T.R.

GEBZE TECHNICAL UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

OPTIMIZATION OF THE PRODUCTION PROCESS AND PURIFICATION OF SARS-COV-2 S1 PROTEIN USING SUMO-FUSION TAG AND DSBC LEADER PEPTIDE

MELİKE YAĞMUR ÜNAL

A THESIS SUBMITTED FOR THE DEGREE OF

MASTER OF SCIENCE

DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS

GEBZE

2022

T.R.

GEBZE TECHNICAL UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

OPTIMIZATION OF THE PRODUCTION PROCESS AND PURIFICATION OF SARS-COV-2 S1 PROTEIN USING SUMO-FUSION TAG AND DSBC LEADER PEPTIDE

MELİKE YAĞMUR ÜNAL

A THESIS SUBMITTED FOR THE DEGREE OF

MASTER OF SCIENCE

DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS

THESIS SUPERVISOR

PROF. DR. TAMER YAĞCI

II. THESIS SUPERVISOR

DR. HASAN ÜMİT ÖZTÜRK

GEBZE

2022

GEBZE TEKNİK ÜNİVERSİTESİ	**YÜKSEK LİSANS JÜRİ ONAY FORMU**

GTÜ Frn Bilimleri Enstitüsü Yönetim Kurulu'nun 07/07/2022 tarih ve 2022/34 sayılı kararıyla oluşturulan jüri tarafından 05/08/2022 tarihinde tez savunma sınavı yapılan Melike Yağmur ÜNAL'ın tez çalışması Moleküler Biyoloji ve Genetik Anabilim Dalında YÜKSEK LİSANS tezi olarak kabul edilmiştir.

JÜRİ

ÜYE (TEZ DANIŞMANI)	: Prof. Dr. Tamer YAĞCI
ÜYE	: Prof. Dr. Ayten KANDİLCİ
ÜYE	: Dr. Öğr. Üyesi Kübra AÇIKALIN COŞKUN

ONAY

Gebze Teknik Üniversitesi Fen Bilimleri Enstitüsü Yönetim Kurulu'nun ...…/…../…… tarih ve ……/…..... sayılı kararı.

İMZA/MÜHÜR

SUMMARY

SARS-CoV-2 is the newest member of the *Coronaviridae* family that causes the Covid-19 disease, which has become a pandemic. Although Covid-19 causes deaths worldwide, it has been determined that it has left permanent damage over the years in people who have had this disease. These reasons show the importance of both Covid-19 treatment and vaccine studies. Within the scope of this thesis, it is aimed to produce recombinantly soluble S1 protein, which is the subunit of SARS-CoV-2 Spike protein, in *Escherichia coli* expression systems to be used in the aforementioned studies.

A problem encountered in the production of recombinant proteins in *E. coli* expression systems is the production of proteins in soluble form. In this study, SUMO fusion tag and DsbC signal peptide were used to produce the SARS-CoV-2 S1 protein in a soluble form.

To examine the effect of the SUMO fusion tag on the production of recombinant soluble protein, the S1 gene was first cloned into the pET SUMO vector. Plasmids containing the S1 gene in the correct orientation were selected by sequencing and transformed into *Escherichia coli* SHuffle, BL21 (DE3) and BL21 Star (DE3) pLysS strains for expression control. The effects of temperature, IPTG concentration and time on the expression of S1 protein were examined. As a result of the experiments, it was determined that S1 protein was produced in a partially soluble form in *E.coli* SHuffle cells after induction of 0.4 mM IPTG after 24 hours of incubation at 16 ℃ with 140 rpm shaking.

In order to examine the effect of DsbC signal peptide on soluble protein production, S1 gene cloned into pET30a (+) vector and expression was performed in *E.coli* SHuffle cells. At the end of different temperatures, IPTG concentrations and time intervals, it was observed that the S1 gene cloned in the pET30a (+) vector was not expressed soluble in SHuffle cells.

Key Words: Recombinant protein expression, Spike protein, SUMO fusion tag, DsbC signal peptide.

ÖZET

SARS-CoV-2 pandemi haline gelmiş olan Covid-19 hastalığına sebep olan *Coronaviridae* ailesinin en yeni üyesidir. Covid-19 dünya genelinde ölümlere sebep olmakla birlikte bu hastalığı geçirmiş olan kişilerde yıllar içerisinde kalıcı hasarlar bıraktığı belirlenmiştir. Bu sebepler hem Covid-19 tedavisi hem de aşı çalışmalarının önemini göstermektedir. Bu tez kapsamında, bahsedilen çalışmalarda da kullanılmak üzere SARS-CoV-2 Spike proteininin alt birimi olan S1 proteinin, rekombinant olarak *Escherichia coli* ekspresyon sistemlerinde çözünür olarak üretiminin yapılması amaçlanmıştır.

E. coli ekspresyon sistemlerinde rekombinant protein üretiminde karşılaşılan bir problem proteinlerin çözünür olarak üretilmesidir. Bu çalışmada, SARS-CoV-2 S1 proteininin çözünür olarak üretilebilmesi amacıyla SUMO füzyon etiketi ve DsbC sinyal peptidi kullanılmıştır.

SUMO füzyon etiketinin, rekombinant çözünür protein üretimine etkisinin bakılması için ilk olarak S1 geni pET SUMO vektörüne klonlanmıştır. Sekanslama ile S1 genini doğru oryantasyonda içeren plazmitler seçilmiş ve ekspresyon kontrolü için *Escherichia coli* SHuffle, BL21 (DE3) ve BL21 Star(DE3) pLysS suşlarına transforme edilmiştir. Sıcaklık, IPTG konsantrasyonu ve zamanın, S1 proteininin ekspresyonu üzerine etkilerine bakılmıştır. Yapılan deneyler sonucunda *E.coli* SHuffle hücrelerinde, 0.4 mM IPTG indüksiyonundan sonra 16 ℃ sıcaklıkta, 140 rpm çalkalama ile 24 saatlik inkübasyonla S1 proteininin kısmi olarak çözünür halde üretildiği belirlenmiştir.

DsbC sinyal peptidinin, rekombinant çözünür protein üretiminin etkisine bakılması için pET30a (+) vektörüne klonlanmış olan S1 geninin, *E.coli* SHuffle hücrelerinde ekspresyon kontrolü yapılmıştır. Farklı sıcaklık, IPTG konsantrasyonu ve zaman denemelerinin sonunda pET30a (+) vektörüne klonlanmış S1 geninin SHuffle hücrelerinde çözünür olarak üretilmediği tespit edilmiştir.

Anahtar Kelimeler: Rekombinant protein ekspresyon, Spike Protein, SUMO Füzyon Etiketi, DsbC sinyal peptidi.

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to my advisor, Prof. Dr. Tamer YAĞCI, for his assistance, understanding, kindness, and patience throughout my education. I would like to thank my co-advisor, Dr. Hasan Ümit ÖZTÜRK, for his encouragement, guidance, and the experience he shared with me throughout tis study. I would like to express my gratitude to Assist. Prof. Dr. Kübra AÇIKALIN COŞKUN, who has always been kind to me and eager to teach eveything what she knows. I would like to thank TUBITAK for supporting my work. I would like to thank everyone who helped me at the TUBITAK, Instution of Genetic Engineering, Immunogenetic Laboratory. In private, I also thank my lab friends MSc. Esin BAYRALI ÜLKER and Tuğba DURAN. They were always pleasant to me and never hesitated to share their knowledge with me.

I would like to thank my fiance, Mustafa GEDİK, for all the help, interest, patience, and love he has shown me at every stage of my work. I'd like to thank my mother Melahat ÜNAL and my father İsmail Hakkı ÜNAL, who have always been there for me and given me strength.

TABLE of CONTENTS

	Page
SUMMARY	iv
ÖZET	v
ACKNOWLEDMENTS	vi
TABLE of CONTENTS	vii
LIST of ABBREVIATIONS and ACRONYMS	viii
LIST of FIGURES	ix
LIST of TABLES	x
1. INTRODUCTION	1
1.1. Aim of the Thesis	1
2. LITERATURE REVIEW	2
2.1. Infection Mechanism of SARS-CoV-2	2
2.2. Recombinant Protein Production	2
2.3. Protein Modification by SUMO	4
2.4. The Role of DsbC in Protein Folding	5
3. MATERIALS & METHODS	7
3.1. Materials	7
3.1.1. Microorganisms	7
3.1.1.1. *E.coli* Mach1Strain	8
3.1.1.2. *E.coli* BL21(DE3) and BL21 Star(DE3) pLysS Strain	8
3.1.1.3. *E.coli* Shuffle Strain	8
3.1.2. Primers	9
3.1.3. Equipment	9
3.1.4. General Kits and Reagents	10
3.1.5. Buffers and Solutions	12
3.1.5.1. Luria-Bertani (LB) Broth Medium: (for 1L)	12
3.1.5.2. Luria-Bertani (LB) Agar Medium: (for 1L)	12
3.1.5.3. 2xYT Broth: (for 1L)	12
3.1.5.4. 2xYT Agar: (for 1L)	12

3.1.5.5. Terrific Broth (TB): (for 1L) 13
3.1.5.6. 50 mg/ml Kanamycin 13
3.1.5.7. 34 mg/ml Chloramphenicol 13
3.1.5.8. 100 mg/ml Spectinomycin 13
3.1.5.9. 100 mM IPTG 14
3.1.5.10. 1000 mM IPTG 13
3.1.5.11. 10X TAE Buffer 13
3.1.5.12. Solutions for SDS – PAGE 14
3.1.5.12.1. 30% Acrylamide/Bisacrylamide Solutions 14
3.1.5.12.2. 1.5 M Tris – HCl pH 8.8 15
3.1.5.12.3. 0.5 M Tris – HCl pH 6.8 15
3.1.5.12.4. 10% SDS 15
3.1.5.12.5. 10X SDS Running Buffer 15
3.1.5.12.6. 2X Sample Buffer 16
3.1.5.12.7. 4X Sample Buffer 16
3.1.5.12.8. Staining Buffer 16
3.1.5.12.9. Destaining Buffer 17
3.1.5.13. Solutions for Western Blot 17
3.1.5.13.1. Transfer Buffer 17
3.1.5.13.2. 10X Phosphate Buffer Saline (PBS) 17
3.1.5.13.3. Triethanolamine Solution pH 7.5 18
3.1.5.13.4. 4-Chloro-1-naphthol (4C1N) Solution 18
3.1.5.13.5. Horseradish Peroxidase (HRP) Substrate 18
3.1.5.13.6. NBT/BCIP Substrate 18
3.2. Methods 19
3.2.1. Cloning of The SARS-CoV-2 S1 gene *E.coli* strains 19
3.2.1.1. Polymerase Chain Reaction (PCR) 19
3.2.1.2. Agarose Gel Electrophoresis 20
3.2.1.3. PCR Fragment Extraction from Agarose Gel 20
3.2.1.4. Quantification of DNA 20
3.2.1.5. Ligation of PCR Products into Bacterial Expression Vector 21
3.2.1.6. Transformation 22
3.2.1.7. Plasmid Isolation from Transformant Colonies 23

