Özkan İLMEK

A Master's Thesis

AGU 2022

IDENTIFICATION OF SINGLE DOMAIN ANTIBODIES AGAINST SARS-CoV-2 OMICRON VARIANT VIA PROTEIN-PROTEIN DOCKING APPROACHES

A THESIS
SUBMITTED TO THE DEPARTMENT OF BIOENGINEERING
AND THE GRADUATE SCHOOL OF ENGINEERING AND
SCIENCE OF ABDULLAH GUL UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

By
Özkan İLMEK
June 2022

IDENTIFICATION OF SINGLE DOMAIN ANTIBODIES AGAINST SARS-CoV-2 OMICRON VARIANT VIA PROTEIN-PROTEIN DOCKING APPROACHES

A THESIS
SUBMITTED TO THE DEPARTMENT OF BIOENGINEERING
AND THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE OF
ABDULLAH GUL UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

By
Özkan İLMEK
June 2022

SCIENTIFIC ETHICS COMPLIANCE

I hereby declare that all information in this document has been obtained in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all materials and results that are not original to this work.

Name-Surname: Özkan İLMEK

Signature :

REGULATORY COMPLIANCE

M.Sc. thesis titled Identification of Single Domain Antibodies Against SARS-CoV-2 Omicron Variant via Protein-Protein Docking Approaches has been prepared in accordance with the Thesis Writing Guidelines of the Abdullah Gül University, Graduate School of Engineering & Science.

Prepared By	Advisor
Özkan İLMEK	Assist. Prof. Dr. Şerife AYAZ GÜNER

Head of the Bioengineering Program
Prof. Dr. Sevil DİNÇER İŞOĞLU

ACCEPTANCE AND APPROVAL

M.Sc. thesis titled Identification of Single Domain Antibodies Against SARS-CoV-2 Omicron Variant via Protein-Protein Docking Approaches and prepared by Özkan İLMEK has been accepted by the jury in the Bioengineering Graduate Program at Abdullah Gül University, Graduate School of Engineering & Science.

20 /06 /2022

(Thesis Defense Exam Date)

JURY:

Advisor : Assist. Prof. Dr. Şerife AYAZ GÜNER .....

Member: Assist. Prof Dr. Emel Başak GENCER AKÇOK .....

Member: Prof. Dr. Servet ÖZCAN .....

APPROVAL:

The acceptance of this M.Sc. thesis has been approved by the decision of the Abdullah Gül University, Graduate School of Engineering & Science, Executive Board dated ….. /….. / …….. and numbered ……..…… .

…….. /…….. / ……..

(Date)

Graduate School Dean
Prof. Dr. İrfan ALAN

ABSTRACT

IDENTIFICATION OF SINGLE DOMAIN ANTIBODIES AGAINST SARS-CoV-2 OMICRON VARIANT VIA PROTEIN-PROTEIN DOCKING APPROACHES

Özkan İLMEK

MSc. in Bioengineering

Advisor: Assist. Prof. Dr. Şerife AYAZ GÜNER

June 2022

Omicron, became the dominant variant in 2022 in terms of spreading rate, has managed to evade from an immune system of patients due to its unique mutations. Single domain antibodies (sdAb) which are functionally important parts of conventional antibodies are commonly used for diagnosis and treatment. Although there are many sdAbs developed to combat coronavirus in recent years, their effectiveness against Omicron variant has not been sufficiently tested and the effect of mutations regarding neutralization level is not clear. In this study, structure modelling of 850 sdAb sequences obtained from previous studies were generated using AlphaFold 2 and effectiveness of these sdAbs against Omicron variant was tested via protein-protein docking approach. In the docking process, within a realistic approach, missing residues were completed into Spike protein PDB structures, and Spike protein homotrimer structure in closed state conformation was used. Finally, top 1000 and top 100 scores are determined as a threshold value for different protein-protein docking scoring functions such as HDOCK, PRODIGY and Bluues. sdAbs that have successful results for Omicron variant were listed. There were 4 sdAbs which exceed the threshold values after 2 different docking experiments against the Omicron variant. The scripting codes and methodological approach developed within this thesis can be used against new SARS-CoV-2 variants that may emerge in the future or other diseases.

Keywords: SARS-CoV-2, Omicron variant, Protein Structure Modelling, sdAb, Protein-Protein Docking

ÖZET

SARS-CoV-2 OMİKRON VARYANTINA ÖZGÜ TEK DOMAİNLİ ANTİKORLARIN PROTEİN-PROTEİN KENETLENMESİ YAKLAŞIMLARIYLA TANIMLANMASI

Özkan İLMEK
Biyomühendislik Anabilim Dalı Yüksek Lisans
Tez Yöneticisi: Dr. Öğr. Üyesi Şerife AYAZ GÜNER
Haziran 2022

2022 yılında baskın varyant olan Omikron, kendine has mutasyonları sayesinde hastaların bağışıklık sisteminden kaçmayı başarmış, yayılma hızı açısından önceki varyantlara göre oldukça başarılı olmuştur. Geleneksel antikorların işlevsel olarak önemli kısmı olan tek domainli antikorlar (sdAb), tanı ve tedavi amacıyla yaygın olarak kullanılmaktadır. Son yıllarda koronavirüs ile mücadele için geliştirilmiş pek çok sdAb olmasına rağmen bunların Omikron varyantına karşı etkinlikleri yeterince test edilmemiş ve mutasyonların nötralizasyon düzeyine etkisi net değildir. Bu çalışmada, AlphaFold 2 kullanılarak önceki çalışmalardan elde edilen 850 sdAb dizisinin yapı modellemesi oluşturulmuş ve bu sdAb'lerin Omicron varyantına karşı etkinliği protein-protein kenetlenmesi yaklaşımı ile analiz edilmiştir. Protein-protein kenetlenmesi işleminde, gerçekçi bir yaklaşımla, Spike proteini PDB yapılarındaki amino asit rezidü eksiklikleri tamamlandı ve SARS-CoV-2'nin Spike protein homotrimer yapısının kapalı konformasyonu kullanıldı. Son olarak, HDOCK, PRODIGY ve Bluues gibi farklı protein-protein kenetlenmesi puanlama fonksiyonları için eşik değerler olarak belirlenen top 1000 ve top 100 puanlarının taranması sonucunda, Omikron varyantına karşı 2 farklı kenetlenme deneyi için toplamda 4 adet yüksek oranda başarılı sdAb tespit edilmiştir. Tez kapsamında geliştirilen kod ve deneysel yaklaşımlar yeni çıkabilecek SARS-CoV-2 varyantlarına veya diğer hastalıklara karşı kullanılabilecektir.

Anahtar kelimeler: SARS-CoV-2, Omikron varyantı, Protein Yapı Modellemesi, sdAb, Protein-Protein Kenetlenmesi

Acknowledgements

Primarily, I would like to express my biggest gratitude to my advisor, Assist. Prof. Dr. Şerife AYAZ GÜNER, for her endless support, patience, and kindness. Her experiences and belief in me have been my most important motivation for all kinds of difficulties which I faced during thesis study.

I would like to express my special appreciation to Bioinformatics Specialist Hüseyin GÜNER for the technical support and extensive opportunities he provided along the thesis process. All extensive studies within this thesis could have remained only a proposal or design without his help. We could have a very good teamwork owing to his experience. We overcame every problem together in this journey and we consistently set new inspiring goals without limits.

I am greatly indebted to my beloved family, my mother Habibe İLMEK, my father Cemil İLMEK, my brother Erkan İLMEK and my sister Derya TANDOĞAN, who are always in my mind and heart for every moment during my life. Their confidence and supports made me stronger for every condition. I got the most important and valuable education in my life from their advice and nurture. Moreover, I want to send lots of love and thanks to my best friends Asel and Arda TANDOĞAN.

Also, I would like to present my gratitude to my jury committee members, Assist. Prof. Dr. Emel Başak GENCER AKÇOK and Prof. Dr. Servet ÖZCAN, for their supports, evaluations, and valuable suggestions for future.

In addition, I would like to thank workmates in Faculty of Nature and Life Sciences and all members of Protein Production and Characterization Laboratory for their contributions.

Finally, I thank to every organism who contributed to preparation of this thesis.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Å	Angstrom
ACE2	Angiotensin Converting Enzyme 2
ALI	Acut Lung Injury
ARDS	Acute Respiratory Distress Syndrome
BCoV	Bovine coronavirus
CcoV	Canine coronavirus
CDR	Complementarity Determining Region
Covid-19	Coronavirus Disease 2019
DMV	Double-Membrane Vesicles
DPP4	Dipeptidyl Peptidase 4
E	Envelope Protein
FIPV	Feline Infectious Peritonitis Virus
HcAb	Heavy Chain Only Antibodies
HcoV	Human Coronavirus
HE	Hemagglutinin Esterase
ICTV	International Committee for Taxonomy of Viruses
ILZ	Isoleucine Zipper Domain
M	Membrane Protein
MBD	Metal Binding Domain
mCEACAM	Murine Carcinoembryonic Antigen-Related Adhesion Molecule 1
MERS-CoV	Middle East Respiratory Syndrome Coronavirus
MHV	Murine hepatitis virus
N protein	Nucleocapsid Protein
Nsps	Non Structural Proteins
orf1ab	Open Reading Frame 1ab
Paratome	Antigen Binding Regions Identification Tool
PDB	Protein Data Bank
PEDV	Porcine epidemic diarrhea virus
PN	aminopeptidase N
pp1ab	Polyprotein 1ab
PRODIGY	The Protein Binding Energy Prediction
RBD	Receptor Binding Domain
RSSB-PDB	Research Collaboratory for Structural Bioinformatics
S protein	Spike Protein
SARS-CoV	Severe Acute Respiratory Syndrome Coronavirus
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
sdAb	Single Domain Antibody
TGEV	Transmissible Gastroenteritis Virus
TMPRSS2	Transmembrane Proteases serine 2
UTR	Untranslated Region
VHH	Variable Region of Antibody
VSG	Variant of Surface Glycoprotein
WHO	World Health Organization

To my family

Chapter 1

INTRODUCTION

1. 1 Coronaviruses

Coronaviruses are member of single-stranded RNA viruses [1] and belong to *Coronavirinae* subfamily of *Coronaviridae* family [2]. The taxonomic classification of the SARS-Cov-2 virus is as follows: *Riboviria* (Realm), *Nidovirales* (Order), *Cornidovirineae* (Suborder), *Coronaviridae* (Family), *Orthocoronavirinae* (Subfamily), *Betacoronavirus* (Genus), *Sarbecovirus* (Subgenus), *Severe acute respiratory syndrome-related coronavirus* (Species) and SARS-CoV-2 (Individuum) [3] (Figure 1.1). Their diameter can change 60 nm to 140 nm [1]. Severe acute respiratory syndrome coronavirus (SARS-CoV), Middle East respiratory syndrome coronavirus (MERS-CoV), H1N1 2009 virus and H5N1 influenza A virus caused crucially important diseases such as acute respiratory distress syndrome (ARDS) and acute lung injury (ALI) in recent decades worldwide [4–8]. In addition to this, it has been reported that coronaviruses can infect humans and animals as well as cause respiratory diseases such as cold, pneumonia in previous studies [9–14].

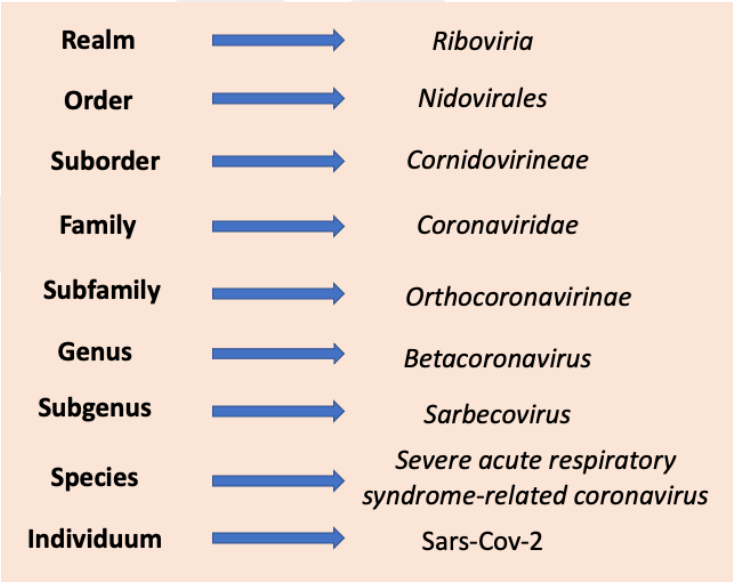

Figure 1.1 Taxonomy of SARS-CoV-2. (This figure is adapted from [3].

1.1.1 SARS-CoV-2 (Covid-19)

In December 2019 new version of coronavirus which highly causes pneumonia in infected patients had occurred in Wuhan, China. After the further genetic analyses, this new version was named by International Committee for Taxonomy of Viruses (ICTV) as a severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), World Health Organization (WHO) called disease as a Coronavirus Disease 2019 (Covid-19) in 11 February 2020 and WHO increased risk assessment of COVID-19 very high a global level [3,15,16] (Figure 2.1). According to WHO, 537,591,764 confirmed cases and 6,319,395 deaths were recorded as of June 20, 2022 [17]. There are four main genera of coronaviruses. alpha (α), beta (β), gamma (γ) and delta (δ). SARS-CoV, MERS-CoV and SARS-CoV-2 which are responsible for respiratory distress syndrome diseases, belong to *Betacoronavirus* genus [3]. It has been determined that SARS-CoV-2 is more pathogenic than previous versions such as SARS-CoV (in 2002) and MERS-CoV (in 2013) [18].

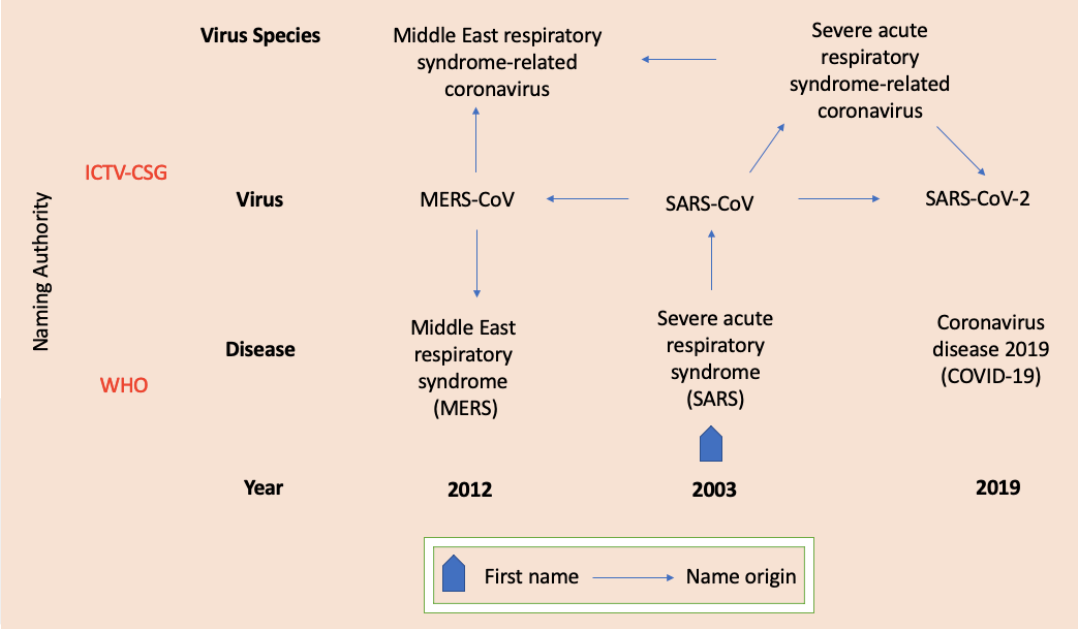

Figure 1.2 Relation between coronavirus species, their taxonomy, and diseases. (This figure is adapted from [3]).

1.1.2 Genome and Structure of SARS-CoV-2

SARS-CoV-2 is a member of *Betacoronavirus* genus [3]. The genomes of MERS-CoV, SARS-CoV, and SARS-CoV-2 have non-segmented positive-sense single-stranded RNA lengths ranging from 26 to 32 kb [19]. The SARS-CoV-2 has 80% identical to previous human coronaviruses regarding genome similarity [20]. There are four major

structural genes in the SARS-CoV-2 genome and these genes are responsible for coding structural genes such as envelope (E), membrane (M), spike (S) and nucleocapsid (N). Polyprotein 1ab (pp1ab) and 15 non-structural proteins (nsps) are encoded by Orf1ab [18,20,21]. In previous studies, significant variations have been noted between SARS-CoV and SARS-CoV-2 genomes. These variations are alterations in amino acid sequence of 3c and 8b proteins and lack of 8a protein [20]. Moreover, spike protein of SARS-Cov-2 use angiotensin-converting enzyme 2 (ACE2) to enter host cells similarly with SARS-CoV [22]. It was shown that mutation at 501. position in amino acid sequence (N501T) of the SARS-CoV-2 genome increases the binding affinity for ACE2 [23].

It has been stated that the 5′UTR and 3′UTRs of *Betacoronaviruses* play important roles in cellular events, such as replication, transcription, intermolecular and intramolecular interactions, and RNA-RNA interactions in different studies [24–26]. ORF1a and ORF1b take up two-thirds of the virus genome and are responsible for the expression of 16 nonstructural proteins, which are significantly important for viral replication and transcription [27]. SARS-CoV, MERS-CoV and SARS-CoV-2 genomes have a specific number of accessory genes with important functions. While 8 functional accessory genes (3a, 3b, 6, 7a, 7b, 8a, 8b, 9b) found in the SARS-CoV genome, 5 accessory genes (3, 4a, 4b, 5, 8b) present in the MERS-CoV genome. The functional accessory gene number is 6 (3, 6, 7a, 7b, 8, 9b) for SARS-CoV-2 [28–30].

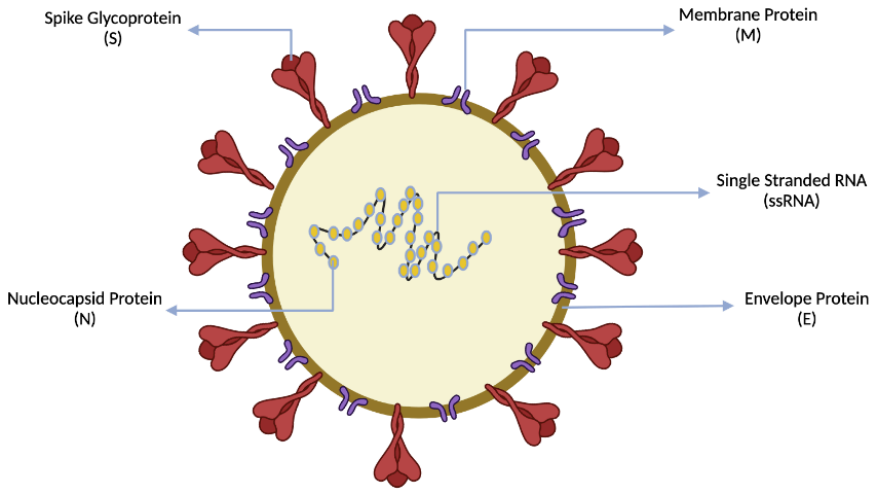

Figure 1.3 Structural proteins of coronaviruses. (Figure was adapted from [31,32] and was designed by BioRender).

Spike (S), envelope (E), membrane (M), and nucleocapsid proteins are commonly found structural proteins in coronaviruses [33]. RNA genome of coronaviruses is located in N protein and other S, E, and M proteins are embedded the viral envelope [30]. The

Spike protein has a very high glycosylation rate and is responsible for entry into host cells as it contains the receptor binding domain (RBD) [30,34]. Spike protein of coronaviruses has highly variable motifs to interact with RBDs of different hosts. Two different glycoproteins which are responsible for infection of host cells can be found in coronaviruses. Hemagglutinin esterase (HE) is specific to some species and Spike glycoprotein trimer (S) is more common among beta coronaviruses. [35–38]. Polybasic cleavage site (RRAR/S) which play an important role during viral infection is found in SARS-CoV and SARS-CoV-2. This cleavage site is cut by host furin-like protease to separate two subunits (S1 and S2) of spike protein [34,39]. Moreover, the ratio of the amount of M protein in the coronaviruses is higher than the E protein because it is essential for the formation of the virus structure [40,41]. On the other hand, E protein blocks the host cell stress response, acts as an ion channel, and ensures the secretion of mature virions from the host cells after the cycle [42].

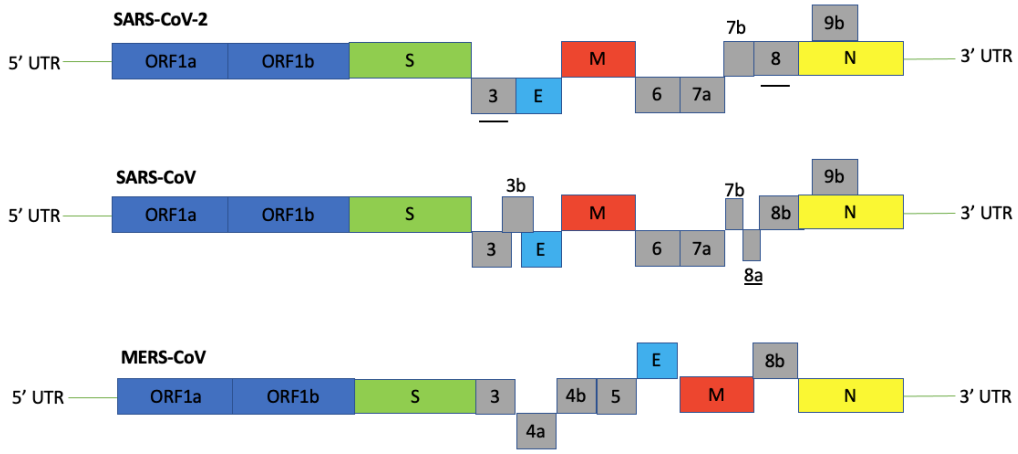

Figure 1.4 Genome organization of *Betacoronaviruses*. This figure was adapted from [1]**.** 5'UTR (untranslated region), ORF1a/b (encoding nonstructural proteins, dark blue box), S (encoding spike protein, green box), numbers (encoding accessory proteins, pink box), E (encoding envelope protein, light blue box), M (encoding membrane protein, red box), N (encoding nucleocapsid protein, yellow box) and 3'UTR are represented. Boxes which are black underlined show important mutation points for SARS-CoV (8a) and SARS-CoV-2 (3 and 8).

There are 16 non-structural proteins which have been discovered before in coronaviruses. These proteins have very important functions such as RNA processing, replication, ssRNA binding, cap methylation of viral mRNAs, prevention of mRNA splicing, protein translation and protein trafficking in host, RNA-dependent RNA polymerase activity for transcription, providing zinc binding domain, endo- and

exoribonuclease activity, methyltransferase activity, transmembrane domain activity [27,43,44].

1.1.3 Spike (S) Glycoprotein

The most important structural protein is the Spike (S) glycoprotein, which is responsible for virus attachment to host cells. In coronaviruses, Spke proteins are class I viral fusion proteins and must be cleaved to be activated. Spike protein is cleaved by a host cell protease into two separate polypeptides, named as S1 and S2. Coronavirus Spike proteins can be cleaved, depending on virus strains, by several host proteases, including Furin, Trypsin, Cathepsins, Transmembrane Protease Serine Protease-2 (TMPRSS-2) and TMPRSS-4 [45]. The presence of these proteases in target cells determines whether the virus can enter to host cells via the cell surface pathway or endocytosis. TMPRSS2 is responsible for the preparation of the Spike protein in SARS-CoV and SARS-CoV-2. Then, ACE2 acts as a suitable receptor for the entry of viruses.

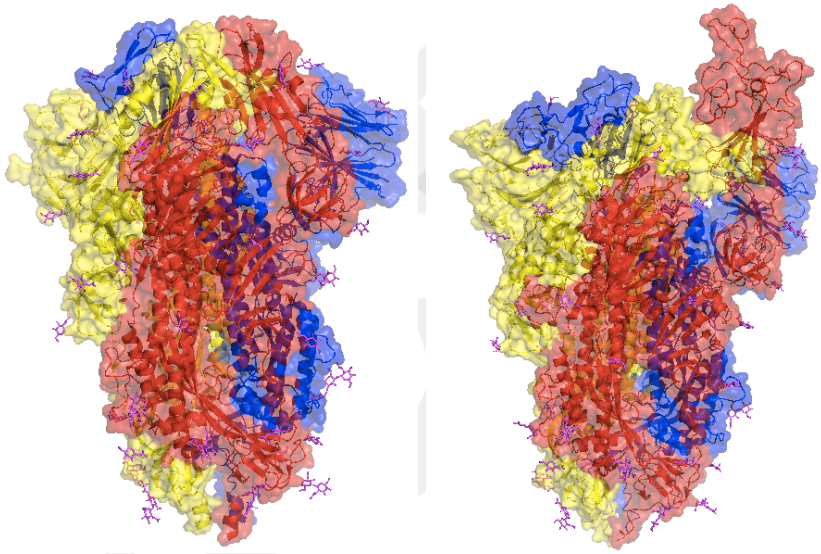

Figure 1.5 SARS-CoV-2 Spike protein trimer conformations. Closed state of (Left, PDB ID:6VXX) and open state (one RBD up) conformation of Spike protein (right, PDB ID:6XM3). Spike protein has homotrimer structure, Chain A (yellow), Chain B (red) and Chain C (blue). NAG structures on spike protein were shown in magenta color.

Spike proteins can be found in two different states in infected organisms (Figure 1.5). The first is the closed state with all RBDs in the down position. In the second state, open conformation, at least one RBD is in up position. Sometimes 2 RBDs or 3 RBDs can be in the up position, too. The virus is in close state conformation for most of the time

when it places in an infected cell, but it passes to open state shortly before interacting with ACE2.

Table 1.1 Coronavirus non-structural proteins (nsps) and their functions. This table was adapted from [27].

