Contents lists available at NCBI

The American Journal of Science and Medical Research

Journal homepage: https://ajsmrjournal.com/

Research Article

Conserved SARS-CoV-2 Viral Peptides as Potential Prophylactic and Therapeutic Targets

Radhakrishna Muttineni[1γ]*, Kalyani Putty[2γ]*, Jhansi Siripuram[1γ], Sreenidhi Ramamoorthy[3γ], Sowmyya T[4], Binitha RN[5], Amoolya Rao Amaravadhi[6], Sandra Kathott Prakash[3γ] , Aravind Vemula[1], Jaslin Panyam[1], Veena K[1], Meer Mushabbir Ali[1], Manasa Manda[1], Mohammad Younus[1], Pavan Kumar Muttineni[7]

[1]*Virus Research Laboratory, Department of Zoology, Osmania University, Hyderabad, India*
[2]*Department of Veterinary Biotechnology, College of Veterinary Science, Rajendranagar, PVNR Telangana Veterinary University, Hyderabad, India*
[3]*Department of Biotechnology and Bioinformatics, School of Life Sciences, University of Hyderabad, India*
[4]*Forensic Science Unit, Department of Chemistry, Osmania University, Hyderabad, India*
[5]*Department of Zoology, Mar Athanasius College (Autonomous), Kothamangalam, Kerala, India*
[6]*Department of General Medicine, Malla Reddy Institute of Medical Sciences, Hyderabad, India*
[7]*Department of General Medicine, Government Medical College and Hospital, Jayashankar Bhupalapally, India*
[γ]*Contributed equally*

Corresponding author:
E-mail: kalyaniputty@gmail.com; muttinenirk@gmail.com

https://dx.doi.org/10.5281/zenodo.18489357

Received: 15 December 2025
Accepted: 28 January 2026
Published: 5 February 2026

ISSN: 2377-6196
Published by AIRA

Keywords: *SARS-CoV-2; COVID-19; Structural proteins; Non-structural proteins*

ABSTRACT

The ongoing evolution of SARS-CoV-2, especially the emergence of heavily mutated variants like Omicron and its sub-lineages, has resulted in antigenic drift that diminishes the effectiveness of current first-generation vaccines, diagnostic tests, and treatments. This study employed a comprehensive immuno-informatics approach to identify highly conserved protein sequences from SARS-CoV-2 isolates reported in India. 1,33,154 complete protein sequences retrieved from the GISAID database between September 2021 and March 2023 were analysed. The analysis revealed a total of 62,94,995 mutations, which include 66,861 unique mutations. Sequences comprising at least eight consecutive amino acids with mutation frequencies below 0.1% were considered conserved regions. This analysis identified 270 conserved sequences across both structural and non-structural proteins. Of these, 73 sequences were found to be antigenic and non-allergenic and were mapped onto their respective crystal structure of proteins to evaluate their functional relevance. Many conserved sequences overlapped with the known functionally significant epitopes conserved across SARS-CoV-2 variants, underscoring their importance. The identified conserved sequences offer valuable targets for developing variant-resilient peptide-based diagnostics, monoclonal antibody therapeutics, and multi-epitope peptide vaccines. This study provides a curated collection of conserved SARS-CoV-2 protein regions identified from Indian clinical isolates and emphasises their potential for diagnostic and therapeutic applications. These findings may contribute to developing universal, variant-proof strategies for SARS-CoV-2 detection, prevention, and treatment.

1. Introduction

The rapid emergence of SARS-CoV-2 variants has challenged the efficacy of first-generation vaccines. For instance, the heavily mutated Omicron variant evades many neutralising antibodies from prior infection or vaccination [1]. This immune escape has led to breakthrough infections and prompted the rollout of variant-updated booster vaccines (e.g. bivalent BA.4/5 and monovalent XBB.1.5 boosters) to restore protection [2]. Emerging COVID-19 vaccines increasingly focus on conserved viral epitopes to achieve broader protection. Strategies include multi-epitope vaccines incorporating conserved B- and T-cell epitopes across the SARS-CoV-2 proteome [3].

The SARS-CoV-2 genome is approximately 30,000 nucleotides long and encodes 29 proteins, with four structural proteins, 16 non-structural proteins (NSPs), and nine accessory proteins [4]. Recent studies revealed that including peptides from internal proteins like nucleocapsid (N) and membrane (M) alongside Spike (S) elicit robust $CD4^+$ and $CD8^+$ T-cell responses to complement neutralising antibodies. These peptides are also reported to induce potent and long-lasting B- and T-cell immunity in clinical trials, suggesting that they confer broad protection against diverse variants [5–8]. T-cell-mediated immunity has proven more resilient to viral variation; studies show that most T-cell epitopes remain unchanged across variants [9]. Notably, >95% of $CD8^+$ T-cell epitopes identified in the original strain are still conserved in Omicron [1]. This preservation of conserved peptide targets helps maintain T-cell recognition and protection against severe disease despite the antigenic drift of SARS-CoV-2 [1]. These reports underscore why next-generation vaccines are pivoting to include conserved elements to confer immunity that endures against current and future variants [9].

Computational immunology has been instrumental in pinpointing conserved SARS-CoV-2 peptide targets. Immuno-informatics leverages the growing genomic data from SARS-CoV-2 variants and related coronaviruses to map regions of high sequence conservation. These conserved epitopes are often functionally important; mutation-intolerant regions are attractive targets for the broad-spectrum vaccine and serve as potential regions for the design of diagnostic and therapeutic approaches. In silico tools also predict the binding of these epitopes to the common human MHC alleles, ensuring broad population coverage [9]. By integrating sequence alignment, structural modelling, and epitope prediction algorithms, researchers can efficiently prioritise viral peptides that remain conserved across variants for inclusion in universal vaccine candidates. This computational screening has unveiled key immunodominant regions likely to induce cross-protective immunity, guiding experimental vaccine design and speeding the development of variant-proof immunotherapies. [10].

Many studies in silico were on the spike protein, and a small subset of SARS-CoV-2 proteins [8]. The primary objective of this study is to identify conserved and immunogenic regions across the SARS-CoV-2 genome using an immunoinformatic approach. The secondary aim of this study is to determine the allergenicity of the conserved epitopes to screen out the best possible, reliable candidates for broad-spectrum vaccines and monoclonal antibody therapy targets.

2. Materials and Methods

2.1. Identification of conserved regions in the SARS-CoV-2 genome

One lakh thirty-three thousand one hundred and fifty-four (1,33,154) SARS-CoV-2 genome sequences from Indian patients deposited from 1st September 2021 to 31st March 2023 were collected from the GISAID database. The total number of mutations was obtained from analysing data through the R package. R libraries like seqinR (4.2.8) (29), protr (1.6.2) (30) and Biostrings (2.58.0) (31) were used to analyse this data. Regions with a minimum of 8 amino acid length that showed no mutation were considered conserved sequences. The reference genome was obtained from the NCBI accession number: NC_045512.2. The frequency of the mutations in any given position was also calculated. Nearly every single position had a mutation, and no conserved regions could be found. To account for that, we then assumed that mutations with less than 0.1% of the maximum frequency of a protein were rare in the population or sequencing errors.

The frequency of mutations for every single position (for each protein) is calculated, and the maximum frequency is identified. 0.1% of this maximum frequency was set as the cut-off, and everything below this was assumed to be zero. The presence or absence of a mutation at any given position was set to binary values, 1 (presence of mutation) and 0 (absence of mutation). Regions with more than eight consecutive zeroes were identified, and their corresponding protein sequence was determined (Figure 1).

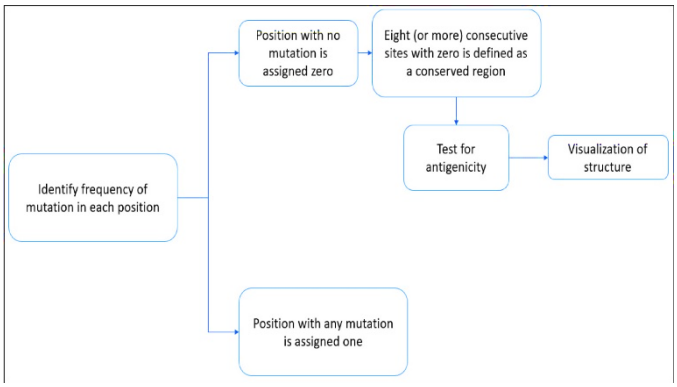


Fig 1. Graphical representation of the analysis workflow

2.2. Immunogenicity and protein modelling

Antigenicity and allergenicity of the identified conserved regions were determined as follows. VaxiJen tool was used to identify immunoprotective sequences based on identifying such sequences from bacterial, viral, tumour, fungi and parasite antigens and calculating their auto-cross covariance values [11]. The threshold for viral models was 0.4 (default value); sequences that scored equal to or below 0.4 were considered non-antigenic, and those that scored above were considered antigenic. The Immune Epitope Database (IEDB) tools were used to find possible T-cell and B-cell epitopes in conserved protein sequences. From the input sequence, it identifies the string most likely to have immunogenicity and assigns a score of likelihood [12]. The NetCTL 1.2 tool was used to identify CTL epitopes in the amino acid sequences of SARS-CoV-2. It supports 12 MHC class 1 supertypes [13]. The allergenicity of the amino acid sequences was identified using the AllerTOP tool. Only those sequences with no allergenic effects were used for further analysis [14]. The conserved antigenic and non-allergenic epitopes were then mapped onto the respective crystal structure of the proteins as follows. Experimentally determined structures of non-structural proteins (NSPs) and structural proteins of SARS-CoV-2 were retrieved from the RSCB Protein Data Bank (PDB) [15]. The structure of conserved peptide sequences was visualised in PyMOL v2.4.1 [16]. The structures of NSP-4 and NSP-6 were predicted using the Alphafold2 v2.1.1, which uses a machine learning approach for modelling even without homologous structures [17]. The best model was selected from the five predicted structures based on the predicted local distance difference test (PLDDT) score. Further, the model was evaluated using the PROCHECK [18] and ERRAT [19] modules of the SAVES v6.0 server to validate the stereochemistry and overall quality factors. The loop regions of the predicted models were refined using the ModLoop server [18] and validated using the SAVES server. Modelled structures were analysed using PyMOL.