3.2.1.8. Polymerase Chain Reaction (PCR) Before Sequencing 24
3.2.1.9. Sequencing of SARS-CoV-2 S1 gene of SARS-Cov-2 25
3.2.2. Expression of SARS-CoV-2 S1 Protein with pET SUMO Expression System 26
3.2.3. Expression of SARS-CoV-2 S1 Protein with DsbC System 27
3.2.4. Lysis of Harvested Culture 27
3.2.5. Purification of His–tagged S1 protein 28
3.2.6. SDS – PAGE Analysis 28
3.2.7. Western Blot Analysis 29
3.2.8. Protein Determination 29
3.2.9. Cleavage of SUMO Fusion Tag 30
4. RESULTS 32
4.1. Cloning, Expression, and Purification of SARS-CoV-2 S1 protein 32
4.1.1. Cloning the SARS-CoV-2 S1 Gene into the pET SUMO Plasmid 32
4.1.2 Sequencing of S1 Gene Ligated pET SUMO Plasmid 35
4.1.3. Transformation of S1 Ligated pET SUMO Vector 36
4.1.4. Expression of S1 Protein with pET SUMO Expression System 39
4.1.5. Purification of S1 Protein 44
4.1.6. Cleavage of SUMO from SUMO Fused S1 Protein 47
4.2. Expression of S1 Protein with DsbC Signal Peptide 47
4.2.1. Transformation of S1 gene Ligated pET30a (+) Plasmid 47
4.2.2. Expression of S1 protein with DsbC Signal Peptide 48
5. DISCUSSION and CONCLUSION 51

REFERENCES 53
BIOGRAPHY 55

LIST OF ABBREVIATIONS AND ACRONYMS

Abbreviations and Acronyms		**Explanations**
ng	:	Nanogram
µg	:	Microgram
mg	:	Milligram
bp	:	Base pair
kb	:	Kilobase
kDa	:	Kilodalton
U	:	Enzyme unit
µM	:	Micromolar
mM	:	Millimolar
M	:	Molar
µl	:	Microliter
ml	:	Milliliter
OD	:	Optical density
kan	:	Kanamycin
dNTP	:	Deoxynucleotide triphosphate
LB	:	Luria-Bertani
dH_2O	:	Distilled water
BSA	:	Bovine serum albumin
APS	:	Ammonium persulfate

LIST OF FIGURES

Figure No:		Page
2.1:	Homology between SUMO and Ubiquitin.	4
3.1:	pET30a (+) plasmid map including SARS-CoV-2 S1 gene.	21
3.2:	Linearized pET SUMO plasmid map.	22
3.3:	Preparation of BCA samples.	30
4.1:	Gradient PCR to the amplification of SARS-CoV-2 S1 gene.	32
4.2:	Colony PCR amplification S1 gene.	33
4.3:	PCR amplification of S1 gene before sequencing.	35
4.4:	S1 gene sequence from pET SUMO plasmid using SUMO forward primer.	36
4.5:	Transformants were obtained from *E.coli* strains.	37
4.6:	Result of colony PCR for the SHuffle strain.	37
4.7:	Result of colony PCR for the BL21 (DE3) strain.	38
4.8:	Result of colony PCR for the BL21 Star(DE3) pLysS strain.	38
4.9:	Total protein samples of different *E.coli* strains.	39
4.10:	S1 expression with SHuffle and BL21 (DE3) strains at 25 ℃, different IPTG concentrations and times.	41
4.11:	S1 expression with SHuffle and BL21 (DE3) strains at 30 ℃, various IPTG concentrations, and different times.	42
4.12:	Total, soluble and insoluble proteins in S1 protein expression.	43
4.13:	Total, soluble and insoluble proteins in S1 protein expression.	44
4.14:	Data collected from AKTA Avant FPLC system.	45
4.15:	Chromatogram of purification of S1 protein (a) and (b).	45
4.16:	SDS-PAGE and Western Blot images of purification samples.	46
4.17:	BCA results of purified S1 protein.	46
4.18:	Result of colony PCR for the SHuffle strain.	48
4.19:	Total protein samples expressed with SHuffle cell.	49
4.20:	Samples of the total soluble and insoluble proteins belonging to 16 ℃ and 30 ℃ expressions.	49

LIST OF TABLES

Table No:		Page
2.1:	Fusion systems and their information about molecular weights and purification methods.	4
3.1:	Microorganisms used in this study.	7
3.2:	Plasmids used in this study.	8
3.3:	List of primers.	9
3.4:	List of equipment.	9
3.5:	List of kits, enzymes, and chemicals.	10
3.6:	PCR Reaction Conditions.	19
3.7:	PCR cycle set up.	20
3.8:	Ligation reaction conditions.	21
3.9:	Colony PCR reaction conditions.	23
3.10:	Colony PCR cycle set up.	23
3.11:	Conditions of PCR reaction before sequencing.	24
3.12:	PCR cycle set up.	24
3.13:	PCR reaction conditions for sequencing.	25
3.14:	PCR cycle set up for sequencing.	25
3.15:	Lysis buffer components.	27
3.16:	Preparation of SDS-PAGE gel.	28
3.17:	BSA dilution for a standard curve.	30
3.18:	Cleavage reaction condition.	31
4.1:	Result of isolated plasmid measurement at 260 nm.	34

1.INTRODUCTION

In December 2019, a group of people in Wuhan city of China's Hubei province were hospitalized with pneumonia of unknown cause. These patients showed symptoms such as back pain, cough, and fever, similar to previous Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) and Middle East Respiratory Syndrome Coronavirus (MERS-CoV) infection. At the end of 2019, the new virus, called SARS-CoV-2, was reported to the World Health Organization in Wuhan, China. SARS-CoV-2 showed 79% genetic similarity to SARS-CoV and 50% to Mers-CoV [Hu et al., 2021].

Due to the spread rate around the world, many researchers have started detection kits and vaccine studies against Covid-19 [Cui et al., 2020] [Zhao et al., 2020]. To facilitate the development of these tools, rapid and effective production of target protein molecules is required. Therefore, expression of SARS-CoV-2 proteins by mammalian and baculovirus expression systems were performed [Tai et al., 2020] [Lan et al., 2020].

In addition to these expression systems, bacterial cells are chosen as a recombinant protein expression because they offer inexpensive growth conditions, short life cycle, well-known genetics, and easy genetic manipulation [Sahdev et al., 2008]. In some studies, proteins from SARS-CoV-2 were produced with the help of bacterial expression systems [Prahlad et al., 2021].

1.1. Aim of the Thesis

This thesis aims to investigate the effects of the SUMO fusion tag and DsbC signal peptide on the expression of SARS-CoV-2 S1 protein in the *E.coli* expression system.

For this purpose, the S1 gene was cloned into pET SUMO and pET30a (+) vectors. Then, vectors containing the S1 gene were transformed into E.coli SHuffle, BL21 (DE3), and BL21 Star(DE3) pLyss strains, and expression control was performed in these strains.

2. LITERATURE REVIEW

2.1. Infection Mechanism of SARS-CoV-2

SARS-CoV-2 virus is composed of four main protein domains: Spike (S), Membrane (M), Envelope (E), and Nucleocapsid (N), which infects human respiratory tract cells by binding to human ACE2 receptors [Indwiani Astuti., 2020]. To enter the host cell, Coronaviruses first bind to the binding receptors on the cell surface and transfer the viral genome into the host cell. The Spike protein, a surface-anchor protein, helps the coronavirus to enter the host cell. In mature viruses, the Spike protein exists as a trimer. Here, there is the S2 protein for membrane fusion, and the S1 protein bound to this protein has three receptor binding properties. The SARS-CoV-2 S1 protein contains the receptor binding domain (RBD) that recognizes human angiotensin-converting enzyme 2 (ACE2) as a receptor [Shang., 2020].

2.2. Recombinant Protein Production

Soluble recombinant protein production is the backbone of structural genomics and proteomics. These proteins are generally desired to be produced in large quantities, and this goal is achieved, especially with the help of the bacterial expression system. Still, this process is rate-limited according to most researchers [Butt et al., 2005].

Escherichia coli is a very useful recombinant protein production system since it is an organism that has been characterized for a long time. It successfully produces many human proteins, except for large and complex proteins (containing many disulfide bonds) or proteins that require post-translational modification for protein activity [Lee et al., 1996].

E.coli is useful for producing recombinant protein due to the cheap production possibility, fast results, and simple processing requirement. However, according to the Southeast Collaboratory for Structural Genomics (SECSG) reports, it can only produce up to 22.7% of the 6386 identified proteins soluble [Butt et al., 2005].

Many proteins produced by *E.coli* result in the formation of insoluble inclusion bodies that are biologically inactive to prevent proteolysis [Correa et al., 2015]. Proteolysis is a frequently applied regulatory factor for misfolded proteins in the cell [Lee et al., 1996].

In general, bacteria do not require recombinant proteins and are degraded by proteases [Baneyx et al., 1999]. In this respect, different strategies have been developed to protect the recombinant protein. These include protease inhibitors, secretion in the periplasm, cell culture, and protective fusion systems. Fusion between the N-terminus and the target protein has been shown to protect the protein from degradation [Arechaga et al., 2003]. *E. coli* bacteria form misfolded and biologically inactive inclusion bodies along the periplasm periphery surrounding the cytosol to protect the overproduced proteins [Butt et al., 2005]. These bodies can be made into active soluble proteins by denaturation and refolding, but these processes are complex, costly, and time-consuming. However, current developments and understanding of the molecular mechanisms of foldases or chaperones during the production of recombinant proteins increase the amount of secreted soluble protein by *E.coli* [Lee et al., 1996].

Much progress has been evolved in improving recombinant protein expression in *E. coli*, including the development of strong promoters, co-expression with chaperones, and the use of protein fusions [Butt et al., 2005]. The T7 promoter acts as a potent trigger for the expression of a high copy number of desired proteins [Guerrero et al., 2015]. Fusion systems for proteins that are difficult to express are very functional, both theoretically and practically [Butt et al., 2005]. For this reason, different fusion partners are preferred (Table 2.1).

Table 2.1: Fusion systems and their information about molecular weights and purification methods.

Fusion Partner	Molecular Weight (kDa)	Purification Method
6× His-SUMO	12.2	Ni–NTA
CTHS	7	NTHS
GST	27.3	Glutathione–Sepharose
6× His-NusA	58.4	Ni–NTA
MBP	45	Amylose resin
TRX	14.3	ThioBond resin
6× HIS-Ub	8.6	Ni–NTA
Flag peptide	1	Anti-Flag antibodies

2.3. Protein Modification by SUMO

Proteins frequently use modifications such as acetylation, phosphorylation, and ubiquitination to rapidly repurpose their functions. Ubiquitins generally bind to protein, target substrate, and initiate the proteolytic process. It generally binds to lysine residues on the substrate. The ubiquitin family has a wide spectrum, and one of them is the SUMO (Small Ubiquitin-Related Modifier) fusion protein [Kuo et al., 2014].