Protein	Function	References
Non-structural protein 1	inhibition host cell translation and innate immune response, mRNA degradation	[46–49]
Non-structural protein 2	Attachment to prohibitin	[50,51]
Non-structural protein 3	function as a transmembrane protein, interaction with Nucleocapsid protein, activation of ADRP, stimulation of cytokine expression inhibition of host innate immune response, cleavage of viral polyprotein via PLPro/Deubiquitinase domain	[52–59]
Non-structural protein 4	function as a transmembrane scaffold protein and has role for proper structure of DMVs	[60,61]
Non-structural protein 5	main protease, cleavage of viral polyprotein,	[62]
Non-structural protein 6	function as a transmembrane scaffold protein	[63]
Non-structural protein 7	complex formation with non-structural protein 8, has role in formation of processivity clamp for RNA polymerase	[64]
Non-structural protein 8	forms hexadecameric complex with nsp7, may act as processivity clamp for RNA polymerase; may act as primase	[64,65]
Non-structural protein 9	Binding to RNA	[66]
Non-structural protein 10	cofactor for non-structural protein 14 and non-structural protein 16 which are responsible for heterodimer formation, stimulation of 2-O-MT and ExoN	[67,68]
Non-structural protein 12	RNA helicase, 5′ triphosphatase	[69]
Non-structural protein 13	RNA-dependent RNA Polymerase	[70,71]
Non-structural protein 14	3′-5′ exoribonuclease, N7 DNA methyltransferase	[72–75]
Non-structural protein 15	viral endoribonuclease (NendoU endoribonuclease)	[76,77]
Non-structural protein 16	mRNA cap 2'-O-ribose methylation, shielding viral RNA from melanoma differentiation-associated protein 5 (MDA5) recognition	[78,79]

1.1.4 Entry Mechanism of SARS-CoV-2 into Cells

SARS-CoV-2 virus usually uses the host cell's ACE2 receptor to enter the cell, and TMPRSS2 serine proteases to prepare the Spike protein for fusion [27]. After the Spike protein binds to the ACE2 receptor, Spike protein undergoes conformational changes to carry out fusion between the viral envelope protein and the membrane of the host cell [80]. As a result of membrane fusion, the virus is covered with the host cell membrane and an intracellular endosome structure is formed [81]. This leads to the release of viral RNA into the host cytoplasm. For the synthesis of structural and non-structural proteins, the virus uses both its own and host cell mechanisms. Viral RNA which enters to the cell after fusion undergoes translation, synthesis of non-structural proteins that are responsible for replication and transcription processes occurs. The necessary RNA and proteins to form the virion are associated together with the help of the endoplasmic reticulum and Golgi of the host cell, and these virions are sent back out of the cell via the vesicles [82,83].

SARS-CoV-2 virus particles perform attachment to host cell and fusion processes via its Spke protein. Spike protein is homotrimer, found into the membrane as multiple and it seems crown-like. Many types of viruses such as Ebola, HIV, and H5N1 influenza A, use their glycoproteins for entry into host cells. After viruses attach to the host cell surface, their glycoproteins are cut, separated into two subunits as transmembrane and extracellular [84–86]. S1 subunit binds to ACE2, while the S2 subunit provides the attachment of Spike protein to the membrane. This event also occurs in SARS-Cov-2 viruses. The Spike protein of SARS-CoV-2 is cleaved by proprotein convertases such as furin of infected organisms and infected cells can release new viruses easily by using this mechanism [87,88].

For viral entry of SARS-CoV-2, two major cleavage events on spike protein have significant role. The first one is occurred in between S1 and S2 subunits when the SARS-CoV-2 attach to ACE2 receptor on cell membrane in the cell surface entry pathway [35]. Another cleavage event is necessary at S2' site, in membrane fusion process, which is a crucially important step to release viral RNA into cell cytoplasm by using surface entry pathway. In addition to this, S2' site is cleaved to release viral RNA after endosome formation by Cathepsin L in the endosomal pathway. S1-S2 cleavage is also important for after maturation to release of new virions in virus-producing cells.

There are two common ways as endosomal and cell surface entry of SARS-CoV-2 into host cells. In the endosomal entry pathway, target host cells do not have enough numbers of transmembrane proteases such as serine 2 (TMPRSS2). Therefore, SARS-CoV-2 tends to clathrin-mediated endocytosis in this condition. First, SARS-CoV-2 binds to ACE2 receptor, and it is engulfed via clathrin-mediated endocytosis by membrane of host cell [89,90]. Next, an endosome structure is formed. When acidic suitable condition is occurred inside of endosome, Cathepsin L cleavages S2' site [91,92]. Viral RNA is released following membrane fusion (Figure 1.6). In the cell surface entry pathway, SARS-CoV-2 can find necessary transmembrane protease (serine 2 TMPRSS2). First virus binds to ACE2. The site which is present between S1 and S2 subunits, S1-S2 boundary, is cleavaged by furin protease [27,35,88]. Some conformational changes occur on Spike protein. Subsequently, serine 2 TMPRSS2 cuts S2' site [93–95]. Membrane fusion occurs, and viral RNA is released into the cytoplasm of host cell (Figure 1.7).

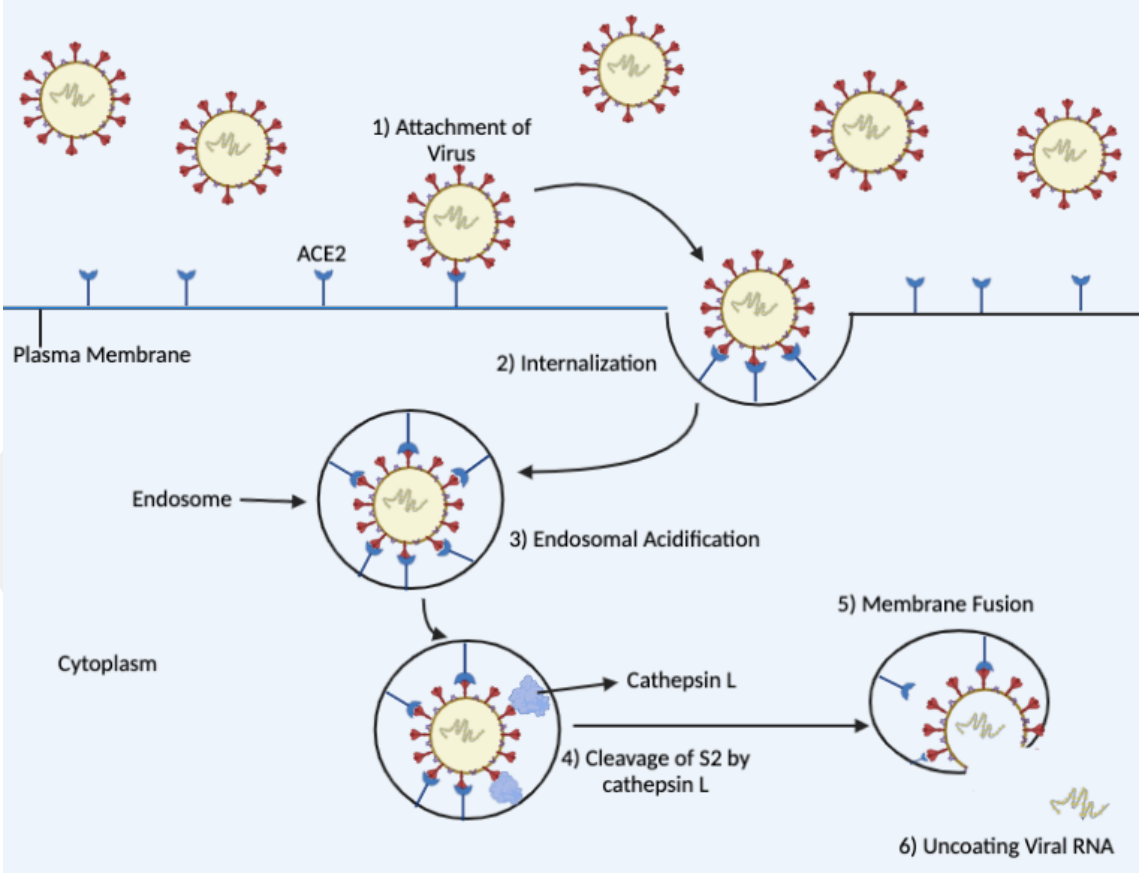

Figure 1.6 Endosomal cell entry pathway of SARS-CoV-2. (Figure was adapted from [96] and was designed by using BioRender).

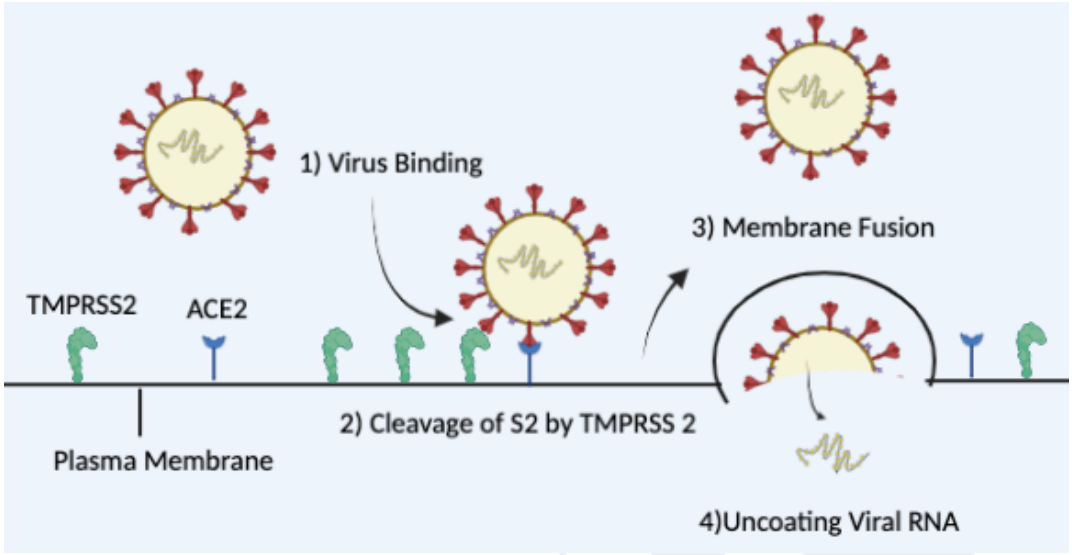

Figure 1.7 Cell surface entry of pathway of SARS-CoV-2. (Figure was adapted from [96] and was designed by uisng BioRender).

1.1.5 SARS-CoV-2 Omicron (B.1.1.529) Variant

Since the SARS-CoV-2 virus emerged and expand worldwide, it has undergone many mutations to escape the immune systems of the host organisms. On November 26, 2021, the World Health Organization declared the highly contagious Omicron (B.1.1.529) variant as a variant of serious concern [97]. For 2022, the dominant variant is Omicron (B.1.1.529).

1.1.6 Omicron (B.1.1.529) Variant Mutations

The location of many Omicron variant mutations on the S gene has played an important role in making it the dominant variant. The presence of 15 new mutations on the RBD of the Omicron variant was detected [98–103]. Some RBD mutation sites, such as Lys417Asn, Ser477Asn, Thr478Lys, Glu484Ala, and Asn501Tyr, have been shown to help SARS-CoV-2 viral escape from the host immune system by increasing its binding affinity for ACE2 [99,104–112]. Six of mutations in RBD is found in receptor binding motif (RBM) of Spike protein. These are Asn440Lys, Gly446Ser, Gln493Arg, Gly496Ser, Gln498Arg, and Tyr505His and are significantly efficient in formation of loop between 470-490 [98]. His557Tyr, Asn679Lys and Pro681His mutations are very close to the cleavage site of the furin protease [113–115]. There are six new mutations Asn764Lys, Asp796Tyr, Asn856Lys, Gln954His, Asn969Lys and Leu981Phe, in S2 subunit of Spike protein [113,114].

Table 1.2 Mutation list of major SARS-CoV-2 variants [102]**.**

Variant of Concern	List of Mutations	Lineage
Omicron	Ala67Val, del69-70, Thr95Ile, Gly142Asp, del143-145, del211, Leu212Ile, ins214GluProGlu, Gly339Asp, Ser371Leu, Ser373Pro, Ser375Phe, Lys417Asn, Asn440Lys, Glu446Ser, Ser477Asn, Thr478Lys, Glu484Ala, Gln493Arg, Gly496Ser, Gln498Arg, Asn501Tyr, Tyr505His, Thr547Lys, Asp614Gly, His655Tyr, Asn679Lys, Pro681His, Asn764Lys, Asp796Tyr, Asn856Lys, Gln954His, Asn969Lys, Leu981Phe	B.1.1.529
Alpha	del69-70, del144, Asn501Tyr, Ala570Asp, Asp614Gly, Pro681His, Thr716Ile, Ser982Ala, Asp1118His	B.1.1.7
Beta	Leu18Phe, Asp80Ala, Asp215Gly, del242-244, Arg246Ile, Lys417Asn, Glu484Lys, Asn501Tyr, Asp614Gly, Ala701Val	B.1.351
Gamma	Leu18Phe, Thr20Asn, Pro26Ser, Asp138Tyr, Arg190Ser, Lys417Thr, Glu484Lys, Asn501Tyr, Asp614Gly, His655Tyr, Thr1027Ile, Val1176Phe	P.1
Delta	Thr19Arg, Gly142Asp, Glu156Gly, del157-158, Leu452Arg, Thr478Lys, Asp614Gly, Pro681Arg, Asp950Asn	B.1.617.2

1.2 Heavy Chain Only Antibodies

Antibodies are biotechnological products that are frequently used for diagnosis and treatment in the field of health and medical biotechnology. Apart from the known antibodies in the *Camelidae* family, there are also antibodies called heavy chain only antibodies (HcAb), which contain only two heavy chains. Hamers-Casterman discovered the presence of antibodies consisting only of the heavy chain, aside from conventional antibodies, in camels and llamas [116]. These antibodies have the potential to make a very important social and economic contribution to their small molecular structure, structural stability, and suitability for high-scale production in industry.

1.2.1 Single Domain Antibodies (sdAbs)

Antibodies are basically proteins that contain two heavy and two light chains in their structure. They are quite important regarding their usage in diagnosis and treatment of various diseases. But because of their big size and affecting large areas in body,

antibodies can cause unwanted immunogenic reaction or side effects in treatment and their long half-life causes a large amount of adverse background formation for molecular imaging in diagnosis [117,118]. *Camelidae* family antibodies do not have light chains. These antibodies are called Heavy Chain Only Antibodies (HcAb) and they consist of two constant regions, a junction region and two variable regions (VHH; antigen binding fragment) so they have only heavy chain [119]. This variable domain (VHH) has all the binding and strong antigen affinity of an antibody.

The part that can bind to the antigens of HcAbs which consists of heavy chains, is also called single domain antibody (sdAb). A normal-sized human antibody has a molecular weight of approximately 160 kDa. While HcAbs are nearly 75 kDa, these sdAbs have only around 15 kDa molecular weight. As can be understood, sdAbs are much smaller than antibodies, they can bind to internal regions and are functional as a whole antibody molecule. Thanks to their small size and long CDR3 domain, they attach to surfaces that cannot be reached by standard antibodies on the protein surface. This allows the capture of toxins or epitopes that could not be detected with conventional antibodies before [120]. They are acceptably successful for tissue penetration and can pass through the blood-brain barrier. [121–125].

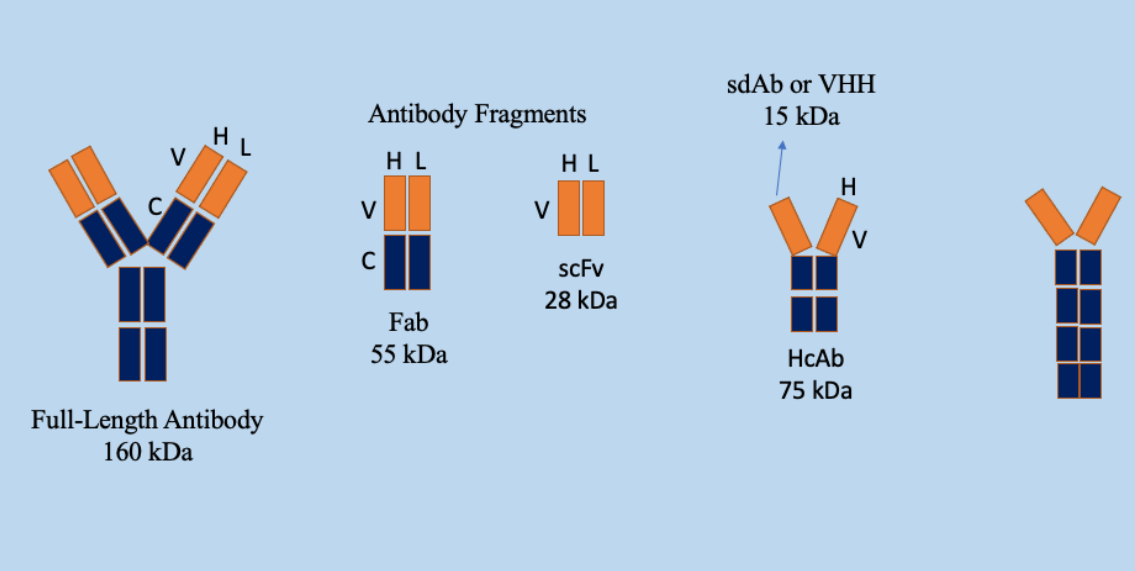

Figure 1.8 Structure models of conventional antibody, antibody fragments and single domain antibody and size comparison. H: Heavy chain, L: Light chain, C: Constant region, V: Variable region, Fab: Fragment antigen-binding, scFv: Single-chain variable fragment, sdAb: Single domain antibody, VHH: Antigen binding fragment HcAb: Heavy chain only antibody.

Single domain antibodies are much more cost effective than antibodies because they can be easily produced in bacteria and fungi [126]. They are quite favorable for engineering and have a high affinity for different targets in large application area. Because the variable parts of sdAbs are more hydrophilic, and have a monomer structure, they are more soluble and more stable than conventional antibodies. High solubility rate also prevents aggregate formation [127]. In addition to this, stability of sdAbs is very good with a remarkable level for high temperature, high pressure, and low pH. They can work well even in the presence of proteases [128–133].

1.2.2 Single Domain Antibody (sdAb) Based Approaches in Infectious Diseases

There are significant studies in the literature regarding the usage of sdAbs for infectious diseases. In a study conducted in 2001, it was determined that the sdAbs obtained from a study in which *Camelus dromedarius* animals were immunized with TEM-1 and BcII beta-lactamases were considerably successful in inhibition of these beta-lactamases and provided an increase of ampicillin sensitivity in bacteria [134]. Moreover, small domain antibody fragments to conserved epitopes in variant of surface glycoproteins (VSG) of the African trypanosome paradigm have been developed. It has been reported that small antibody fragments have the ability to penetrate the VSG coat for targeting epitopes in condensed lectin structure [135]. sdAbs produced in llamas targeting the cell wall protein Malf1 of *Malassezia furfur* fungi, which cause dandruff formation in the hair, provided successful results in high urea concentration and presence of shampoo that has extreme ionic conditions [131]. Rotavirus-induced diarrhea-targeted Llama VHH fragments were produced by *Lactobacillus paracasei* bacteria and it was determined that these sdAbs decreased level of diarrhea infection in cell culture and mouse model experiments [126]. In another study, an avian influenza virus A/Mallard/Pennsylvania/10218/84 (H5N2) was used for immunization of *Camelus bactrianus*. After immunization, new type of sdAb fragments which involves the addition of the isoleucine zipper domain (ILZ) was created. Due to this modification, it has been recorded that improvement in virus neutralization [136]. Produced sdAbs, anti-flagellin VHHs, against to flagellum of *P. aeruginosa* had acceptable results in vitro assays for prevention of swimming and biofilm formation [137]. It was shown in vitro that the single-domain antibody fragments bound and inhibited H5N1 virus infection which

influenza virus neuraminidase (NA) was used as a target for alpaca (*Vicugna pacos*) immunization [138].

1.2.3 Single Domain Antibody (sdAb) Based Diagnosis Approaches

Single domain antibodies (sdAb) are important partners for molecular imaging. Because of their small size, they allow multiple targeting by combining multiple sdAbs, which is an important property in the imaging field. In addition, kidney glomeruli, which have 30-50 kDa filtering feature, can rapidly remove sdAbs with an average size of 13 kDa from the blood circulatory system. Rapid removal of sdAbs significantly reduces their likelihood of causing toxicity [139]. Molecular optical imaging using monoclonal antibodies has a low tumor to background ratio. Scientist developed anti-HER2 VHH-IR for imaging of human breast tumor xenografts. Anti-HER2 VHH-IR has a detection rate 20 times faster than traditional methods with a much higher contrast between tumor and background tissue [140]. Researchers performed fusion reaction between VHH fragments of HcAbs in alpacas and fluorescent proteins, generated Chromobody molecule which can diffuse into subcellular compartments of living cells. They expressed in living cells and visualized S phase and mitosis [141]. Furthermore, transient interactions between metal binding domains (MBDs) and how copper ions regulate subcellular localization of ATP7B, a copper transporter, were elucidated by using a sdAb-based design [142].

1.2.4 Single Domain Antibody (sdAb) Based Approaches in Covid-19 Disease

Neutralization potential of sdAbs was tested against SARS-CoV-2 recently in various studies. It was established that isolated sdAb from llamas can attach to RBD of Spike protein and prevent interaction between ACE2 and SARS-CoV-2 [143]. TY1 also is a successful sdAb candidate which binds to RBD with high affinity for the SARS-CoV-2. It has been indicated that isolated and characterized TY1 was targeting receptor binding site of the SARS-Cov-2 spike protein and inhibiting its interaction with ACE2 in the study which SARS-CoV-2 virus and alpaca (*Lama pacos*) were used [144]. Moreover, it was shown using in vitro experiments that two sdAbs, H11-H4 and H11-D4, can block RBD-ACE2 interaction [145]. There are many sdAbs that have achieved successful results against SARS-CoV-2, which is also mentioned in the experimental procedure part of the thesis (Chapter 2).

1.3 Structure Modelling of Antibody Derivatives

Antibodies are proteins that have crucially important functions in the immune system. Many of the available antibody structures have been obtained under laboratory conditions using experimental processes such as X-ray crystallography, nuclear magnetic resonance (NMR), and cryo-electron microscopy (cryo-EM). Besides the economic drawbacks of these methods, it is not possible to use for prediction of new antibodies. It is not suitable for large-scale libraries or screening studies that have been created as a result of information obtained from previous studies and have been developed using new approaches, such as site-specific rational mutations. Therefore, the usage of antibody structure prediction approaches in creation structural modelling of antibody derivatives is inevitable for biological research. Because sdAbs have relatively small structures compared to antibodies, approaches which are used in antibody modelling are favorable for sdAb structure modelling too.

1.3.1 RosettaAntibody

RosettaAntibody is a server which is used to predict antibody variable region structures depending on amino acid sequences of heavy and light chains. RosettaAntibody uses homologous template structures of light and heavy chains for this process. Software also has a high resolution protocol which can optimize torsion angles of complementarity determining region (CDR) for steric clashes [146].

1.3.2 AlphaFold

AlphaFold is a leader program that attempts to predict and visualize the 3D structures of proteins. The software workflow attempts to improve the accuracy of structure prediction based on the PDB structures of existing proteins and using state-of-the-art machine learning techniques referred to as DeepMind including utilization of multiple alignments.

1.4 Protein-Protein Docking

Proteins are molecules that carry out many tasks for organisms. They are responsible for various tasks, such as structural processes, cellular signal transduction, performing of intracellular chemical reactions, cellular defense, and transport of

molecules. While the responsible proteins perform all these tasks, they interact with other proteins in cellular processes. The elucidation of these interactions is essential for understanding the mechanism of cellular events and behaviors. Protein-protein docking experiments used for this purpose are very useful approaches in terms of shedding light on these processes. Protein-protein docking processes are usually carried out with two main members of the experiment. The first is the target protein with a large conformation called the receptor, while the second is the ligand proteins that have been tested for their ability to bind and neutralize a target slightly smaller than the receptor.

1.4.1 HDOCK

HDOCK server is a useful and comprehensive software for homology-based studies, macromolecule structure prediction, template based modelling and macromolecular docking experiments. Amino acid sequences or PDB (Protein Data Bank) structures can be directly used in the HDOCK server. The software is used a specific hybrid algorithm which depends on template-based and template-free docking to predict structures of ligand and receptor and to determine interaction between them. Furthermore, the server has a special workflow and scoring strategy for protein-RNA/DNA docking experiments besides protein-protein docking. A normal docking experiment is completed in approximately 30 minutes. After the docking experiment is concluded, the docking model with top 100 scores, which shows interaction between receptor and ligand can be downloaded from the software and the top 10 of them can be visualized on the server [147,148].

1.4.2 Paratome

The target regions of the receptor molecules used for protein-protein docking experiments were usually determined, but antibody or sdAb loop regions which are functional in docking should be specified to target these regions. These interaction loop regions on antibodies or sdAbs are called as a paratope (or antigen binding region) which are interacting with the epitope of target antigens. Antigen Binding Regions Identification Tool (Paratome) is a beneficial server to find paratope sites by using special algorithm. Software uses alignment consensus regions of antibody-antigen complexes as a template in paratope determination process [149]. Important interaction sites generally overlap with complementarity determining regions (CDRs). However, it was reported that 22%

of antigen-binding sites are found outside of CDRs [150]. Paratome software can estimate paratope sites depends on 3D structure or amino acid sequence of antibody or sdAb. Moreover, it has been reported that Paratome web server can predict antigen binding residues with 94% accuracy [150].

1.5 Protein-Protein Docking Scoring and Analyzing Tools

1.5.1 Bluues

Bluues scoring program uses electrostatic properties of proteins to find out the best accurate scoring for docking experiments. The blues program relies on seven main approaches to fully reveal its scoring. These are generalized Born radius of each atom, pH-dependent properties, pKa of all ionizable groups, electrostatic potential in a volume surrounding the molecule, electrostatic solvation free energy, electrostatic forces on each atom and electrostatic potential at the surface of the molecule [151].

1.5.2 Protein Binding Energy Prediction (PRODIGY)

The PROtein binDIng enerGY prediction (PRODIGY) web server is a useful tool for estimating binding affinities and the predisposition of protein-protein interactions after docking experiments. In this respect, the PRODIGY can be preferred for many purposes, such as determining the biological functions of small ligand molecules or proteins related to diseases. Server achieves this via its strong predictive ability for interactions between protein-protein complexes. Moreover, special calculations made by the software about intermolecular interactions are factors that increase the accuracy of prediction rate. Software utilizes 3D protein structures of interaction complexes when estimating binding energies [152].

