X-31 H3N2)
7	P03449	HEMA_I71A1	Hemagglutinin	Influenza A virus (strain A/Memphis/1/1971 H3N2)
8	P03452	HEMA_I34A1	Hemagglutinin	Influenza A virus (strain A/Puerto Rico/8/1934 H1N1)
9	P03453	HEMA_177AB	Hemagglutinin	Influenza A virus (strain A/USSR/90/1977 H1N1)

10	P03459	HEMA_I34A0	Hemagglutinin	Influenza A virus (strain A/Fowl plague virus/Rostock/8/1934 H7N1)
11	P03466	NCAP_I34A1	Nucleoprotein	Influenza A virus (strain A/Puerto Rico/8/1934 H1N1)
12	P04923	CRA_PLAFA	Circumsporozoite protein-related antigen	Plasmodium falciparum
13	P04926	EXP1_PLAFA	Malaria protein EXP-1	Plasmodium falciparum
14	P04934	MSP1_PLAFC	Merozoite surface protein 1	Plasmodium falciparum (isolate Camp/Malaysia)
15	P04958	TETX_CLOTE	Tetanus toxin	Clostridium tetani (strain Massachusetts/E88)
16	P08089	M6B_STRPY	M protein, serotype 6	Streptococcus pyogenes
17	P09621	CH10_MYCTU	10 kDa chaperonin	Mycobacterium tuberculosis (strain ATCC 25618/H37Rv)
18	P09621	CH10_MYCTO	10 kDa chaperonin	Mycobacterium tuberculosis (strain CDC 1551/Oshkosh)
19	P0A0L2	ETXA_STAAU	Enterotoxin type A	Staphylococcus aureus
20	P0A4V2	A85A_MYCTU	Diacylglycerol acyltransferase	Mycobacterium tuberculosis (strain ATCC 25618/H37Rv)
21	P0A4V2	A85A_MYCTO	Diacylglycerol acyltransferase	Mycobacterium tuberculosis (strain CDC 1551/Oshkosh)
22	P0A550	DPO1_MYCTU	DNA polymerase I	Mycobacterium tuberculosis (strain ATCC 25618/H37Rv)
23	P0A550	DPO1_MYCTO	DNA polymerase I	Mycobacterium tuberculosis (strain CDC 1551/Oshkosh)
24	P0A564	ESXA_MYCTU	6 kDa early secretory antigenic target	Mycobacterium tuberculosis (strain ATCC 25,618 / H37Rv)

25	P0A564	ESXA_MYCTO	6 kDa early secretory antigenic tar	Mycobacterium tuberculosis (strain CDC 1551/Oshkosh)
26	P0A5J0	LPQH_MYCTU	Lipoprotein LpqH	Mycobacterium tuberculosis (strain ATCC 25618/H37Rv)
27	P0A5J0	LPQH_MYCTO	Lipoprotein LpqH	Mycobacterium tuberculosis (strain CDC 1551/Oshkosh)
28	P0A5Q4	MP64_MYCTU	Immunogenic protein MPT64	Mycobacterium tuberculosis (strain ATCC 25618/H37Rv)
29	P0A5Q4	MP64_MYCTO	Immunogenic protein MPT64	Mycobacterium tuberculosis (strain CDC 1551/Oshkosh)
30	P11089	FLA1_BORBU	Flagellar filament 41 kDa core protein	Borrelia burgdorferi (strain ATCC 35210/B31/CIP 102532/DSM 4680)
31	P13423	PAG_BACAN	Protective antigen	Bacillus anthracis
32	P13830	RESA_PLAFF	Ring-infected erythrocyte surface antigen	Plasmodium falciparum (isolate FC27/Papua New Guinea)
33	P14013	OSPA_BORBU	Outer surface protein A	Borrelia burgdorferi (strain ATCC 35210/B31/CIP 102532/DSM 4680)
34	P14013	OSPA_BORBN	Outer surface protein A	Borrelia burgdorferi (strain N40)
35	P14013	OSPA_BORBZ	Outer surface protein A	Borrelia burgdorferi (strain ZS7)
36	P14916	URE23_HELPY	Urease subunit alpha	Helicobacter pylori (strain ATCC 700392/26695) (Campylobacter pylori)
37	P15712	PSTS1_MYCTU	Phosphate-binding protein PstS 1	Mycobacterium tuberculosis (strain ATCC 25618/H37Rv)
38	P15712	PSTS1_MYCTO	Phosphate-binding protein PstS 1	Mycobacterium tuberculosis (strain CDC 1551/Oshkosh)
39	P15917	LEF_BACAN	Lethal factor	Bacillus anthracis