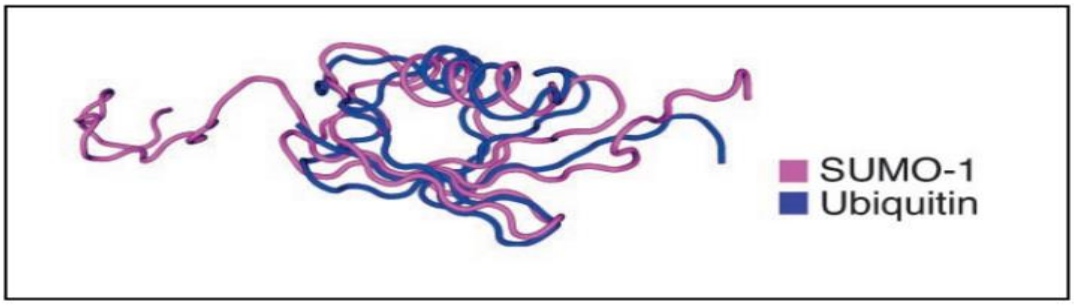


Figure 2.1: Homology between SUMO and Ubiquitin.

SUMOs form a highly conserved family of proteins found in most eukaryotic cells that are vital for their survival and activity. Since its discovery in 1996, it has been attached to more than 50 proteins, including proteins such as androgen receptor,

c-jun, histone deacetylases, and proteins associated with processes such as DNA repair and transcription.

SUMO is a protein of approximately 11 kDa but appears to be 20 kDa in SDS-PAGE analyses, which has 20 amino acids more than ubiquitin. All SUMO genes encode a precursor carrying a short C-terminal peptide to produce the mature Gly-Gly C-terminus found in most Ubiquitin-Like Proteins (Figure 2.1) [Johnson., 2004].

The patterns of SUMO conjugates are dynamic and change with stimuli. SUMO-cleavage enzymes (also known as isopeptidase) perform two important functions; first, they are removed from the protein and thereby lead to refolding of the protein, and secondly, they ensure the presence of free and reusable SUMO proteins after cleavage. All known SUMO cleavage enzymes contain approximately 200 amino acids with C-terminal cleavage activity [Johnson., 2004].

The SUMO fusion system has recently been identified to increase recombinant protein expression. Membrane and nucleocapsid proteins of severe acute respiratory syndrome coronavirus (SARS-Cov) were expressed recombinantly using SUMO fusion systems. The SUMO fusion tag leads to increased solubility and expression of recombinant proteins. The 6xHis-SUMO tag is useful for expression and purification by Ni-NTA chromatography. Fusion systems have been shown to act as solubility enhancers, although the mechanisms for improving solubility have not been fully elucidated [Marblestone et al., 2006].

2.4. The Role of DsbC in Protein Folding

Disulfide bonds are very important in the folding of secreted proteins. At this point, disulfide bond formation is conceptually divided into two parts; oxidation of cysteine residues to form disulfide bonds and isomerization of malformed disulfide bonds. For proteins containing more than two cysteines, possible mismatched disulfide bonds begin to form rapidly. Depending on the cysteine residues, at which point the cysteine residues that should be correctly matched remain constant as formed. Thus, disulfide isomerase plays a vital role in the correct coupling of multiple cysteine residues. In recent studies, one mechanism for the proper disulfide bond formation of proteins is related to a protein called DsbC, which involves disulfide bond isomerization in prokaryotes. Oxidative protein folding is triggered by

Dsb proteins found in the prokaryote periplasm. Here, DsbA acts as a direct disulfide bond former, while DsbC provides isomerization of mismatched disulfide bonds. DsbC mutants are sensitive to copper, a redox metal that acts more than DsbA and provides disulfide bond formation. This sensitivity of DsbC mutants may reveal copper-mediated false disulfide bond formation. These results suggest that DsbC has a primary role in correcting the misformed disulfide bonds formed during a certain level of oxidative stress [Gleiter et al., 2008].

3. MATERIALS & METHODS

3.1. Materials

3.1.1. Bacterial Strains

Tables 3.1 and 3.2 list the microorganisms and primers used in this study.

Table 3.1: Bacterial strains used in this work.

***Escherichiacoli* Strains**	**Genotype**	**Source/Reference**
Mach1 $T1^R$	F^- φ80*lacZ*ΔM15 Δ*lac*X74 *hsd*R(r_K^-, m_K^+) Δ*rec*A1398 *end*A1 *ton*A	Invitrogen by Thermo Fisher Scientific
BL21(DE3)	F^-*omp*T *hsd*S_B (r_B^-, m_B^-) *gal dcm* (DE3)	Invitrogen by Thermo Fisher Scientific
BL21 Star(DE3) pLysS	F^-*omp*T *hsd*S_B (r_B^-, m_B^-) *gal dcm rne*131 (DE3) pLysS(Cam^R)	Invitrogen by Thermo Fisher Scientific
Shuffle® T7	F′ *lac, pro, lacI*^{q} / *Δ(ara-leu)7697 araD139 fhuA2 lacZ::T7 gene1 Δ(phoA)PvuII phoR ahpC* galE (or U) galK λatt*::pNEB3-r1-*cDsbC* ($Spec^R$, *lacI*^{q}) *ΔtrxB rpsL150*(Str^R) *Δgor Δ(malF)3*	New England BioLabs

Table 3.2: Plasmids used in this work.

Plasmid	Construct	Source/References
pET SUMO		Invitrogen by Thermo Fisher Scientific
pET30a (+)		Novagen

3.1.1.1. *E.coli* Mach1 Strain

The Mach1 *E.coli* strain is used for cloning studies and has a faster doubling time compared to other cloning strains. The *end*A1 mutations in cells increase plasmid yield and quality. The *ton*A mutation renders cells resistant to T1 and T5 phages.

3.1.1.2. *E.coli* BL21(DE3) and BL21 Star(DE3) pLysS Strain

BL21(DE3) and BL21 Star(DE3) pLysS are both used for protein expression. Due to the absence of the outer membrane protease *omp*T and the cytoplasmic protease *lon*, BL21(DE3) and its derivatives are utilized extensively in the production of recombinant proteins. Both proteases have been shown to inhibit protein synthesis and purification. In both strains, the gene encoding the target protein is provided on a plasmid under the control of a T7-based promoter.

The BL21(DE3) strain does not carry a plasmid that expresses T7 lysozyme, whereas the BL21(DE3)pLysS strain does as a pLysS plasmid. Expression of T7 lysozyme decreases the basal expression level of the gene of interest.

3.1.1.3. *E.coli* SHuffle Strain

E. coli Shuffle strain promotes disulfide bond formation in the cytoplasm due to disulfide bond isomerase (DsbC). The gene encoding the target protein is provided on a plasmid controlled by a T7-based promoter. In addition, *lacI*^{q}'s tight expression control enables the cloning of potentially toxic genes.

3.1.2. Primers

Table 3.3: List of primers.

Primer name	Primer Sequence	Purpose of use
S1 Fwd	5' – ATG AGC CAA TGC GTG AAT CTG ACC ACC CGC – 3'	For cloning and PCR control
S1 NHis Rev	5' – TTA ACG CGC ACG ACG CGG GCT GT – 3'	For cloning and PCR control
SUMO Forward	5' – AGA TTC TTG TAC GAC GGT ATT AG – 3'	For sequencing
T7 Reverse	5' – TAG TTA TTG CTC AGC GGT GG – 3'	For sequencing
NHis pET30a Forward	5' – GGA ACA GCA ACA ACC TGG ATA G – 3'	For cloning and PCR control
NHis pET30a Reverse	5' – AAT TTC TTG TTG CTC TCG GTC A – 3'	For cloning and PCR control

3.1.3. Equipment

Table 3.4: List of equipment.

Name of Equipment	Brand, Country
Thermocycler	Bio-Rad, USA
Heater	Techne Dri-Block, DB-2A, UK
Electrophoresis System	PowerPac HC, Bio-Rad, USA
Power Supply	Bio-Rad, USA
NanoDrop	ND-ONEC-W, Thermo Fisher Scientific, USA
Biological Safety Cabinet	Esco Class II, Singapore
PCR Cabinet	Biosan, Latvia
Ice Machine	F125 Compact, Icematic, Italy
Vortex	Dragon Lab, China
15 ml and 50 ml Falcon Tubes	Thermo Scientific, USA
Petri	Fıratmed, Turkey
Water Bath	Lauda LCB 0726, Germany

Table 3.4: Table of continued.

pH Meter	Mettler Toledo, USA
Autoclave	Sistek, Turkey
Shaker	Innova 4230, New Brunswick Scientific, USA
Chromatography System	Cytiva, USA
Genetic Analyzer	CEQ800, Beckman Coulter, USA
Incubator	Nüve, Turkey
Spectrophotometer	DU730, Beckman Coulter, USA
Centrifuges	5417R Eppendorf, USA Avanti J26 XPI, Beckman Coulter, USA
0,5 ml and 1,5 ml Microcentrifuge Tubes	Axygen, USA
Gel Doc Imaging Systems	Azure Biosystem C600, USA

3.1.4. General Kits and Reagents

Table 3.5: List of kits, enzymes, and chemicals.

Name of Kits and Reagents	Code, Country
ZymoPURE™ Midiprep Kit	D4200S, Zymo Research, USA
GeneRuler DNA Ladder Mix	SM0312, Thermo Scientific, USA
PageRuler™ Plus Prestained Protein Ladder	26619, Thermo Scientific, USA
Kanamycin	60615, Sigma-Aldrich, USA
Spectinomycin	S9007, Sigma-Aldrich, USA
Chloramphenicol	C0378, Sigma-Aldrich, USA
Taq DNA Polymerase (Recombinant)	EP0401, Thermo Scientific, USA
MyTaq DNA Polymerase	BIO 21105, Bioline, USA
Agarose	A5093, Sigma, USA
Acrylamide	161-0100, Bio-Rad, USA
Bis-acrylamide	161-0125, Bio-Rad, USA

Table 3.5: Table of continued.