1.5.3 LigPlot+

LigPlot+ is a visualization system for ligand-protein or protein-protein interactions. System uses 3D coordinates of complexes to occur multiple 2D diagrams of ligand-protein interactions. The diagrams show hydrophobic or hydrogen-bond interactions between side or main chain of protein and ligand. The list of interactions can be obtained

as a file from the system. In addition, the interactions of all ligands that bind to the same protein at the same time can also be viewed at once [153].

1.5.4 PyMOL

PyMOL is a program which is used for visualization of protein interactions in 3D structures of complexes. The interaction regions can be focused and the bonds between atoms can be visualized on the 3D structure by using the program. It is an ideal application for labeling, detailed visualization and cleaning processes for all small parts of 3D structures, from chain differences to atomic residues [154].

1.6 Aim of The Study

The coronavirus outbreak which started end of the 2019, has a status of a pandemic worldwide thus far, because of the virus has a high infectivity rate, undergone various mutations and evolved different variants. Due to each new variant of the SARS-CoV-2 virus has different mutations, the success of many antibody-based approaches developed for therapeutic purposes is declining. Omicron, a dominant variant in 2022, has managed to evade from many antibodies and sdAbs due to its unique mutations, in terms of spreading rate, it has been quite successful compared to previous variants. Although there are many sdAbs developed to combat coronavirus in recent years, their effectiveness against the new Omicron variant has not been sufficiently tested and the effect of mutations regarding neutralization level is not clear.

Therefore, testing sdAbs known sequences and have proven efficacy against previous variants of the SARS-CoV-2 virus is a very rational approach. In this study, the effectiveness of sdAbs, which were successful for previous variants obtained from the literature and databases, against the Omicron variant, was evaluated via protein-protein docking experiments. Even if the tested sdAbs do not show the expected effect against the Omicron variant, their efficacy can be increased by making some mutations on possible interaction points. The important thing here is to establish a rational and repeatable method for sdAbs selection. In this way, time-consuming and difficult processes, including animal experiments, can be shortened. The main aim of this study is developing a new strategy for screening and analyzing effective nanobody candidates against to various pathogens including new SARS-CoV-2 variants in the future.

Chapter 2

EXPERIMENTAL PROCEDURE

2.1 Collection of Single Domain Antibody (sdAb) Sequences

625 sdAb sequences with neutralizing attributes against SARS-CoV-2 variants were obtained from Coronavirus Antibody Database [155] by using specific searching parameters on 17 February 2022. Another sequences were taken from different studies; 117 from [156], 50 from [157], 99 from [158] and 31 sequences listed from Protein Data Bank (PDB) [159]. 72 redundant sdAbs sequences were removed from the collection.

2.2 Structure Modelling of Single Domain Antibodies (sdAbs)

850 sdAb sequences were prepared in FASTA format for structure modelling. sdAb sequences obtained from previous studies or databases were used as a template. 5 model structure files (total 4250) in PDB format for each sdAb sequence were created via AlphaFold 2.1 [160]. After structure modelling, the PDB file was selected out of top 5 highest ranked models which were produced by AlphaFold 2.1 for each sdAb to use in protein-protein docking experiments. AlphaFold 2.1 was installed and run on local High-Performance Computing (HPC) cluster. Parameters in software were set as default. Standard procedure was used. After prediction of first models by deep learning in AlphaFold 2.1 algorithm, relaxation, and energy minimization for PDB structures were performed in AMBER force field to obtain the final top 5 ranked model.

2.3 First Round Protein-Protein Docking Experiments

PDB structures for SARS-CoV-2 wilt type and Omicron variant were obtained from Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (PDB) [159]. PDB ID:6VXX [161] for the wild-type Spike protein and PDB ID:7WP9 [109] for the Omicron variant were selected (Table 2.1). Active residues were determined based on Omicron variant mutation sites (Table 2.2). Amino acid sequences were obtained from the PDB website [159] and sequence alignment was performed on Clustal Omega web server [162,163] to determine the exact position of residues. Paratome web server [149] was used to define active residues (paratope regions) (Figure 2.1 and Table 2.4) of sdAbs whose PDB structures were generated in AlphaFold [160] for the docking experiment. Although the spike protein trimer was loaded onto the server as a receptor, only one chain of active sites was given to the system for each running. Because of three chains of Spike proteins, totally six running (three for 6VXX and three for 7WP9) were performed for PDB structures. HADDOCK server was tested for 3 chains, but software did not permit docking all chain in same time so realistic approach was not applied Spike homotrimer. Therefore, the HDOCK program was used for protein-protein docking [147,148]. All PDB structures were visually inspected via PyMOL before starting docking experiment [154]. Files which have a suitable format for HDOCK server were prepared for active residues. Standard running parameters were used for the docking process.

Table 2.1 Model structures on PDB website and their properties.

Variant	PDB ID	State (Open/Closed)	Method	Resolution
Wild Type	6VXX	Closed	Electron Microscopy	2.80 Å
Wild Type	6VYB	Open	Electron Microscopy	3.20 Å
Wild Type	6VSB	Open	Electron Microscopy	3.46 Å
Wild Type	6XM3	Open	Electron Microscopy	2.90 Å
Omicron	7WK3	Open	Electron Microscopy	3.40 Å
Omicron	7TGW	Open	Electron Microscopy	3.00 Å
Omicron	7WP9	Closed	Electron Microscopy	2.56 Å
Omicron	7QO7	Open	Electron Microscopy	3.02 Å
Omicron	7WPA	Open	Electron Microscopy	2.77 Å

Table 2.2 Active residues which were used in 1st round docking experiments.

Spike Protein PDB ID	HDOCK Protein-Protein Docking Active Residues
6VXX	157-165:A, 385-400:A, 431-441:A, 453-470:A, 491-529:A, 561-571:A, 628-638:A, 669-679:A, 692-710:A, 778-788:A, 810-820:A, 870-880:A, 968-978:A, 983-993:A, 995-1005:A, 157-165:B, 385-400:B, 431-441:B, 453-470:B, 491-529:B, 561-571:B, 628-638:B, 669-679:B, 692-710:B, 778-788:B, 810-820:B, 870-880:B, 968-978:B, 983-993:B, 995-1005:B, 157-165:C, 385-400:C, 431-441:C, 453-470:C, 491-529:C, 561-571:C, 628-638:C, 669-679:C, 692-710:C, 778-788:C, 810-820:C, 870-880:C, 968-978:C, 983-993:C, 995-1005:C
7WP9	135-145:A, 362-377:A, 409-419:A, 431-449:A, 469-507:A, 539-549:A, 606-616:A, 647-657:A, 670-688:A,755-765:A, 788-798:A, 848-858:A, 946-956:A, 961-971:A, 973-983:A, 135-145:B, 362-377:B, 409-419:B, 431-449:B, 469-507:B, 539-549:B, 606-616:B, 647-657:B, 670-688:B,755-765:B, 788-798:B, 848-858:B, 946-956:B, 961-971:B, 973-983:B, 135-145:C, 362-377:C, 409-419:C, 431-449:C, 469-507:C, 539-549:C, 606-616:C, 647-657:C, 670-688:C,755-765:C, 788-798:C, 848-858:C, 946-956:C, 961-971:C, 973-983:C

Table 2.3 Active residues which were used in 2nd round docking experiments.

Spike Protein PDB ID	HDOCK Protein-Protein Docking Active Residues
6VXX	138-146:A, 366-381:A, 412-422:A, 434-451:A, 472-510:A, 542-552:A, 609-619:A, 650-660:A, 673-691:A, 759-769:A, 791-801:A 851-861:A, 949-959:A, 964-974:A, 976-986:A, 138-146:B, 366-381:B, 412-422:B, 434-451:B, 472-510:B, 542-552:B, 609-619:B, 650-660:B, 673-691:B, 759-769:B, 791-801:B, 851-861:B, 949-959:B, 964-974:B, 976-986:B, 138-146:C, 366-381:C, 412-422:C, 434-451:C, 472-510:C, 542-552:C, 609-619:C, 650-660:C, 673-691:C, 759-769:C, 791-801:C, 851-861:C, 949-959:C, 964-974:C, 976-986:C
7WP9	135-145:A, 362-377:A, 409-419:A, 431-449:A, 469-507:A, 539-549:A, 606-616:A, 647-657:A, 670-688:A,755-765:A, 788-798:A, 848-858:A, 946-956:A, 961-971:A, 973-983:A, 135-145:B, 362-377:B, 409-419:B, 431-449:B, 469-507:B, 539-549:B, 606-616:B, 647-657:B, 670-688:B,755-765:B, 788-798:B, 848-858:B, 946-956:B, 961-971:B, 973-983:B, 135-145:C, 362-377:C, 409-419:C, 431-449:C, 469-507:C, 539-549:C, 606-616:C, 647-657:C, 670-688:C,755-765:C, 788-798:C, 848-858:C, 946-956:C, 961-971:C, 973-983:C

Table 2.4 Paratope regions in amino acid sequences of sdAb NM1223, sdAb-2F2, S1-24 and VHH E.

sdAb Name	CDR1	CDR 2	CDR 3
NM1223	AFSSVSMS (27-35)	WVAEIDRDGGNGNYE (47-61)	RLGTRDHIMSG (97-107)
sdAb-2F2	LAQSKWAYG (27-35)	AVAAIDVATGPWYY (47-60)	AHHIPTKHPAFPDFRDY (98-114)
S1-24	STTTNYHMG (27-35)	LVAAINAGGITNYA (47-60)	NIGGGWDYRNSYYIPRVDS (96-114)
VHH E	VTLDYYAIG (27-35)	GVSCIGSSDGRTYY (47-60)	LTVGTYYSGNYHYTCSDDMDY (98-118)

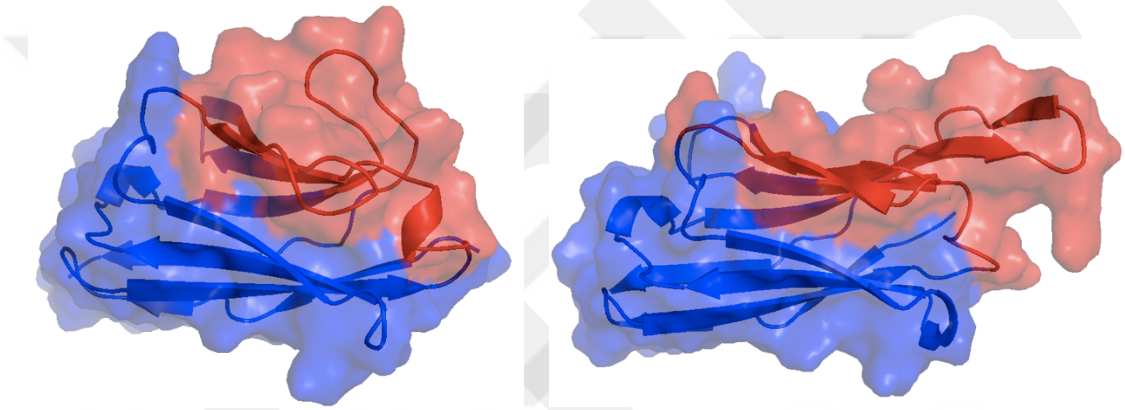

Figure 2.1 Active residues for sdAbs NM1223 and S1-24 used in docking experiments. Paratope regions (red) and rigid body structure (blue) were represented. Structure alignment was performed via PyMOL (surface representation and 50% transparency).

2.4 Second Round Protein-Protein Docking Experiments

PDB structures for SARS-CoV-2 wilt type and Omicron variant were obtained from Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (PDB) [159]. PDB ID:6VXX [161] for the wild-type Spike protein and PDB ID:7WP9 [109] for the Omicron variant were selected (Table 2.1). PDB models had some deficiencies about missing residues, so various tools were tried to a solve problem. PDBFixer, MODELLER, CHARMM-GUI PDB Reader and GalaxyFill were tested. CHARMM-GUI PDB Reader caused some error in the HDOCK server regarding chain ID of Spike proteins which were formed after fixation. All these tools can be used for preparation of proteins in docking experiment or molecular dynamics (MD) simulations such as missing residue problem,

adding hydrogens, setting pH, deficiency in crystal structure. PDBFixer, MODELLER and CHARMM-GUI PDB Reader use amino acid sequence in the PDB website [109,161] as a template to fill missing residues but there were some mistakes in the sequences compared to natural ones. Unlike these tools, GalaxyFill uses the amino acid sequence provided by users. Therefore, PDB structure data files were analyzed via text editor program and engineered mutation sites were determined (residue P983K and P984V in PDB ID:7WP9, residue P986K and P987V in PDB ID:6VXX). These sites were removed manually in text and were given to GalaxyFill as a missing residue. FASTA file was downloaded from PDB website for 7WP9 and necessary substitutions depending on literature readings were occurred in sequences. For the 6VXX, PDB file was downloaded from CHARMM-GUI Covid-19 archive [164] and FASTA format was obtained by using PDB to FASTA converter web server [165]. Membrane-associated parts of Spike proteins (PDB ID:6VXX and 7WP9) after 1145 residue number for PDB ID:7WP9 and 1148 residue number for PDB ID:6VXX in the C terminal of amino acid sequences were removed in FASTA files. NAG structures on PDB models of Spike proteins (PDB ID:6VXX and PDB ID:7WP9) were removed. Active residues were determined based on Omicron variant mutation sites (Table 2.3). Amino acid sequence alignment was performed on Clustal Omega web server [162,163] to determine the exact position of residues. Paratome web server [149] was used to define active residues (paratope regions) (Figure 2.1 and Table 2.4) of sdAbs whose PDB structures were generated in AlphaFold [160] for the docking experiment. Prepared FASTA files were used as a template and PDB structures which have extra missing residues (P983K and P984K in PDB ID:6VXX, P986K and P987V in PDB ID:7WP9) were given to GalaxyFill (Table 2.5 and Table 2.6). After residue filling process, Spike protein structure models obtained from CHARMM-GUI PDB Reader, PDBFixer and GalaxyFill were aligned and analyzed via PyMOL (Figure 2.2 and Figure 2.3). As a result of the analysis, it was decided to use the structural modelling obtained by using GalaxyFill in next protein-protein docking experiments. Molecular Dynamic Simulation (MD) was performed for structure models of Spike proteins obtained from GalaxyFill to energy minimization and relaxation of newly added loops in edge parts of structures (Figure 2.4 and Figure 2.5). GROMACS force field standard protocols were used for 1 nanosecond (ns) on local cluster. All PDB structures were visually inspected via PyMOL before starting docking experiment. [154].

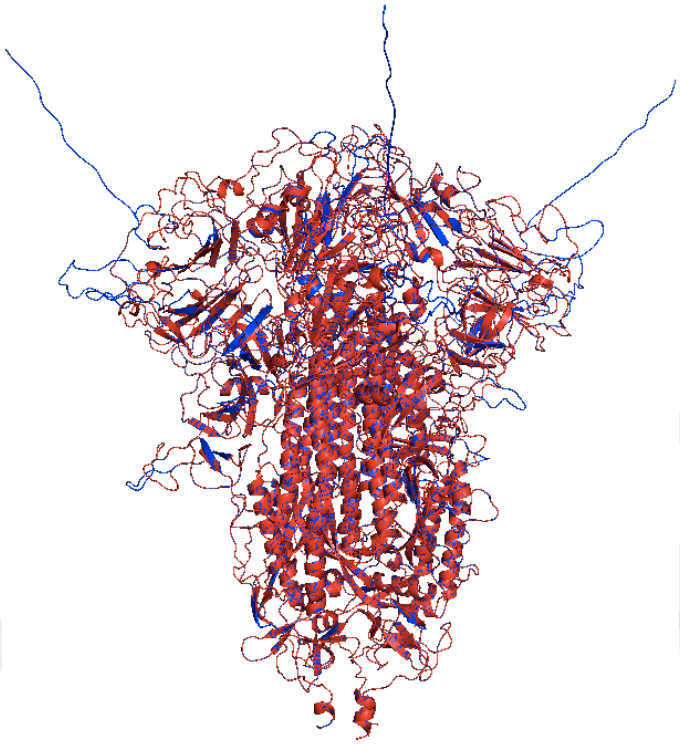

Figure 2.2 Spike protein model structure alignment (PDB ID:7WP9) after addition of missing residues and PDB fixation with GalaxyFill and PDBFixer. GalaxyFill (red) and PDBFixer (blue) were used for fixation. Structure alignment was performed via PyMOL (cartoon representation). RMSD = 0.008 Å (24465 to 24465 atoms).

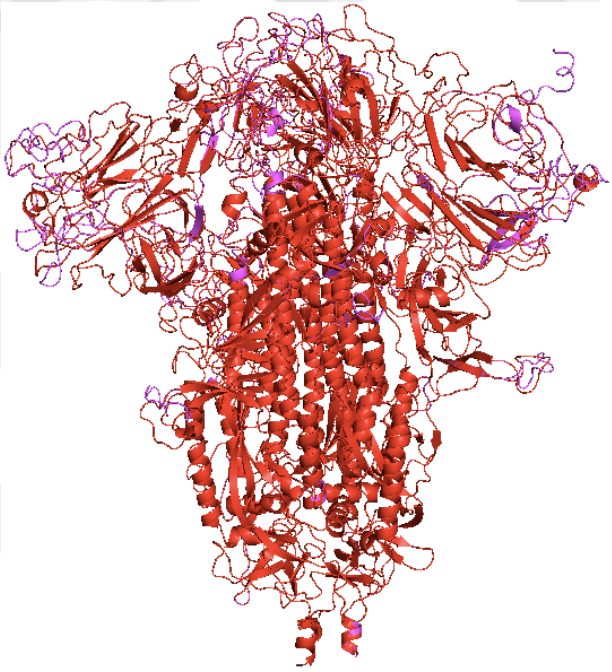

Figure 2.3 Spike protein model structure alignment (PDB ID:7WP9) after the addition of missing residues and PDB fixation with GalaxyFill and CHARMM-GUI. GalaxyFill (red) and CHARMM-GUI PDB Reader (magenta) were used for fixation. Structure alignment was performed via PyMOL (cartoon representation). RMSD = 0.000 Å (23871 to 23871 atoms).

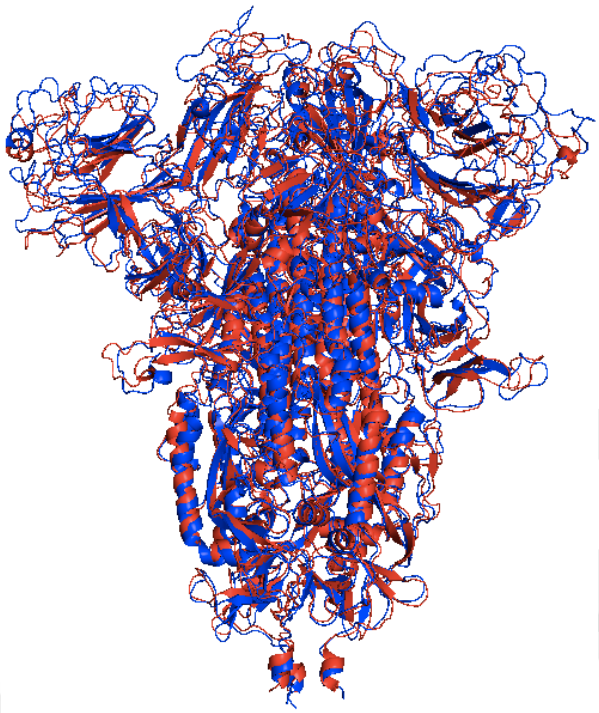

Figure 2.4 Spike protein model structure alignment (PDB ID:7WP9) after molecular dynamics (MD) simulation for 1 nanosecond. Before MD (red) and after MD (blue) models were represented. Structure alignment was performed via PyMOL (cartoon representation). RMSD = 1.877 Å (26691 to 26691 atoms).

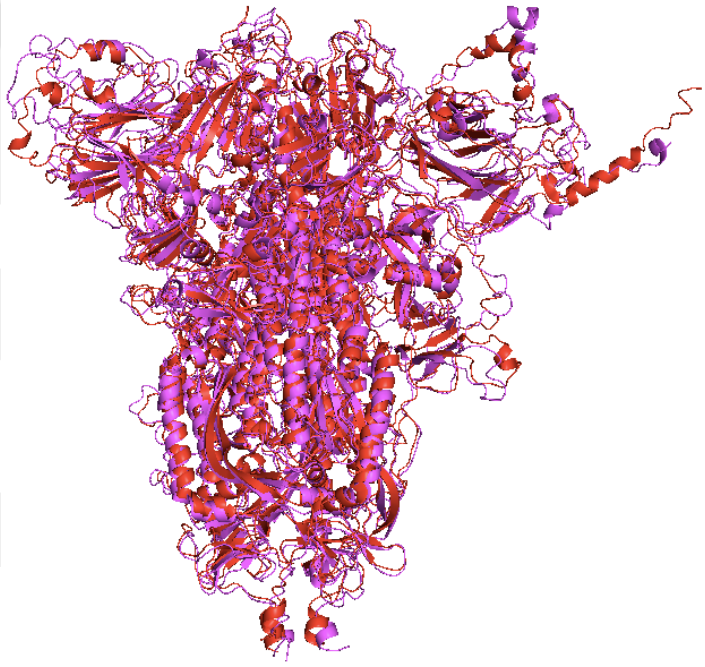

Figure 2.5 Spike protein model structure alignment for PDB ID:6VXX) after molecular dynamics (MD) simulation for 1 nanosecond. Before MD (red) and after MD (magenta) models were represented. Structure alignment was performed via PyMOL (cartoon representation). RMSD = 2.107 Å (26890 to 26890 atoms).

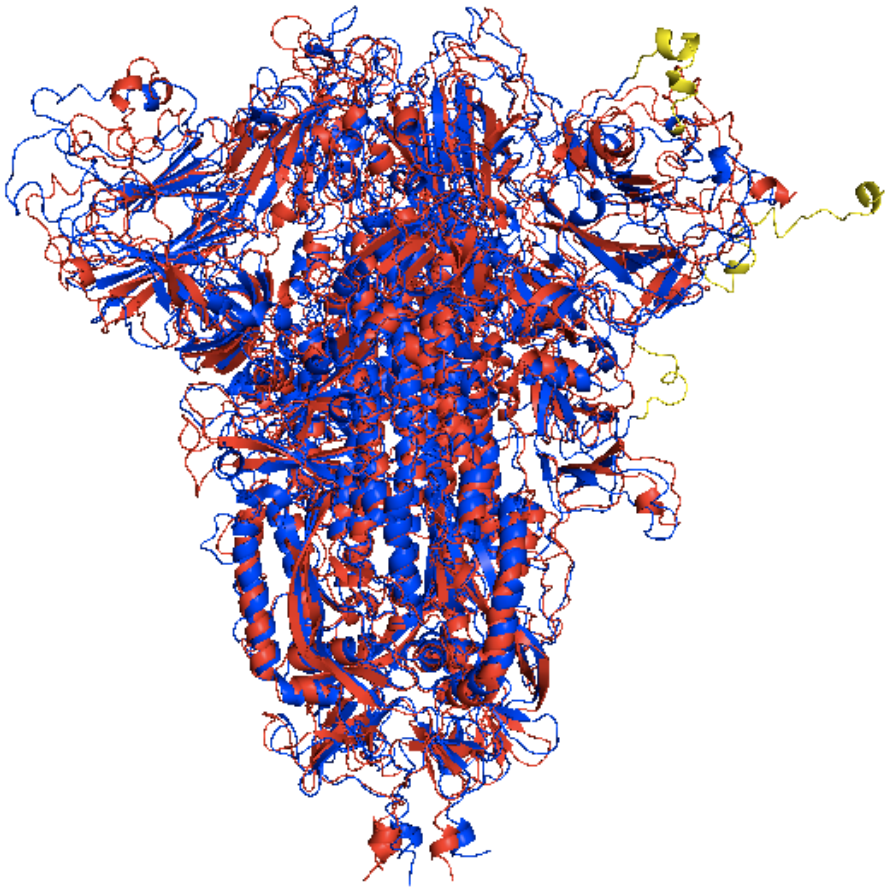

Figure 2.6 Spike protein model structure alignment for receptor proteins which were used in 2nd round protein-protein docking experiment. (Blue:PDB ID:6VXX, Red:PDB ID:7WP9 and Yellow: Unexpected loops comprised of chain C of 6VXX after missing residue addition in GalaxyFill and molecular dynamics (MD) simulation for 1 nanosecond application. RMSD = 2.640 Å (42027 to 42027 atoms).

Table 2.5 Missing residues for PDB ID:7WP9.*

Missing Residues	Chain ID
1-16	A, B, C
24	B
67-77	A, B
141-150	A, B, C
174-180	A, B, C
240-260	A, B, C
369-373	A
402	B
408-413	B
412	A
416	B
420	B
465-467	B
496-502	B
497-500	C
497-502	A,
675-684	A, B, C
834-843	A, B, C
1145-1205	A, B, C

***(1145-1205 residues were not added in GalaxyFill.)**

Table 2.6 Missing Residues for PDB ID:6VXX.*

Missing Residues	Chain ID
1-26	A, B, C
70-79	A, B, C
144-164	A, B, C
173-185	A, B, C
246-262	A, B, C
445-446	A, B, C
455-461	A, B, C
469-488	A, B, C
502	A, B, C
621-640	A, B, C
677-688	A, B, C
828-853	A, B, C
1148-1262	A, B, C

***(1148-1262 residues were not added in GalaxyFill.)**

Paratome web server [149] was used to define active residues (paratope regions) (Figure 2.1 and Table 2.4) of sdAbs whose PDB structures were generated in AlphaFold [160] for the docking experiment. Although the spike protein trimer was loaded onto the server as a receptor, only one chain of active sites was given to system for each running. Because of three chains of Spike proteins, totally six running (three for 6VXX and three for 7WP9) were performed for PDB structures. HADDOCK server was tested for 3 chains, but software did not permit docking all chain in same time so realistic approach was not applied Spike homotrimer. Therefore, the HDOCK program was used for protein-protein docking [147,148]. All PDB structures were visually inspected via PyMOL before starting docking experiment [154]. Files which have suitable format for HDOCK server were prepared for active residues. Standard running parameters were used for the docking process.