3. Results and Discussion

The analysis revealed 62,94,995 total mutations, including 66,861 unique mutations across the SARS-CoV-2 genome in the Indian strains (Table 1). When data was analysed directly, no conserved regions were detected. Then, an algorithm was employed, assuming that the mutations below a certain "level" would be zero. Several trials were conducted, with cut-off numbers, namely "below 5 (whole number)", "below 10 (whole number)", "0.1% of highest frequency", and "1% of highest frequency" across multiple clades and lineages during the preliminary stage.

Table 1. Number of conserved, antigenic and non-allergenic sequences in each of the SARS-CoV-2 proteins.

S.No	Protein	Length of the protein (aa)	No. of conserved sequences	No. of antigenic and non-allergenic sequences
1	Spike	1273	49	11
2	Envelope	75	2	1
3	Membrane	222	7	4
4	Nucleocapsid	419	14	2
5	NSP-1	180	7	2
6	NSP-2	638	0	0
7	NSP-3	1945	59	21
8	NSP-4	500	20	6
9	NSP-5	306	11	2
10	NSP-6	290	9	4
11	NSP-7	83	0	0
12	NSP-8	198	1	0
13	NSP-9	113	0	0
14	NSP-10	139	0	0
15	NSP-12	932	31	6
16	NSP-13	601	24	6
17	NSP-14	527	22	6
18	NSP-15	346	14	2
19	NSP-16	298	0	0

It was observed that "0.1% of highest frequency" gave a reasonable number of conserved sequences. Therefore, the algorithm used was that all sequences comprising at least eight consecutive amino acids with mutation frequencies above 0.1% were classified as conserved regions in this study (Supplementary Table 1).

SARS-CoV-2 viral proteins were then analysed for conserved sequences to identify areas of potential clinical significance. The analysis focused on identifying highly conserved regions across different variants to ensure broad applicability. A total of 270 conserved sequences were identified throughout the SARS-CoV-2 genome (Supplementary Table 1). These sequences were then assessed for their antigenicity and allergenic properties. Among these, 73 antigenic and non-allergenic residues were identified throughout the proteome (Table 2, Supplementary Table 1, Supplementary Table 2). These findings provide insights into the stability and evolutionary conservation, suggesting their probable crucial role in developing robust detection assays with targeted therapeutic and preventive interventions.

In the present study, conserved peptide regions are identified from the target proteins involved in the viral RNA synthesis and replication, viral invasion via structural proteins, and act as virulence effectors. The conserved, antigenic, and non-allergenic residues identified in each SARS-CoV-2 protein were subsequently mapped onto corresponding crystal structures and predicted homology models. The functional significance of each residue is detailed below.

Table 2. Number of mutations and unique mutations observed across the SARS-CoV-2 genome

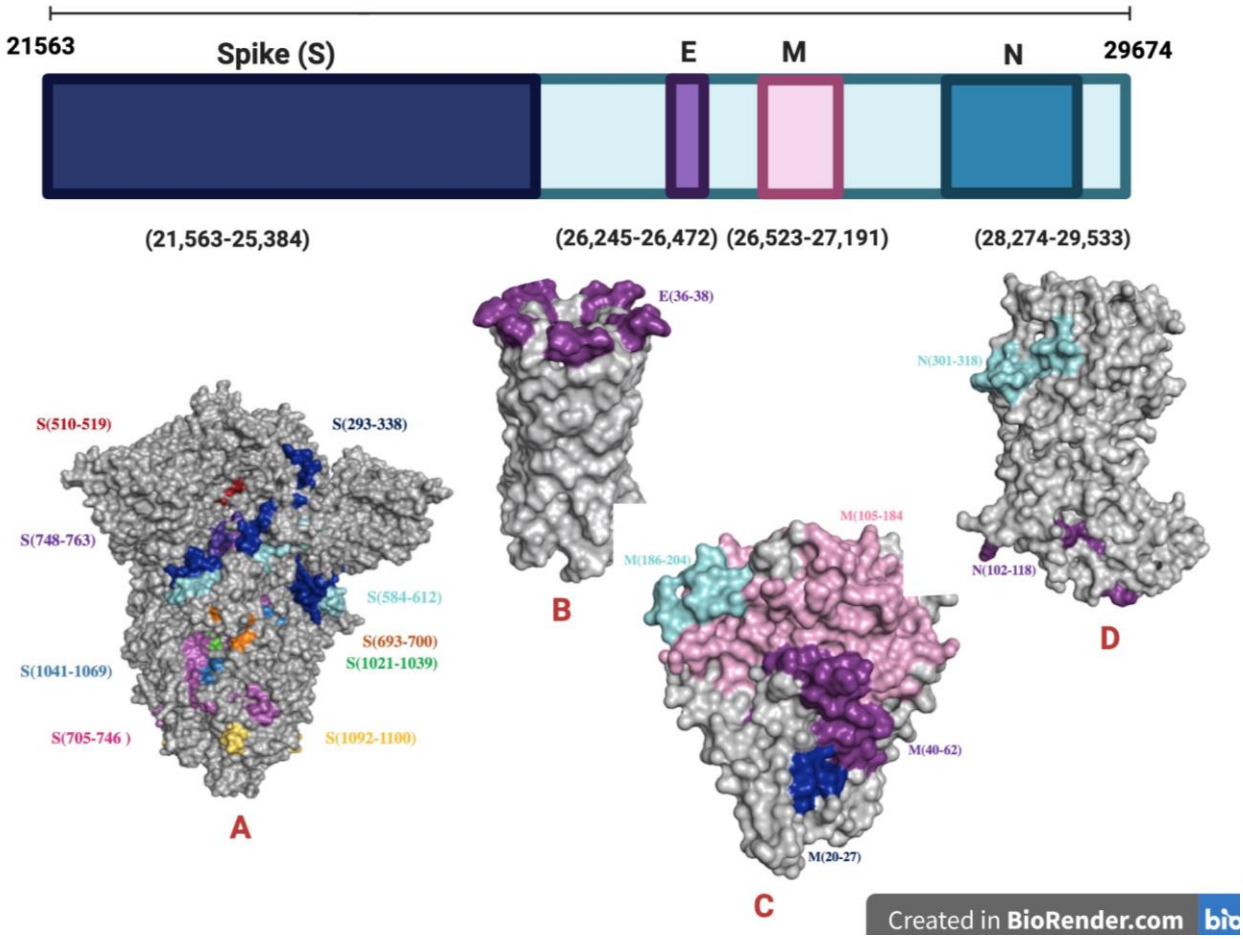
SNo	Protein	Length (AA)	Total No. of mutations	No. of Unique mutations	AA Position with most mutations and the mutation	Frequency of highest mutation
1	Spike	1273	2875012	12929	D614G	131707
2	Envelope	75	108563	481	T9I	87662
3	Membrane	222	237591	1801	A63T	88148
4	Nucleocapsid	419	679718	3811	P13L	86041
5	NSP1	180	94446	1177	S135R	76750
6	NSP2	638	45316	4556	P129L	1581
7	NSP3	1945	458711	18134	G489S	79123
8	NSP4	500	428822	4556	T492I	119745
9	NSP5	306	117695	2131	P132H	89191
10	NSP6	290	262791	2080	G107del	57094
11	NSP7	83	1625	317	M75I	74
12	NSP8	198	20068	926	N118S	12832
13	NSP9	113	3826	380	T35I	375
14	NSP10	139	6023	610	M122V	1209
15	NSP12	932	224130	5698	P323L	129979
16	NSP13	601	168140	3740	R392C	80369
17	NSP14	527	153419	3371	I42V	86910
18	NSP15	346	92205	2562	T112I	73896
19	NSP16	298	11520	2092	K160R	800

3.1. Spike Protein:

Our analysis revealed that the Spike (S) protein, which is 1273 amino acids long, contains 49 conserved regions (Table 2, Supplementary Table 1). Of these, 11 were found to be antigenic and non-allergenic peptides (Table 2, Supplementary Table 1, Figure 2A). Identifying conserved regions within the S protein is pivotal for developing broad-spectrum antivirals and vaccines. Despite the high mutation rate observed in various S protein domains, certain regions remain highly conserved, underscoring their functional importance. The N-terminal conserved domain (NTD) peptide (293–338) is implicated in host-cell recognition and has been identified as a target for neutralising antibodies [19]. Mutations in this region can contribute to immune evasion; however, conserved epitopes within the NTD may serve as viable targets for vaccine development [19].