Methanol	34860, Sigma, USA
Acetic acid	27225, Sigma, USA
Bacto Agar	214010, BD
Bacto Tryptone	211705, BD
Bovine Serum Albumin (BSA)	A7638, Sigma, USA
Coomassie Brillant Blue R-250	112553, Sigma, USA
Ethylenediaminetetracetic acid (EDTA)	E9884, Sigma, USA
Glycine	G8898, Sigma, USA
2-Propanol	100995, Merck, Germany
Sodium Dodecyl Sulfate (SDS)	L3771, Sigma, USA
TEMED (N, N, N', N' tetramethylethylenediamine)	1610801, Bio-Rad, USA
Trizma Base	77861, Sigma, USA
Yeast Extract	21929, BD
Sodium Hydroxide	S8045, Sigma, USA
Bromophenol blue	B5525, Sigma, USA
2-Mercaptoethanol Electrophoresis Reagent	805740, Sigma, USA
Potassium Phosphate Monobasic	P5655, Sigam, USA
Potassium Phosphate Dibasic	P2222, Sigma, USA
IPTG (Isopropyl β-D-1-thiogalactopyranoside)	R0392, Thermo Scientific, USA
DNase	90083, Thermo Scientific, USA
Lysozyme	90082, Thermo Scientific, USA
B-per	78243, Thermo Scientific, USA
Protease Inhibitor Cocktail	S8820, Sigma, USA
Ammonium persulfate	161-0700, Bio-Rad, USA
Triethanolamine	90279, Sigma, USA
4-Chloro-1-naphthol (4C1N)	C6788, Sigma, USA
Syber Green	4364344, Thermo Scientific, USA

3.1.5. Buffers and Solutions

3.1.5.1. Luria-Bertani (LB) Broth Medium: (for 1L)

- 10 g Bacto Tryptone
- 5 g Yeast Extract
- 10 g NaCl
- Distilled water is added up to 1L and autoclaved at 121℃ for 15 minutes.

3.1.5.2. Luria-Bertani (LB) Agar Medium: (for 1L)

- 10 g Bacto Tryptone
- 5 g Yeast Extract
- 10 g NaCl
- 15 g Bacto Agar
- Distilled water is added up to 1L and autoclaved at 121℃ for 15 minutes.

3.1.5.3. 2xYT Broth: (for 1L)

- 16 g Bacto Tryptone
- 10 g Yeast Extract
- 5 g NaCl
- Distilled water is added up to 1L and autoclaved at 121℃ for 15 minutes.

3.1.5.4. 2xYT Agar: (for 1L)

- 16 g Bacto Tryptone
- 10 g Yeast Extract
- 5 g NaCl
- 15 g Bacto Agar
- Distilled water is added up to 1L and autoclaved at 121℃ for 15 minutes.

3.1.5.5. Terrific Broth (TB): (for 1L)

The TB medium consists of two components. Following autoclaving, the components are combined. The first component of TB medium:

- 12 g Bacto Tryptone
- 24 g Yeast Extract
- 5 ml Glycerol
- Distilled water is added up to 800 ml and autoclaved at 121°C for 15 minutes.

The second component of TB medium:

- 4.62 g KH_2PO_4
- 25 g K_2HPO_4
- Distilled water is added up to 200 ml and autoclaved at 121°C for 15 minutes.

3.1.5.6. 50 mg/ml Kanamycin: (for 10 ml)

- 0.5 g of kanamycin is dissolved in 10 ml of ddH_2O and then filtered through a 0.22 μm filter.
- The kanamycin solution is aliquoted as 1 ml and stored at -20°C.

3.1.5.7. 34 mg/ml Chloramphenicol: (for 10 ml)

- 0.34 g of chloramphenicol is dissolved in 10 ml of ddH_2O and then filtered through a 0.22 μm filter.
- The chloramphenicol solution is aliquoted as 1 ml and stored at -20°C.

3.1.5.8. 100 mg/ml Spectinomycin: (for 10 ml)

- 1 g of spectinomycin is dissolved in 10 ml of ddH_2O and then filtered through a 0.22 μm filter.
- The spectinomycin solution is aliquoted as 0.5 ml and stored at -20°C.

3.1.5.9. 100 mM IPTG: (for 10 ml)

- In 10 ml of distilled water, 0.2383 g of IPTG is dissolved and filtered through a 0.22 µm filter.
- The stock solution of IPTG is aliquoted to 1 ml and stored at -20℃.

3.1.5.10. 1000 mM IPTG: (for 10 ml)

- In 10 ml of distilled water, 2.383 g of IPTG is dissolved and filtered through a 0.22 µm filter.
- The stock solution of IPTG is aliquoted to 1 ml and stored at -20℃.

3.1.5.11. 10X TAE Buffer: (for 1L)

- 48.4 g Tris base
- 3.7 g EDTA
- 11.4 ml Acetic acid
- Distilled water is added up to 1L.
- 900 ml of distilled water is added to 100 ml of 10X TAE buffer to dilute to 1X TAE buffer.

3.1.5.12. Solutions for SDS – PAGE (Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis)

3.1.5.12.1. 30% Acrylamide/Bisacrylamide Solutions: (for 1L)

- 292 g Acrylamide
- 8 g Bis acrylamide
- Distilled water is added up to 1L and then filtered through a 0.45 µm filter.

3.1.5.12.2. 1.5 M Tris – HCl pH 8.8: (for 1L)

- 181.53 g of Trizma Base is dissolved with distilled water.
- HCl is used to bring the pH level to 8.8.
- The solution volume is adjusted to 1L.

3.1.5.12.3. 0,5 M Tris – HCl pH 6.8: (for 1L)

- 60 g of Trizma Base is dissolved with distilled water.
- HCl is used to bring the pH level to 6.8.
- The solution volume is adjusted to 1L.

3.1.5.12.4. 10% SDS: (for 100 ml)

- 10 g of SDS (Sodium Dodecyl Sulfate) is dissolved with distilled water and then the volume is adjusted to 100 ml.

3.1.5.12.5. 10X SDS Running Buffer: (for 1L)

- 30.30 g Trizma Base
- 144.10 g Glycine
- 10 g SDS
- Distilled water is added up to 1L.

900 ml of distilled water is added to 100 ml of 10X SDS running buffer to dilute to 1X SDS running buffer.

3.1.5.12.6. 2X Sample Buffer: (for 30 ml)

- 3.75 ml 0.5 M Tris – HCl
- 15 ml Glycerol
- 0,003 g Bromophenol Blue
- 6 ml 10% SDS solution
- Distilled water is added up to 30 ml.
- The stock solution of 2X sample buffer is aliquoted to 950 µl and stored at -20°C.
- 5% 2-mercaptoethanol must be added just before using the solution.

3.1.5.12.7. 4X Sample Buffer: (for 10 ml)

- 2.4 ml 0.5 M Tris – HCl
- 4 ml Glycerol
- 0,004 g Bromophenol Blue
- 0.8 g SDS
- Distilled water is added up to 10 ml.
- The stock solution of 4X sample buffer is aliquoted to 950 µl and stored at -20°C. 5% 2-mercaptoethanol must be added just before using the solution.

3.1.5.12.8. Staining Buffer: (for 1L)

- 450 ml ddH_2O
- 450 ml Methanol
- 100 ml Acetic acid
- 1 g Coomassie Brilliant Blue R-250
- Every component of the buffer is combined.

3.1.5.12.9. Destaining Buffer: (for 1L)

- 800 ml ddH_2O
- 100 ml Methanol
- 100 ml Acetic acid
- Every component of the buffer is combined.

3.1.5.13. Solutions for Western Blotting

3.1.5.13.1. Transfer Buffer: (for 1L)

- 3.0285 g Trizma Base
- 14.413 g Glycine
- 200 ml Methanol
- In 800 ml of distilled water, trizma base and glycine are dissolved before 200 ml of methanol is added.

3.1.5.13.2. 10X Phosphate Buffer Saline (PBS): (for 1L)

- 80 g NaCl
- 2 g KH_2PO_4
- 11.5 g Na_2HPO_4
- 2 g KCl
- Distilled water is added up to 1L and then filtered through a 0.45 µm filter.

3.1.5.13.3. Triethanolamine Solution pH 7.5: (for 1L)

- 7.5 g NaCl
- 2.8 ml Triethanolamine
- The components are dissolved in distilled water. The pH is adjusted to 7.5, and distilled water is added up to 1L.

3.1.5.13.4. 4-Chloro-1-naphthol (4C1N) Solution:

- 1 tablet (30 mg) 4C1N
- 10 ml methanol
- One 4C1N tablet is dissolved in 10 ml of methanol.

3.1.5.13.5. Horseradish Peroxidase (HRP) Substrate:

- 10 ml Triethanolamine solution
- 2 ml 4C1N solution
- 150 µl H_2O_2
- All solutions are mixed. The mixture must be used immediately.

3.1.5.13.6. NBT/BCIP Substrate:

- 1 tablet NBT/BCIP
- 10 ml dH_2O
- One NBT/BCIP tablet is dissolved in 10 ml of dH_2O.

3.2. METHODS

3.2.1. Cloning of the SARS-CoV-2 S1 Gene into *Escherichia coli* Strains

3.2.1.1. Polymerase Chain Reaction (PCR)

In a previous study in our lab, the SARS-CoV-2 S gene was cloned onto the pET30a (+) plasmid. This gene served as the template for the current study. Primer oligonucleotides were purchased from Sentebiolab Biotechnology. The forward and reverse primers for the amplification of the S1 fragment to be cloned are S1 Fwd and S1 NHis Rev, respectively (Table 3.3). During the gradient PCR, several different annealing temperatures were investigated to find out which produced the best results. The annealing temperature was found with the help of the NEB Tm Calculator. The reaction details, including its ingredients and its conditions, are given in Tables 3.6 and 3.7, respectively.

Table 3.6: PCR Reaction Conditions.

PCR ingredients	**Stock concentration**	**Final amount**	**Volume (20 µl)**
Template (S gene pET30a (+))	132 ng/µl	25 ng	0,25 µl
Taq Buffer with $(NH_4)_2SO_4$	10X	1X	2 µl
S1 Fwd	10 µM	0.5 µM	1 µl
S1 NHis Rev	10 µM	0.5 µM	1 µl
dNTP mix	10 mM	200 µM	0.4 µl
$MgCl_2$	25 mM	250 µM	2 µl
Taq DNA polymerase	5 U/µl	1 U/µl	0.2 µl
ddH_2O	-	-	13.15 µl

Table 3.7: PCR cycle set up.

Step		Temperature	Time
Initial Denaturation		95 ℃	2 min
32 cycle	Denaturation	95 ℃	30 sec
	Annealing	65 – 73 ℃	30 sec
	Extension	72 ℃	2 min
Final extension		72 ℃	10 min
Final hold		+4 ℃	∞

3.2.1.2. Agarose Gel Electrophoresis

Initially, agarose gel electrophoresis was utilized to validate PCR reaction results. After PCR, the PCR products were analyzed by electrophoresis on an agarose gel containing 1.2% agarose. Before it was cast, 2 µl of Syber Green was added to the liquid agarose gel mixture. The gel was loaded by mixing 20 µl of PCR product with 4 µl of gel loading buffer. The gel was run at 120V for 30 minutes. After electrophoresis, the agarose gel was visualized under UV light, and the DNA separation pattern was photographed. The size of the PCR fragments was determined by comparing them to a commercial size marker.