2.5 Visualization and Analysis of First Round HDOCK Docking Experiment Results

After the docking experiment, 4392 HDOCK binding modes were obtained for each complex. PDB model structures which have top 10 HDOCK scores were created for 850 sdAbs and 3 chains, so the total number of complexes was 25500. Bluues and PRODIGY

scoring tools need file in PQR format, so after docking experiments, PDB2PQR was used to covert PDP files into PQR. Consensus score was used for the selection of docking experiment results. Three docking scoring functions, HDOCK, Bluues and PRODIGY were scripted and implemented as a consensus and score of results were sorted. Common ones between the top 1000 scores and 3 chains for each Spike protein were listed. Proportional Venn diagrams were created for listed docking results (Figure 3.3 and Figure 3.4). LigPlot+ was used to examine the interactions between sdAbs and Omicron Spike protein (Figure 3.6 and Figure 3.7).

2.6 Visualization and Analysis of Second Round HDOCK Docking Experiment Results

After the docking experiment, 4392 HDOCK binding modes were obtained for each complex. PDB model structures which have top 220 HDOCK scores were created for 850 sdAbs and 3 chains, so the total number of complexes was 561000. Before the scoring with HDOCK, Bluues and PRODIGY, newly developed strategy by us was used to decrease number of complexes. This strategy is based on evolutionary information about interaction sites between target protein and antibody derivatives. Necessary codes were scripted, pre-scanning and visually inspection analyses were performed on large result datasets. Minimum 10 close contacts (3.5 Å distance between sdAb and Spike proteins) in determined interaction sites were selected as a parameter. 80% coverage of these results also were considered. Total number was reduced to 13871 for 6VXX and 7WP9 samples. Minimum scores were determined based on the thousandth or hundredth ranked complexes for each scoring function, chain, and variants. Bluues and PRODIGY scoring tools need file in PQR format, so after docking experiments PDB2PQR was used to covert PDP files into PQR. Consensus score was used for filtration of docking experiment results. Three docking scoring functions, HDOCK, Bluues and PRODIGY were scripted and implemented as a consensus and score of results were sorted. Common ones between the top 1000 and top 100 scores with 3 chains for each Spike protein were listed. Proportional Venn diagrams were created for listed docking results (Figure 3.8, Figure 3.9, Figure 3.10 and Figure 3.11) LigPlot+ was used to examine the interactions between sdAbs and Omicron Spike protein (Figure 3.13, Figure 3.14, Figure 3.16, Figure 3.18 and Figure 3.19).

Chapter 3

RESULTS

3.1 Structure Modelling Results of Single Domain Antibodies (sdAbs)

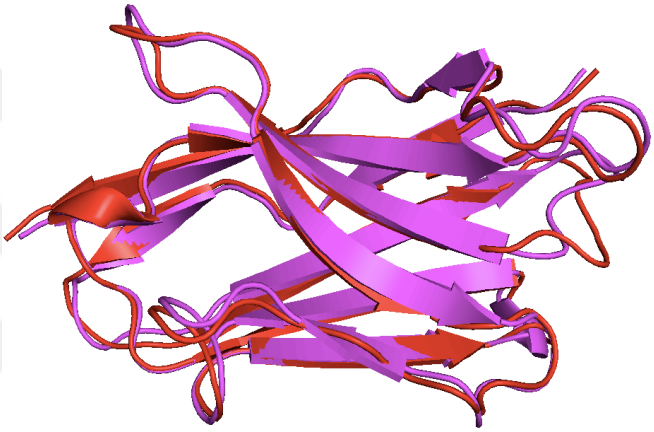

Figure 3.1 Structure modelling experiment results of sdAb Nb20. Red: PDB ID:7JVB, Nb20 and magenta: Nb20 structure was created AlphaFold2. Structure alignment was performed via PyMOL (cartoon representation). RMSD = 0.697 Å (713 to 713 atoms).

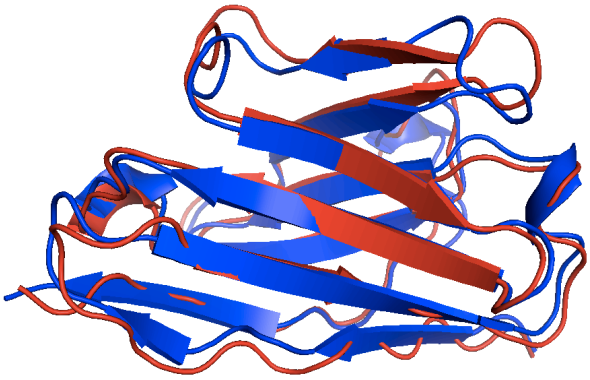

Figure 3.2 Structure modelling experiment results of sdAb Ty1. Red: PDB ID:6ZXN, Ty1 and blue: Ty1 structure was created AlphaFold2. Structure alignment was performed via PyMOL (cartoon representation). RMSD = 1.842 Å (1514 to 1514 atoms).

According to the obtained results (Figure 3.1 and Figure 3.2), AlphaFold 2 is quite useful for the structural modelling of proteins which have small amino acid sequences like sdAb. Structure estimation of functional loops is a more complex issue for many structures modelling tools. When the alignment was visually examined in detail, It can be easily seen that AlphaFold 2 gave similar results to 3D model structures obtained in laboratory conditions for loop-like structures which have especially important functions for interaction with the target proteins. If the root-mean-square deviation (RMSD) value is close to 0, which is obtained as a result after structure alignment of the models, is considered more appropriate. 0.697 Å for sdAb Nb20 (713 atoms) and 1.842 Å for sdAb Ty1 (1514 atoms) values were observed in PyMOL structure alignment. These results can be considered acceptable.

3.2 Protein-Protein First Round Docking Experiment Results

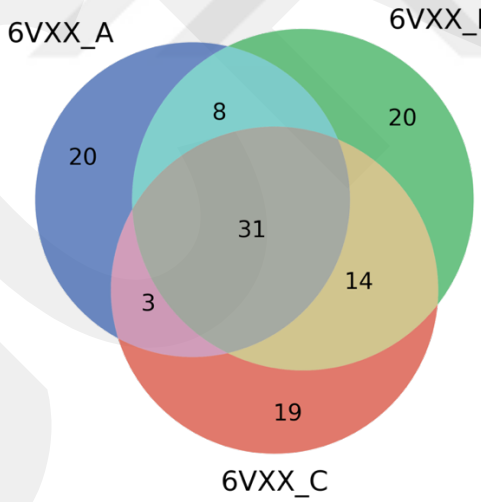

Figure 3.3 Common sdAb numbers in top 1000 scores for wild-type SARS-CoV-2 after 1st round docking experiment by chains.

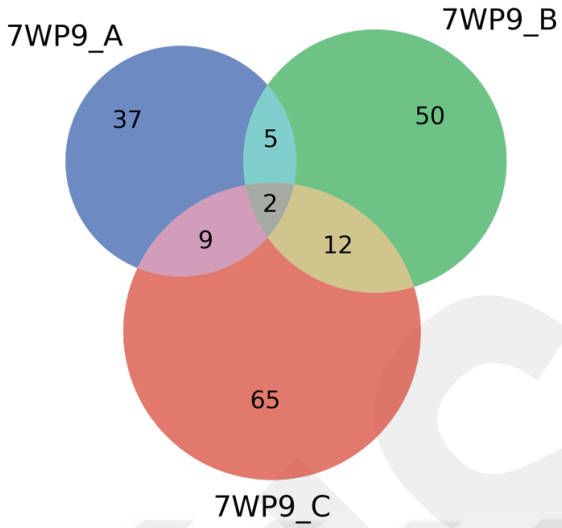

Figure 3.4 Common sdAb numbers in top 1000 scores for Omicron variant SARS-CoV-2 after 1st round docking experiment by chains.

Top 1000 scores of HDOCK, Bluues and PRODIGY were used for the selection of successful sdAbs. Then, comparison was performed to find common sdAb numbers in three chains of spike proteins (Figure 3.3 and Figure 3.4). In total, out of 25500 complex models present for Omicron Spike protein, 2 sdAbs which can be a successful candidate against Omicron variant when tested in the laboratory, were detected. According to the first Venn diagram (Figure 3.3), the number of sdAbs, which have a probability of success for 3 chains and successful results were obtained for at least 2 chains, is 56 for wild-type Spike protein. In addition, when the Venn diagram formed for Omicron variant Spike protein was investigated, it was seen that this number is 28. At the end of the scanning occurred with a rational approach, the number of sdAb that gave successful results for the wild type was 31, while this number decreased to 2 for the Omicron variant. Theses result are explaining why Omicron variant can easily evade from existing antibodies or sdAbs, and became dominant variants.

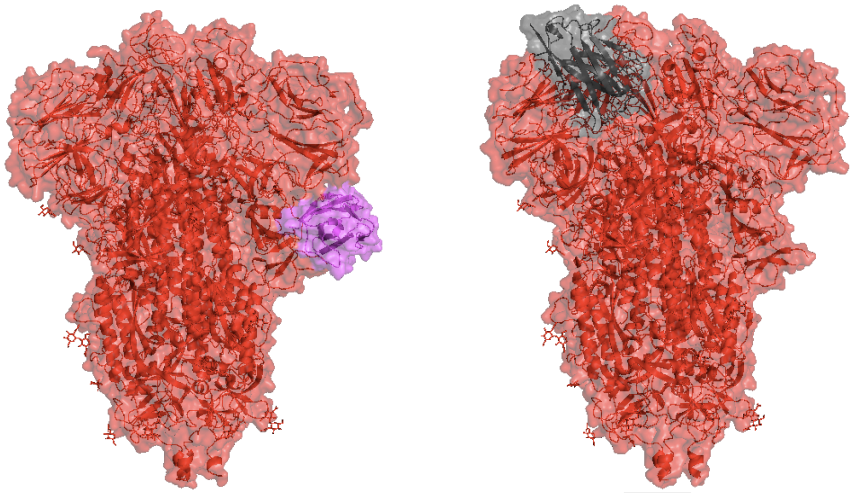

Figure 3.5 Docking sites for sdAb NM1223 and sdAb 2F2 with SARS-CoV-2 Omicron variant in 1st docking experiment. Red: PDB ID:7WP9 Spike protein, magenta: sdAb NM1223, gray: sdAb 2F2. Visualization was performed via PyMOL, (surface representation and 50% transparency).

In Figure 3.5, it was appeared that sdAb NM1223 may be bound to a site close to the possible furin protease cleavage site. On the other hand, the strong binding affinity of sdAb 2F2 to the RBD region of the Spike protein was revealed by visual analysis of the complex molecule in PyMOL. In the 1st round protein docking experiment, 3D structures of the Spike protein available on the PDB website were used. Besides, it should be considered that those PDB models can have some structural deficiencies regarding missing residues.

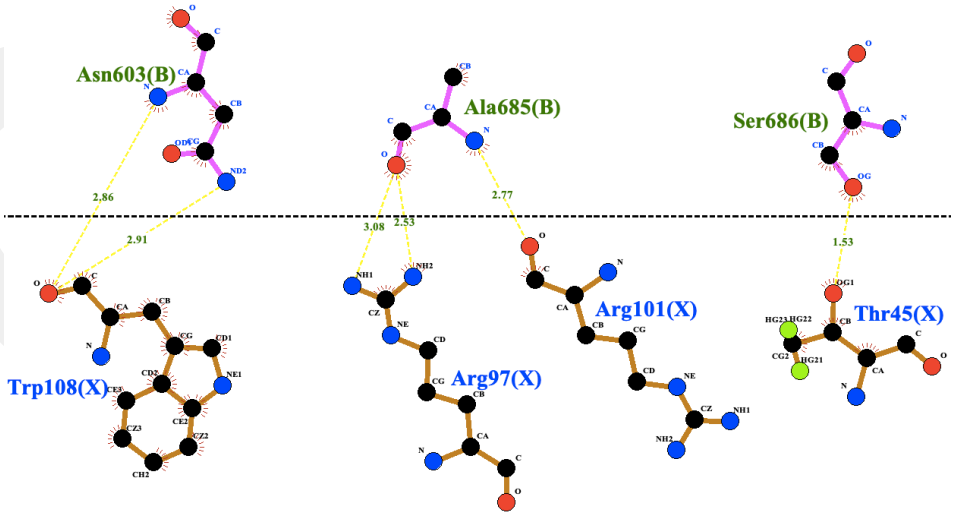

Figure 3.6 LigPlot+ 2D interaction analysis between sdAb NM1223 and Omicron Spike protein after 1st round docking experiment. (Chain X: NM1223 and Chain A: Omicron spike protein; yellow lines: hydrogen bonds, Distances; Asn603(B)-1-Trp108(X):2.86Å, Asn603(B)-2-Trp108(X):2.91Å, Ala685(B)-1-Arg97(X):3.08Å, Ala685(B)-2-Arg97(X):2.53Å, Ala685(B)-Arg101(X):2.77Å, Ser686(B)-Thr45(X):1.53Å.

When the LigPlot+ analyses are examined in detail, it was seen that the NM1223 sdAb interacts with the Asn603(2 interactions, 2.86Å, 2.91Å), Ala685(3 interactions, 2.53Å, 2.77Å, 3.08Å) and Ser686 (1.53Å) residues of the Omicron spike protein by using hydrogen bonds (Figure 3.6). The Asn603 position is very close to the D614G mutation (611. Residue in 7WP9), which has contributed to the acceleration of infection of the Omicron variant, as indicated in many studies [166–168]. Furthermore, residues Ala685 and Ser686 of Omicron spike protein, which are among the interaction sites of sdAb NM1223, are located near the furin protease cleavage site, which is crucially important for cell entry of SARS-CoV-2 virus [35,87,96,169,170].

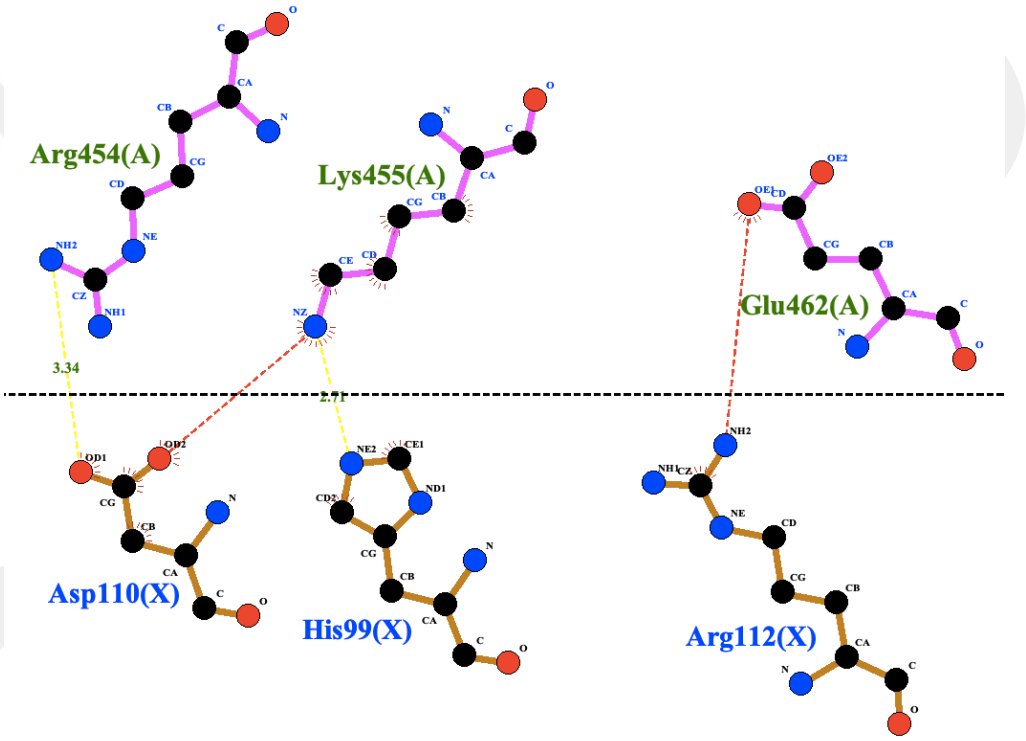

Figure 3.7 LigPlot+ 2D interaction analysis between sdAb 2F2 and Omicron Spike protein after 1st round docking experiment. (Chain X: sdAb 2F2 and Chain A: Omicron spike protein; yellow lines: hydrogen bonds, red lines: salt bridges, Distances; Arg454(A)-Asp110(X):3.34Å, Lys455(A)-His99(X): 2.71Å, Lys455(A)-Asp110(X):2.97Å, Glu462(A)-Arg(X):3.18Å.

On the other hand, sdAb 2F2 was established strong interactions with Arg454, Lys455 and Glu462 residues of Omicron spike protein (Figure 3.7). While it interacted with Arg454 with 3.34Å distance by using hydrogen bonding, it formed a salt bridge interaction with Glu462 at a distance of 3.18Å. Moreover, sdAb 2F2 has both hydrogen bond (2.97Å) and salt bridge (2.71Å) interaction with Lys455 residue of Omicron spike

protein. All three interaction sites of Omicron variants, Arg454, Lys455 and Glu462, are found in RBD, which is responsible for attachment to ACE2 receptor of host cells.

Table 3.1 HDOCK protein-protein 1st docking experiment scores for Omicron Spike protein (PDB ID:7WP9) and common sdAbs.

Spike Protein PDB ID	Chain Name	sdAb Name	HDOCK Score	Bluues Score	PRODIGY Score
7WP9	A	NM1223	-397.68	-131.38	-15.6
7WP9	B	NM1223	-393.34	-111.36	-15.6
7WP9	C	NM1223	-383.53	-104.54	-16.8
7WP9	A	sdAb-2F2	-381.9	-115.15	-15.8
7WP9	B	sdAb-2F2	-382.28	-116.26	-15.9
7WP9	C	sdAb-2F2	-363.28	-103.69	-13.5

Table 3.2 Amino acid sequences of two selected successful sdAbs for SARS-CoV-2 Omicron variant in 1st round docking experiment.

sdAb Name	Amino Acid Sequences
NM1223 [171]	EVQLVESGGGLVQPGGSLRLSCSASGFAFSSVSMSWVRLLPGKGTEWVAE IDRDGGNGNYEDSVKGRFTISRDNAKNTLFLQMNSLVPEDTALYYCRLGTR DHIMSGWGPGAPVTVSS
sdAb-2F2 [172]	QVQLVESGGGLVQPGGSLRLSCAASGLAQSKWAYGWFRQAPGKGLEAVA AIDVATGPWYYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCAAH HIPTKHPAFPDFRDYWGQGTQVTVSS

3.3 Protein-Protein Second Round Docking Experiment Results

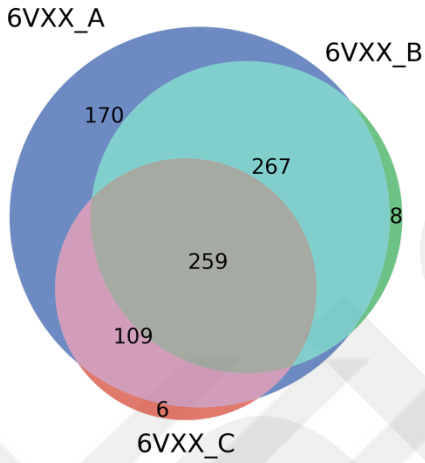

Figure 3.8 Common sdAb numbers in top 1000 scores for wild-type SARS-CoV-2 after 2nd round docking experiment by chains.

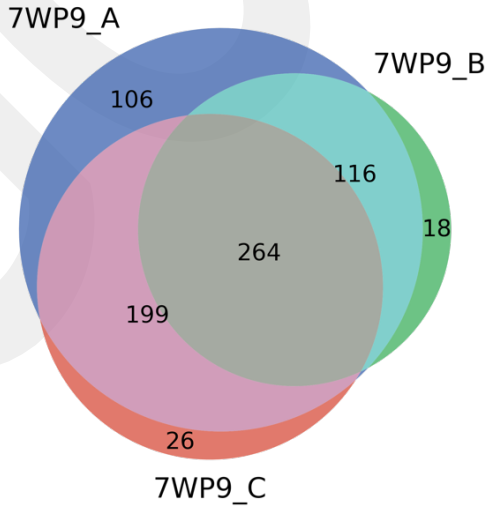

Figure 3.9 Common sdAb numbers in top 1000 scores for SARS-CoV-2 Omicron variant after 2nd round docking experiment by chains.

First, newly designed approach was used for pre-filtration in 2nd round docking experiment results. Total number of docking result complexes was reduced to 13871 (7429 for PDB ID:6VXX and 6442 for PDB ID:7WP9) from 561000. Minumum values in the top 1000 scores of HDOCK, Bluues and PRODIGY were used as a threshold for the selection of successful sdAbs. Then, comparison was performed to find common sdAb numbers in three chains of spike proteins (Figure 3.8 and Figure 3.9). In total, out of 6442 complex models present for Omicron Spike protein, 2 sdAbs which could be a successful candidate against Omicron variant when tested in the laboratory were determined. According to Venn diagram in Figure 3.8, the number of sdAbs, which have a possibility of success for 3 chains and gave successful results for at least 2 chains, is 635 for wild-type Spike protein. Besides, when Venn diagram created for Omicron variant Spike protein was investigated, it was seen that this number is 579 (Figure 3.9). Finally, after scanning carried out with a rational approach, the number of sdAb that gave successful results for the wild type was 259, while this number decreased to 264 for the Omicron variant. These results were shown that the number of candidates sdAbs that could be successful against Omicron and wild type Spike protein was similar.

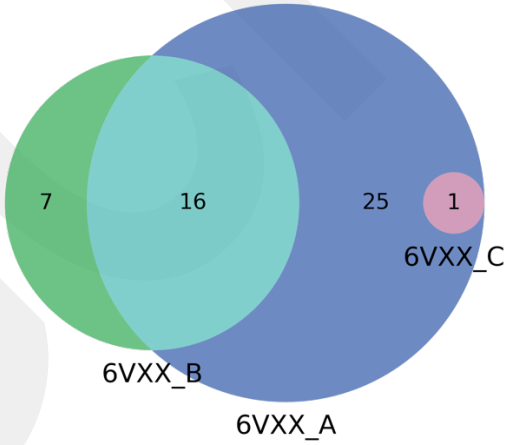

Figure 3.10 Common sdAb numbers in top 100 scores for wild-type SARS-CoV-2 after 2nd round docking experiment by chains.

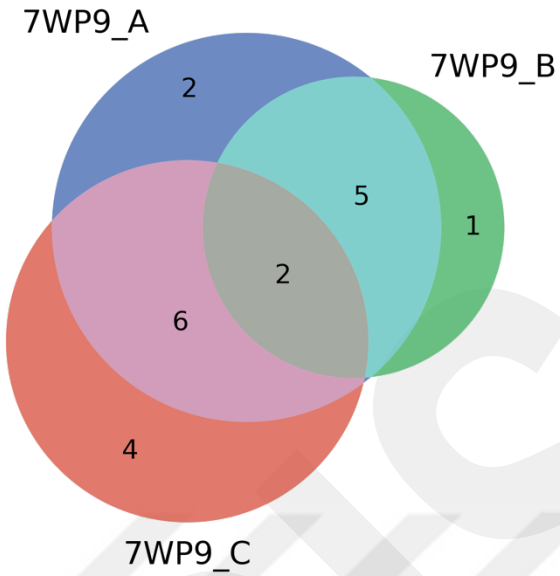

Figure 3.11 Common sdAb numbers in top 100 scores for SARS-CoV-2 Omicron variant 2nd round docking experiment by chains.

After pre-filtering, which was carried out with newly added approach for 2nd round docking experiment results, total number of docking result complexes was decreased to 7429 for wild type Spike protein and 6442 for Omicron Spike protein (totally 13871) from 561000. Minimum values of 100th complexes for HDOCK, Bluues and PRODIGY scoring functions were determined as a threshold in the selection of successful sdAbs. Comparison was performed to find common sdAb numbers in three chains of spike proteins in next step. 2 sdAbs which could be a successful candidate against Omicron variant when tested in the laboratory were detected as a result of scanning in 6442 complex models present for Omicron Spike protein (Figure 3.11). Number of sdAbs, which have a probability of success for 3 chains and gave successful results for at least 2 chains, is 13 for Omicron Spike protein in Figure 11. But there was only 1 sdAb was common between chain A and chain B of wild type Spike Protein, compared to Omicron variant. Common sdAb with other chains could not be detected for the chain C when diagram was examined. Unexpected loops comprised after missing residue addition can be a reason for insufficient number of sdAbs in chain C of wild type (PDB ID:6VXX) docking experiment (Figure 2.6).

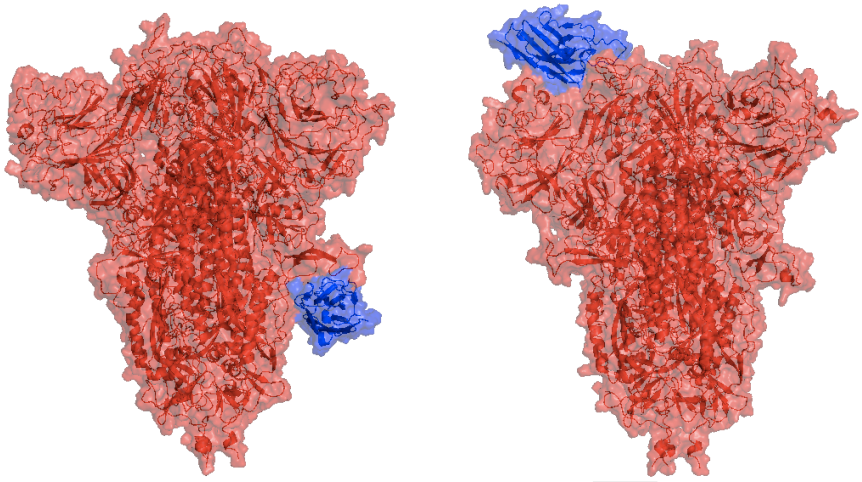

Figure 3.12 Docking sites between sdAb VHH E and SARS-CoV-2 Omicron variant in 2nd docking experiment. Red: PDB ID:7WP9 Spike protein, blue: sdAb VHH E. Left interaction sites on Spike Protein: Arg682 (Chain A). Right interaction sites on Spike Protein: Ser440 (Chain B), Ser491, Gly444 (Chain A). Visualization was performed via PyMOL, (surface representation and 50% transparency).

sdAb VHH E shown affinity for different sites on Omicron Spike protein. In Figure 3.12, VHH E has 3 strong interactions with chain A (2) and chain B (1) for RBD region. It also has 1 interaction for potentially for furin protease site which is important cell entry of SARS-CoV-2. Its ability to bind to different sites indicates that it has considerable flexibility regarding affinity. Loop structures of sdAbs' have crucially important role for flexibility to reach difficult areas on target proteins. sdAb VHH has large affinity for at least two important sites.