The S protein's Receptor Binding Domain (RBD) is critical for binding to the ACE2 receptor, facilitating viral entry into host cells [20]. While the RBD is subject to frequent mutations, specific conserved sequences within this domain are essential for maintaining its structural integrity and function. One of the identified conserved peptides is mapped (510–519) in the RBD. Also, previous study has found peptide (510–519) as a potential target for binding the S2P26 antiviral peptide and can restrict from forming interaction with ACE2 receptor [21]. Notably, the N501Y mutation has strengthened the binding affinity between

Fig 2. Mapping of immunogenic conserved peptide residues onto SARS-CoV-2 structural protein homology models.
Antigenic and non-allergenic residues identified were mapped onto Spike (A), Envelope (B), Membrane (C), and Nucleocapsid (D) homology models. Residues mapped on the surface of the homology models are depicted



the spike protein and ACE2 receptor, enhancing viral infectivity [20]. The H519N mutation significantly decreases SARS-CoV-2 replication in human lung epithelial cells and reduces infectivity in pseudotyped virus [22]. Targeting these conserved regions could lead to the development of broad-spectrum neutralising antibodies [23]. The fusion peptide of the spike (S) protein plays a critical role in mediating the fusion between the viral and host cell membranes, an essential step in viral entry [24]. The high conservation of fusion peptide sequences across various coronavirus strains suggests their potential as universal vaccine targets. Peptide-based fusion inhibitors derived from these conserved sequences have demonstrated efficacy against emerging coronaviruses [25]. We have identified conserved regions (693–700, 705–746, 748–763) in this region, suggesting these peptides' functional significance. S protein's heptad repeat (HR) regions are integral to the conformational changes required for membrane fusion and viral entry [26]. Mimetic proteins structurally imitating the HR1 region in a trimeric coiled-coil conformation showed potential in inhibiting SARS-CoV-2 infection *in vitro* [27]. Our analysis revealed three conserved peptides (1021–1039, 1041–1069, 1092–1100) in this region. Targeting these conserved regions within the S protein is a strategic approach to developing effective therapeutics and vaccines across multiple coronavirus strains. Focusing on these regions makes it possible to design interventions that maintain efficacy even as the virus undergoes mutations in other regions.

3..2. Envelope Protein:

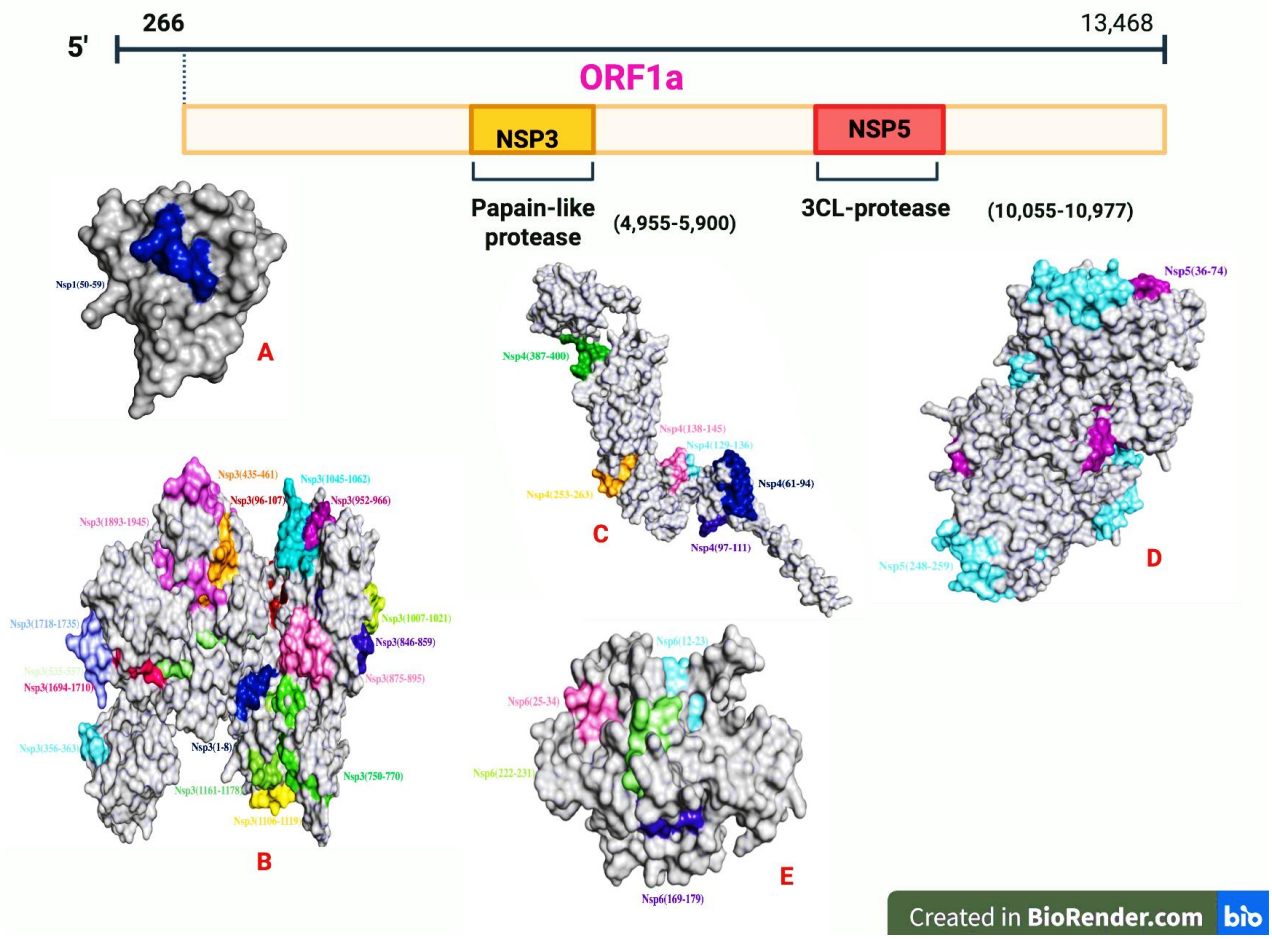
The SARS-CoV-2 envelope (E) protein, comprising 75 amino acids, plays a crucial role in virus assembly, budding, and pathogenesis. The conserved antigenic and non-allergenic sequence (residues 36–48) is located within the hydrophobic transmembrane domain (TMD) of the E protein, which is essential for anchoring it to the viral envelope (Figure 2B, Table 2) [28]. Structural studies have shown that the E protein's TMD forms a pentameric helical bundle, creating a cation-selective ion channel critical for viral pathogenicity, highlighting its role as a potential target for antiviral drugs [28,29]. This underscores the TMD's potential as a therapeutic intervention target for hindering virus assembly and release. Targeting this region could lead to the design of inhibitors that disrupt E protein function, thereby mitigating SARS-CoV-2 infectivity and pathogenicity.

3.3. Membrane Protein:

The SARS-CoV-2 membrane (M) protein, comprising 222 amino acids, plays a vital role in viral assembly, morphogenesis, and pathogenesis [30]. It is the most abundant structural component of the virus and interacts with other essential proteins, such as S and E proteins, to regulate viral budding and genome encapsulation [30]. Four of the seven conserved regions observed in the current study were found to be antigenic and non-allergenic (Figure 2C, Table 2). The transmembrane location of the conserved region (residues 20-27) could contribute to its anchoring within the viral envelope and is essential for membrane curvature during viral budding [31].

Fig 3. Mapping of immunogenic conserved peptide residues onto SARS-CoV-2 non-structural protein (located on the Orf1-a of the SARS-CoV-2 genome) homology models.

Antigenic and non-allergenic residues identified were mapped onto NSP-1 (A), NSP-3 (B), NSP-4 (C), NSP-5 (D), and NSP-6 (E) homology models. Residues mapped on the surface of the homology models are depicted



Hydrophobic interactions within this region facilitate lipid bilayer integration, ensuring the stability of the viral structure [32]. Another conserved region (residues 40–62) is known to be critical for protein-protein interactions involved in viral assembly [30]. Studies suggest that mutations in this domain disrupt virion formation, producing non-infectious viral particles [33].

The conserved region (residues 105–184) is known to interact with the nucleocapsid (N) protein, which plays a key role in encapsulating the viral RNA genome [33]. Disrupting this region has been shown to impair viral replication and assembly, making it a key target for antiviral drug development [33]. The host-cell interaction domain region (residues 186–206) plays a fundamental role in binding to host cell membranes and is involved in the budding process of new virions [33]. Computational docking studies suggest that peptides mimicking this domain could act as viral assembly inhibitors, preventing new virions from exiting the host cell [31]. Given its role in viral structure and function, the M protein is a key therapeutic intervention target. Conserved sequences within the M protein offer promising opportunities for developing antiviral agents that can disrupt viral assembly and budding [31]. In addition, targeting the M protein interactions with S and E proteins could prevent the formation of fully functional virions, reducing viral spread within infected individuals [34]. Moreover, the immunogenic properties of the M protein suggest its potential use in vaccine formulations. The conserved nature of the M protein ensures broad-spectrum vaccine effectiveness, even against emerging SARS-CoV-2 variants.

3.4. Nucleocapsid Protein:

The nucleocapsid (N) protein of SARS-CoV-2, a crucial structural component, is involved in RNA binding, genome packaging, and viral replication [35]. Comprising 419 amino acids, the N protein exhibited 14 conserved regions, with two areas being antigenic and non-allergenic, underscoring its evolutionary stability across different variants. This conservation suggests a crucial functional significance, making the N protein a prime target for diagnostic and therapeutic applications. Both the conserved antigenic residues (residues 102–118 and residues 301–318) (Figure 2D, Table 2) are particularly noteworthy due to their potential involvement in RNA binding and protein-protein interactions [36]. The first conserved peptide sequence (residues 102–118) falls within a highly structured region, which may contribute to the protein's ability to interact with viral and host components [36]. Similarly, the second conserved peptide sequence (residues 301–318) is positioned within a flexible domain that may facilitate dynamic conformational changes essential for nucleocapsid function [36,37].

The high conservation of these regions across SARS-CoV-2 variants suggests their indispensable role in viral assembly and propagation. The N protein is highly immunogenic so that these

conserved sequences may serve as potential epitopes for monoclonal antibody development and vaccine design [38]. Additionally, the involvement of the identified conserved sequences in viral RNA replication processes suggests that targeting them with small molecules or peptide inhibitors could disrupt viral replication, offering a promising therapeutic avenue [39]. The N protein remains a primary target for rapid antigen detection tests in diagnostic applications due to its abundance during infection. Monoclonal antibodies directed against the identified conserved regions could enhance the sensitivity and specificity of lateral flow assays, aiding in early and reliable SARS-CoV-2 detection [40]. In conclusion, the remarkable conservation of these two regions within the SARS-CoV-2 N protein highlights their functional importance in viral pathogenesis. Future studies should focus on elucidating their structural and functional roles in greater detail, paving the way for improved antiviral strategies and diagnostic tools.