3.2.1.3. PCR Fragment Extraction from Agarose Gel

The amplified S1 gene fragment was extracted from the agarose gel using the MN Gel Extraction and Purification Kit.

3.2.1.4. Quantification of DNA

Spectrophotometry was used to determine the amount of isolated DNA. The concentration of DNA was determined using the absorption reading at 260 nm.

3.2.1.5. Ligation of PCR Products into Bacterial Expression Vector

The agarose gel-extracted S1 gene PCR fragment was ligated into the expression vector. For maximum ligation efficiency, a vector-to-insert molar ratio of 1:1 was chosen. The ligation reaction was incubated overnight at 15 ℃. Table 3.8 provides information about the reaction, including its ingredients and conditions.
pET30a (+) and pET SUMO vectors were used for cloning of SARS-CoV-2 S1 gene and vector map was shown in Figure 3.1 and Figure 3.2, respectively.

Table 3.8: Ligation reaction conditions.

Ligation ingredients	**Stock concentration**	**Final amount**	**Volume (10 µl)**
Insert (S1 gene PCR product)	50 ng/µl	5 ng	1 µl
Ligation buffer	10X	1X	1 µl
Vector	25 ng/µl	5 ng	2 µl
T4 DNA Ligase	4 U/µl	0.4 U/µl	1 µl
ddH_2O	-	-	5 µl

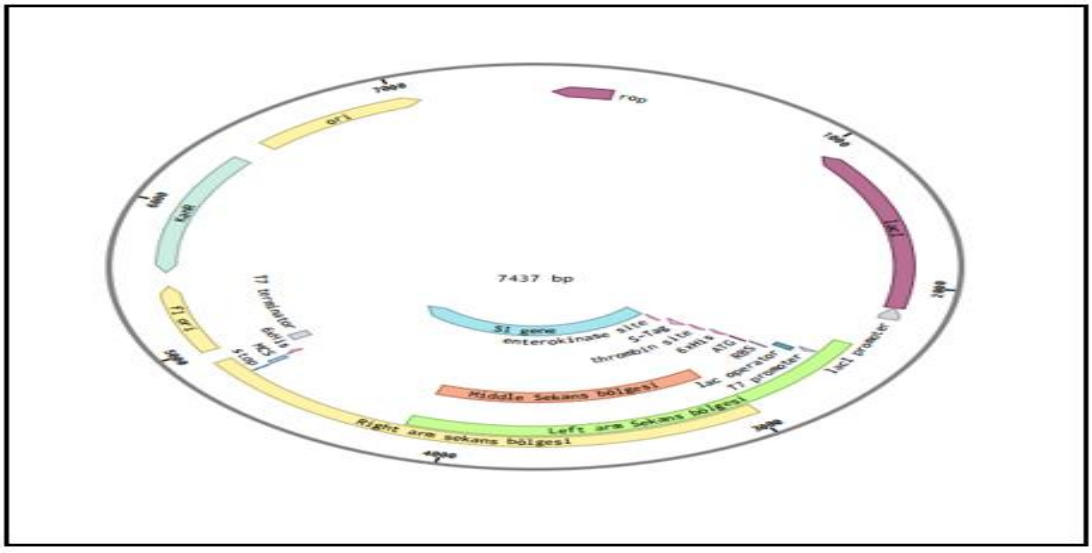


Figure 3.1: pET30a (+) plasmid map including SARS-CoV-2 S1 gene.

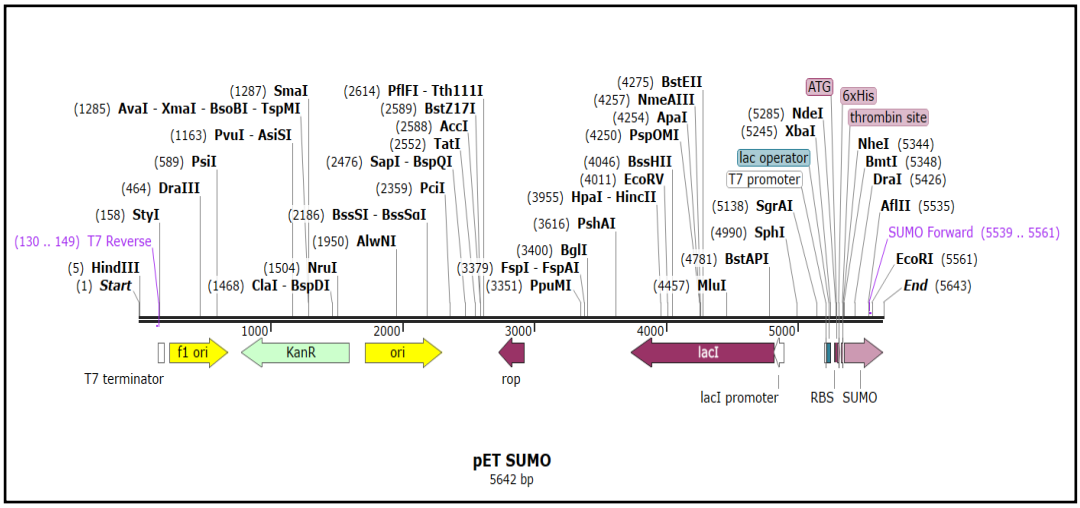


Figure 3.2: Linearized pET SUMO plasmid map.

3.2.1.6. Transformation

For transformation, 2 µl ligation product or 100 ng appropriate plasmid DNA was added into the competent *E. coli* cells and mixed gently. The mixture was incubated on ice for 30 minutes. After a 45-second heat shock at 42°C, the sample was incubated on ice for 3 minutes. Then, 500 µl of S.O.C medium was added to the mixture, which was then incubated for 60 minutes at 37°C with horizontal agitation (200 rpm). Each transformation was plated 150 µl on a selective medium containing the appropriate antibiotics. The plate was incubated overnight at 37 °C, and Following incubation, colonies were observed. These colonies were chosen at random for testing and sequencing. 10 colonies were chosen for colony PCR testing. In 12.5 µl of PCR mixture, selected colonies were dissolved and used as a template. PCR reaction was conducted with S1 forward and S1 Nhis reverse primers. The components and conditions for colony PCR are listed in Tables 3.9 and 3.10, respectively. After PCR, the PCR products were analyzed by electrophoresis on an agarose gel containing 1.2% agarose.

Table 3.9: Colony PCR reaction conditions.

PCR ingredients	Stock concentration	Final amount	Volume (12.5 µl)
Single colony (as a template)	-	-	-
MyTaq Reaction Buffer	5X	1X	2.5 µl
S1 Fwd	10 µM	0.4 µM	0.5 µl
S1 NHis Rev	10 µM	0.4 µM	0.5 µl
MyTaq DNA polymerase	5 U/µl	1 U/µl	0.25 µl
ddH_2O	-	-	8.75 µl

Table 3.10. Colony PCR cycle set up.

Step		Temperature	Time
Initial Denaturation		95 ℃	10 min
35 cycle	Denaturation	95 ℃	15 sec
	Annealing	66.6 ℃	15 sec
	Extension	72 ℃	1 min
Final extension		72 ℃	10 min
Final hold		+4 ℃	∞

3.2.1.7. Plasmid Isolation from Transformant Colonies

For plasmid isolation, LB medium containing the appropriate antibiotics was inoculated with PCR-verified colonies. Cultures were incubated for overnight at 37°C. The 'MN Plasmid Isolation Kit' was utilized for plasmid isolation. Spectrophotometric analysis was used to determine the concentration of isolated plasmids.

3.2.1.8. Polymerase Chain Reaction (PCR) Before Sequencing

Plasmid DNA was amplified using sequencing primers (SUMO Forward and T7 Reverse) after it was extracted from colonies that had been verified. The annealing temperature was found with the help of the NEB Tm Calculator. In Tables 3.11 and 3.12, respectively, are listed the reaction's products and conditions. Prior to sequencing, PCR products were purified utilizing the ZR DNA Sequencing Clean up Kit.

Table 3.11: Conditions of PCR reaction before sequencing.

PCR ingredients	**Stock concentration**	**Final amount**	**Volume (20 µl)**
Fresh PCR product	Variable	25 ng	Variable
Taq Buffer with $(NH_4)_2SO_4$	10X	1X	2 µl
S1 Fwd	10 µM	0.5 µM	1 µl
S1 NHis Rev	10 µM	0.5 µM	1 µl
dNTP mix	10 mM	200 µM	0.4 µl
$MgCl_2$	25 mM	250 µM	2 µl
Taq DNA polymerase	5 U/µl	1 U/µl	0.2 µl
ddH_2O	-	-	Up to 20 µl

Table 3.12: PCR cycle set up.

Step		**Temperature**	**Time**
Initial Denaturation		95 °C	2 min
32 cycle	Denaturation	95 °C	30 sec
	Annealing	47 °C	30 sec
	Extension	72 °C	2 min
Final extension		72 °C	10 min
Final hold		+4 °C	∞

3.2.1.9. Sequencing of SARS-CoV-2 S1 Gene of SARS-CoV-2

The employed sequencing technique was based on the Sanger-sequencing technique. This technique utilizes dideoxynucleotide chain termination. This involves the addition of dideoxynucleotides to the DNA's regular nucleotides. Rather than a hydroxyl group, dideoxynucleotides contain a hydrogen group on the 3' carbon. When integrated into a sequence, these modified nucleotides prevent the addition of additional nucleotides or halt further elongation. Due to the inability to form a phosphodiester bond between the dideoxynucleotide, DNA chain elongation is halted [Brown., 2007].

All plasmids were amplified using the S1 gene-specific primers listed in Table 3.3 and the PCR components and protocol for the S1 gene listed in Table 3.13 and Table 3.14, respectively. For S1 gene sequencing, a Beckman Coulter Dye Terminator Cycle Sequencing (DTCS) Quick Start Kit was utilized. All samples were subjected to ethanol precipitation following PCR amplification.

Table 3.13: PCR reaction conditions for sequencing.

PCR ingredients	**Stock concentration**	**Final amount**	**Volume (20 µl)**
Clean PCR product	Variable	33 ng	Variable
SUMO Forward	10 µM	1.5 µM	3 µl
DTCS Quick Start Master Mix	-	-	8 µl
ddH_2O	-	-	Up to 20 µl

Table 3.14: PCR cycle set up for sequencing.

Step		**Temperature**	**Time**
30 cycle	Denaturation	96 ℃	20 sec
	Annealing	50 ℃	20 sec
	Extension	60 ℃	4 min
Final hold		+4 ℃	∞

The BLAST program at the National Centre for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/BLAST) was used for analysis. The sequences of the S1 gene were compared to the nucleotide database.

For protein expression, competent *E. coli* cells were transformed with plasmids confirmed by sequencing.