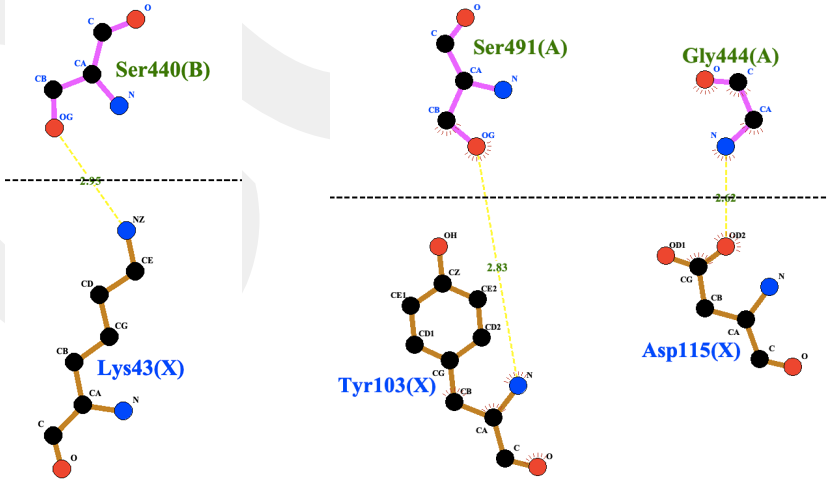

Figure 3.13 LigPlot+ 2D interaction analysis between sdAb VHH E and RBD of Omicron Spike protein after 2nd round docking experiment. (Chain X: sdAb VHH E and Chain A, B: Omicron spike protein; yellow lines: hydrogen bonds. Distances; Ser440(B)-Lys43(X):2.95Å, Ser491(A)-Tyr103(X): 2.83Å, Gly444(A)-Asp115(X):2.62Å.

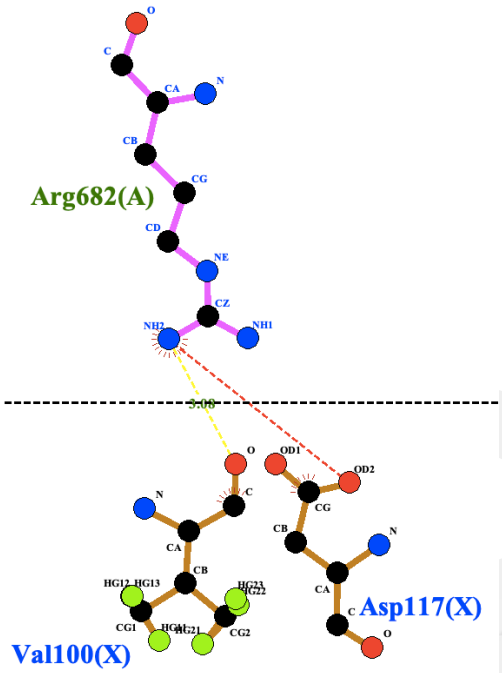

Figure 3.14 LigPlot+ 2D interaction analysis between sdAb VHH E and protease cleavage sites of Omicron Spike protein after 2nd round docking experiment. (Chain X: sdAb VHH E and Chain A: Omicron spike protein; yellow lines: hydrogen bonds, red lines: salt bridges. Distances; Arg482(A)-Val100(X):3.08Å, Arg482(A)-Asp117(X): 3.22Å.

sdAb VHH E has various interactions on RBD region of Omicron Spike protein. sdAb VHH E interacts with Ser440 (Chain B) at 2.95Å distance, Gly444 (Chain A)at 2.62Ådistance and Ser491 (Chain A) at 2.83Å distance by forming hydrogen bonds. sdAb VHH E used Lys43(for Ser440), Asp115 (for Gly444) and Tyr103 (for Ser491) paratope regions for interactions (Figure 3.13). When the LigPlot+ analysis is examined in detail, sdAb VHH E interacts with the Arg482 by using 2 paratope regions. It was observed that sdAb VHH E formed hydrogen bond using Val100 residue at 3.08 Å distance and salt bridge was formed with Asp117 residue (3.22Å distance) (Figure 3.14). 685-686 positions are furin protease cleavage sites for Omicron variant of SARS-CoV-2 (PDB ID). Disruption of interaction between furin protease and this site is significant to prevent the spreading of SARS-CoV-2. Moreover, sdAb VHH E is docking to furin protease site with two strong bonds [35,83,96,170,173].

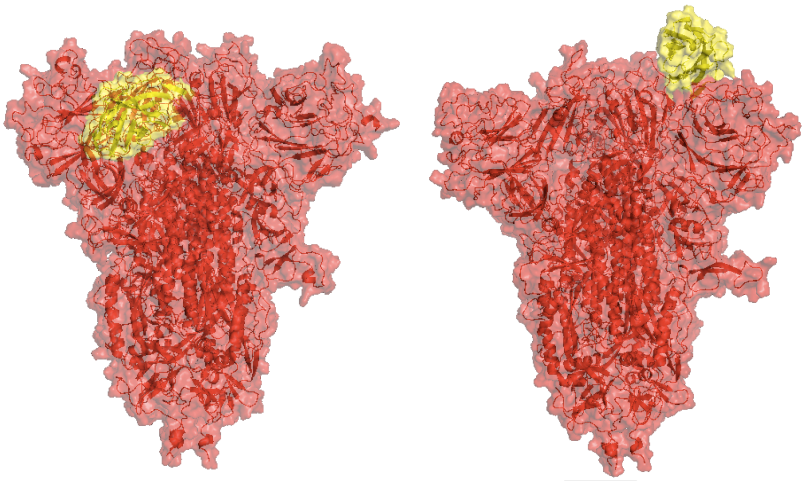

Figure 3.15 Docking sites between sdAb S1-24 and RBD of SARS-CoV-2 Omicron variant in 2nd docking experiment. Red: PDB ID:7WP9 Spike protein, yellow: sdAb S1-24. Left interaction sites on Spike Protein: Ser456 (Chain A), Asn457 (Chain A), Lys459 (Chain A). Right interaction sites on Spike Protein: Tyr348 (Chain C). Visualization was performed via PyMOL, (surface representation and 50% transparency).

sdAb S1-24 has interactions on different sites of the RBD region. Interactions occurred in various regions with a closed state of Omicron Spike Protein. Normally, sdAbs or antibody derivatives can reach and bind easily to the RBD region in the open state of Spike protein, which at least one RBD of protein up, but sdAb movement can be affected differently during docking experiment.

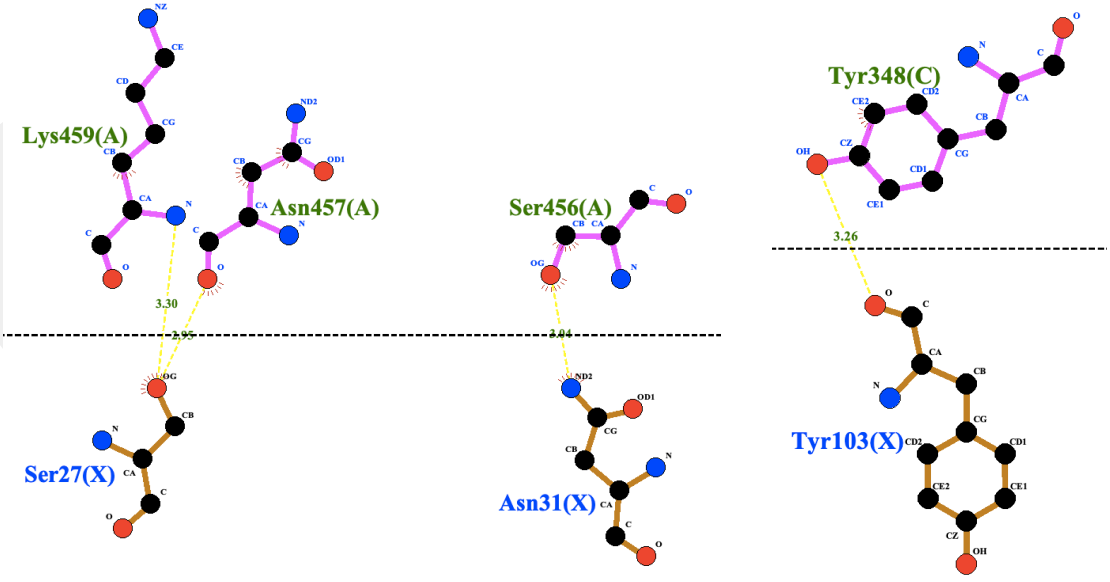

Figure 3.16 LigPlot+ 2D interaction analysis between sdAb S1-24 and RBD of Omicron Spike protein after 2nd round docking experiment. (Chain X: sdAb S1-24 and Chain A,C: Omicron spike protein; yellow lines: hydrogen bonds. Distances; Ser456(A)-Asn31(X):3.04Å, Asn457(A)-Ser27(X): 2.95Å, Lys459(A)-Ser27(X):3.30Å, Tyr348(C)-Tyr103(X):3.26Å.

RBD region of SARS-CoV-2 is the first target region for neutralization. sdAb S1-24's 4 determined interactions were shown in Figure 3.16. sdAb S1-24 interacts with Ser456(Chain A) at 3.04Å distance, Asn457 (Chain A) at 2.95Å distance, Asn459 (Chain A) at 3.30Å distance and Tyr348 (Chain C) at 3.26Å distance via hydrogen bonds. S1-24 used Ser27 residue for Lys459 and Asn457, Asn31 residue for Ser456 and Tyr103 residue for Tyr348. In addition, LigPlot+ analysis is showing that S1-24 has successful results for different chains.

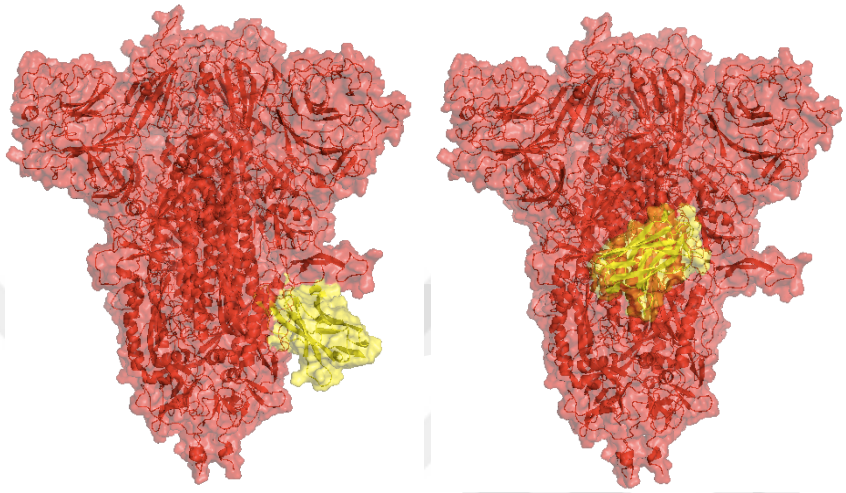

Figure 3.17 Docking sites between sdAb S1-24 and protease cleavage sites of SARS-CoV-2 Omicron variant in 2nd docking experiment. Red: PDB ID:7WP9 Spike protein, yellow: sdAb S1-24. Left interaction sites on Spike Protein: Glu658 (Chain A), Cys659 (Chain A), Ser695 (Chain A). Right interaction sites on Spike Protein: Asp817 (Chain A), Asp834 (Chain A). Visualization was performed via PyMOL. (Surface representation and 50% transparency).

Although, SARS-CoV-2 can entry to host cells via endosomal pathway, generally viruses need to proteases depending on abundance of proteases in environment. Therefore, protease cleavage sites can be targeted by antibody derivatives. sdAb S1-24. Docking visualization results shown that sdAb S1-24 may be also effective for protease cleavage regions in addition to its high affinity in the RBD region. There are some interactions near furin cleavage sites at Glu658 (Chain A), Cys659(Chain A) and Ser695 (Chain A) positions. There are some interactions near furin cleavage sites at Glu658 (Chain A), Cys659 (Chain A) and Ser695 (Chain A) positions for sdAb S1-24 (Figure 3.18) but these interaction sites are quite far from furin cleavage site. Therefore, efficacy of S1-24 can be lower than other sdAbs which have specific interactions for this region, such as VHH E (Figure 3.14).

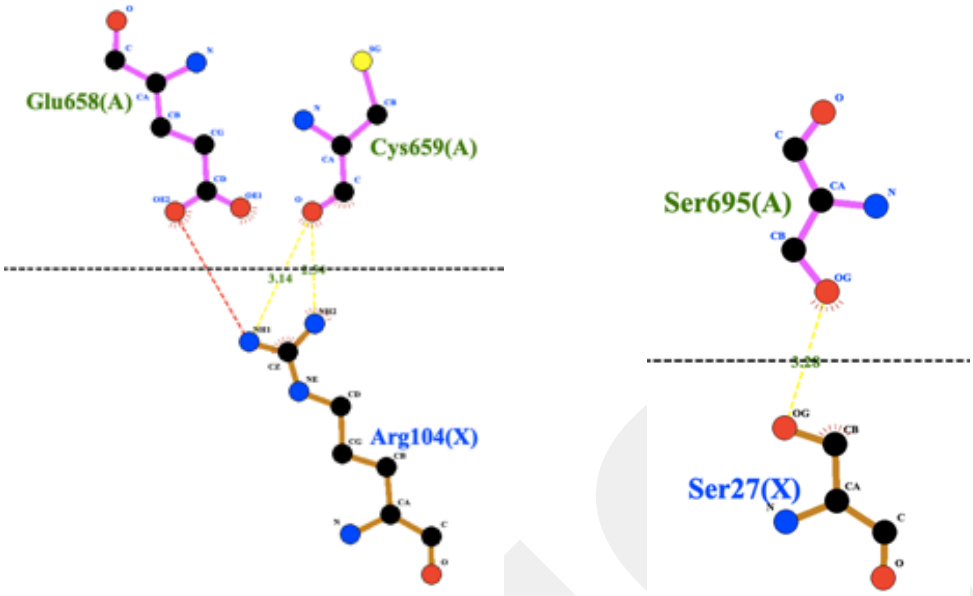

Figure 3.18 LigPlot+ 2D interaction analysis between sdAb S1-24 and protease cleavage sites of Omicron Spike protein after 2nd round docking experiment. (Chain X: sdAb S1-24 and Chain A: Omicron spike protein; yellow lines: hydrogen bonds, red lines: salt bridges. Distances; Glu658(A)-Arg104(X):2.98Å, Cys659(A)-Arg104(X): 2.54Å and 3.14 Å, Ser695(A)-Ser27(X):3.28Å.

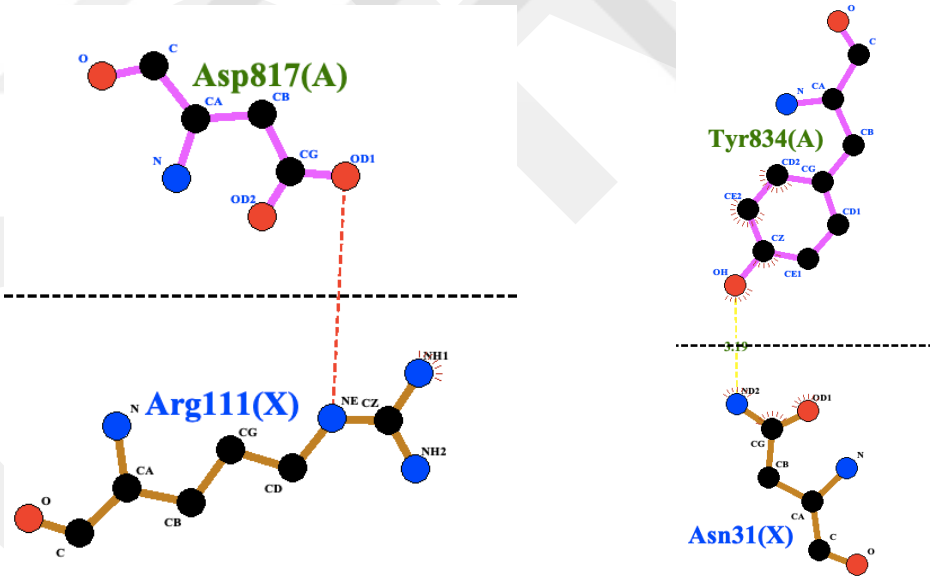

Figure 3.19 LigPlot+ 2D interaction analysis between sdAb S1-24 and protease cleavage sites of SARS-CoV-2 Omicron variant Spike protein after 2nd round docking experiment. (Chain X: sdAb S1-24 and Chain A: Omicron spike protein; yellow lines: hydrogen bonds, red lines: salt bridges. Distances; Asp817(A)-Asp111(X):3.46Å, Tyr834(A)-Asn31(X): 3.19Å.

The 818th and 819th residues (815-816 in wild type) are the serine protease TMPRSS2 recognition sites used by SARS-CoV-2 during cell infection. When looking at the interactions in Figure 3.19, it is seen that sdAb S1-24 docked by forming at 3.46Å distance salt bridge with Asp817(Chain A). It uses the Asp111 residue in the paratope region for this interaction. Also, in support of this, there is a hydrogen bond interaction at 3.19Å for the Tyr834 (Chain A) residue. Although this region is not close to the cleavage region of the TMPRSS2 protease, it is anticipated that it may help sdAb S1-24 to bind more strongly into this docking position. The sdAb residue responsible for this interaction is Asn31.

Table 3.3 HDOCK protein-protein 2nd experiment docking scores for Omicron Spike protein (PDB ID:7WP9) and common sdAbs.

Spike Protein PDB ID	Chain Name	sdAb Name	HDOCK Score	Bluues Score	PRODIGY Score
7WP9	A	VHH E	-753.34	-121.92	-13.8
7WP9	B	VHH E	-721.28	-134.21	-14.1
7WP9	C	VHH E	-755.12	-123.06	-13.9
7WP9	A	S1-24	-697.59	-127.01	-13.0
7WP9	B	S1-24	-683.77	-121.29	-12.4
7WP9	C	S1-24	-697.71	-123.13	-12.8

There is proper correlation between all scores applied for sdAbs and each chain of Omicron Spike protein. 2nd round experiment docking scores for HDOCK and Bluues are more than 1st round but PRODIGY score of 1st round is higher than 2nd ones (Table 3.1 and Table 3.3). The reason for high Bluues and HDOCK scores may be the addition of missing residues in the 2nd round and the removing of NAG structures from the models taken from the PDB website.

Table 3.4 Amino acid sequences of two selected successful sdAbs for SARS-CoV-2 Omicron variant in 2nd round docking experiment.

sdAb Name	Amino Acid Sequences
VHH E [174]	QVQLVETGGGFVQPGGSLRLSCAASGVTLDYYAIGWFRQAPGKEREGVSCI GSSDGRTYYSDSVKGRFTISRDNAKNTVYLQMNSLKPEDTAVYYCALTVGT YYSGNYHYTCSDDMDYWGKGTQVTVSS
S1-24 [175]	QVQPVESGGGLVQAGGSLRLSCVASGSTTTNYHMGWYRQTPGEQRELVAAI NAGGITNYADSVKGRFTISRDNAKNTMYLQMNNLRFEDTAVYYCNIGGGWD YRNSYYIPRVDSWGQGTQVTVS

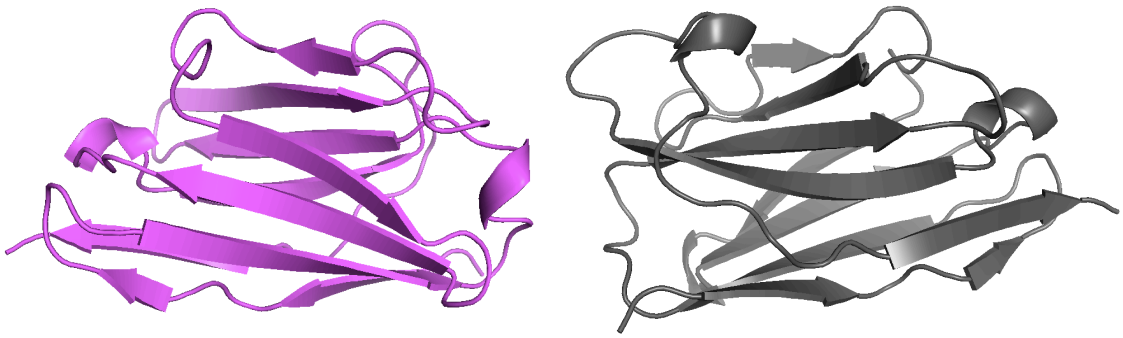

Figure 3.20 Successful sdAb candidates for Omicron variants common between 3 chains after 1st round docking experiment. (sdAb NM1223 (left, magenta) and sdAb 2F2 (right, gray)). Visualization was created via PyMOL (cartoon representation).

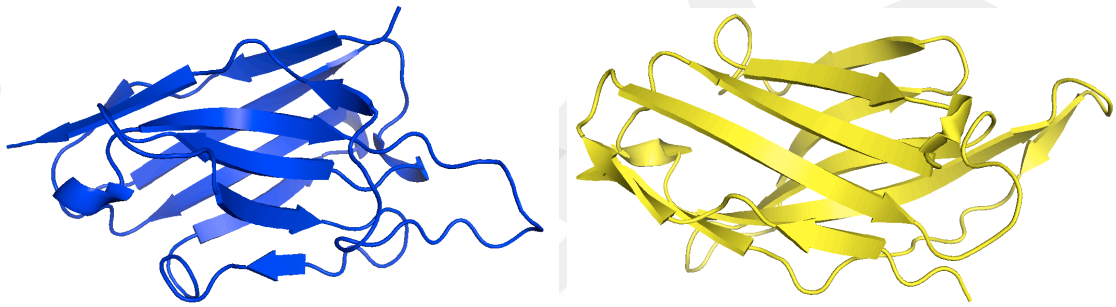

Figure 3.21 Successful sdAb candidates for Omicron variants common between 3 chains after 2nd round docking experiment. (sdAb VHH E (left, blue) and sdAb S1-24 (right, yellow)). Visualization was created via PyMOL (cartoon representation).

```
S1-24       QVQPVESGGGLVQAGGSLRLSCVASGSTTTNYHMGWYRQTPGEQRELVAAIN-AGGITNY   59
NM1223      EVQLVESGGGLVQPGGSLRLSCSASGFAFSSVSMSWVRLLPGKGTEWVAEIDRDGGNGNY   60
2F2         QVQLVESGGGLVQPGGSLRLSCAASGLAQSKWAYGWFRQAPGKGLEAVAAIDVATGPWYY   60
VHH-E       QVQLVETGGGFVQPGGSLRLSCAASGVTLDYYAIGWFRQAPGKEREGVSCIGSSDGRTYY   60
            :** **:***:** ******** *** :      .* *  **:  * *: *.   *   *

S1-24       ADSVKGRFTISRDNAKNTMYLQMNNLRFEDTAVYYCNIGGGWDYRNSYY---IPRVDSWG   116
NM1223      EDSVKGRFTISRDNAKNTLFLQMNSLVPEDTALYYCRLGTRD----------HIMSGWG   109
2F2         ADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCAAHHIP----TKHPAFPDFRDYWG   116
VHH-E       SDSVKGRFTISRDNAKNTVYLQMNSLKPEDTAVYYCALTVGTYYSGNYHYTCSDDMDYWG   120
             ************:***::****.*  ****:***                    . **

S1-24       QGTQVTVS-   124
NM1223      PGAPVTVSS   118
2F2         QGTQVTVSS   125
VHH-E       KGTQVTVSS   129
             *: ****
```

Figure 3.22 Amino acid sequence alignment for obtained 4 candidate sdAbs. Alignment was performed Clustal Omega web server [162,163].

Chapter 4

CONCLUSIONS AND FUTURE PROSPECTS

4.1 Conclusions

The usage of sdAbs against infectious diseases has been among the trend approaches in recent years. sdAbs are antibody fragment molecules containing only the heavy chain, synthesized in camels. These fragments are called as a variable regions and are responsible for recognition of target antigen epitopes [119,125]. Once the sequence information of the targeted sdAbs gene is obtained, it provides a continuous, reliable source by allowing these sdAbs to be resynthesized at any time and the desired modifications to be made [176,177].

Before 1st round docking experiments, in order to find appropriate PDB structures were screened RSSB-PDB database. PDB ID:6VXX [161] was chosen for the SARS-CoV-2 wild type and PDB ID:7WP9 [178] for the Omicron variant because these two PDB structures belong to the closed (all RBD's down) state of the SARS-Cov-2 Spike protein and their resolutions are quite high in comparison with others. There are 2 common forms of Spike protein which was determined when SARS-CoV-2 infects the organism [98,104,111,179–181]. The first one is open form, which has at least one RBD of Spike protein is in the up position, and this is the form usually used for docking experiments. Besides many studies currently use the RBD region as a target and it is easier to reach the RBD region in open form. Closed state from among conformations of SARS-CoV-2 Spike proteins was used in docking experiments. After infection of host organisms, SARS-CoV-2 Spike proteins are in open state conformation for a very short time interval. Open state conformation, which at least one RBD is in up position, only can be seen immediately before ACE2-RBD interactions. The Spike protein is in the closed conformation in infected organisms during the rest of the time. Therefore, a more realistic state of the Spike protein in the natural environment was targeted. When the spike

protein is in the closed state, the candidate sdAbs interacting on many RBD regions can prevent the transition of the RBDs from the down position to the up position. This may block the expected interaction between ACE2 and the Spike protein (Figure 3.5, Figure 3.12, and Figure 3.15).