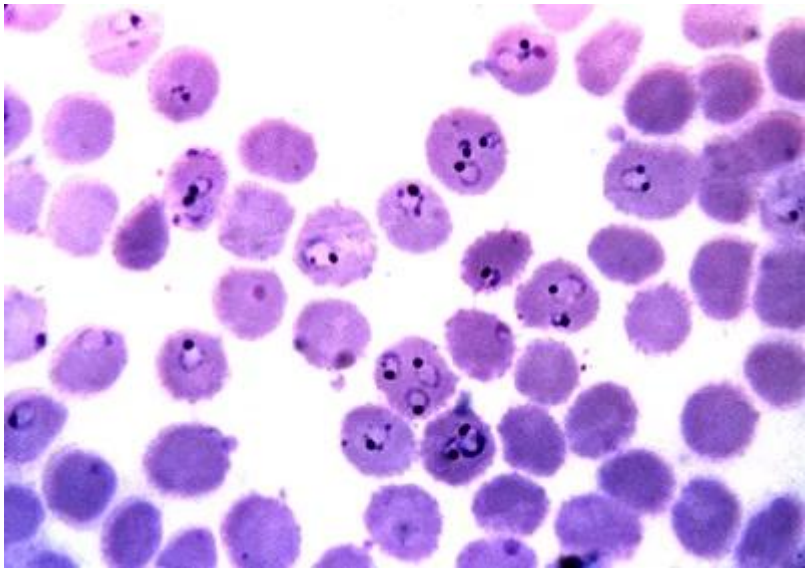
40	P16893	TRAP_PLAFA	Thrombospondin-related anonymous protein	Plasmodium falciparum
41	P19214	EBA1_PLAFC	Erythrocyte-binding antigen 175	Plasmodium falciparum (isolate Camp/Malaysia)
42	P21171	P60_LISMO	Probable endopeptidase p60	Listeria monocytogenes serovar 1/2a (strain ATCC BAA-679/EGD-e)
43	P23024	TCPA_VIBCL	Toxin coregulated pilin	Vibrio cholera
44	P24301	CH10_MYCLE	10 kDa chaperonin	Mycobacterium leprae (strain TN)
45	P25893			
46	P26948	CAF1_YERPE	F1 capsule antigen	Yersinia pestis
47	P31951	A85B_MYCLE	Diacylglycerol acyltransferase	Mycobacterium leprae (strain TN)
48	P31952	A85B_MYCTA	Diacylglycerol acyltransferase	Mycobacterium tuberculosis (strain ATCC 25177/H37Ra)
49	P31952	A85B_MYCTU	Diacylglycerol acyltransferase	Mycobacterium tuberculosis (strain ATCC 25618/H37Rv)
50	P31952	A85B_MYCTO	Diacylglycerol acyltransferase	Mycobacterium tuberculosis (strain CDC 1551/Oshkosh)
51	Q07337	OSPC_BORBU	Outer surface protein C	Borrelia burgdorferi (strain ATCC 35210/B31/CIP 102532/DSM 4680)
52	Q25893	Q25893_PLAFA	Liver stage antigen-1	Plasmodium falciparum
53	Q9U0P0	Q9U0P0_PLAFA	Liver stage antigen-3	Plasmodium falciparum
54	Q60153	TCPA_VIBCL	Toxin coregulatedpilin	Vibrio cholerae serotype O1 (strain ATCC 39315/El Tor Inaba N16961)

Most of the antigenic determinants found in pathogenic microorganisms that showed molecular similarity to the RBD's antigenic sites were toxins and factors with a diverse

functional role in pathogens. The antigenic determinants predicted from *Mycobacterium tuberculosis* play a role in immune recognition escape and ensure persistence in the host body, possibly in the latent stage. 10-kDa co-chaperonin (cpn10) led *in vitro* to bone weakness and fragility, while the MPT64 is involved in the persistence and survival of pathogens in host cells and tissues.