3.5. Non-structural protein 1:

NSP-1 is a key virulence factor in suppressing host immune responses and facilitating viral replication [41]. Of the seven conserved regions in the NSP-1, two peptides were found to be antigenic and non-allergenic. One of the peptides (residues 50–59) was mapped onto the surface of the NSP-1 (Figure 3A, Table 2). This region (residues 50–59) within the N-terminal domain (NTD) of SARS-CoV-2 NSP-1 may play a role in host immune suppression and viral replication. Structurally, this region forms part of the β2–β3 loop, contributing to the electrostatic surface properties of the NTD and stabilising its interactions with host factors [42]. The acidic patch formed by residues E55 and E57 is located adjacent to the α1-helix and may play a role in the selective binding of viral RNA, while simultaneously inhibiting host mRNA translation [41]. Functionally, NSP-1 inhibits host gene expression by blocking translation and suppressing mRNA transport [43]. Mutational studies indicated that alterations in the 50–59 loop impair NSP-1's ability to suppress host defences, leading to attenuation of viral replication [44]. Deletion or mutation of residues in this region has been linked to reduced cytotoxicity and loss of host shutoff function, suggesting its importance in NSP-1's virulence mechanism [45]. Additionally, attenuated viral strains with targeted mutations in this region may serve as candidates for live-attenuated vaccine development [44]. This conserved region presents a promising therapeutic target for antiviral strategies and vaccine design by enabling immune evasion and viral replication.

3.6. Non-structural protein 3:

The NSP-3 of SARS-CoV-2 is reported to play a role in viral replication and the modulation of host immune responses [46]. In our study, NSP-3 is the protein identified with the maximum number of conserved residues with antigenic significance (N=59, N=21, respectively) (Table 2, Figure 3B). Other researchers analysed several key amino acid sequences within NSP-3 for their structural and functional significance in viral pathogenesis [47]. The N-terminal region, consisting of residues 1–8, plays a crucial role in the early stages of viral replication and is also associated with ubiquitin-like domain 1. This domain has been implicated in protein-protein interactions that may modulate host immune responses [48]. Similarly, residues 96–107 are located within the SARS-Unique Domain (SUD), a region unique to SARS-related coronaviruses, which interacts with host cell proteins and may influence viral replication efficiency [46]. The papain-like protease (PLPro) domain, containing residues 356–363, plays a pivotal role in cleaving the viral polyprotein and has de-ubiquitinating functions that help SARS-CoV-2 evade immune detection [49]. Structural studies have highlighted the significance of PLPro in counteracting host antiviral responses, making it a key therapeutic target [50].

Additionally, the macrodomain 1 of NSP-3, which includes residues 435–461, has been identified as essential in interfering with the host's ADP-ribosylation signalling. This domain is vital for viral replication and immune evasion, with structural analyses emphasising its potential as a drug target [51]. The transmembrane regions of NSP-3, represented by residues 535–557 and 846–859, contribute to double-membrane vesicles (DMVs), which provide a protected niche for viral RNA synthesis [52]. These domains interact with NSP-4 and NSP-6, highlighting their role in viral replication complex formation. The SUD contains multiple functionally relevant sequences, residues 740–748 and 750–770, which are implicated in host-pathogen interactions and may influence the efficiency of viral replication [53]. Studies suggest that SUD enhances viral pathogenicity through interactions with cellular proteins, further supporting its significance in SARS-CoV-2 infection dynamics. Further downstream, residues 1844–1866 and 1893–1945 are located in regions that may contribute to NSP-3's overall structural stability and interactions with other viral and host components[46]. Their functional implications warrant further investigation, particularly in viral replication and immune evasion mechanisms. The comprehensive analysis of these NSP-3 sequences provides valuable insights into their functional roles in SARS-CoV-2 pathogenesis. Given their involvement in viral replication, immune modulation, and host interactions, these sequences are potential targets for antiviral drug development. Further structural and biochemical studies must elucidate their full mechanistic contributions and therapeutic potential.

3.7. Non-structural protein 4:

SARS-CoV-2 non-structural protein 4 (NSP-4) is crucial for membrane remodelling during viral replication, interacting with NSP-3 to promote the formation of double-membrane vesicles (DMVs) [54,55]. Our analysis revealed 20 conserved residues, with 6 having antigenic potential (Table 2, Figure 3C). These six conserved motifs in NSP-4, including residues 61–94, 97–111, 129–136, 138–145, 253–263, 387–400, have been shown to possess functional importance [56]. Structurally, these conserved sequences map to key domains, including luminal loops involved in membrane curvature (61–145, 253–263), transmembrane helices forming the DMV-spanning pore (387–400), and the cytosolic C-terminal tail critical for replication complex assembly (493–500) [57,58]. Notably, N-glycosylation at Asn131 and multiple disulfide bonds within luminal loops suggest an additional layer of regulation influencing NSP-4 stability and interaction with NSP-3 [59]. Despite NSP-4's high conservation, the T492I mutation, which emerged in Delta and Omicron variants, enhanced viral replication by increasing 3CLpro processing efficiency [60]. This highlights how even minor mutations in NSP-4 can impact viral fitness. Given its critical role in DMV formation, NSP-4 represents an attractive antiviral target, particularly disrupting the NSP–3–NSP–4 interface or DMV-spanning pore [61]. Future studies should explore small molecules or antibodies targeting NSP4, given its low mutation rate compared to other viral proteins, making it a viable broad-spectrum target across coronaviruses.

3.8. Non-structural protein 5:

The NSP-5 main protease (Mpro) of SARS-CoV-2 plays a pivotal role in viral replication by cleaving polyproteins (pp1a/pp1ab) into functional non-structural proteins [62]. The enzyme is highly conserved, with key sequence segments 36–74 and 248–259 demonstrating strong purifying selection across SARS-CoV-2 variants and related coronaviruses [56]. Interestingly, in our study of the 11 conserved residues identified, these two have antigenic potential (Table 2, Figure 3D). The 36–74 region contributes to the catalytic domain (domain I) and harbours His41, a critical residue forming the His41–Cys145 catalytic dyad essential for substrate cleavage and viral replication [63]. The previous study has reported anti-asthmatic drug montelukast binds with NSP-5 protein with low affinity, involving His41 forming a hydrophobic pocket to accommodate the drug molecule [64]. This segment also defines the substrate-binding pocket and interacts with known inhibitors such as Paxlovid (nirmatrelvir) [65]. The 248–259 region within domain III is crucial for dimerisation, an essential step in NSP-5 activation [66]. Disrupting this interface destabilises the enzyme, leading to a loss of function and viral replication inhibition [67].

Despite the emergence of SARS-CoV-2 variants of concern, NSP-5 remains highly conserved, with minimal mutations in its active site or dimerisation interface [68]. Mutational studies confirm that substitutions in these regions impair viral fitness, highlighting their functional importance [69]. Given the stability of these sequences, Mpro remains a prime antiviral target. Covalent inhibitors like nirmatrelvir and peptidomimetics have been designed to irreversibly bind the active site, effectively blocking viral replication [65]. Additionally, allosteric inhibitors targeting the dimerisation interface have shown promise, offering an alternative antiviral approach [70]. Given these segments' extreme conservation and functional indispensability, targeting NSP5 with direct-acting antivirals provides a robust therapeutic strategy against SARS-CoV-2 and future coronavirus outbreaks [62]. Future research should focus on dimer interface inhibitors to complement existing active-site-targeting drugs, ensuring broad-spectrum efficacy against coronaviruses.

3.9. Non-structural protein 6:

The NSP-6 of SARS-CoV-2 promotes the virus's ability to remodel host cell membranes, facilitating DMVs essential for viral RNA replication. This function is conserved across coronaviruses, underscoring the importance of specific sequences within NSP-6 [71]. We identified nine conserved residues in this protein, with 4 exhibiting antigenic potential (Table 2, Figure 3E). The sequences with residues 12–23 and residues 25–34 are located within the transmembrane domains of NSP-6, which are highly conserved among SARS-CoV-2 variants and related coronaviruses, including SARS-CoV and MERS-CoV. This conservation suggests a fundamental role in membrane association and DMV formation [72]. Similarly, the sequences with residues 169–179 and residues 222–231 are preserved across various coronavirus species, indicating their critical role in viral replication [73]. The transmembrane regions encompassing residues 12–34 are integral to NSP-6's ability to anchor to the endoplasmic reticulum membrane. This anchorage is vital for inducing autophagosome formation and subsequent DMV development in viral replication [74].

The structural conservation of these sequences ensures the integrity required for proper membrane curvature and vesicle formation. Residues 169–179 and 222–231 are implicated in protein-protein interactions within the viral replication complex, facilitating the coordination necessary for efficient RNA synthesis [73]. Mutations within these conserved regions can significantly impact NSP-6 function. For instance, alterations in the transmembrane domains may disrupt membrane association, hindering DMV formation and attenuating viral replication [75]. Targeting these conserved sequences with antiviral agents could impair NSP-6's function, offering a potential therapeutic strategy [35]. Compounds that disrupt NSP-6's membrane interactions or their role in autophagosome formation could effectively inhibit viral replication [71]. The conserved sequences within NSP-6 are integral to modifying host cell membranes for viral replication. Their preservation across coronavirus species highlights their essential role and presents opportunities for targeted antiviral interventions. Given these sequences' high conservation and functional importance, NSP-6 remains a viable antiviral target, particularly for strategies that disrupt its interactions with host cell membranes and interfere with DMV formation [35].