The receptor molecule used for SARS-CoV-2 neutralization interactions is usually the RBD of the Spike protein. Because SARS-CoV-2 acts as a ligand that binds to the ACE2 receptors of cells by using the RBD to infect host cells. Therefore, targeting the RBD of SARS-CoV-2 and neutralizing this region has been the primary approach in various studies [143–145]. In order to find neutralizing molecules against SARS-CoV-2, regions with different functions on the Spike protein, such as the cleavage site of proteases that have an important role in cell infections, can be targeted. Within the scope of the current thesis, other biologically important regions were also accepted as a target by using the homotrimer structure of the spike protein as a receptor.

Furthermore, mutation sites of the SARS-CoV-2 Omicron variant were used, while active residues were determining for the Spike protein during the docking experiments. In this way, it was aimed to investigate the effects of mutations of the Omicron variant on sdAbs activities (Table 2.2 and Table 2.3). When the results of the 1st round docking experiment are examined, it is appeared that the Omicron variant mutations significantly decrease the number of successful sdAbs for 3 chains. While the common successful sdAb number for the 3 chains in strong interaction with the wild type Spike protein was 31, this number was observed as 2 in the docking experiment using Omicron variant (Figure 3.3 and Figure 3.4). A healthy comparison could not be made for the second-round docking experiment, because after the addition of missing residues, unexpected loops were formed far away from the rigid body of wild type Spike protein structure. These loops could not have a favorable position as a result of 1 nanosecond molecular dynamic simulation (Figure 2.5 and Figure 2.6). Therefore, chain C of the wild type Spike protein which was used in the second round experiment has a problematic structure. As a result, this situation significantly decreased the number of successful sdAb for chain C (Figure 3.10). However, in chain A and chain B, the number of successful sdAbs decreased from 48 (wild type) to 16 (Omicron variant) (Figure 3.10 and Figure 3.11). Here, the effect of mutations in the Omicron variant is seen, too.

In this thesis study, as a result of using Spike protein as a complete structure instead of only RBD or single chain, binding of sdAbs to regions that are normally inaccessible

was prevented. This situation presented a more realistic experiment approach. Otherwise, the region which a sdAb effectively binds could be inside the Spike homotrimer structure for docking experiments that single chain is used as a receptor. Likewise, adding missing residues before the 2nd round docking experiments enabled the structural deficiencies of the models on the PDB website to be eliminated. This is clearly seen when the common successful sdAb numbers for 3 chains obtained as a result of the 1st and 2nd round docking experiments are examined. According to the results where TOP 1000 scores were used as threshold, the number of common successful sdAbs for 3 chains for the 2 variants (wild type and Omicron) in the 1st round docking experiments was totally 33, while it was 523 in the 2nd round (Figure 3.3, Figure 3.4, Figure 3.8 and Figure 3.9). The effect of the newly developed filtration approach, which is used before scoring the 2nd round docking experiments and based on sdAb interaction sites and distance between receptor-ligand as parameters, should not be ignored. This filtration technique also contributed to a significant increase in the number of determined common successful sdAbs for 3 chains after the 2nd round docking experiments. Considering the results obtained, the addition of missing residues significantly increased the number of common sdAbs as a result of docking experiments (Figure 3.3, Figure 3.4, Figure 3.8, Figure 3.9, Figure 3.10 and Figure 3.11).

In addition, several studies have reported that glycosylation affects the interactions of the Spike protein with other proteins [94,179,182,183]. Therefore, it is not a realistic approach to use only the RBD region of Spike protein without glycans. Although there is a fully glycosylated model for wild type Spike protein [164], it was not present in any model for the Omicron variant. In 2nd round docking experiment, glycosylation patterns were researched to create a model. In literature, there are some studies about glycosylation patterns of Spike proteins [184,185]. Glycosylation patterns were listed and an extra new site for O-glycan was added from another study [186]. Some glycosylation modelling tools [164,187,188] were tested but because of complex structure of Spike protein and inadequacy of softwares, it could not be performed. Therefore, only missing residues were completed for Spike protein PDB models.

Regions out of the RBD that may be important for SARS-CoV-2 infections can be overlooked. For example, in cell entry of SARS-CoV-2 and releasing of maturated new virions process, there are significantly important proteases and cleavage sites of these proteases are not in the RBD region. Within the frame of this thesis study, both the use of

the entire trimer structure of the spike protein is an important approach to find new biologically functional neutralization sites on the spike protein. Moreover, addition of missing residues ensured realistic experimental design because 3D models on the PDB website have structural deficiency and effect of this application was seen in 2nd round docking results. In addition, if the sdAbs interact with 2 chains of Spike protein at the same time, interactions can weaken. In addition, if the sdAbs interact with 2 chains of Spike protein at the same time, interactions can weaken. This process also provided to test repeatability of the docking experiment because each chain of Spike protein has the same sequence and structure. Detection of successful common sdAbs for all 3 chains in two different docking experiments is the biggest indicator of this. Determination of active residues based on the mutation sites of the Omicron variant for docking experiments was ensured possibility to observe effects of the mutations on sdAbs which are successful for wild type.

As a result, all approaches used during the thesis study provided a more realistic, rational, and repeatable screening method for creating large libraries for sdAb-large complex protein interactions. Structure models of 850 sdAbs were created for the new structure library. This library is also important for candidate sdAb screening processes for new variants are anticipated to emerge in the future. In addition, a consensus scoring approach was used which are using chemistry and mathematical algorithms such as HDOCK, Bluues and PRODIGY. Moreover, the new strategy developed in the thesis study and used for pre-filtration before the 2nd round scoring is based on molecular biology and evolutionary properties, unlike other scoring functions such as HDOCK, Bluues, PRODIGY. This application was very helpful to increase the accuracy of results. Omicron variant mutations effects on protein-protein docking experiments was observed. New regions, such as protease cleavage sites can be targeted in SARS-CoV-2 infections were determined. The usage of new sites as a target together with RBD regions, can increase neutralization of SARS-CoV-2 with sdAbs. In addition to these, it is among the findings obtained as a result of the thesis that the deficiencies in protein model structures on the PDB website can have serious negative effects on docking experiments. Finally, it is not possible to test the 561000 complexes obtained after the 2nd round docking experiment in the laboratory environment, both in terms of time and cost. As a result of this thesis, this number has been reduced to 2 owing to rational design and molecular biology-based approaches.

4.2 Societal Impact and Contribution to Global Sustainability

The selection of heavy chain only antibodies (HcAb) or sdAbs, which are known for their suitability for high-scale production, have a great potential to contribute to fight against different diseases and developing high value-added domestic biotechnological drugs. Identification of SARS-CoV-2-specific HcAb or sdAbs with proteomics methods will enable faster development of high value-added diagnostic kits and alternative treatment methods (such as smart drugs, passive vaccines) against new coronavirus variants. These sdAbs have the potential to make a very important social and economic contribution to their small molecular structure, structural stability, and suitability for high-scale production in industry. Traditional methods such as phage display and hybridoma have been used for antibody selection that require extreme effort and time, but the method which was used in this thesis makes the process easier. The roadmap has been obtained in terms of approach and method will enable the selection of candidate sdAbs from extensive databases for new variants of SARS-CoV-2 that will emerge in the future. Determined sequences for new variants after fast screening may be cloned to microorganisms such as *Escherichia coli* bacteria under laboratory conditions, and new sdAbs which have a successful neutralization rate can be produced for diagnosis or treatment. Moreover, some of the selected sdAbs can be involved in sdAb maturation processes and may be increased their affinity by performing mutations in their sequences. The sdAb sequence repertoire can be revealed by transcriptomic approaches, and new sdAbs specific to the target epitope can be identified by using proteomic methods. Due to all this process, time, material, and labor that would normally be spent in a laboratory environment can be saved. Most importantly, as a result of the identification of candidate sequences without the use of any animal from the *Camelidae* family, which is necessary for the production of sdAbs, these animals are prevented from being damaged in terms of health.

4.3 Future Prospects

Single domain antibodies (sdAb) with their several advantages due to size probably will lead the field which antibody-based neutralization is foreseeable in thefuture because they have all functional properties of antibodies. Especially, in diagnosis area antibodies

have serious a negative, dirty background problem which is undergoing because of their half-life. In such cases, it is inevitable to prefer sdAbs. In addition, the ability of sdAbs to reach some hard-to-attain regions makes them very useful for potential new targets. The 4 sdAbs obtained as a result of the thesis study will also use these advantages when testing against the Omicron variant.

Obtained 4 sdAb structures can be tested by using molecular dynamics simulation with a long time against to Omicron variant. New strategies can be researched for glycan model of SARS-CoV-2 Spike protein Omicron variant. Possible glycan models can provide to perform more realistic protein-protein docking or molecular dynamics simulation experiments. Also, the activity of sdAbs against to Omicron variant can be increased by making some mutations in the complementary determining regions.

The names and sequence information of these 4 sdAbs are already known. Therefore, they can be easily cloned into microorganism and produced in their expression system. Neutralization levels can then be verified in with experiments under laboratory conditions. After the necessary tests and approvals, it can be offered to humanity as a drug or sera. In addition, the scripting codes and methodological approach developed within the scope of this thesis can be used against new SARS-CoV-2 variants that may emerge in the future or other diseases. In silico studies can be conducted as more realistic when glycan structure models are included to the method process.

BIBLIOGRAPHY

[1] M.A. Shereen, S. Khan, A. Kazmi, N. Bashir, R. Siddique, COVID-19 infection: Origin, transmission, and characteristics of human coronaviruses, J. Adv. Res. 24 (2020) 91–98. https://doi.org/10.1016/j.jare.2020.03.005.

[2] S.B. Kadam, G.S. Sukhramani, P. Bishnoi, A.A. Pable, V.T. Barvkar, SARS-CoV-2, the pandemic coronavirus: Molecular and structural insights, J. Basic Microbiol. 61 (2021) 180–202. https://doi.org/10.1002/jobm.202000537.

[3] A.E. Gorbalenya, S.C. Baker, R.S. Baric, R.J. de Groot, C. Drosten, A.A. Gulyaeva, B.L. Haagmans, C. Lauber, A.M. Leontovich, B.W. Neuman, D. Penzar, S. Perlman, L.L.M. Poon, D. V. Samborskiy, I.A. Sidorov, I. Sola, J. Ziebuhr, The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2, Nat. Microbiol. 5 (2020) 536–544. https://doi.org/10.1038/s41564-020-0695-z.

[4] L.F. Wang, C. Cowled, Bats and Viruses: A New Frontier of Emerging Infectious Diseases, Bats Viruses A New Front. Emerg. Infect. Dis. (2015) 1–368. https://doi.org/10.1002/9781118818824.

[5] G.E. Antonio, K.T. Wong, D.S.C. Hui, N. Lee, E.H.Y. Yuen, A. Wu, S.S.C. Chung, J.J.Y. Sung, A.T. Ahuja, Imaging of severe acute respiratory syndrome in Hong Kong, Am. J. Roentgenol. 181 (2003) 11–17. https://doi.org/10.2214/ajr.181.1.1810011.

[6] R.J. de Groot, S.C. Baker, R.S. Baric, C.S. Brown, C. Drosten, L. Enjuanes, R.A.M. Fouchier, M. Galiano, A.E. Gorbalenya, Z.A. Memish, S. Perlman, L.L.M. Poon, E.J. Snijder, G.M. Stephens, P.C.Y. Woo, A.M. Zaki, M. Zambon, J. Ziebuhr, Commentary: Middle East Respiratory Syndrome Coronavirus (MERS-CoV): Announcement of the Coronavirus Study Group, J. Virol. 87 (2013) 7790–7792. https://doi.org/10.1128/jvi.01244-13.

[7] S.J. Sullivan, R.M. Jacobson, W.R. Dowdle, G.A. Poland, 2009 H1N1 influenza, Mayo Clin. Proc. 85 (2010) 64–76. https://doi.org/10.4065/mcp.2009.0588.

[8] K. Subbarao, C. Luke, H5N1 viruses and vaccines, PLoS Pathog. 3 (2007) 5–7. https://doi.org/10.1371/journal.ppat.0030040.

[9] V.N. Chouljenko, X.Q. Lin, J. Storz, K.G. Kousoulas, A.E. Gorbalenya, Comparison of genomic and predicted amino acid sequences of respiratory and enteric bovine coronaviruses isolated from the same animal with fatal shipping pneumonia, J. Gen. Virol. 82 (2001) 2927–2933. https://doi.org/10.1099/0022-1317-82-12-2927.

[10] P.G. Halbur, P.S. Paul, E.M. Vaughn, J.J. Andrews, Experimental reproduction of pneumonia in gnotobiotic pigs with porcine respiratory coronavirus isolate AR310, J. Vet. Diagnostic Investig. 5 (1993) 184–188. https://doi.org/10.1177/104063879300500207.

[11] J.S.M. Peiris, S.T. Lai, L.L.M. Poon, Y. Guan, L.Y.C. Yam, W. Lim, J. Nicholls, W.K.S. Yee, W.W. Yan, M.T. Cheung, V.C.C. Cheng, K.H. Chan, D.N.C. Tsang, R.W.H. Yung, T.K. Ng, K.Y. Yuen, Coronavirus as a possible cause of severe acute respiratory syndrome, Lancet. 361 (2003) 1319–1325. https://doi.org/10.1016/S0140-6736(03)13077-2.

[12] J.S.M. Peiris, C.M. Chu, V.C.C. Cheng, K.S. Chan, I.F.N. Hung, L.L.M. Poon, K.I. Law, B.S.F. Tang, T.Y.W. Hon, C.S. Chan, K.H. Chan, J.S.C. Ng, B.J. Zheng, W.L. Ng, R.W.M. Lai, Y. Guan, K.Y. Yuen, Clinical progression and viral load in

a community outbreak of coronavirus-associated SARS pneumonia: A prospective study, Lancet. 361 (2003) 1767–1772. https://doi.org/10.1016/S0140-6736(03)13412-5.

[13] J. Storz, X. Lin, C.W. Purdy, V.N. Chouljenko, K.G. Kousoulas, F.M. Enright, W.C. Gilmore, R.E. Briggs, R.W. Loan, Coronavirus and Pasteurella infections in bovine shipping fever pneumonia and Evans' criteria for causation, J. Clin. Microbiol. 38 (2000) 3291–3298. https://doi.org/10.1128/jcm.38.9.3291-3298.2000.

[14] K. Subbarao, J. McAuliffe, L. Vogel, G. Fahle, S. Fischer, K. Tatti, M. Packard, W.-J. Shieh, S. Zaki, B. Murphy, Prior Infection and Passive Transfer of Neutralizing Antibody Prevent Replication of Severe Acute Respiratory Syndrome Coronavirus in the Respiratory Tract of Mice, J. Virol. 78 (2004) 3572–3577. https://doi.org/10.1128/jvi.78.7.3572-3577.2004.

[15] S. Ludwig, A. Zarbock, Coronaviruses and SARS-CoV-2: A Brief Overview, Anesth. Analg. 131 (2020) 93–96. https://doi.org/10.1016/j.meegid.2020.104502.

[16] B. Hu, H. Guo, P. Zhou, Z.L. Shi, Characteristics of SARS-CoV-2 and COVID-19, Nat. Rev. Microbiol. 19 (2021) 141–154. https://doi.org/10.1038/s41579-020-00459-7.

[17] WHO Coronavirus (COVID-19) Dashboard, (n.d.). https://covid19.who.int/ (accessed April 6, 2022).

[18] Y. Chen, Q. Liu, D. Guo, Emerging coronaviruses: Genome structure, replication, and pathogenesis, J. Med. Virol. 92 (2020) 418–423. https://doi.org/10.1002/jmv.25681.

[19] L. Mousavizadeh, S. Ghasemi, Genotype and phenotype of COVID-19: Their roles in pathogenesis, J. Microbiol. Immunol. Infect. 54 (2020) 159–163. https://doi.org/10.1016/j.jmii.2020.03.022.

[20] A. Wu, Y. Peng, B. Huang, X. Ding, X. Wang, P. Niu, J. Meng, Z. Zhu, Z. Zhang, J. Wang, J. Sheng, L. Quan, Z. Xia, W. Tan, G. Cheng, T. Jiang, Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China, Cell Host Microbe. 27 (2020) 325–328. https://doi.org/10.1016/j.chom.2020.02.001.

[21] R. Lu, X. Zhao, J. Li, P. Niu, B. Yang, H. Wu, W. Wang, H. Song, B. Huang, N. Zhu, Y. Bi, X. Ma, F. Zhan, L. Wang, T. Hu, H. Zhou, Z. Hu, W. Zhou, L. Zhao, J. Chen, Y. Meng, J. Wang, Y. Lin, J. Yuan, Z. Xie, J. Ma, W.J. Liu, D. Wang, W. Xu, E.C. Holmes, G.F. Gao, G. Wu, W. Chen, W. Shi, W. Tan, Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding, Lancet. 395 (2020) 565–574. https://doi.org/10.1016/S0140-6736(20)30251-8.

[22] X. Xu, C. Ping, W. Jingfang, F. Jiannan, Z. Hui, L. Xuan, Z. Wu, H. Pei, Evolution Of The Novel Coronavirus From The Ongoing Wuhan Outbreak And Modeling Of Its Spike Protein For Risk Of Human Transmission, Sci. China Life Sci. 63 (2020) 457–460.

[23] Y. Wan, J. Shang, R. Graham, R.S. Baric, F. Li, Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus, J. Virol. 94 (2020). https://doi.org/10.1128/jvi.00127-20.

[24] Z. Miao, A. Tidu, G. Eriani, F. Martin, Secondary structure of the SARS-CoV-2 5'-UTR, RNA Biol. 18 (2021) 447–456. https://doi.org/10.1080/15476286.2020.1814556.

[25] R. Madhugiri, M. Fricke, M. Marz, J. Ziebuhr, RNA structure analysis of

alphacoronavirus terminal genome regions, Virus Res. 194 (2014) 76–89. https://doi.org/10.1016/j.virusres.2014.10.001.
[26] D. Yang, J.L. Leibowitz, The structure and functions of coronavirus genomic 3' and 5' ends, Virus Res. 206 (2015) 120–133. https://doi.org/10.1016/j.virusres.2015.02.025.
[27] P.S. Fehr A.R., Coronaviruses: An Overview of Their Replication and Pathogenesis, in: B.P. Maier H., Bickerton E. (Ed.), Coronaviruses. Methods Mol. Biol. Vol 1282, Humana Press, New York, 2015: pp. 1–23. https://doi.org/10.1007/978-1-4939-2438-7_1.
[28] N. Redondo, S. Zaldívar-López, J.J. Garrido, M. Montoya, SARS-CoV-2 Accessory Proteins in Viral Pathogenesis: Knowns and Unknowns, Front. Immunol. 12 (2021) 1–8. https://doi.org/10.3389/fimmu.2021.708264.
[29] D.X. Liu, T.S. Fung, K.K.L. Chong, A. Shukla, R. Hilgenfeld, Accessory proteins of SARS-CoV and other coronaviruses, Antiviral Res. 109 (2014) 97–109. https://doi.org/10.1016/j.antiviral.2014.06.013.
[30] P. Zhou, X. Lou Yang, X.G. Wang, B. Hu, L. Zhang, W. Zhang, H.R. Si, Y. Zhu, B. Li, C.L. Huang, H.D. Chen, J. Chen, Y. Luo, H. Guo, R. Di Jiang, M.Q. Liu, Y. Chen, X.R. Shen, X. Wang, X.S. Zheng, K. Zhao, Q.J. Chen, F. Deng, L.L. Liu, B. Yan, F.X. Zhan, Y.Y. Wang, G.F. Xiao, Z.L. Shi, A pneumonia outbreak associated with a new coronavirus of probable bat origin, Nature. 579 (2020) 270–273. https://doi.org/10.1038/s41586-020-2012-7.
[31] I. Seah, X. Su, G. Lingam, Revisiting the dangers of the coronavirus in the ophthalmology practice, Eye. 34 (2020) 1155–1157. https://doi.org/10.1038/s41433-020-0790-7.
[32] J. Yang, S.J.L. Petitjean, M. Koehler, Q. Zhang, A.C. Dumitru, W. Chen, S. Derclaye, S.P. Vincent, P. Soumillion, D. Alsteens, Molecular interaction and inhibition of SARS-CoV-2 binding to the ACE2 receptor, Nat. Commun. 11 (2020). https://doi.org/10.1038/s41467-020-18319-6.
[33] E.J. Snijder, P.J. Bredenbeek, J.C. Dobbe, V. Thiel, J. Ziebuhr, L.L.M. Poon, Y. Guan, M. Rozanov, W.J.M. Spaan, A.E. Gorbalenya, Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage, J. Mol. Biol. 331 (2003) 991–1004. https://doi.org/10.1016/S0022-2836(03)00865-9.
[34] B.J. Bosch, R. van der Zee, C.A.M. de Haan, P.J.M. Rottier, The Coronavirus Spike Protein Is a Class I Virus Fusion Protein: Structural and Functional Characterization of the Fusion Core Complex, J. Virol. 77 (2003) 8801–8811. https://doi.org/10.1128/jvi.77.16.8801-8811.2003.
[35] S. Belouzard, V.C. Chu, G.R. Whittaker, Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites, Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 5871–5876. https://doi.org/10.1073/pnas.0809524106.
[36] Y. Lang, W. Li, Z. Li, D. Koerhuis, A.C.S. Van Den Burg, E. Rozemuller, B.J. Bosch, F.J.M. Van Kuppeveld, G.J. Boons, E.G. Huizinga, H.M. Van Der Schaar, R.J. De Groot, Coronavirus hemagglutinin-esterase and spike proteins coevolve for functional balance and optimal virion avidity, Proc. Natl. Acad. Sci. U. S. A. 117 (2020) 25759–25770. https://doi.org/10.1073/pnas.2006299117.
[37] M. Zandi, E. Behboudi, S. Soltani, Role of Glycoprotein Hemagglutinin-Esterase in COVID-19 Pathophysiology?, Stem Cell Rev. Reports. 17 (2021) 2359–2360. https://doi.org/10.1007/s12015-021-10210-1.
[38] J. Cui, F. Li, Z.L. Shi, Origin and evolution of pathogenic coronaviruses, Nat. Rev.

Microbiol. 17 (2019) 181–192. https://doi.org/10.1038/s41579-018-0118-9.
[39] N. Nao, J. Yamagishi, H. Miyamoto, M. Igarashi, R. Manzoor, A. Ohnuma, Y. Tsuda, W. Furuyama, A. Shigeno, M. Kajihara, N. Kishida, R. Yoshida, A. Takada, Genetic Predisposition To Acquire a Polybasic Cleavage Site for Highly, MBio. 8 (2017) e02298-16.
[40] Y.L. Siu, K.T. Teoh, J. Lo, C.M. Chan, F. Kien, N. Escriou, S.W. Tsao, J.M. Nicholls, R. Altmeyer, J.S.M. Peiris, R. Bruzzone, B. Nal, 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. 82 (2008) 11318–11330. https://doi.org/10.1128/jvi.01052-08.
[41] B. Nal, C. Chan, F. Kien, L. Siu, J. Tse, K. Chu, J. Kam, sabelle Staropoli, B. Crescenzo-Chaigne, N. Escriou, S. van der Wef, K.Y. Yuen, R. Altmeyer, Differential maturation and subcellular localization of severe acute respiratory syndrome coronavirus surface proteins S, M and E, J. Gen. Virol. 86 (2005) 1423–1434. https://doi.org/10.1099/vir.0.80671-0.
[42] T.R. Ruch, C.E. Machamer, The coronavirus E protein: Assembly and beyond, Viruses. 4 (2012) 363–382. https://doi.org/10.3390/v4030363.
[43] A.K. Banerjee, M.R. Blanco, E.A. Bruce, D.D. Honson, L.M. Chen, A. Chow, P. Bhat, N. Ollikainen, S.A. Quinodoz, C. Loney, J. Thai, Z.D. Miller, A.E. Lin, M.M. Schmidt, D.G. Stewart, D. Goldfarb, G. De Lorenzo, S.J. Rihn, R.M. Voorhees, J.W. Botten, D. Majumdar, M. Guttman, SARS-CoV-2 Disrupts Splicing, Translation, and Protein Trafficking to Suppress Host Defenses, Cell. 183 (2020) 1325-1339.e21. https://doi.org/10.1016/j.cell.2020.10.004.
[44] A.A.T. Naqvi, K. Fatima, T. Mohammad, U. Fatima, I.K. Singh, A. Singh, S.M. Atif, G. Hariprasad, G.M. Hasan, M.I. Hassan, Insights into SARS-CoV-2 genome, structure, evolution, pathogenesis and therapies: Structural genomics approach, Biochim. Biophys. Acta - Mol. Basis Dis. 1866 (2020) 165878. https://doi.org/10.1016/j.bbadis.2020.165878.
[45] X. Ou, Y. Liu, X. Lei, P. Li, D. Mi, L. Ren, L. Guo, R. Guo, T. Chen, J. Hu, Z. Xiang, Z. Mu, X. Chen, J. Chen, K. Hu, Q. Jin, J. Wang, Z. Qian, Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV, Nat. Commun. 11 (2020). https://doi.org/10.1038/s41467-020-15562-9.
[46] C. Huang, K.G. Lokugamage, J.M. Rozovics, K. Narayanan, B.L. Semler, S. Makino, Alphacoronavirus Transmissible Gastroenteritis Virus nsp1 Protein Suppresses Protein Translation in Mammalian Cells and in Cell-Free HeLa Cell Extracts but Not in Rabbit Reticulocyte Lysate, J. Virol. 85 (2011) 638–643. https://doi.org/10.1128/jvi.01806-10.
[47] W. Kamitani, C. Huang, K. Narayanan, K.G. Lokugamage, S. Makino, A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein, Nat. Struct. Mol. Biol. 16 (2009) 1134–1140. https://doi.org/10.1038/nsmb.1680.
[48] W. Kamitani, K. Narayanan, C. Huang, K. Lokugamage, T. Ikegami, N. Ito, H. Kubo, S. Makino, Severe acute respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by promoting host mRNA degradation, Proc. Natl. Acad. Sci. U. S. A. 103 (2006) 12885–12890. https://doi.org/10.1073/pnas.0603144103.
[49] T. Tanaka, W. Kamitani, M.L. DeDiego, L. Enjuanes, Y. Matsuura, Severe Acute Respiratory Syndrome Coronavirus nsp1 Facilitates Efficient Propagation in Cells