The antigenic sequences predicted from malaria parasites were circumsporozoite protein-related antigen, malaria protein EXP-1, thrombospondin-related anonymous protein, liver stage antigens -1 and -3, erythrocyte-binding antigen 175, and ring-infected erythrocyte surface antigen. The circumsporozoite proteins play an important role in the invasion of liver cells in humans. The antigenic determinants predicted from *Bacillus anthracis* are two of the three proteins that form the anthrax toxin. Five antigenic determinants were predicted from spirochete *Borrelia burgdorferi*. The antigenic sequences predicted from *Clostridium perfrigens* and *Clostridium tetani* were toxin proteins, as well as the enterotoxin A from *Staphylococcus aureus*, M protein from *Streptococcus pyogenes*, and toxins from *Vibrio cholerae*. Two proteins predicted from *Mycobacterium leprae* play a role in bacterial virulence and act as chaperones to prevent membrane lysis. One predicted antigenic site in SARS-CoV-2 RBD had molecular similarity with antigenic determinant from *Yersisnia pestis,* a capsule-like antigen, fraction 1 (F1) which has been implicated to be involved in the ability of bacteria to prevent uptake by macrophages. Certain antigenic determinants from influenza virus A play a role in the assembly of newly budded virions.

Interestingly, a study that examined the association between the cross-reactivity to SARS-CoV-2 S protein/RBD and exposure to *Plasmodium falciparum* infection in 741 pre-pandemic samples from eight malaria-endemic and non-endemic countries showed that people with confirmed acute malaria had more pronounced cross-reactivity than those previously exposed to malaria but without acute *Plasmodium falciparum* infection. Also, IgM but not IgG cross-reactivity was higher among uninfected individuals exposed to infection in malaria-endemic areas than among people from non-endemic settings. Importantly, there was no cross-reactivity between acute *Plasmodium falciparum* infection and other human coronaviruses or other SARS-CoV-2 proteins. (Lapidus, S *et al. Plasmodium* infection is associated with cross-reactive antibodies to carbohydrate epitopes on the SARS-CoV-2 Spike protein. *Sci Rep* 2022; 12, 22175.) https://doi.org/10.1038/s41598-022-26709-7



Plasmodium falciparum

The author of the present study stated that the findings of the molecular similarities between antigenic sites in the SARS-CoV-2 RBD and antigenic determinants of fifteen pathogens (bacterial, malarial, and viral species) are seriously alarming because their presence makes the SARS-CoV-2 more pathogenic than any other previously known coronavirus. Especially, antigenic determinants unique to SARS-CoV-2. As all seven predicted antigenic sites in SARS-CoV-2 RBD had molecular similarity with antigenic determinants from *Mycobacterium tuberculosis* and *Plasmodium falciparum*, the author suggested that COVID-19 patients should have at least symptoms of these two diseases.

The researcher also stated that molecular similarities between antigenic determinants of the SARS-CoV-2 RBD and highly potent antigenic determinants found in fifteen pathogens might explain numerous pathophysiological complications, like exacerbated innate and adaptive immune responses leading to hyper-activation of B-cells, T-cells (both cytotoxic T-cells and T-helper cells), dendritic cells, natural killer cells, and macrophage/monocyte lineage cells, the excessive release of cytokines that adversely affect numerous vital organs and multiple organ failure.

The author also speculated that individuals previously infected with or immunized/vaccinated against malaria, tuberculosis (TB), or other diseases caused by fifteen microorganisms with molecular similarities between their antigenic determinants and SARS-CoV-2 RBD, are expected to display a considerable degree of resistance against SARS-CoV-2 infection. The results of a recent study that investigated the immune response in people from malaria-endemic regions of Ghana, who were exposed to *Plasmodium*

falciparum and positive for SARS-CoV-2 has shown that modulation of immune response to SARS-CoV-2 after exposure to *Plasmodium falciparum* may contribute to reduced severity of COVID-19 among people living in malaria-endemic regions.
https://discovermednews.com/immune-response-to-sars-cov-2-after-exposure-plasmodium-falciparum/

One possible explanation is that the memory B or T cells previously generated by these microorganisms would be reactivated upon SARS-CoV-2 infection due to similar antigenic specificity and common antigenic determinants. The authors also suggested that the over-representation of antigenic determinants from *Mycobacterium tuberculosis* and *Plasmodium falciparum* in all antigenic sites suggests that anti-malarial and anti-TB drugs can be clinically beneficial for COVID-19 treatment.

This article was published in Immunobiology.

Journal Reference

Dakal TC. Antigenic sites in SARS-CoV-2 spike RBD show molecular similarity with pathogenic antigenic determinants and harbors peptides for vaccine development. Immunobiology 2021 Sep; 226(5): 152091. (Open Access)
https://www.sciencedirect.com/science/article/pii/S0171298521000395