3.2.2. Expression of SARS-CoV-2 S1 Protein with pET SUMO Expression System

E. coli Shuffle strain carrying pET SUMO vector were grown in 20 ml TB medium containing Kanamycin (50 µg/ml) and Spectinomycin (100 µg/ml) at 37℃ shaking with 220 rpm overnight. The next day, the overnight culture volume was calculated for addition to the expression culture volume. Calculation was given below (3.1).

$$OD600\ value\ of\ overnight\ culture\ X\ Volume = Expected\ OD600\ value\ X\ Volume\ of\ expression\ culture \quad (3.1)$$

For expression culture, the desired OD_{600} value was 0.05. The overnight culture was inoculated in 50 ml TB containing 50 µg/ml kanamycin and 100 µg/ml spectinomycin. The expression culture was grown until the absorbance at 600 nm reached 0.6, at which point IPTG was added to an exponentially growing *E. coli* culture at a final concentration of 0.4 mM. The cultures were incubated at 16℃ shaking with 140 rpm overnight. The induced cultures were harvested by centrifugation at 10,000 g for 30 minutes.

3.2.3. Expression of SARS-CoV-2 S1 Protein with DsbC System

E.coli Shuffle strain harboring the pET30a (+) expression vector were grown overnight at 37°C and 220 rpm in 20 ml of LB broth containing 50 µg/ml of kanamycin and 100 µg/ml spectinomycin. The following day, expression culture was initiated by adding the appropriate volume of overnight culture to 50 ml of LB broth containing Kanamycin (50 mg/ml) and Spectinomycin (100 µg/ml). The culture was grown until the absorbance at 600 nm reached at 0.06, after which IPTG was added at a final concentration of 0,4 mM. The cultures were incubated at 30℃ for 5 hours. The induced cultures were harvested by centrifugation at 8,000 g for 15 minutes.

3.2.4. Lysis of Harvested Culture

Using both mechanical and chemical methods, the harvested cells were lysed. Sonication was utilized for the mechanical disruption of cells, whereas lysis buffer (Thermo Scientific B-per lysis buffer) was utilized for the chemical disruption of the cells. Components of the lysis buffer were optimized for 1 g of bacterial pellet. The cells were sonicated with a 40% amplitude, pulsed for 10 seconds ON and 10 seconds OFF in a 10 cycle. Table 3.15 describes the lysis buffer component.

Table 3.15: Lysis buffer components.

Components	Volume (for 4 ml)
B-per	4 ml
DNase	8 µl
Lysozyme	8 µl
Protease Inhibitor Coctail	400 µl

1 g of the bacterial pellet was resuspended in lysis buffer and incubated on ice with shaking for 1 hour. The resuspended pellet was subjected to a sonication protocol. Following sonication, the cell lysates were centrifuged for 15 minutes at 10,000 g to remove insoluble material. After centrifugation, the pellet contains insoluble proteins while the supernatant contains soluble proteins. Afterwards, the

samples, which contained soluble proteins, were purified with Ni-charged resin (EconoFit Nuvia IMAC Column).

3.2.5. Purification of His – tagged S1 Protein

Akta Avant Fast Performance Liquid Chromatography system was used for purification studies. EconoFit Nuvia IMAC column (1 ml) was prefered for protein purification that includes 6XHis tag.

Firstly, dH_2O was passed from the system for the washing step. To remove %20 ethanol from the column, five-column volume (CV) dH_2O passed from the column. After that 5 CV equilibration buffer (20 mM Na-P, 30 mM Imidazole, 300 mM NaCl, ph:7.0) was loaded to system. Then, the protein sample was loaded into the system and unbound proteins were collected from the outline. After sample load, 3 CV more equilibration buffer was passed from the system of the device and bound proteins were eluted from column by elution buffer (20 mM Na-P, 300 mM Imidazole, 300 mM NaCl, ph:7.0).

After the purification step, purified protein samples were analysed with SDS-PAGE electroporesis and protein concentration was determined by BCA assay.

3.2.6. SDS – PAGE Analysis

In all studies, 10% SDS-PAGE gel was used. Table 3.16 provides the gel content.

Table 3.16: Preparation of SDS-PAGE gel.

Solution components	**Separation gel (10%)**	**Stacking gel (4%)**
dH_2O	1.9 ml	0.68 ml
1.5 M Tris-HCl, pH 8.8	1.3 ml	----
0.5 M Tris-HCl, pH 6.8	----	0.13 ml
30% Acrylamide/Bisacrylamide	1.7 ml	0.17 ml
10% SDS	0.05 ml	0.01 ml
10% Ammonium persulfate (APS)	0.05 ml	0.01 ml
TEMED	0.002 ml	0.001 ml

To test the presence of extracted protein samples, a 10% polyacrylamide gel was used. Cellular lysates were combined with sample loading dye just prior to SDS–PAGE analysis. After 5 minutes of heating at 95 °C, they were loaded onto the gel. So that the protein samples could run on the stacking gel, the voltage was set to 100 V. In 10 minutes, the voltage was increased to 120 V. Next, polyacrylamide gels were stained for 30 minutes at room temperature with SDS-PAGE staining solution, washed with destain buffer, and then visualized.

3.2.7. Western Blot Analysis

10% SDS-PAGE gel was used to separate proteins. The gel was equilibrated for 5 minutes in transfer buffer, and protein bands were transferred to a nitrocellulose membrane at 25 V for 30 minutes. The membrane was then blocked in blocking buffer (1% BSA and 0.1% Tween 20-PBS) for 1 hour at room temperature. The gel was incubated overnight at 4°C with the primary antibody, rabbit monoclonal anti-Spike S1 antibody (dilution 1:5000; GeneTex). After three washes with PBS-T, the second antibody, Alkaline phosphatase-antigoat rabbit IgG (1:5000 dilution; Thermo Scientific), was added and incubated at room temperature for 1 hour. After three washes with PBS-T, NBT/BCIP solution was used to visualize the protein bands.

For His-tag detection, a mouse monoclonal anti-polyhistidine-peroxidase antibody (1/10000 dilution; Sigma) was used instead of the protocol described above. Triethanolamine solution was used to visualize protein bands.

3.2.8. Protein Determination

The Pierce^{TM} BCA Protein Assay Kit was used to determine the concentration of purified proteins. Figure 3.4 presents the steps of the BCA method. Preparation of diluted albumin (BSA) standards are given in Table 3.17.

Table 3.17: BSA dilution for standard curve.

Reagent	Final BSA concentration (µg/ml)								
	2000	**500**	**1000**	**750**	**500**	**250**	**25**	**5**	**0**
2 mg/ml BSA	1000 µl	750 µl	500 µl	375 µl	250 µl	125 µl	62.5 µl	12.5 µl	0 µl
dH_2O	0 µl	250 µl	500 µl	625 µl	750 µl	875 µl	937.5 µl	987.5 µl	1000 µl

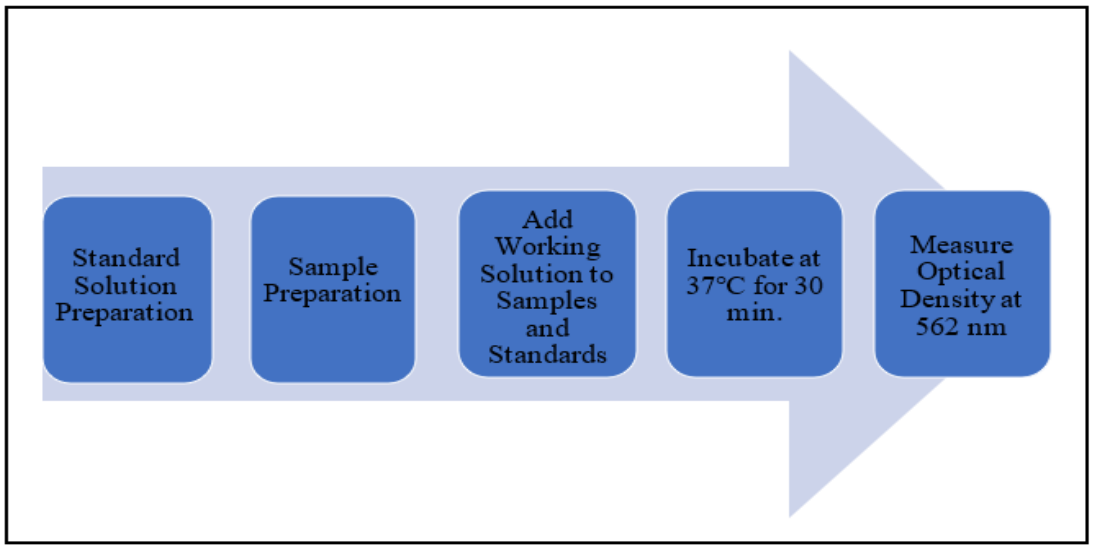


Figure 3.3: Preparation of BCA samples.

3.2.9. Cleavage of SUMO Fusion Tag

Purified SARS-CoV-2 S1 protein was cleaved with SUMO Protease in order to obtain native protein. The N-terminal peptide containing the 6xHis tag and SUMO is cleaved by the SUMO Protease. The reaction products are listed in Table 3.18. The reaction was incubated at 30°C for 6 hours, and 20 µl samples were collected at 1, 2, 4, and 6 hours. PAGE analysis was performed for each sample.

Table 3.18: Cleavage reaction condition.

Ingredients	**Stock concentration**	**Final amount**	**Volume (200 µl)**
Fusion protein	variable	20 µg	-
SUMO Protease buffer (150 mM NaCl)	10X	1X	20 µl
SUMO Protease	1U/µl	10U	10 µl
dH_2O	-	-	Up to 200 µl

4. RESULTS

4.1. Cloning, Expression and Purification of SARS-CoV-2 S1 protein

4.1.1. Cloning the SARS-CoV-2 S1 gene into the pET SUMO Plasmid

In this study, the SARS-CoV-2 S gene that was previously cloned into the pET-30a (+) plasmid was used as a template. The S1 gene of SARS-CoV-2 was amplified with PCR via this template. During gradient PCR, various annealing temperatures were tested to determine the optimal working temperature. The PCR results were examined on the agarose gel (Figure 4.1).