3.10. Non-structural protein 12:

The SARS-CoV-2 NSP-12 functions as the RNA-dependent RNA polymerase (RdRp), a crucial enzyme for viral genome replication [76]. We identified 31 conserved residues in this protein, with 6 exhibiting antigenic potential (Table 2, Figure 4A). These six conserved sequence regions (residues 186–226, 294–322, 401–462, 464–486, 672–693, and 720–738) have been shown to play essential roles in RdRp function, structural stability, and antiviral drug interactions. Residues 186–226 of NSP-12 facilitate nucleotidyl transfer reactions necessary for RNA synthesis and viral RNA capping [77]. Key residues within this region participate in ATP and GTP binding, making it an attractive antiviral target [78] Region 294–322 contains a highly conserved Cys/His-rich zinc-binding motif stabilising NSP-12 by coordinating Zn^{2+} ions [48]. Structural studies indicate that disrupting these zinc-coordinating residues impairs polymerase activity [79]. Polymerase core regions with residues 401–486, 672–693, 720–738 form key subdomains essential for RNA synthesis. The fingers domain 401–486 contributes to RNA binding and cofactor interactions, particularly with NSP-8, which enhances processivity [76]. The palm domain 672–693 contains motif B, a flexible loop that adjusts during nucleotide incorporation, facilitating efficient RNA synthesis [80].

The 720–738 region lies adjacent to the active site and stabilises the polymerase's catalytic domain [81]. The high conservation of these sequences makes them ideal targets for antiviral drugs. Remdesivir, a nucleotide analogue, binds to the polymerase active site and disrupts RNA synthesis by causing delayed chain termination [56]. Structural studies have shown that remdesivir interacts with conserved residues in the fingers and palm domains, stabilising an inactive polymerase complex [82]. The studies also revealed that remdesivir forms an interaction with the RNA-binding channel of RdRp involving Asn497, Arg569, and Asp684 to form polar contact and Leu576, Ala685, and Tyr687 to form a hydrophobic pocket [64]. This report aligns with our findings that interacting residues Asp684, Ala685, and Tyr687 belong to the predicted conserved peptide 672–693. Molnupiravir and favipiravir also exploit conserved residues to induce lethal mutagenesis [83]. These regions remain highly conserved across SARS-CoV, MERS-CoV, and other coronaviruses, indicating strong evolutionary constraints [84]. However, mutations such as P323L (interface domain) and G671S (motif B loop) have emerged in SARS-CoV-2 variants, potentially enhancing viral replication efficiency [85]. Nevertheless, remdesivir resistance mutations (e.g., F480L,

V557L) remain rare due to the functional constraints on these conserved regions [83] Overall, the conserved sequences in NSP-12 are essential for viral replication and represent key targets for developing antiviral drugs. Their limited mutational tolerance reinforces their potential as drug-binding sites, highlighting the importance of continued structural and functional studies to optimise therapeutic strategies.

3.11. Non-structural protein 13:

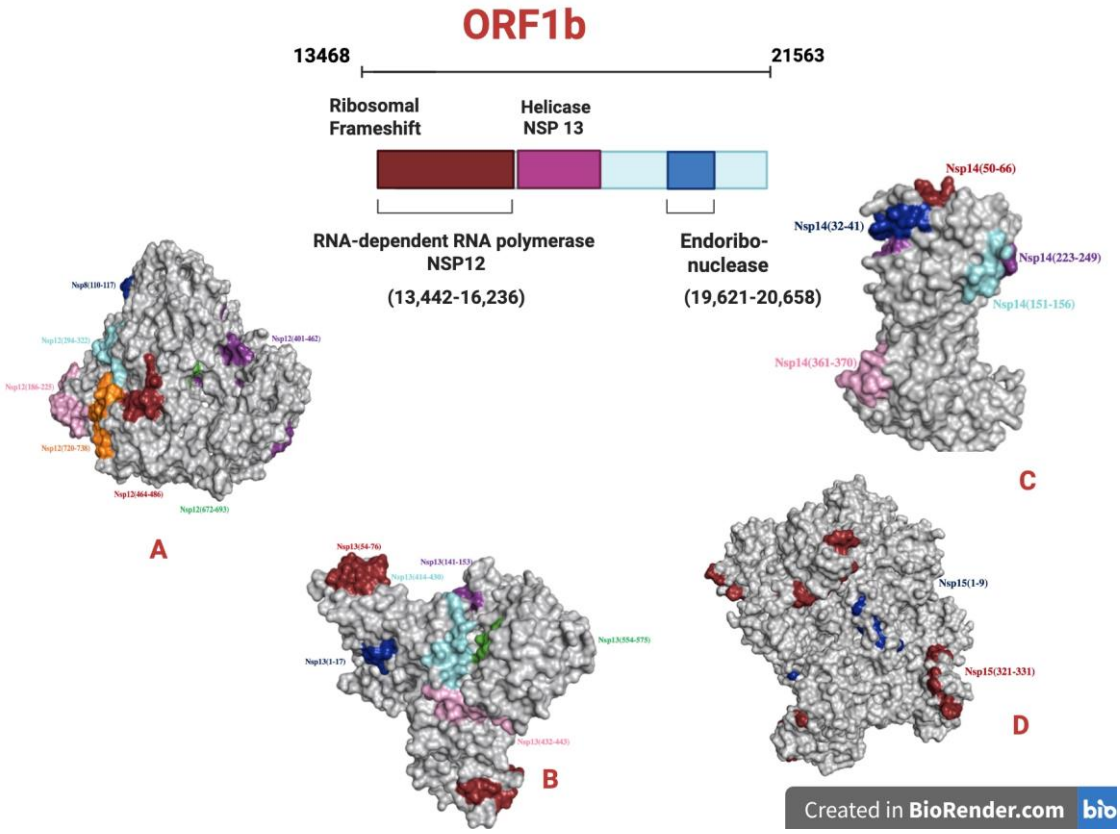
The SARS-CoV-2 NSP-13 is a highly conserved helicase that plays a critical role in viral replication by unwinding RNA and hydrolysing ATP [86]. We identified 24 conserved residues in this protein, with 6 exhibiting antigenic potential (Table 2, Figure 4B). These conserved sequences in NSP-13, including residues 1-17, 54-76, 141-153, 414-430, 432-443, and 554-575, correspond to essential functional motifs required for enzymatic activity, structural integrity, and interaction with other viral components. The zinc-binding domain (ZBD), comprising the 1–17 and 54–76 sequences, coordinates Zn^{2+} or Fe–S clusters necessary for NSP-13 stability and RNA unwinding. Studies have shown that mutations in these cysteine-rich motifs disrupt metal coordination and abolish enzymatic function [87]. The stalk domain (141–153) supports the helicase core and facilitates interactions with NSP-12 and NSP-8 in the replication-transcription complex [88]. The helicase core, including 414–430 (RecA1 domain), 432–443 (RecA2 domain), and 554–575 (motif VI), contains essential ATP-binding and RNA-binding motifs. Motif VI, particularly Arg567 within the 554–575 segment, functions as an "arginine finger," crucial for ATP hydrolysis and energy transduction [89]. Structural studies have revealed that these conserved sequences undergo conformational changes upon ATP binding, facilitating RNA unwinding [86]. NSP13 is a promising antiviral target due to its high conservation across coronaviruses. Small-molecule inhibitors such as bismuth complexes, flavonoids (myricetin, scutellarein), and SSYA10-001 have been shown to target these conserved motifs, disrupting ATPase and helicase activity [90]. Additionally, the ZBD is a druggable site, as bismuth-based compounds destabilise its zinc-finger motifs, effectively inhibiting the helicase [91]. Comparative sequence analysis across SARS-CoV, MERS-CoV, and other coronaviruses highlights the strong evolutionary conservation of these motifs, underscoring their indispensable role in viral replication, making them attractive targets for therapeutics and vaccines.

3.12. Non-structural protein 14:

The NSP-14 of SARS-CoV-2 is integral to the virus's replication fidelity and immune evasion strategies [92,93]. Our sequence analysis has identified 22 conserved regions within NSP-14, including six antigenic residues: 32–41, 50–66, 126–139, 145–156, 223–249, and 361–370 (Table 2, Figure 4C). The conservation of these sequences across various SARS-CoV-2 isolates suggests their critical role in maintaining the structural integrity and enzymatic functions of NSP-14 [94]. Notably, NSP-14 of SARS-CoV and SARS-CoV-2 share over 95% amino acid sequence similarity, underscoring their evolutionary conservation and potential as a therapeutic target [95]. The ExoN activity of NSP-

Fig 4. Mapping of immunogenic conserved peptide residues onto SARS-CoV-2 non-structural protein (located on the Orf1-b of the SARS-CoV-2 genome) homology models.

Antigenic and non-allergenic residues identified were mapped onto NSP-12 (A), NSP-13 (B), NSP-14 (C), and NSP-15 (D) homology models. Residues mapped on the surface of the homology models are depicted.



14 is essential for correcting errors during RNA synthesis, thereby reducing the mutation rate and contributing to the stability of the viral genome [96]. Additionally, the N7-MTase domain's role in mRNA capping protects viral RNA from host immune responses, facilitates efficient translation, and ensures viral proliferation [97]. The crystal structure of SARS-CoV-2 NSP-14 demonstrates that the ExoN domain's enzymatic activity is metal ion-dependent, preferably utilising Mg^{2+}, and that the N7-MTase domain harbours a conserved DxG motif for S-adenosyl-L-methionine binding, characteristic of coronaviruses [94].