through a Specific Translational Shutoff of Host mRNA, J. Virol. 86 (2012) 11128–11137. https://doi.org/10.1128/jvi.01700-12.
[50] R.L. Graham, A.C. Sims, R.S. Baric, M.R. Denison, The nsp2 proteins of mouse hepatitis virus and SARS coronavirus are dispensable for viral replication, Adv. Exp. Med. Biol. 581 (2006) 67–72. https://doi.org/10.1007/978-0-387-33012-9_10.
[51] C.T. Cornillez-Ty, L. Liao, J.R. Yates, P. Kuhn, M.J. Buchmeier, Severe Acute Respiratory Syndrome Coronavirus Nonstructural Protein 2 Interacts with a Host Protein Complex Involved in Mitochondrial Biogenesis and Intracellular Signaling, J. Virol. 83 (2009) 10314–10318. https://doi.org/10.1128/jvi.00842-09.
[52] A. Chatterjee, M.A. Johnson, P. Serrano, B. Pedrini, J.S. Joseph, B.W. Neuman, K. Saikatendu, M.J. Buchmeier, P. Kuhn, K. Wüthrich, Nuclear Magnetic Resonance Structure Shows that the Severe Acute Respiratory Syndrome Coronavirus-Unique Domain Contains a Macrodomain Fold, J. Virol. 83 (2009) 1823–1836. https://doi.org/10.1128/jvi.01781-08.
[53] M.-P. Egloff, H. Malet, Á. Putics, M. Heinonen, H. Dutartre, A. Frangeul, A. Gruez, V. Campanacci, C. Cambillau, J. Ziebuhr, T. Ahola, B. Canard, Structural and Functional Basis for ADP-Ribose and Poly(ADP-Ribose) Binding by Viral Macro Domains, J. Virol. 80 (2006) 8493–8502. https://doi.org/10.1128/jvi.00713-06.
[54] K.K. Eriksson, L. Cervantes-Barragán, B. Ludewig, V. Thiel, Mouse Hepatitis Virus Liver Pathology Is Dependent on ADP-Ribose-1″-Phosphatase, a Viral Function Conserved in the Alpha-Like Supergroup, J. Virol. 82 (2008) 12325–12334. https://doi.org/10.1128/jvi.02082-08.
[55] M. Frieman, K. Ratia, R.E. Johnston, A.D. Mesecar, R.S. Baric, Severe Acute Respiratory Syndrome Coronavirus Papain-Like Protease Ubiquitin-Like Domain and Catalytic Domain Regulate Antagonism of IRF3 and NF-κB Signaling, J. Virol. 83 (2009) 6689–6705. https://doi.org/10.1128/jvi.02220-08.
[56] B.W. Neuman, J.S. Joseph, K.S. Saikatendu, P. Serrano, A. Chatterjee, M.A. Johnson, L. Liao, J.P. Klaus, J.R. Yates, K. Wüthrich, R.C. Stevens, M.J. Buchmeier, P. Kuhn, Proteomics Analysis Unravels the Functional Repertoire of Coronavirus Nonstructural Protein 3, J. Virol. 82 (2008) 5279–5294. https://doi.org/10.1128/jvi.02631-07.
[57] P. Serrano, M.A. Johnson, A. Chatterjee, B.W. Neuman, J.S. Joseph, M.J. Buchmeier, P. Kuhn, K. Wüthrich, Nuclear Magnetic Resonance Structure of the Nucleic Acid-Binding Domain of Severe Acute Respiratory Syndrome Coronavirus Nonstructural Protein 3, J. Virol. 83 (2009) 12998–13008. https://doi.org/10.1128/jvi.01253-09.
[58] P. Serrano, M.A. Johnson, M.S. Almeida, R. Horst, T. Herrmann, J.S. Joseph, B.W. Neuman, V. Subramanian, K.S. Saikatendu, M.J. Buchmeier, R.C. Stevens, P. Kuhn, K. Wüthrich, Nuclear Magnetic Resonance Structure of the N-Terminal Domain of Nonstructural Protein 3 from the Severe Acute Respiratory Syndrome Coronavirus, J. Virol. 81 (2007) 12049–12060. https://doi.org/10.1128/jvi.00969-07.
[59] J. Ziebuhr, V. Thiel, A.E. Gorbalenya, The Autocatalytic Release of a Putative RNA Virus Transcription Factor from Its Polyprotein Precursor Involves Two Paralogous Papain-like Proteases that Cleave the Same Peptide Bond, J. Biol. Chem. 276 (2001) 33220–33232. https://doi.org/10.1074/jbc.M104097200.
[60] M.A. Clementz, A. Kanjanahaluethai, T.E. O'Brien, S.C. Baker, Mutation in

murine coronavirus replication protein nsp4 alters assembly of double membrane vesicles, Virology. 375 (2008) 118–129. https://doi.org/10.1016/j.virol.2008.01.018.
[61] M.J. Gadlage, J.S. Sparks, D.C. Beachboard, R.G. Cox, J.D. Doyle, C.C. Stobart, M.R. Denison, Murine Hepatitis Virus Nonstructural Protein 4 Regulates Virus-Induced Membrane Modifications and Replication Complex Function, J. Virol. 84 (2010) 280–290. https://doi.org/10.1128/jvi.01772-09.
[62] Y. Lu, X. Lu, M.R. Denison, Identification and characterization of a serine-like proteinase of the murine coronavirus MHV-A59, J. Virol. 69 (1995) 3554–3559. https://doi.org/10.1128/jvi.69.6.3554-3559.1995.
[63] M. Oostra, M.C. Hagemeijer, M. van Gent, C.P.J. Bekker, E.G. te Lintelo, P.J.M. Rottier, C.A.M. de Haan, Topology and Membrane Anchoring of the Coronavirus Replication Complex: Not All Hydrophobic Domains of nsp3 and nsp6 Are Membrane Spanning, J. Virol. 82 (2008) 12392–12405. https://doi.org/10.1128/jvi.01219-08.
[64] Y. Zhai, F. Sun, X. Li, H. Pang, X. Xu, M. Bartlam, Z. Rao, Insights into SARS-CoV transcription and replication from the structure of the nsp7-nsp8 hexadecamer, Nat. Struct. Mol. Biol. 12 (2005) 980–986. https://doi.org/10.1038/nsmb999.
[65] I. Imbert, J.C. Guillemot, J.M. Bourhis, C. Bussetta, B. Coutard, M.P. Egloff, F. Ferron, A.E. Gorbalenya, B. Canard, A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus, EMBO J. 25 (2006) 4933–4942. https://doi.org/10.1038/sj.emboj.7601368.
[66] M.P. Egloff, F. Ferron, V. Campanacci, S. Longhi, C. Rancurel, H. Dutartre, E.J. Snijder, A.E. Gorbalenya, C. Cambillau, B. Canard, The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 3792–3796. https://doi.org/10.1073/pnas.0307877101.
[67] M. Bouvet, C. Debarnot, I. Imbert, B. Selisko, E.J. Snijder, B. Canard, E. Decroly, In vitro reconstitution of sars-coronavirus mRNA cap methylation, PLoS Pathog. 6 (2010) 1–13. https://doi.org/10.1371/journal.ppat.1000863.
[68] E. Decroly, C. Debarnot, F. Ferron, M. Bouvet, B. Coutard, I. Imbert, L. Gluais, N. Papageorgiou, A. Sharff, G. Bricogne, M. Ortiz-Lombardia, J. Lescar, B. Canard, Crystal structure and functional analysis of the SARS-coronavirus RNA cap 2′-o-methyltransferase nsp10/nsp16 complex, PLoS Pathog. 7 (2011). https://doi.org/10.1371/journal.ppat.1002059.
[69] X. Xu, Y. Liu, S. Weiss, E. Arnold, S.G. Sarafianos, J. Ding, Molecular model of SARS coronavirus polymerase: Implications for biochemical functions and drug design, Nucleic Acids Res. 31 (2003) 7117–7130. https://doi.org/10.1093/nar/gkg916.
[70] K.A. Ivanov, V. Thiel, J.C. Dobbe, Y. van der Meer, E.J. Snijder, J. Ziebuhr, Multiple Enzymatic Activities Associated with Severe Acute Respiratory Syndrome Coronavirus Helicase, J. Virol. 78 (2004) 5619–5632. https://doi.org/10.1128/jvi.78.11.5619-5632.2004.
[71] K.A. Ivanov, J. Ziebuhr, Human Coronavirus 229E Nonstructural Protein 13: Characterization of Duplex-Unwinding, Nucleoside Triphosphatase, and RNA 5′-Triphosphatase Activities, J. Virol. 78 (2004) 7833–7838. https://doi.org/10.1128/jvi.78.14.7833-7838.2004.
[72] L.D. Eckerle, M.M. Becker, R.A. Halpin, K. Li, E. Venter, X. Lu, S. Scherbakova,

R.L. Graham, R.S. Baric, T.B. Stockwell, D.J. Spiro, M.R. Denison, Infidelity of SARS-CoV Nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing, PLoS Pathog. 6 (2010) 1–15. https://doi.org/10.1371/journal.ppat.1000896.

[73] L.D. Eckerle, X. Lu, S.M. Sperry, L. Choi, M.R. Denison, High Fidelity of Murine Hepatitis Virus Replication Is Decreased in nsp14 Exoribonuclease Mutants, J. Virol. 81 (2007) 12135–12144. https://doi.org/10.1128/jvi.01296-07.

[74] E. Minskaia, T. Hertzig, A.E. Gorbalenya, V. Campanacci, C. Cambillau, B. Canard, J. Ziebuhr, Discovery of an RNA virus 3′→5′ exoribonuclease that is critically involved in coronavirus RNA synthesis, Proc. Natl. Acad. Sci. U. S. A. 103 (2006) 5108–5113. https://doi.org/10.1073/pnas.0508200103.

[75] Y. Chen, H. Cai, J. Pan, N. Xiang, P. Tien, T. Ahola, D. Guo, Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase, Proc. Natl. Acad. Sci. U. S. A. 106 (2009) 3484–3489. https://doi.org/10.1073/pnas.0808790106.

[76] K. Bhardwaj, J. Sun, A. Holzenburg, L.A. Guarino, C.C. Kao, RNA Recognition and Cleavage by the SARS Coronavirus Endoribonuclease, J. Mol. Biol. 361 (2006) 243–256. https://doi.org/10.1016/j.jmb.2006.06.021.

[77] K.A. Ivanov, T. Hertzig, M. Rozanov, S. Bayer, V. Thiel, A.E. Gorbalenya, J. Ziebuhr, Major genetic marker of nidoviruses encodes a replicative endoribonuclease, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 12694–12699. https://doi.org/10.1073/pnas.0403127101.

[78] E. Decroly, I. Imbert, B. Coutard, M. Bouvet, B. Selisko, K. Alvarez, A.E. Gorbalenya, E.J. Snijder, B. Canard, Coronavirus Nonstructural Protein 16 Is a Cap-0 Binding Enzyme Possessing (Nucleoside-2′ O)-Methyltransferase Activity, J. Virol. 82 (2008) 8071–8084. https://doi.org/10.1128/jvi.00407-08.

[79] R. Züst, L. Cervantes-Barragan, M. Habjan, R. Maier, B.W. Neuman, J. Ziebuhr, K.J. Szretter, S.C. Baker, W. Barchet, M.S. Diamond, S.G. Siddell, B. Ludewig, V. Thiel, Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5, Nat. Immunol. 12 (2011) 137–143. https://doi.org/10.1038/ni.1979.

[80] B. Coutard, C. Valle, X. de Lamballerie, B. Canard, N.G. Seidah, E. Decroly, The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade, Antiviral Res. 176 (2020) 104742. https://doi.org/10.1016/j.antiviral.2020.104742.

[81] D. Yesudhas, A. Srivastava, M.M. Gromiha, COVID-19 outbreak: history, mechanism, transmission, structural studies and therapeutics, Infection. 49 (2021) 199–213. https://doi.org/10.1007/s15010-020-01516-2.

[82] T.A. Santantonio, G. Messina, Update on Coronavirus Disease 2019 (COVID-19), Open Neurol. J. 14 (2020) 4–5. https://doi.org/10.2174/1874205x02014010004.

[83] M. Hoffmann, H. Kleine-Weber, S. Schroeder, N. Krüger, T. Herrler, S. Erichsen, T.S. Schiergens, G. Herrler, N.H. Wu, A. Nitsche, M.A. Müller, C. Drosten, S. Pöhlmann, SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor, Cell. 181 (2020) 271-280.e8. https://doi.org/10.1016/j.cell.2020.02.052.

[84] K. Sakai, Y. Ami, N. Nakajima, K. Nakajima, M. Kitazawa, M. Anraku, I. Takayama, N. Sangsriratanakul, M. Komura, Y. Sato, H. Asanuma, E. Takashita, K. Komase, K. Takehara, M. Tashiro, H. Hasegawa, T. Odagiri, M. Takeda, TMPRSS2 Independency for Haemagglutinin Cleavage in Vivo Differentiates

Influenza B Virus from Influenza A Virus, Sci. Rep. 6 (2016) 1–10. https://doi.org/10.1038/srep29430.
[85] L.W. Shen, H.J. Mao, Y.L. Wu, Y. Tanaka, W. Zhang, TMPRSS2: A potential target for treatment of influenza virus and coronavirus infections, Biochimie. 142 (2017) 1–10. https://doi.org/10.1016/j.biochi.2017.07.016.
[86] H. Limburg, A. Harbig, D. Bestle, D.A. Stein, H.M. Moulton, J. Jaeger, H. Janga, K. Hardes, J. Koepke, L. Schulte, A.R. Koczulla, B. Schmeck, H.-D. Klenk, E. Böttcher-Friebertshäuser, TMPRSS2 Is the Major Activating Protease of Influenza A Virus in Primary Human Airway Cells and Influenza B Virus in Human Type II Pneumocytes, J. Virol. 93 (2019). https://doi.org/10.1128/jvi.00649-19.
[87] M. Hoffmann, H. Kleine-Weber, S. Pöhlmann, A Multibasic Cleavage Site in the Spike Protein of SARS-CoV-2 Is Essential for Infection of Human Lung Cells, Mol. Cell. 78 (2020) 779-784.e5. https://doi.org/10.1016/j.molcel.2020.04.022.
[88] J. Shang, Y. Wan, C. Luo, G. Ye, Q. Geng, A. Auerbach, F. Li, Cell entry mechanisms of SARS-CoV-2, Proc. Natl. Acad. Sci. U. S. A. 117 (2020). https://doi.org/10.1073/pnas.2003138117.
[89] Y. Inoue, N. Tanaka, Y. Tanaka, S. Inoue, K. Morita, M. Zhuang, T. Hattori, K. Sugamura, Clathrin-Dependent Entry of Severe Acute Respiratory Syndrome Coronavirus into Target Cells Expressing ACE2 with the Cytoplasmic Tail Deleted, J. Virol. 81 (2007) 8722–8729. https://doi.org/10.1128/jvi.00253-07.
[90] A. Bayati, R. Kumar, V. Francis, P.S. McPherson, SARS-CoV-2 infects cells after viral entry via clathrin-mediated endocytosis, J. Biol. Chem. 296 (2021) 100306. https://doi.org/10.1016/j.jbc.2021.100306.
[91] I.C. Huang, B.J. Bosch, F. Li, W. Li, H.L. Kyoung, S. Ghiran, N. Vasilieva, T.S. Dermody, S.C. Harrison, P.R. Dormitzer, M. Farzan, P.J.M. Rottier, H. Choe, SARS coronavirus, but not human coronavirus NL63, utilizes cathepsin L to infect ACE2-expressing cells, J. Biol. Chem. 281 (2006) 3198–3203. https://doi.org/10.1074/jbc.M508381200.
[92] G. Simmons, D.N. Gosalia, A.J. Rennekamp, J.D. Reeves, S.L. Diamond, P. Bates, Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 11876–11881. https://doi.org/10.1073/pnas.0505577102.
[93] S. Matsuyama, N. Nagata, K. Shirato, M. Kawase, M. Takeda, F. Taguchi, Efficient Activation of the Severe Acute Respiratory Syndrome Coronavirus Spike Protein by the Transmembrane Protease TMPRSS2, J. Virol. 84 (2010) 12658–12664. https://doi.org/10.1128/jvi.01542-10.
[94] I. Glowacka, S. Bertram, M.A. Müller, P. Allen, E. Soilleux, S. Pfefferle, I. Steffen, T.S. Tsegaye, Y. He, K. Gnirss, D. Niemeyer, H. Schneider, C. Drosten, S. Pöhlmann, Evidence that TMPRSS2 Activates the Severe Acute Respiratory Syndrome Coronavirus Spike Protein for Membrane Fusion and Reduces Viral Control by the Humoral Immune Response, J. Virol. 85 (2011) 4122–4134. https://doi.org/10.1128/jvi.02232-10.
[95] A. Shulla, T. Heald-Sargent, G. Subramanya, J. Zhao, S. Perlman, T. Gallagher, A Transmembrane Serine Protease Is Linked to the Severe Acute Respiratory Syndrome Coronavirus Receptor and Activates Virus Entry, J. Virol. 85 (2011) 873–882. https://doi.org/10.1128/jvi.02062-10.
[96] C.B. Jackson, M. Farzan, B. Chen, H. Choe, Mechanisms of SARS-CoV-2 entry into cells, Nat. Rev. Mol. Cell Biol. 23 (2022) 3–20. https://doi.org/10.1038/s41580-021-00418-x.

[97] I. Nemet, L. Kliker, Y. Lustig, N. Zuckerman, O. Erster, C. Cohen, Y. Kreiss, S. Alroy-Preis, G. Regev-Yochay, E. Mendelson, M. Mandelboim, Third BNT162b2 Vaccination Neutralization of SARS-CoV-2 Omicron Infection, N. Engl. J. Med. 386 (2022) 492–494. https://doi.org/10.1056/nejmc2119358.
[98] L. Wang, G. Cheng, Sequence analysis of the emerging SARS-CoV-2 variant Omicron in South Africa, J. Med. Virol. 94 (2022) 1728–1733. https://doi.org/10.1002/jmv.27516.
[99] J. Chen, R. Wang, N.B. Gilby, G.W. Wei, Omicron Variant (B.1.1.529): Infectivity, Vaccine Breakthrough, and Antibody Resistance, J. Chem. Inf. Model. 62 (2022) 412–422. https://doi.org/10.1021/acs.jcim.1c01451.
[100] J. Fantini, N. Yahi, P. Colson, H. Chahinian, B. La Scola, D. Raoult, The puzzling mutational landscape of the SARS-2-variant Omicron, J. Med. Virol. 94 (2022) 2019–2025. https://doi.org/10.1002/jmv.27577.
[101] G. Tiecco, S. Storti, M.D. Antoni, E. Focà, F. Castelli, E. Quiros-Roldan, Omicron Genetic and Clinical Peculiarities That May Overturn SARS-CoV-2 Pandemic: A Literature Review, Int. J. Mol. Sci. 23 (2022). https://doi.org/10.3390/ijms23041987.
[102] D. Ni, K. Lau, P. Turelli, C. Raclot, B. Beckert, S. Nazarov, F. Pojer, A. Myasnikov, H. Stahlberg, D. Trono, Structural analysis of the Spike of the Omicron SARS-COV-2 variant by cryo-EM and implications for immune evasion, BioRxiv. (2021) 2021.12.27.474250. https://www.biorxiv.org/content/10.1101/2021.12.27.474250v1%0Ahttps://www.biorxiv.org/content/10.1101/2021.12.27.474250v1.abstract.
[103] M.S.A. Parvez, M.K. Saha, M. Ibrahim, Y. Araf, M.T. Islam, G. Ohtsuki, M.J. Hosen, Insights from a computational analysis of the SARS-CoV-2 Omicron variant: Host-pathogen interaction, pathogenicity and possible therapeutics, (2022). https://arxiv.org/abs/2201.08176.
[104] C.K. Wibmer, F. Ayres, T. Hermanus, M. Madzivhandila, P. Kgagudi, B. Oosthuysen, B.E. Lambson, T. de Oliveira, M. Vermeulen, K. van der Berg, T. Rossouw, M. Boswell, V. Ueckermann, S. Meiring, A. von Gottberg, C. Cohen, L. Morris, J.N. Bhiman, P.L. Moore, SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma, Nat. Med. 27 (2021) 622–625. https://doi.org/10.1038/s41591-021-01285-x.
[105] Q. Li, J. Wu, J. Nie, L. Zhang, H. Hao, S. Liu, C. Zhao, Q. Zhang, H. Liu, L. Nie, H. Qin, M. Wang, Q. Lu, X. Li, Q. Sun, J. Liu, L. Zhang, X. Li, W. Huang, Y. Wang, The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity, Cell. 182 (2020) 1284-1294.e9. https://doi.org/10.1016/j.cell.2020.07.012.
[106] T.N. Starr, A.J. Greaney, S.K. Hilton, D. Ellis, K.H.D. Crawford, A.S. Dingens, M.J. Navarro, J.E. Bowen, M.A. Tortorici, A.C. Walls, N.P. King, D. Veesler, J.D. Bloom, Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding, Cell. 182 (2020) 1295-1310.e20. https://doi.org/10.1016/j.cell.2020.08.012.
[107] A.J. Greaney, A.N. Loes, K.H.D. Crawford, T.N. Starr, K.D. Malone, H.Y. Chu, J.D. Bloom, Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies, Cell Host Microbe. 29 (2021) 463-476.e6. https://doi.org/10.1016/j.chom.2021.02.003.
[108] P. Wang, M.S. Nair, L. Liu, S. Iketani, Y. Luo, Y. Guo, M. Wang, J. Yu, B. Zhang,

P.D. Kwong, B.S. Graham, J.R. Mascola, J.Y. Chang, M.T. Yin, M. Sobieszczyk, C.A. Kyratsous, L. Shapiro, Z. Sheng, Y. Huang, D.D. Ho, Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7, Nature. 593 (2021) 130–135. https://doi.org/10.1038/s41586-021-03398-2.

[109] W. Yin, Y. Xu, P. Xu, X. Cao, C. Wu, C. Gu, X. He, X. Wang, S. Huang, Q. Yuan, K. Wu, W. Hu, Z. Huang, J. Liu, Z. Wang, F. Jia, K. Xia, P. Liu, X. Wang, B. Song, J. Zheng, H. Jiang, X. Cheng, Y. Jiang, S.-J. Deng, H.E. Xu, Structures of the Omicron spike trimer with ACE2 and an anti-Omicron antibody, Science (80-.). 375 (2022) 1048–1053. https://doi.org/10.1126/science.abn8863.

[110] D. Planas, N. Saunders, P. Maes, F. Guivel-Benhassine, C. Planchais, J. Buchrieser, W.H. Bolland, F. Porrot, I. Staropoli, F. Lemoine, H. Péré, D. Veyer, J. Puech, J. Rodary, G. Baele, S. Dellicour, J. Raymenants, S. Gorissen, C. Geenen, B. Vanmechelen, T. Wawina -Bokalanga, J. Martí-Carreras, L. Cuypers, A. Sève, L. Hocqueloux, T. Prazuck, F.A. Rey, E. Simon-Loriere, T. Bruel, H. Mouquet, E. André, O. Schwartz, Considerable escape of SARS-CoV-2 Omicron to antibody neutralization, Nature. 602 (2022) 671–675. https://doi.org/10.1038/s41586-021-04389-z.

[111] M. Mccallum, N. Czudnochowski, L.E. Rosen, S.K. Zepeda, J.E. Bowen, A.C. Walls, K. Hauser, A. Joshi, C. Stewart, J.R. Dillen, A.E. Powell, T.I. Croll, J. Nix, H.W. Virgin, D. Corti, Structural basis of SARS-CoV-2 Omicron immune evasion and receptor engagement, Science (80-.). 375 (2022) 864–868.

[112] Y. Cao, J. Wang, F. Jian, T. Xiao, W. Song, A. Yisimayi, W. Huang, Q. Li, P. Wang, R. An, J. Wang, Y. Wang, X. Niu, S. Yang, H. Liang, H. Sun, T. Li, Y. Yu, Q. Cui, S. Liu, X. Yang, S. Du, Z. Zhang, X. Hao, F. Shao, R. Jin, X. Wang, J. Xiao, Y. Wang, X.S. Xie, Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies, Nature. 602 (2022) 657–663. https://doi.org/10.1038/s41586-021-04385-3.

[113] Emma Hodcroft, Variant: 21L (Omicron), (n.d.). https://covariants.org/variants/21L.Omicron.

[114] D. of V.D. National Center for Immunization and Respiratory Diseases (NCIRD), Science Brief: Omicron (B.1.1.529) Variant, (n.d.). https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/scientific-brief-omicron-variant.html#:~%7B%7D:text=On December 1%2C 2021%2C the,the days preceding symptom onset (accessed April 23, 2022).

[115] X. He, W. Hong, X. Pan, G. Lu, X. Wei, SARS-CoV-2 Omicron variant: Characteristics and prevention, MedComm. 2 (2021) 838–845. https://doi.org/10.1002/mco2.110.

[116] C. Hamers-Casterman, T. Atarhouch, S. Muyldermans, G. Robinson, C. Hammers, E.B. Songa, N. Bendahman, R. Hammers, Naturally occurring antibodies devoid of light chains, Nature. 363 (1993) 446–448. https://doi.org/10.1038/363446a0.

[117] N.S. Lipman, L.R. Jackson, L.J. Trudel, F. Weis-Garcia, Monoclonal versus polyclonal antibodies: Distinguishing characteristics, applications, and information resources, ILAR J. 46 (2005) 258–267. https://doi.org/10.1093/ilar.46.3.258.

[118] T. Olafsen, A.M. Wu, Antibody Vectors for Imaging, Semin. Nucl. Med. 40 (2010) 167–181. https://doi.org/10.1053/j.semnuclmed.2009.12.005.

[119] I. Van Audenhove, J. Gettemans, Nanobodies as Versatile Tools to Understand, Diagnose, Visualize and Treat Cancer, EBioMedicine. 8 (2016) 40–48. https://doi.org/10.1016/j.ebiom.2016.04.028.