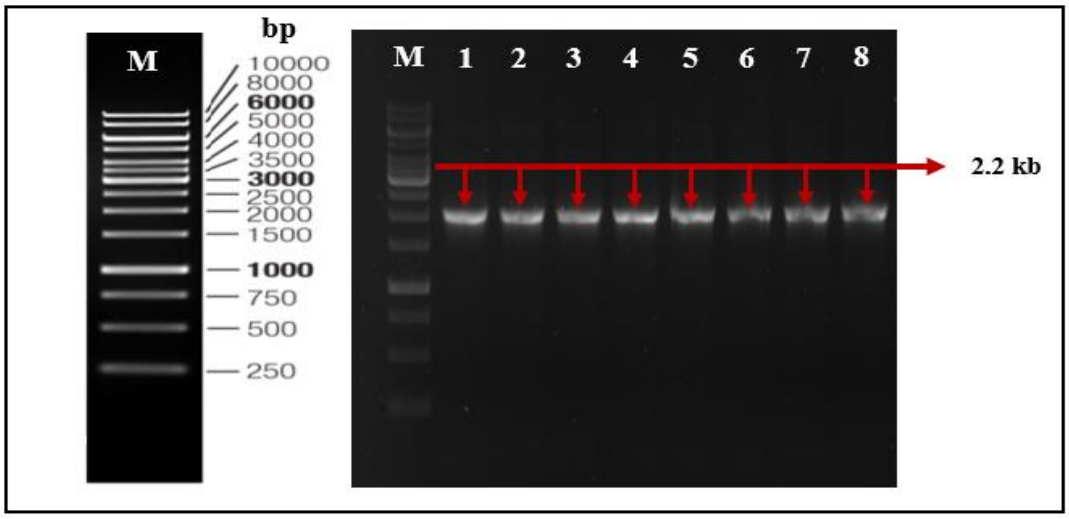


Figure 4.1: Gradient PCR used for the amplification of SARS-CoV-2 S1 gene. (M) GeneRuler 1 kb DNA Ladder, (1) 72 ℃, (2) 71.5 ℃, (3) 70.7 ℃, (4) 69.2 ℃, (5) 67.7 ℃, (6) 66.4 ℃, (7) 65.5 ℃, (8) 65 ℃.

It was determined through a visual inspection of the agarose gel that there was not a discernible temperature effect between the different PCR samples. All samples were extracted from agarose gel and purified for the ligation reaction. The concentration of purified PCR products was determined at 260 nm and the result was recorded as 182 ng/µl. The ligation of the SARS-CoV-2 S1 gene to the pET SUMO plasmid was performed. The ligation product was transformed into *E.coli* Mach1

cells using heat shock method. After transformation, transformants were chosen from LB-kanamycin containing plates. Before isolating the plasmid, transformants were confirmed by colony PCR and PCR results were visualized on an agarose gel (Figure 4.2).

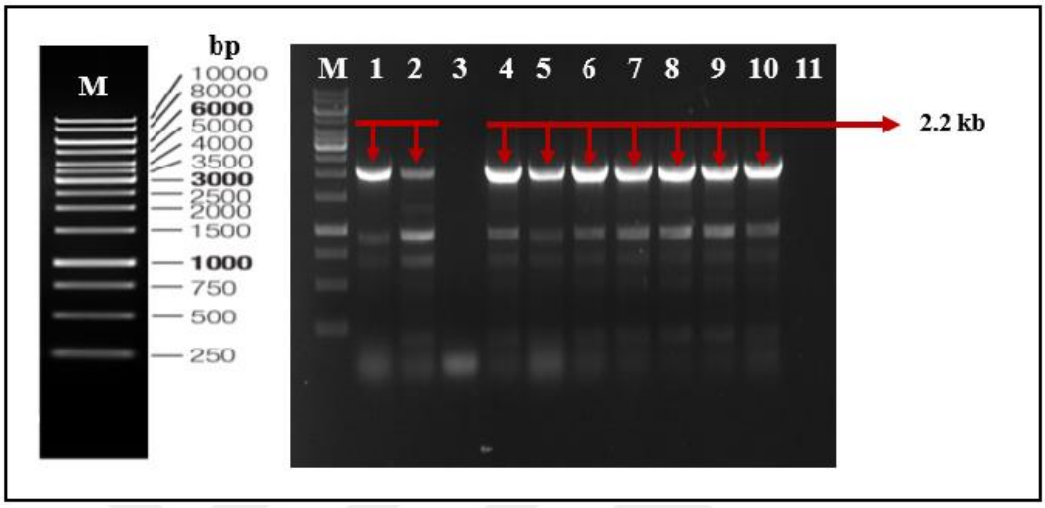


Figure 4.2: Colony PCR amplification result of S1 gene. (M) GeneRuler 1 kb DNA Ladder, (1, 2, 3, 4, 5, 6, 7, 8, 9, 10) PCR products, (11) control (w/o template).

As a result of colony PCR, 2.2 kb PCR products were obtained for all samples except the third one. For plasmid isolation, colonies containing the S1 gene were grown in LB–kanamycin broth media. From nine colonies, plasmids were isolated and their concentrations were measured at 260 nm. The measurement outcomes are presented in Table 4.1.

Table 4.1: Result of isolated plasmid measurement at 260 nm.

Samples	**Concentration**	**A260/280**	**A260/230**
Colony 1	92 ng/µl	1.89	1.81
Colony 2	90 ng/µl	1.91	1.75
Colony 4	95.6 ng/µl	1.90	1.80
Colony 5	87.8 ng/µl	1.92	1.81
Colony 6	98.8 ng/µl	1.88	1.85
Colony 7	102.4 ng/µl	1.84	1.89
Colony 8	116 ng/µl	1.82	1.89
Colony 9	100 ng/µl	1.87	1.94
Colony 10	91.4 ng/µl	1.82	1.87

4.1.2 Sequencing of S1 Gene Ligated pET SUMO Plasmid

Prior to sequencing, the presence of the SARS-CoV-2 S1 gene in purified plasmids was determined by PCR amplification. PCR samples were controlled on agarose gel (Figure 4.3).

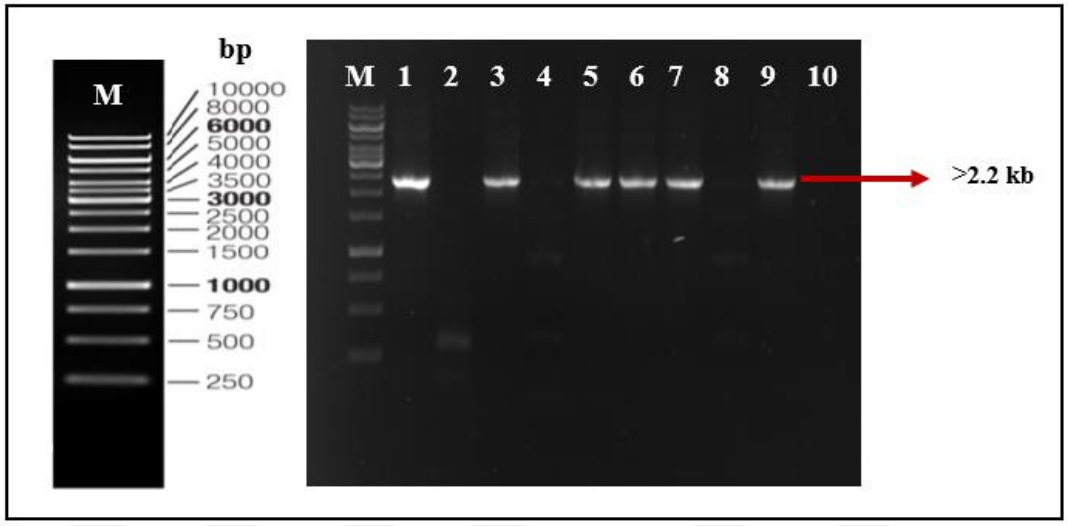


Figure 4.3: PCR amplification of S1 gene before sequencing. (M) GeneRuler 1 kb DNA Ladder, (1, 2, 3, 4, 5, 6, 7, 8, 9) PCR products, (10) control (w/o template).

According to the Figure 4.3., the amplified S1 gene was detected in a different location than usual due to the use of sequencing primers. Six positive PCR samples were obtained and extracted from agarose gel. PCR was used to amplify purified samples with fluorescently-labeled ddNTPs. The only forward primer was used to confirm the correct orientation of the S1 gene. The corresponding PCR products were sequenced to confirm the cloning of the S1 gene into pET SUMO. The resulting sequences were aligned using the National Centre for Biotechnology Information's BLAST tool (NCBI). Sequence results showed that only one of the plasmids includes the S1 gene in the correct orientation (Figure 4.4).

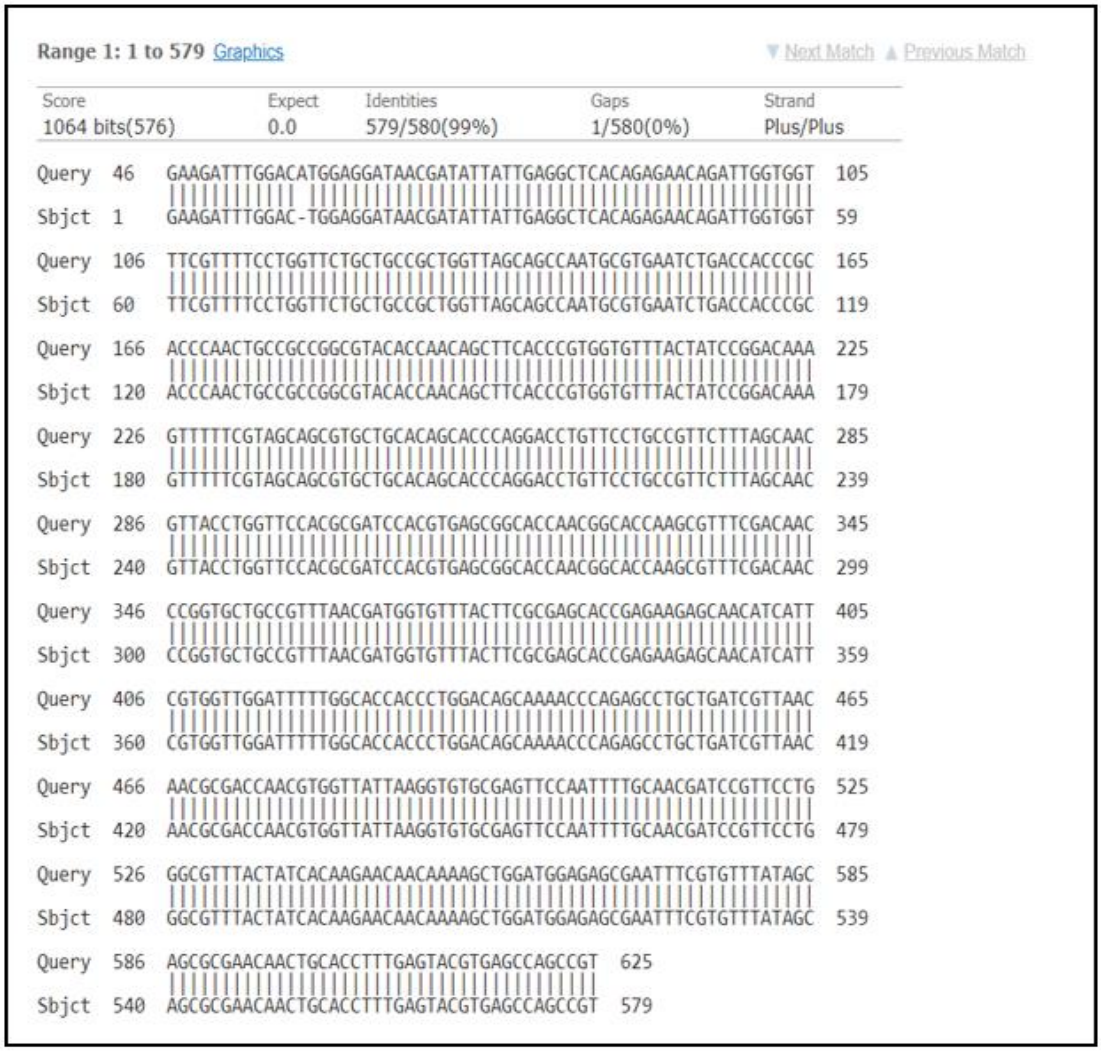


Figure 4.4: S1 gene sequence from pET SUMO plasmid using SUMO forward primer. Query line represents original S1 sequence. Subject line represents sequencing sample.