3.13. Non-structural protein 15:

The SARS-CoV-2 NSP-15 is an endoribonuclease highly conserved among coronaviruses and plays a crucial role in viral RNA processing and immune evasion [98]. It cleaves viral RNA at uridine sites, preventing the accumulation of immunostimulatory dsRNA, thereby helping the virus evade host immune responses [99]. We identified 14 conserved residues in this protein, with 2 exhibiting antigenic potential (Table 2, Figure 4D). The conserved residues 1–9 and 321–336 are functionally significant as they contribute to NSP-15's hexameric structure and catalytic activity, making them potential antiviral drug targets ([100]. The peptide with residues 1-9 at the N-terminus is part of the oligomerisation domain, essential for forming the active hexameric complex [98]. Mutations in this region impair hexamerization and lead to loss of enzymatic activity, suggesting its importance in structural integrity [101] The conservation of this motif across coronaviruses highlights its critical role in viral replication and potential as a drug target [99].

The peptide with residues 321-336 in the C-terminal region is located near the active site of NSP-15 and is highly conserved across beta-coronaviruses [100]. It contributes to substrate recognition and enzymatic function, ensuring efficient RNA cleavage [100]. Structural studies show that this sequence forms part of the uridine-binding pocket, crucial for its RNA endonuclease activity [102, 103]. Mutations in this region disrupt catalytic efficiency, accumulating dsRNA and increasing host immune activation [99]. Viruses lacking a functional NSP-15 enzyme show reduced replication efficiency and are more susceptible to host immune responses. The conserved residues 1–9 and 321–336 in SARS-CoV-2 NSP-15 are structurally and functionally crucial. They contribute to enzyme oligomerisation, substrate recognition, and RNA processing, ensuring viral replication and immune evasion. Their high conservation makes them ideal targets for antiviral drug development.

4. Conclusion

Our findings provide insights into the stability and evolutionary conservation of SARS-CoV-2 protein sequences, suggesting their probable crucial role in developing robust detection assays and targeted therapeutic and preventive interventions. These conserved sequences within SARS-CoV-2 are integral for viral replication and immune evasion. Mutations within these sequences are rare due to their functional constraints, and the reports of experimental mutagenesis studies confirm that disruptions in these regions would lead to loss of viral replication efficiency, making them attractive therapeutic and preventive targets for broad-spectrum SARS-CoV-2 variants. Future research may focus on developing inhibitors that exploit these conserved sites to disrupt viral replication, effectively providing effective therapeutic and prophylactic strategies against SARS-CoV-2 and related coronaviruses.

Acknowledgements:

The authors would like to acknowledge Osmania University, and PVNRTVU, Hyderabad, for providing resources to conduct the study.

Declarations

Ethics approval and consent to participate: Not applicable

Availability of data and material:

The datasets used/analysed during the current study are available in the NISAID repository, https://gisaid.org

Competing interests:

The authors declare that they have no competing interests

Funding: No funding was received for this study

Authors' contributions: *RM, KP, JS, and SR contributed to the study conception and design. JS, SR, ST, BRN, ARA, SKP, AV, JP, VK, MMA, MM, MY, and PKM performed material preparation and data collection. RM, KP, analysed the data. The first draft of the manuscript was written by RM, KP, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.*

Supplementary data available at:

Table-1: https://doi.org/10.5281/zenodo.18489299
Tavle-2: https://doi.org/10.5281/zenodo.18489329

References

[1] Choi SJ, Kim DU, Noh JY, et al. T cell epitopes in SARS-CoV-2 proteins are substantially conserved in the Omicron variant. *Cellular & Molecular Immunology 2022 19:3*. 2022;19(3):447-448. doi:10.1038/s41423-022-00838-5

[2] Muik A, Quandt J, Lui BG, et al. Immunity against conserved epitopes dominates after two consecutive exposures to SARS-CoV-2 Omicron BA.1. *Cell Rep*. 2024;43(8). doi:10.1016/j.celrep.2024.114567

[3] Wang CY, Peng WJ, Kuo BS, et al. Toward a pan-SARS-CoV-2 vaccine targeting conserved epitopes on spike and non-spike proteins for potent, broad and durable immune responses. *PLoS Pathog*. 2023;19(4):e1010870. doi:10.1371/JOURNAL.PPAT.1010870

[4] Bai C, Zhong Q, Gao GF. Overview of SARS-CoV-2 genome-encoded proteins. *Science China Life Sciences 2022 65:2*. 2021;65(2):280-294. doi:10.1007/S11427-021-1964-4

[5] Wang CY, Peng WJ, Kuo BS, et al. Toward a pan-SARS-CoV-2 vaccine targeting conserved epitopes on spike and non-spike proteins for potent, broad and durable immune responses. *PLoS Pathog*. 2023;19(4). doi:10.1371/JOURNAL.PPAT.1010870

[6] Prakash S, Dhanushkodi NR, Zayou L, et al. Cross-protection induced by highly conserved human B, CD4+, and CD8+ T-cell epitopes-based vaccine against severe infection, disease, and death caused by multiple SARS-CoV-2 variants of concern. *Front Immunol*. 2024;15:1328905. doi:10.3389/FIMMU.2024.1328905/BIBTEX

[7] Hsieh CL, Leist SR, Miller EH, et al. Prefusion-stabilized SARS-CoV-2 S2-only antigen provides protection against SARS-CoV-2 challenge. *Nat Commun*. 2024;15(1). doi:10.1038/S41467-024-45404-X

[8] Sankhala RS, Dussupt V, Chen WH, et al. Antibody targeting of conserved sites of vulnerability on the SARS-CoV-2 spike receptor-binding domain. *Structure*. 2024;32(2):131-147.e7. doi:10.1016/J.STR.2023.11.015

[9] Prakash S, Dhanushkodi NR, Zayou L, et al. Cross-protection induced by highly conserved human B, CD4+, and CD8+ T-cell epitopes-based vaccine against severe infection, disease, and death caused by multiple SARS-CoV-2 variants of concern. *Front Immunol*. 2024;15:1328905. doi:10.3389/FIMMU.2024.1328905/BIBTEX

[10] Deng H, Li Y, Wang G, Li R. Comprehensive Analysis of the Immune Response to SARS-CoV-2 Epitopes: Unveiling Potential Targets for Vaccine Development. *Biology (Basel)*. 2025;14(1):67. doi:10.3390/BIOLOGY14010067/S1

[11] Doytchinova IA, Flower DR. VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics*. 2007;8. doi:10.1186/1471-2105-8-4

[12] Vita R, Blazeska N, Marrama D, et al. The Immune Epitope Database (IEDB): 2024 update. *Nucleic Acids Res*. 2025;53(D1):D436-D443. doi:10.1093/NAR/GKAE1092

[13] Larsen M V., Lundegaard C, Lamberth K, Buus S, Lund O, Nielsen M. Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. *BMC Bioinformatics*. 2007;8. doi:10.1186/1471-2105-8-424

[14] Dimitrov I, Bangov I, Flower DR, Doytchinova I. AllerTOP v.2--a server for in silico prediction of allergens. *J Mol Model*. 2014;20(6). doi:10.1007/S00894-014-2278-5

[15] Berman HM, Bhat TN, Bourne PE, et al. The Protein Data Bank and the challenge of structural genomics. *Nat Struct Biol*. 2000;7(SUPPL.):957-959. doi:10.1038/80734,

[16] Seeliger D, De Groot BL. Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *J Comput Aided Mol Des*. 2010;24(5):417-422. doi:10.1007/S10822-010-9352-6,

[17] Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583-589. doi:10.1038/S41586-021-03819-2,

[18] Fiser A, Sali A. ModLoop: Automated modeling of loops in protein structures. *Bioinformatics*. 2003;19(18):2500-2501. doi:10.1093/BIOINFORMATICS/BTG362,

[19] Xia X. Domains and Functions of Spike Protein in SARS-Cov-2 in the Context of Vaccine Design. *Viruses 2021, Vol 13, Page 109*. 2021;13(1):109. doi:10.3390/V13010109

[20] Tian F, Tong B, Sun L, et al. N501y mutation of spike protein in sars-cov-2 strengthens its binding to receptor ace2. *Elife*. 2021;10. doi:10.7554/ELIFE.69091

[21] Essa RZ, Wu Y seng, Batumalaie K, Sekar M, Poh C laa. Antiviral peptides against SARS-CoV-2: therapeutic targets, mechanistic antiviral activity, and efficient delivery. *Pharmacological Reports*. 2022;74(6):1166-1181. doi:10.1007/s43440-022-00432-6

[22] Cereghino C, Michalak K, DiGiuseppe S, et al. Evolution at Spike protein position 519 in SARS-CoV-2 facilitated adaptation to humans. *npj Viruses 2024 2:1*. 2024;2(1):1-9. doi:10.1038/s44298-024-00036-2

[23] Huan X, Zhan J, Gao H. Research progress of spike protein mutation of SARS-CoV-2 mutant strain and antibody development. *Front Immunol*. 2024;15:1407149. doi:10.3389/FIMMU.2024.1407149/XML/NLM

[24] Hoffmann M, Kleine-Weber H, Schroeder S, et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell*. 2020;181(2):271-280.e8. doi:10.1016/J.CELL.2020.02.052

[25] Xia S, Liu Q, Wang Q, et al. Middle East respiratory syndrome coronavirus (MERS-CoV) entry inhibitors targeting spike protein. *Virus Res*. 2014;194:200-210. doi:10.1016/J.VIRUSRES.2014.10.007

[26] Xia S, Yan L, Xu W, et al. A pan-coronavirus fusion inhibitor targeting the HR1 domain of human coronavirus spike. *Sci Adv*. 2019;5(4). doi:10.1126/SCIADV.AAV4580

[27] Zhu Y, Yu D, Yan H, Chong H, He Y. Design of Potent Membrane Fusion Inhibitors against SARS-CoV-2, an Emerging Coronavirus with High Fusogenic Activity. *J Virol*. 2020;94(14):e00635-20. doi:10.1128/JVI.00635-20

[28] Schoeman D, Fielding BC. Coronavirus envelope protein: current knowledge. *Virology Journal 2019 16:1*. 2019;16(1):1-22. doi:10.1186/S12985-019-1182-0