[120] M. Arbabi-Ghahroudi, Camelid single-domain antibodies: Historical perspective and future outlook, Front. Immunol. 8 (2017) 1–8. https://doi.org/10.3389/fimmu.2017.01589.
[121] S.M. Ghahroudi, Arbabi, A. Desmyter, L. Wyns, R. Hamers, Selection and identification of single domain antibody fragments from camel, FEBS Lett. 414 (1997) 521–526. https://doi.org/10.1016/S0014-5793(97)01062-4.
[122] M.R.J. Clokie, A.M. Kropinski, Bacteriophages Methods and Protocols Volume 2: Molecular and Applied Aspects, Methods Mol. Biol. 502 (2009) xxii, 373 p. https://doi.org/10.1007/978-1-60327-565-1.
[123] S. Muyldermans, T.N. Baral, V.C. Retamozzo, P. De Baetselier, E. De Genst, J. Kinne, H. Leonhardt, S. Magez, V.K. Nguyen, H. Revets, U. Rothbauer, B. Stijlemans, S. Tillib, U. Wernery, L. Wyns, G. Hassanzadeh-Ghassabeh, D. Saerens, Camelid immunoglobulins and nanobody technology, Vet. Immunol. Immunopathol. 128 (2009) 178–183. https://doi.org/10.1016/j.vetimm.2008.10.299.
[124] S. Muyldermans, Nanobodies: Natural single-domain antibodies, Annu. Rev. Biochem. 82 (2013) 775–797. https://doi.org/10.1146/annurev-biochem-063011-092449.
[125] E. Pardon, T. Laeremans, S. Triest, S.G.F. Rasmussen, A. Wohlkönig, A. Ruf, S. Muyldermans, W.G.J. Hol, B.K. Kobilka, J. Steyaert, A general protocol for the generation of Nanobodies for structural biology, Nat. Protoc. 9 (2014) 674–693. https://doi.org/10.1038/nprot.2014.039.
[126] N. Pant, A. Hultberg, Y. Zhao, L. Svensson, Q. Pan-Hammarström, K. Johansen, P.H. Pouwels, F.M. Ruggeri, P. Hermans, L. Frenken, T. Borén, H. Marcotte, L. Hammarström, Lactobacilli expressing variable domain of llama heavy-chain antibody fragments (lactobodies) confer protection against rotavirus-induced diarrhea, J. Infect. Dis. 194 (2006) 1580–1588. https://doi.org/10.1086/508747.
[127] G. Hassanzadeh-Ghassabeh, N. Devoogdt, P. De Pauw, C. Vincke, S. Muyldermans, Nanobodies and their potential applications, Nanomedicine. 8 (2013) 1013–1026. https://doi.org/10.2217/nnm.13.86.
[128] M. Lauwereys, M.A. Ghahroudi, A. Desmyter, J. Kinne, W. Hölzer, E. De Genst, L. Wyns, S. Muyldermans, Potent enzyme inhibitors derived from dromedary heavy-chain antibodies, EMBO J. 17 (1998) 3512–3520. https://doi.org/10.1093/emboj/17.13.3512.
[129] J.M.J. Pérez, J.G. Renisio, J.J. Prompers, C.J. Van Platerink, C. Cambillau, H. Darbon, L.G.J. Frenken, Thermal unfolding of a llama antibody fragment: A two-state reversible process, Biochemistry. 40 (2001) 74–83. https://doi.org/10.1021/bi0009082.
[130] M. Dumoulin, K. Conrath, A. Van Meirhaeghe, F. Meersman, K. Heremans, L.G.J. Frenken, S. Muyldermans, L. Wyns, A. Matagne, Single-domain antibody fragments with high conformational stability, Protein Sci. 11 (2009) 500–515. https://doi.org/10.1110/ps.34602.
[131] E. Dolk, M. Van Der Vaart, D.L. Hulsik, G. Vriend, H. De Haard, S. Spinelli, C. Cambillau, L. Frenken, T. Verrips, Isolation of llama antibody fragments for prevention of dandruff by phage display in shampoo, Appl. Environ. Microbiol. 71 (2005) 442–450. https://doi.org/10.1128/AEM.71.1.442-450.2005.
[132] G. Hussack, T. Hirama, W. Ding, R. MacKenzie, J. Tanha, Engineered single-domain antibodies with high protease resistance and thermal stability, PLoS One. 6 (2011). https://doi.org/10.1371/journal.pone.0028218.

[133] M.M. Harmsen, C.B. Van Solt, H.P.D. Fijten, M.C. Van Setten, Prolonged in vivo residence times of llama single-domain antibody fragments in pigs by binding to porcine immunoglobulins, Vaccine. 23 (2005) 4926–4934. https://doi.org/10.1016/j.vaccine.2005.05.017.
[134] K.E. Conrath, M. Lauwereys, M. Galleni, A. Matagne, J.M. Frère, J. Kinne, L. Wyns, S. Muyldermans, β-Lactamase inhibitors derived from single-domain antibody fragments elicited in the Camelidae, Antimicrob. Agents Chemother. 45 (2001) 2807–2812. https://doi.org/10.1128/AAC.45.10.2807-2812.2001.
[135] B. Stijlemans, K. Conrath, V. Cortez-Retamozo, H. Van Xong, L. Wyns, P. Senter, H. Revets, P. De Baetselier, S. Muyldermans, S. Magez, Efficient targeting of conserved cryptic epitopes of infectious agents by single domain antibodies: African trypanosomes as paradigm, J. Biol. Chem. 279 (2004) 1256–1261. https://doi.org/10.1074/jbc.M307341200.
[136] S. V. Tillib, T.I. Ivanova, L.A. Vasilev, M. V. Rutovskaya, S.A. Saakyan, I.Y. Gribova, I.L. Tutykhina, E.S. Sedova, A.A. Lysenko, M.M. Shmarov, D.Y. Logunov, B.S. Naroditsky, A.L. Gintsburg, Formatted single-domain antibodies can protect mice against infection with influenza virus (H5N2), Antiviral Res. 97 (2013) 245–254. https://doi.org/10.1016/j.antiviral.2012.12.014.
[137] H. Adams, W.M. Horrevoets, S.M. Adema, H.E.V. Carr, R.E. van Woerden, M. Koster, J. Tommassen, Inhibition of biofilm formation by Camelid single-domain antibodies against the flagellum of Pseudomonas aeruginosa, J. Biotechnol. 186 (2014) 66–73. https://doi.org/10.1016/j.jbiotec.2014.06.029.
[138] F.M. Cardoso, L.I. Ibañez, S. Van den Hoecke, S. De Baets, A. Smet, K. Roose, B. Schepens, F.J. Descamps, W. Fiers, S. Muyldermans, A. Depicker, X. Saelens, Single-Domain Antibodies Targeting Neuraminidase Protect against an H5N1 Influenza Virus Challenge, J. Virol. 88 (2014) 8278–8296. https://doi.org/10.1128/jvi.03178-13.
[139] L.O. Tchouate Gainkam, V. Caveliers, N. Devoogdt, C. Vanhove, C. Xavier, O. Boerman, S. Muyldermans, A. Bossuyt, T. Lahoutte, Localization, mechanism and reduction of renal retention of technetium-99m labeled epidermal growth factor receptor-specific nanobody in mice, Contrast Media Mol. Imaging. 6 (2011) 85–92. https://doi.org/10.1002/cmmi.408.
[140] M. Kijanka, F.J. Warnders, M. El Khattabi, M. Lub-De Hooge, G.M. Van Dam, V. Ntziachristos, L. De Vries, S. Oliveira, P.M.P. Van Bergen En Henegouwen, Rapid optical imaging of human breast tumour xenografts using anti-HER2 VHHs site-directly conjugated to IRDye 800CW for image-guided surgery, Eur. J. Nucl. Med. Mol. Imaging. 40 (2013) 1718–1729. https://doi.org/10.1007/s00259-013-2471-2.
[141] U. Rothbauer, K. Zolghadr, S. Tillib, D. Nowak, L. Schermelleh, A. Gahl, N. Backmann, K. Conrath, S. Muyldermans, M.C. Cardoso, H. Leonhardt, Targeting and tracing antigens in live cells with fluorescent nanobodies, Nat. Methods. 3 (2006) 887–889. https://doi.org/10.1038/nmeth953.
[142] Y. Huang, S. Nokhrin, G. Hassanzadeh-Ghassabeh, C.H. Yu, H. Yang, A.N. Barry, M. Tonelli, J.L. Markley, S. Muyldermans, O.Y. Dmitriev, S. Lutsenko, Interactions between metal-binding domains modulate intracellular targeting of Cu(I)-ATPase ATP7B, as revealed by nanobody binding, J. Biol. Chem. 289 (2014) 32682–32693. https://doi.org/10.1074/jbc.M114.580845.
[143] T.J. Esparza, N.P. Martin, G.P. Anderson, E.R. Goldman, D.L. Brody, High affinity nanobodies block SARS-CoV-2 spike receptor binding domain interaction with human angiotensin converting enzyme, Sci. Rep. 10 (2020) 1–13.

https://doi.org/10.1038/s41598-020-79036-0.
[144] L. Hanke, L. Vidakovics Perez, D.J. Sheward, H. Das, T. Schulte, A. Moliner-Morro, M. Corcoran, A. Achour, G.B. Karlsson Hedestam, B.M. Hällberg, B. Murrell, G.M. McInerney, An alpaca nanobody neutralizes SARS-CoV-2 by blocking receptor interaction, Nat. Commun. 11 (2020) 1–9. https://doi.org/10.1038/s41467-020-18174-5.
[145] J. Huo, A. Le Bas, R.R. Ruza, H.M.E. Duyvesteyn, H. Mikolajek, T. Malinauskas, T.K. Tan, P. Rijal, M. Dumoux, P.N. Ward, J. Ren, D. Zhou, P.J. Harrison, M. Weckener, D.K. Clare, V.K. Vogirala, J. Radecke, L. Moynié, Y. Zhao, J. Gilbert-Jaramillo, M.L. Knight, J.A. Tree, K.R. Buttigieg, N. Coombes, M.J. Elmore, M.W. Carroll, L. Carrique, P.N.M. Shah, W. James, A.R. Townsend, D.I. Stuart, R.J. Owens, J.H. Naismith, Neutralizing nanobodies bind SARS-CoV-2 spike RBD and block interaction with ACE2, Nat. Struct. Mol. Biol. 27 (2020) 846–854. https://doi.org/10.1038/s41594-020-0469-6.
[146] A. Sircar, E.T. Kim, J.J. Gray, RosettaAntibody: Antibody variable region homology modeling server, Nucleic Acids Res. 37 (2009) 474–479. https://doi.org/10.1093/nar/gkp387.
[147] Y. Yan, D. Zhang, P. Zhou, B. Li, S.Y. Huang, HDOCK: A web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy, Nucleic Acids Res. 45 (2017) W365–W373. https://doi.org/10.1093/nar/gkx407.
[148] Y. Yan, H. Tao, J. He, S.Y. Huang, The HDOCK server for integrated protein–protein docking, Nat. Protoc. 15 (2020) 1829–1852. https://doi.org/10.1038/s41596-020-0312-x.
[149] V. Kunik, S. Ashkenazi, Y. Ofran, Paratome: An online tool for systematic identification of antigen-binding regions in antibodies based on sequence or structure, Nucleic Acids Res. 40 (2012) 521–524. https://doi.org/10.1093/nar/gks480.
[150] V. Kunik, B. Peters, Y. Ofran, Structural consensus among antibodies defines the antigen binding site, PLoS Comput. Biol. 8 (2012). https://doi.org/10.1371/journal.pcbi.1002388.
[151] F. Fogolari, A. Corazza, V. Yarra, A. Jalaru, P. Viglino, G. Esposito, Bluues: A program for the analysis of the electrostatic properties of proteins based on generalized Born radii, BMC Bioinformatics. 13 (2012). https://doi.org/10.1186/1471-2105-13-S4-S18.
[152] L.C. Xue, J.P. Rodrigues, P.L. Kastritis, A.M. Bonvin, A. Vangone, PRODIGY: A web server for predicting the binding affinity of protein-protein complexes, Bioinformatics. 32 (2016) 3676–3678. https://doi.org/10.1093/bioinformatics/btw514.
[153] R.A. Laskowski, M.B. Swindells, LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery, J. Chem. Inf. Model. 51 (2011) 2778–2786. https://doi.org/10.1021/ci200227u.
[154] L. Schrodinger, The PyMOL Molecular Graphics System, Version 2.0, Schrödinger, LLC., (n.d.).
[155] M.I.J. Raybould, A. Kovaltsuk, C. Marks, C.M. Deane, CoV-AbDab: The coronavirus antibody database, Bioinformatics. 37 (2021) 734–735. https://doi.org/10.1093/bioinformatics/btaa739.
[156] Y. Xiang, S. Nambulli, Z. Xiao, H. Liu, Z. Sang, W.P. Duprex, D. Schneidman-Duhovny, C. Zhang, Y. Shi, Versatile and multivalent nanobodies efficiently neutralize SARS-CoV-2, Science (80-.). 370 (2020) 1479–1484.

https://doi.org/10.1126/science.abe4747.
[157] P. Pymm, A. Adair, L.J. Chan, J.P. Cooney, F.L. Mordant, C.C. Allison, E. Lopez, E.R. Haycroft, M.T. O'Neill, L.L. Tan, M.H. Dietrich, D. Drew, M. Doerflinger, M.A. Dengler, N.E. Scott, A.K. Wheatley, N.A. Gherardin, H. Venugopal, D. Cromer, M.P. Davenport, R. Pickering, D.I. Godfrey, D.F.J. Purcell, S.J. Kent, A.W. Chung, K. Subbarao, M. Pellegrini, A. Glukhova, W.H. Tham, Nanobody cocktails potently neutralize SARS-CoV-2 D614G N501Y variant and protect mice, Proc. Natl. Acad. Sci. U. S. A. 118 (2021) 1–12. https://doi.org/10.1073/pnas.2101918118.
[158] D.L. Tingting Li, Hongmin Cai, Hebang Yao, Bingjie Zhou, Yapei Zhao, Wenming Qin, Cedric A.J. Hutter, Yanling Lai, Juan Bao, Jiaming Lan, Gary Wong, Markus Seeger, Dimitri Lavillette, Potent synthetic nanobodies against SARS-CoV-2 and molecular basis for neutralization, (2020) 1–18. https://doi.org/10.1101/2020.06.09.143438.
[159] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, Nucleic Acids Res. 28 (2000) 235–242. http://www.rcsb.org/pdb/status.html (accessed April 26, 2022).
[160] K. Tunyasuvunakool, J. Adler, Z. Wu, T. Green, M. Zielinski, A. Žídek, A. Bridgland, A. Cowie, C. Meyer, A. Laydon, S. Velankar, G.J. Kleywegt, A. Bateman, R. Evans, A. Pritzel, M. Figurnov, O. Ronneberger, R. Bates, S.A.A. Kohl, A. Potapenko, A.J. Ballard, B. Romera-Paredes, S. Nikolov, R. Jain, E. Clancy, D. Reiman, S. Petersen, A.W. Senior, K. Kavukcuoglu, E. Birney, P. Kohli, J. Jumper, D. Hassabis, Highly accurate protein structure prediction for the human proteome, Nature. 596 (2021) 590–596. https://doi.org/10.1038/s41586-021-03828-1.
[161] A.C. Walls, Y.J. Park, M.A. Tortorici, A. Wall, A.T. McGuire, D. Veesler, Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein, Cell. 181 (2020) 281-292.e6. https://doi.org/10.1016/j.cell.2020.02.058.
[162] F. Sievers, A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J.D. Thompson, D.G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, Mol. Syst. Biol. 7 (2011). https://doi.org/10.1038/msb.2011.75.
[163] F. Madeira, M. Pearce, A.R.N. Tivey, P. Basutkar, J. Lee, O. Edbali, N. Madhusoodanan, A. Kolesnikov, R. Lopez, Search and sequence analysis tools services from EMBL-EBI in 2022, Nucleic Acids Res. (2022) 1–4. https://doi.org/10.1093/nar/gkac240.
[164] CHARMM-GUI Archive - COVID-19 Proteins Library, (2022). https://www.charmm-gui.org/?doc=archive&lib=covid19 (accessed June 1, 2022).
[165] Convert PDB structure to FASTA sequence, (2022). https://zhanggroup.org/pdb2fasta/ (accessed June 1, 2022).
[166] T.J. Yang, P.Y. Yu, Y.C. Chang, S. Te Danny Hsu, D614G mutation in the SARS-CoV-2 spike protein enhances viral fitness by desensitizing it to temperature-dependent denaturation, J. Biol. Chem. 297 (2021) 101238. https://doi.org/10.1016/j.jbc.2021.101238.
[167] C.B. Jackson, L. Zhang, M. Farzan, H. Choe, Functional importance of the D614G mutation in the SARS-CoV-2 spike protein, Biochem. Biophys. Res. Commun. 538 (2021) 108–115. https://doi.org/10.1016/j.bbrc.2020.11.026.
[168] J.A. Plante, Y. Liu, J. Liu, H. Xia, B.A. Johnson, K.G. Lokugamage, X. Zhang, A.E. Muruato, J. Zou, C.R. Fontes-Garfias, D. Mirchandani, D. Scharton, J.P.

Bilello, Z. Ku, Z. An, B. Kalveram, A.N. Freiberg, V.D. Menachery, X. Xie, K.S. Plante, S.C. Weaver, P.Y. Shi, Spike mutation D614G alters SARS-CoV-2 fitness, Nature. 592 (2021) 116–121. https://doi.org/10.1038/s41586-020-2895-3.

[169] B.A. Johnson, X. Xie, B. Kalveram, K.G. Lokugamage, A. Muruato, J. Zou, X. Zhang, T. Juelich, J.K. Smith, L. Zhang, N. Bopp, C. Schindewolf, M. Vu, A. Vanderheiden, D. Swetnam, J.A. Plante, P. Aguilar, K.S. Plante, B. Lee, S.C. Weaver, M.S. Suthar, A.L. Routh, P. Ren, Z. Ku, Z. An, K. Debbink, P.Y. Shi, A.N. Freiberg, V.D. Menachery, Furin Cleavage Site Is Key to SARS-CoV-2 Pathogenesis., BioRxiv Prepr. Serv. Biol. (2020). https://doi.org/10.1101/2020.08.26.268854.

[170] M. Örd, I. Faustova, M. Loog, The sequence at Spike S1/S2 site enables cleavage by furin and phospho-regulation in SARS-CoV2 but not in SARS-CoV1 or MERS-CoV, Sci. Rep. 10 (2020) 1–10. https://doi.org/10.1038/s41598-020-74101-0.

[171] T.R. Wagner, E. Ostertag, P.D. Kaiser, M. Gramlich, N. Ruetalo, D. Junker, J. Haering, B. Traenkle, M. Becker, A. Dulovic, H. Schweizer, S. Nueske, A. Scholz, A. Zeck, K. Schenke-Layland, A. Nelde, M. Strengert, J.S. Walz, G. Zocher, T. Stehle, M. Schindler, N. Schneiderhan-Marra, U. Rothbauer, NeutrobodyPlex—monitoring SARS-CoV-2 neutralizing immune responses using nanobodies, EMBO Rep. 22 (2021) 1–16. https://doi.org/10.15252/embr.202052325.

[172] X. Chi, X. Liu, C. Wang, X. Zhang, X. Li, J. Hou, L. Ren, Q. Jin, J. Wang, W. Yang, Humanized single domain antibodies neutralize SARS-CoV-2 by targeting the spike receptor binding domain, Nat. Commun. 11 (2020) 2–8. https://doi.org/10.1038/s41467-020-18387-8.

[173] B.A. Johnson, X. Xie, A.L. Bailey, B. Kalveram, K.G. Lokugamage, A. Muruato, J. Zou, X. Zhang, T. Juelich, J.K. Smith, L. Zhang, N. Bopp, C. Schindewolf, M. Vu, A. Vanderheiden, E.S. Winkler, D. Swetnam, J.A. Plante, P. Aguilar, K.S. Plante, V. Popov, B. Lee, S.C. Weaver, M.S. Suthar, A.L. Routh, P. Ren, Z. Ku, Z. An, K. Debbink, M.S. Diamond, P.Y. Shi, A.N. Freiberg, V.D. Menachery, Loss of furin cleavage site attenuates SARS-CoV-2 pathogenesis, Nature. 591 (2021) 293–299. https://doi.org/10.1038/s41586-021-03237-4.

[174] P.A. Koenig, H. Das, H. Liu, B.M. Kümmerer, F.N. Gohr, L.M. Jenster, L.D.J. Schiffelers, Y.M. Tesfamariam, M. Uchima, J.D. Wuerth, K. Gatterdam, N. Ruetalo, M.H. Christensen, C.I. Fandrey, S. Normann, J.M.P. Tödtmann, S. Pritzl, L. Hanke, J. Boos, M. Yuan, X. Zhu, J.L. Schmid-Burgk, H. Kato, M. Schindler, I.A. Wilson, M. Geyer, K.U. Ludwig, B.M. Hällberg, N.C. Wu, F.I. Schmidt, Structure-guided multivalent nanobodies block SARS-CoV-2 infection and suppress mutational escape, Science (80-.). 371 (2021). https://doi.org/10.1126/science.abe6230.

[175] F.D. Mast, P.C. Fridy, N.E. Ketaren, J. Wang, E.Y. Jacobs, J.P. Olivier, T. Sanyal, K.R. Molloy, F. Schmidt, M. Rutkowska, Y. Weisblum, L.M. Rich, E.R. Vanderwall, N. Dambrauskas, V. Vigdorovich, S. Keegan, J.B. Jiler, M.E. Stein, P.D.B. Olinares, L. Herlands, T. Hatziioannou, D.N. Sather, J.S. Debley, D. Fenyö, A. Sali, P.D. Bieniasz, J.D. Aitchison, B.T. Chait, M.P. Rout, Highly synergistic combinations of nanobodies that target SARS-CoV-2 and are resistant to escape, Elife. 10 (2021) 1–56. https://doi.org/10.7554/eLife.73027.

[176] A. Gray, A.R. M Bradbury, A. Knappik, A. Plückthun, C.A. K Borrebaeck, S. Dübel, Animal-free alternatives and the antibody iceberg, Nat. Biotechnol. 38 (2020) 1234–1241. https://doi.org/10.1038/s41587-020-0687-9.

[177] A. de Marco, Recombinant expression of nanobodies and nanobody-derived

immunoreagents, Protein Expr. Purif. 172 (2020) 105645. https://doi.org/10.1016/j.pep.2020.105645.
[178] W. Yin, Y. Xu, P. Xu, X. Cao, C. Wu, C. Gu, X. He, X. Wang, S. Huang, Q. Yuan, K. Wu, W. Hu, Z. Huang, J. Liu, Z. Wang, F. Jia, K. Xia, P. Liu, X. Wang, B. Song, J. Zheng, H. Jiang, X. Cheng, Y. Jiang, S.J. Deng, H.E. Xu, Structures of the Omicron spike trimer with ACE2 and an anti-Omicron antibody, Science (80-.). 375 (2022) 1048–1053. https://doi.org/10.1126/science.abn8863.
[179] L. Casalino, Z. Gaieb, J.A. Goldsmith, C.K. Hjorth, A.C. Dommer, A.M. Harbison, C.A. Fogarty, E.P. Barros, B.C. Taylor, J.S. Mclellan, E. Fadda, R.E. Amaro, Beyond shielding: The roles of glycans in the SARS-CoV-2 spike protein, ACS Cent. Sci. 6 (2020) 1722–1734. https://doi.org/10.1021/acscentsci.0c01056.
[180] M. Gur, E. Taka, S.Z. Yilmaz, C. Kilinc, U. Aktas, M. Golcuk, Conformational transition of SARS-CoV-2 spike glycoprotein between its closed and open states, J. Chem. Phys. 153 (2020). https://doi.org/10.1063/5.0011141.
[181] T. Güttler, M. Aksu, A. Dickmanns, K.M. Stegmann, K. Gregor, R. Rees, W. Taxer, O. Rymarenko, J. Schünemann, C. Dienemann, P. Gunkel, B. Mussil, J. Krull, U. Teichmann, U. Groß, V.C. Cordes, M. Dobbelstein, D. Görlich, Neutralization of SARS-CoV-2 by highly potent, hyperthermostable, and mutation-tolerant nanobodies, EMBO J. 40 (2021) 1–26. https://doi.org/10.15252/embj.2021107985.
[182] J. Zhang, Y. Cai, C.L. Lavine, H. Peng, H. Zhu, K. Anand, P. Tong, A. Gautam, M.L. Mayer, S. Rits-Volloch, S. Wang, P. Sliz, D.R. Wesemann, W. Yang, M.S. Seaman, J. Lu, T. Xiao, B. Chen, Structural and functional impact by SARS-CoV-2 Omicron spike mutations, 2022.
[183] F. Labroussaa, J. Jores, SARS-CoV-2 nanobodies 2.0, Signal Transduct. Target. Ther. 6 (2021) 2–4. https://doi.org/10.1038/s41392-021-00632-1.
[184] Y. Watanabe, J.D. Allen, D. Wrapp, J.S. McLellan, M. Crispin, Site-specific glycan analysis of the SARS-CoV-2 spike, Science (80-.). 369 (2020) 330–333. https://doi.org/10.1126/science.abb9983.
[185] A. Shajahan, N.T. Supekar, A.S. Gleinich, P. Azadi, Deducing the N- And O-glycosylation profile of the spike protein of novel coronavirus SARS-CoV-2, Glycobiology. 30 (2020) 981–988. https://doi.org/10.1093/glycob/cwaa042.
[186] D.S. Roberts, M. Mann, B.H. Li, D. Kim, A.R. Brasier, S. Jin, Y. Ge, Distinct Core Glycan and O-Glycoform Utilization of SARS-CoV-2 Omicron Variant Spike Protein RBD Revealed by Top-Down Mass Spectrometry, BioRxiv. (2022) 2022.02.09.479776. http://biorxiv.org/content/early/2022/02/10/2022.02.09.479776.abstract.
[187] Woods Group. (2005-2022) GLYCAM Web. Complex Carbohydrate Research Center, University of Georgia, Athens, GA., (n.d.). https://glycam.org/ (accessed June 1, 2022).
[188] T. Lemmin, C. Soto, Glycosylator: A Python framework for the rapid modeling of glycans, BMC Bioinformatics. 20 (2019) 1–7. https://doi.org/10.1186/s12859-019-3097-6.

CURRICULUM VITAE

2014 – 2019	B.Sc., Molecular Biology and Genetics, Gebze Technical University, Kocaeli, TURKEY
2020 – Present	Research Assitant, Molecular Biology and Genetics, Abdullah Gül University, Kayseri, TURKEY