[29] Hong M, Mandala V, McKay M, Shcherbakov A, Dregni A, Kolocouris A. Structure and Drug Binding of the SARS-CoV-2 Envelope Protein in Phospholipid Bilayers. *Res Sq*. Published online September 24, 2020. doi:10.21203/RS.3.RS-77124/V1

[30] Zhang Z, Nomura N, Muramoto Y, et al. Structure of SARS-CoV-2 membrane protein essential for virus assembly. *Nature Communications 2022 13:1*. 2022;13(1):1-12. doi:10.1038/s41467-022-32019-3

[31] Neuman BW, Kiss G, Kunding AH, et al. A structural analysis of M protein in coronavirus assembly and morphology. *J Struct Biol*. 2011;174(1):11-22. doi:10.1016/J.JSB.2010.11.021

[32] Dolan KA, Dutta M, Kern DM, Kotecha A, Voth GA, Brohawn SG. Structure of SARS-CoV-2 M protein in lipid nanodiscs. *Elife*. 2022;11. doi:10.7554/ELIFE.81702

[33] Siu YL, Teoh KT, Lo J, et al. The M, E, and N Structural Proteins of the Severe Acute Respiratory Syndrome Coronavirus Are Required for Efficient Assembly, Trafficking, and Release of Virus-Like Particles. *J Virol*. 2008;82(22):11318-11330. doi:10.1128/JVI.01052-08/SUPPL_FILE/SUPPL_DATA.ZIP

[34] Boson B, Legros V, Zhou B, et al. The SARS-CoV-2 envelope and membrane proteins modulate maturation and retention of the spike protein, allowing assembly of virus-like particles. *Journal of Biological Chemistry*. 2021;296:100111. doi:10.1074/JBC.RA120.016175

[35] Kang S, Yang M, Hong Z, et al. Crystal structure of SARS-CoV-2 nucleocapsid protein RNA binding domain reveals potential unique drug targeting sites. *Acta Pharm Sin B*. 2020;10(7):1228-1238. doi:10.1016/J.APSB.2020.04.009

[36] Cubuk J, Alston JJ, Incicco JJ, et al. The SARS-CoV-2 nucleocapsid protein is dynamic, disordered, and phase separates with RNA. *Nature Communications 2021 12:1*. 2021;12(1):1-17. doi:10.1038/s41467-021-21953-3

[37] Dinesh DC, Chalupska D, Silhan J, et al. Structural basis of RNA recognition by the SARS-CoV-2 nucleocapsid phosphoprotein. *PLoS Pathog*. 2020;16(12):e1009100. doi:10.1371/JOURNAL.PPAT.1009100

[38] Dutta NK, Mazumdar K, Gordy JT. The Nucleocapsid Protein of SARS–CoV-2: a Target for Vaccine Development. *J Virol*. 2020;94(13). doi:10.1128/JVI.00647-20

[39] Iserman C, Roden CA, Boerneke MA, et al. Genomic RNA Elements Drive Phase Separation of the SARS-CoV-2 Nucleocapsid. *Mol Cell*. 2020;80(6):1078-1091.e6. doi:10.1016/J.MOLCEL.2020.11.041

[40] Grant BD, Anderson CE, Alonzo LF, et al. A SARS-CoV-2 coronavirus nucleocapsid protein antigen-detecting lateral flow assay. *PLoS One*. 2021;16(11):e0258819. doi:10.1371/JOURNAL.PONE.0258819

[41] Thoms M, Buschauer R, Ameismeier M, et al. Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2. *Science*. 2020;369(6508):1249-1256. doi:10.1126/SCIENCE.ABC8665

[42] Li G, Hilgenfeld R, Whitley R, De Clercq E. Therapeutic strategies for COVID-19: progress and lessons learned. *Nature Reviews Drug Discovery 2023 22:6*. 2023;22(6):449-475. doi:10.1038/s41573-023-00672-y

[43] Schubert K, Karousis ED, Jomaa A, et al. SARS-CoV-2 Nsp1 binds ribosomal mRNA channel to inhibit translation. Published online July 7, 2020. doi:10.1101/2020.07.07.191676

[44] Lapointe CP, Grosely R, Johnson AG, Wang J, Fernández IS, Puglisi JD. Dynamic competition between SARS-CoV-2 NSP1 and mRNA on the human ribosome inhibits translation initiation. *Proc Natl Acad Sci U S A*. 2021;118(6). doi:10.1073/PNAS.2017715118/-/DCSUPPLEMENTAL

[45] Kamitani W, Huang C, Narayanan K, Lokugamage KG, Makino S. A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein. *Nat Struct Mol Biol*. 2009;16(11):1134-1140. doi:10.1038/NSMB.1680,

[46] Lei J, Kusov Y, Hilgenfeld R. Nsp3 of coronaviruses: Structures and functions of a large multi-domain protein. *Antiviral Res*. 2018;149:58-74. doi:10.1016/J.ANTIVIRAL.2017.11.001

[47] Low ZY, Zabidi NZ, Yip AJW, Puniyamurti A, Chow VTK, Lal SK. SARS-CoV-2 Non-Structural Proteins and Their Roles in Host Immune Evasion. *Viruses*. 2022;14(9). doi:10.3390/V14091991

[48] Gao Y, Yan L, Huang Y, et al. Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science*. 2020;368(6492):779-782. doi:10.1126/SCIENCE.ABB7498

[49] Bailey-Elkin BA, Knaap RCM, Kikkert M, Mark BL. Structure and Function of Viral Deubiquitinating Enzymes. *J Mol Biol*. 2017;429(22):3441-3470. doi:10.1016/j.jmb.2017.06.010

[50] Shin D, Mukherjee R, Grewe D, et al. Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity. *Nature*. 2020;587(7835):657-662. doi:10.1038/S41586-020-2601-5,

[51] Frick DN, Virdi RS, Vuksanovic N, Dahal N, Silvaggi NR. Molecular Basis for ADP-Ribose Binding to the Mac1 Domain of SARS-CoV-2 nsp3. *Biochemistry*. 2020;59(28):2608-2615. doi:10.1021/ACS.BIOCHEM.0C00309,

[52] Wolff G, Limpens RWAL, Zevenhoven-Dobbe JC, et al. A molecular pore spans the double membrane of the coronavirus replication organelle. *Science (1979)*. 2020;369(6509):1395-1398. doi:10.1126/SCIENCE.ABD3629,

[53] Tan AT, Linster M, Tan CW, et al. Early induction of functional SARS-CoV-2-specific T cells associates with rapid viral clearance and mild disease in COVID-19 patients. *Cell Rep*. 2021;34(6). doi:10.1016/j.celrep.2021.108728

[54] Angelini MM, Akhlaghpour M, Neuman BW, Buchmeier MJ. Severe acute respiratory syndrome coronavirus nonstructural proteins 3, 4, and 6 induce double-membrane vesicles. *mBio*. 2013;4(4). doi:10.1128/MBIO.00524-13

[55] Oudshoorn D, Rijs K, Limpens RWAL, et al. Expression and Cleavage of Middle East Respiratory Syndrome Coronavirus nsp3-4 Polyprotein Induce the Formation of Double-Membrane Vesicles That Mimic Those Associated with Coronaviral RNA Replication. *mBio*. 2017;8(6). doi:10.1128/MBIO.01658-17

[56] Gordon CJ, Tchesnokov EP, Woolner E, et al. Remdesivir is a direct-acting antiviral that inhibits RNA-dependent RNA polymerase from severe acute respiratory syndrome coronavirus 2 with high potency. *J Biol Chem*. 2020;295(20):6785-6797. doi:10.1074/JBC.RA120.013679

[57] Neuman BW. Bioinformatics and functional analyses of coronavirus nonstructural proteins involved in the formation of replicative organelles. *Antiviral Res*. 2016;135:97-107. doi:10.1016/j.antiviral.2016.10.005

[58] Zimmermann L, Zhao X, Makroczyova J, et al. SARS-CoV-2 nsp3 and nsp4 are minimal constituents of a pore spanning replication organelle. *Nat Commun*. 2023;14(1). doi:10.1038/S41467-023-43666-5,

[59] Oostra M, Hagemeijer MC, van Gent M, et al. Topology and Membrane Anchoring of the Coronavirus Replication Complex: Not All Hydrophobic Domains of nsp3 and nsp6 Are Membrane Spanning. *J Virol*. 2008;82(24):12392-12405. doi:10.1128/JVI.01219-08,

[60] Lin X, Fu B, Xiong Y, et al. Unconventional secretion of unglycosylated ORF8 is critical for the cytokine storm during SARS-CoV-2 infection. *PLoS Pathog*. 2023;19(1). doi:10.1371/JOURNAL.PPAT.1011128,

[61] V'kovski P, Kratzel A, Steiner S, Stalder H, Thiel V. Coronavirus biology and replication: implications for SARS-CoV-2. *Nat Rev Microbiol*. 2021;19(3):155-170. doi:10.1038/S41579-020-00468-6

[62] Zhang L, Lin D, Sun X, et al. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science*. 2020;368(6489):409-412. doi:10.1126/SCIENCE.ABB3405

[63] Zhou Y, Hou Y, Shen J, Huang Y, Martin W, Cheng F. Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Discov*. 2020;6(1). doi:10.1038/S41421-020-0153-3

[64] Wu C, Liu Y, Yang Y, et al. Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharm Sin B*. 2020;10(5):766-788. doi:10.1016/j.apsb.2020.02.008

[65] Owen DR, Allerton CMN, Anderson AS, et al. An oral SARS-CoV-2 Mpro inhibitor clinical candidate for the treatment of COVID-19. *Science*. 2021;374(6575):1586-1593. doi:10.1126/SCIENCE.ABL4784

[66] Kneller DW, Phillips G, O'Neill HM, et al. Structural plasticity of SARS-CoV-2 3CL Mpro active site cavity revealed by room temperature X-ray crystallography. *Nat Commun*. 2020;11(1). doi:10.1038/S41467-020-16954-7

[67] Su Y, Yuan D, Chen DG, et al. Multiple early factors anticipate post-acute COVID-19 sequelae. *Cell*. 2022;185(5):881-895.e20. doi:10.1016/j.cell.2022.01.014

[68] Sacco MD, Ma C, Lagarias P, et al. Structure and inhibition of the SARS-CoV-2 main protease reveal strategy for developing dual inhibitors against Mpro and cathepsin L. *Sci Adv*. 2020;6(50). doi:10.1126/SCIADV.ABE0751

[69] Jin Z, Du X, Xu Y, et al. Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors. *Nature 2020 582:7811*. 2020;582(7811):289-293. doi:10.1038/s41586-020-2223-y

[70] Kubra B, Badshah SL, Faisal S, et al. Inhibition of the predicted allosteric site of the SARS-CoV-2 main protease through flavonoids. *J Biomol Struct Dyn*. 2023;41(18):9103-9120. doi:10.1080/07391102.2022.2140201,

[71] Benvenuto D, Angeletti S, Giovanetti M, et al. Evolutionary analysis of SARS-CoV-2: how mutation of Non-Structural Protein 6 (NSP6) could affect viral autophagy. *Journal of Infection*. 2020;81(1):e24-e27. doi:10.1016/J.JINF.2020.03.058

[72] Gordon DE, Jang GM, Bouhaddou M, et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature 2020 583:7816*. 2020;583(7816):459-468. doi:10.1038/s41586-020-2286-9

[73] Cottam EM, Whelband MC, Wileman T. Coronavirus NSP6 restricts autophagosome expansion. *Autophagy*. 2014;10(8):1426-1441. doi:10.4161/AUTO.29309

[74] Gordon DE, Jang GM, Bouhaddou M, et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature*. 2020;583(7816):459-468. doi:10.1038/S41586-020-2286-9

[75] Yang Y, Yang M, Shen C, et al. Evaluating the accuracy of different respiratory specimens in the laboratory diagnosis and monitoring the viral shedding of 2019-nCoV infections. *medRxiv*. Published online February 17, 2020:2020.02.11.20021493. doi:10.1101/2020.02.11.20021493

[76] Hillen HS, Kokic G, Farnung L, Dienemann C, Tegunov D, Cramer P. Structure of replicating SARS-CoV-2 polymerase. *Nature 2020 584:7819*. 2020;584(7819):154-156. doi:10.1038/s41586-020-2368-8

[77] Slanina H, Madhugiri R, Bylapudi G, et al. Coronavirus replication-transcription complex: Vital and selective NMPylation of a conserved site in nsp9 by the NiRAN-RdRp subunit. *Proc Natl Acad Sci U S A*. 2021;118(6). doi:10.1073/PNAS.2022310118/-/DCSUPPLEMENTAL

[78] Subissi L, Posthuma CC, Collet A, et al. One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. *Proc Natl Acad Sci U S A*. 2014;111(37):E3900-E3909. doi:10.1073/PNAS.1323705111

[79] Wang Q, Wu J, Wang H, et al. Structural Basis for RNA Replication by the SARS-CoV-2 Polymerase. *Cell*. 2020;182(2):417-428.e13. doi:10.1016/J.CELL.2020.05.034

[80] Shannon A, Le NTT, Selisko B, et al. Remdesivir and SARS-CoV-2: Structural requirements at both nsp12 RdRp and nsp14 Exonuclease active-sites. *Antiviral Res*. 2020;178. doi:10.1016/j.antiviral.2020.104793

[81] Yin W, Mao C, Luan X, et al. Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science*. 2020;368(6498):1499-1504. doi:10.1126/SCIENCE.ABC1560

[82] Kokic G, Hillen HS, Tegunov D, et al. Mechanism of SARS-CoV-2 polymerase stalling by remdesivir. *Nature Communications 2021 12:1*. 2021;12(1):1-7. doi:10.1038/s41467-020-20542-0

[83] Agostini ML, Andres EL, Sims AC, et al. Coronavirus Susceptibility to the Antiviral Remdesivir (GS-5734) Is Mediated by the Viral Polymerase and the Proofreading Exoribonuclease. *mBio*. 2018;9(2). doi:10.1128/MBIO.00221-18

[84] Pachetti M, Marini B, Benedetti F, et al. Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant. *J Transl Med*. 2020;18(1). doi:10.1186/S12967-020-02344-6

[85] Kim SM, Kim EH, Casel MAB, et al. SARS-CoV-2 variants with NSP12 P323L/G671S mutations display enhanced virus replication in ferret upper airways and higher transmissibility. *Cell Rep*. 2023;42(9):113077. doi:10.1016/J.CELREP.2023.113077

[86] Jia Z, Yan L, Ren Z, et al. Delicate structural coordination of the Severe Acute Respiratory Syndrome coronavirus Nsp13 upon ATP hydrolysis. *Nucleic Acids Res*. 2019;47(12):6538-6550. doi:10.1093/NAR/GKZ409,

[87] Maio N, Heffner AL, Rouault TA. Iron-sulfur clusters in viral proteins: Exploring their elusive nature, roles and new avenues for targeting infections. *Biochim Biophys Acta Mol Cell Res*. 2024;1871(5). doi:10.1016/j.bbamcr.2024.119723

[88] Chen J, Malone B, Llewellyn E, et al. Structural Basis for Helicase-Polymerase Coupling in the SARS-CoV-2 Replication-Transcription Complex. *Cell*. 2020;182(6):1560-1573.e13. doi:10.1016/j.cell.2020.07.033

[89] Hao W, Wojdyla JA, Zhao R, et al. Crystal structure of Middle East respiratory syndrome coronavirus helicase. *PLoS Pathog*. 2017;13(6). doi:10.1371/JOURNAL.PPAT.1006474,

[90] Adedeji AO, Singh K, Calcaterra NE, et al. Severe acute respiratory syndrome coronavirus replication inhibitor that interferes with the nucleic acid unwinding of the viral helicase. *Antimicrob Agents Chemother*. 2012;56(9):4718-4728. doi:10.1128/AAC.00957-12,

[91] Shu T, Huang M, Wu D, et al. SARS-Coronavirus-2 Nsp13 Possesses NTPase and RNA Helicase Activities That Can Be Inhibited by Bismuth Salts. *Virol Sin*. 2020;35(3):321-329. doi:10.1007/S12250-020-00242-1,

[92] Chen Y, Cai H, Pan J, et al. Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. *Proc Natl Acad Sci U S A*. 2009;106(9):3484-3489. doi:10.1073/PNAS.0808790106

[93] Ogando NS, Zevenhoven-Dobbe JC, van der Meer Y, Bredenbeek PJ, Posthuma CC, Snijder EJ. The Enzymatic Activity of the nsp14 Exoribonuclease Is Critical for Replication of MERS-CoV and SARS-CoV-2. *J Virol*. 2020;94(23). doi:10.1128/JVI.01246-20

[94] Saikatendu KS, Joseph JS, Subramanian V, et al. Structural Basis of Severe Acute Respiratory Syndrome Coronavirus ADP-Ribose-1″-Phosphate Dephosphorylation by a Conserved Domain of nsP3. *Structure*. 2005;13(11):1665-1675. doi:10.1016/J.STR.2005.07.022

[95] Robson F, Khan KS, Le TK, et al. Coronavirus RNA Proofreading: Molecular Basis and Therapeutic Targeting. *Mol Cell*. 2020;79(5):710-727. doi:10.1016/j.molcel.2020.07.027

[96] Bouvet M, Debarnot C, Imbert I, et al. In Vitro Reconstitution of SARS-Coronavirus mRNA Cap Methylation. *PLoS Pathog*. 2010;6(4):e1000863. doi:10.1371/JOURNAL.PPAT.1000863

[97] Lin S, Chen H, Ye F, et al. Crystal structure of SARS-CoV-2 nsp10/nsp16 2′-O-methylase and its implication on antiviral drug design. *Signal Transduction and Targeted Therapy 2020 5:1*. 2020;5(1):1-4. doi:10.1038/s41392-020-00241-4

[98] Kim Y, Jedrzejczak R, Maltseva NI, et al. Crystal structure of Nsp15 endoribonuclease NendoU from SARS-CoV-2. *Protein Sci*. 2020;29(7):1596-1605. doi:10.1002/PRO.3873

[99] Deng X, Hackbart M, Mettelman RC, et al. Coronavirus nonstructural protein 15 mediates evasion of dsRNA sensors and limits apoptosis in macrophages. *Proc Natl Acad Sci U S A*. 2017;114(21):E4251-E4260. doi:10.1073/PNAS.1618310114

[100] Saramago M, Costa VG, Souza CS, et al. The nsp15 Nuclease as a Good Target to Combat SARS-CoV-2: Mechanism of Action and Its Inactivation with FDA-Approved Drugs. *Microorganisms*. 2022;10(2):342. doi:10.3390/MICROORGANISMS10020342

[101] Ivanov KA, Hertzig T, Rozanov M, et al. Major genetic marker of nidoviruses encodes a replicative endoribonuclease. *Proc Natl Acad Sci U S A*. 2004;101(34):12694-12699. doi:10.1073/PNAS.0403127101/SUPPL_FILE/03127FIG8.PDF

[102] Gurrapu, S., & Mamidala, E. (2017). In Vitro HIV-1 reverse transcriptase inhibition of andrographolide isolated from Andrographis paniculata. European Journal of Biomedical, 4(12), 516-522.

[103] Hackbart M, Deng X, Baker SC. Coronavirus endoribonuclease targets viral polyuridine sequences to evade activating host sensors. *Proc Natl Acad Sci U S A*. 2020;117(14):8094-8103. doi:10.1073/PNAS.1921485117,