4.1.3. Transformation of S1 Ligated into pET SUMO Vector

100 ng of isolated plasmid was transferred to various strains of *E. coli*. SHuffle, BL21 (DE3) and BL21 Star(DE3) pLysS *E.coli* strains were prefered for expression studies. The cells were spread on LB and 2xYT agar media, and transformants were obtained after overnight incubation at 37 °C (Figure 4.5).

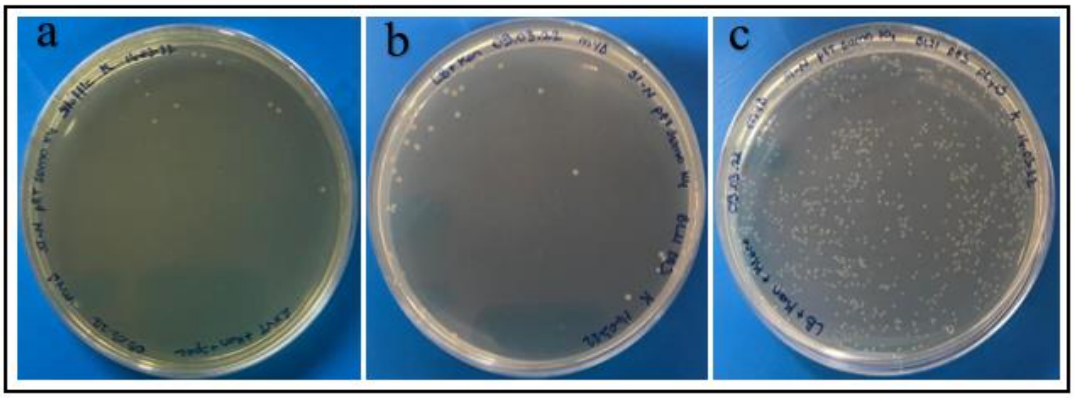


Figure 4.5: Transformants obtained from *E.coli* strains. (a) SHuffle Strain, (b) BL21 (DE3) Strain, (c) BL21 Star(DE3) pLysS Strain.

On each plate, colonies were selected and confirmed by colony PCR. Figures 4.6, 4.7, and 4.8 represent the results of colony PCR.

Figure 4.6: Result of colony PCR for the SHuffle strain. (M) GeneRuler 1 kb DNA Ladder, (1, 2, 3, 4, 5, 6, 7, 8, 9, 10) PCR products, (11) Control (w/o template).

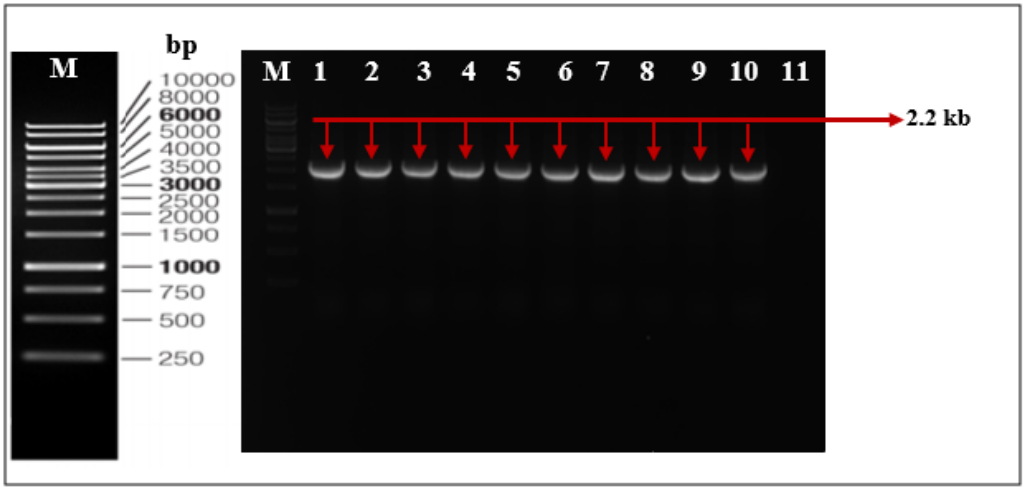


Figure 4.7: Result of colony PCR for the BL21 (DE3) strain. (M) GeneRuler 1 kb DNA Ladder, (1, 2, 3, 4, 5, 6, 7, 8, 9, 10) PCR products, (11) Control (w/o template).

Figure 4.8: Result of colony PCR for the BL21 Star(DE3) pLysS strain. (M) GeneRuler 1 kb DNA Ladder, (1, 2, 3, 4, 5, 6, 7, 8, 9, 10) PCR products, (11) Control (w/o template).

According to colony PCR results, it was observed that selected colonies from all strains are positive. In the next step, all of the selected colonies were evaluated to determine SARS-CoV-2 S1 protein expression.

4.1.4. Expression of S1 Protein with pET SUMO Expression System

PCR-selected transformants were utilized to determine S1 protein expression. For that purpose, transformants were grown overnight in 5 ml of LB medium at 37 °C, and expression cultures were then grown in 20 ml of LB medium. Cultures were incubated at 30 ℃ after 0.4 mM IPTG induction. After 5 hours, cells were harvested by cell lysis to obtain total protein. Soluble and insoluble proteins were separated from each other with centrifugation. The obtained total protein samples were analyzed using SDS-PAGE (Figure 4.9).

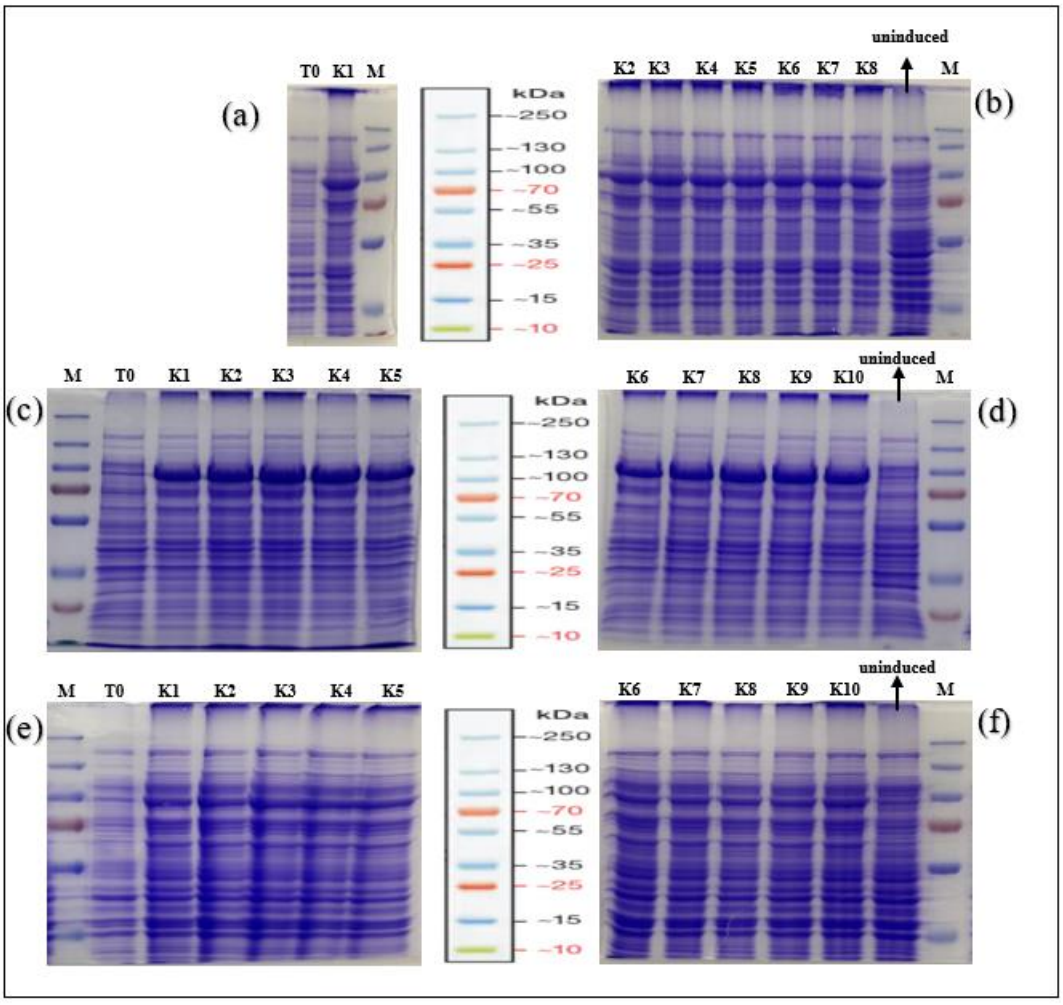


Figure 4.9: Total protein samples of different *E.coli* strain. Protein samples were collected at 5h. T0 and uninduced samples also analyzed. Expressed total protein by SHuffle (a) and (b), expressed total protein by BL21 (DE3) (c) and (d), expressed total protein by BL21 Star(DE3) pLysS (e) and (f).

When the SDS-PAGE results were examined, it was observed that the target protein was obtained at the expected size (~83 kDa). First, each strain was evaluated individually. All gels demonstrated that S1 protein expression is negligible level prior to IPTG induction. Following 5 hours of IPTG induction, there was an increase in protein expression. Evaluation of uninduced samples revealed a low level of S1 expression. This situation can be explained with leaky expression without IPTG induction. Comparing all of the strains, the SHuffle and BL21 (DE3) strains were found to have the highest levels of S1 protein expression.

The effect of different temperature (25 ℃ and 30 ℃), IPTG concentration (0.1, 0.5 and 1 mM) and time intervals (3 and 5 hours) on the expression of S1 protein was investigated in SHuffle and BL21 (DE3) strains, which were found to have high expression. All experiments were conducted in 3 repeats. Figure 4.10 and Figure 4.11 represent results of SDS-PAGE analysis.

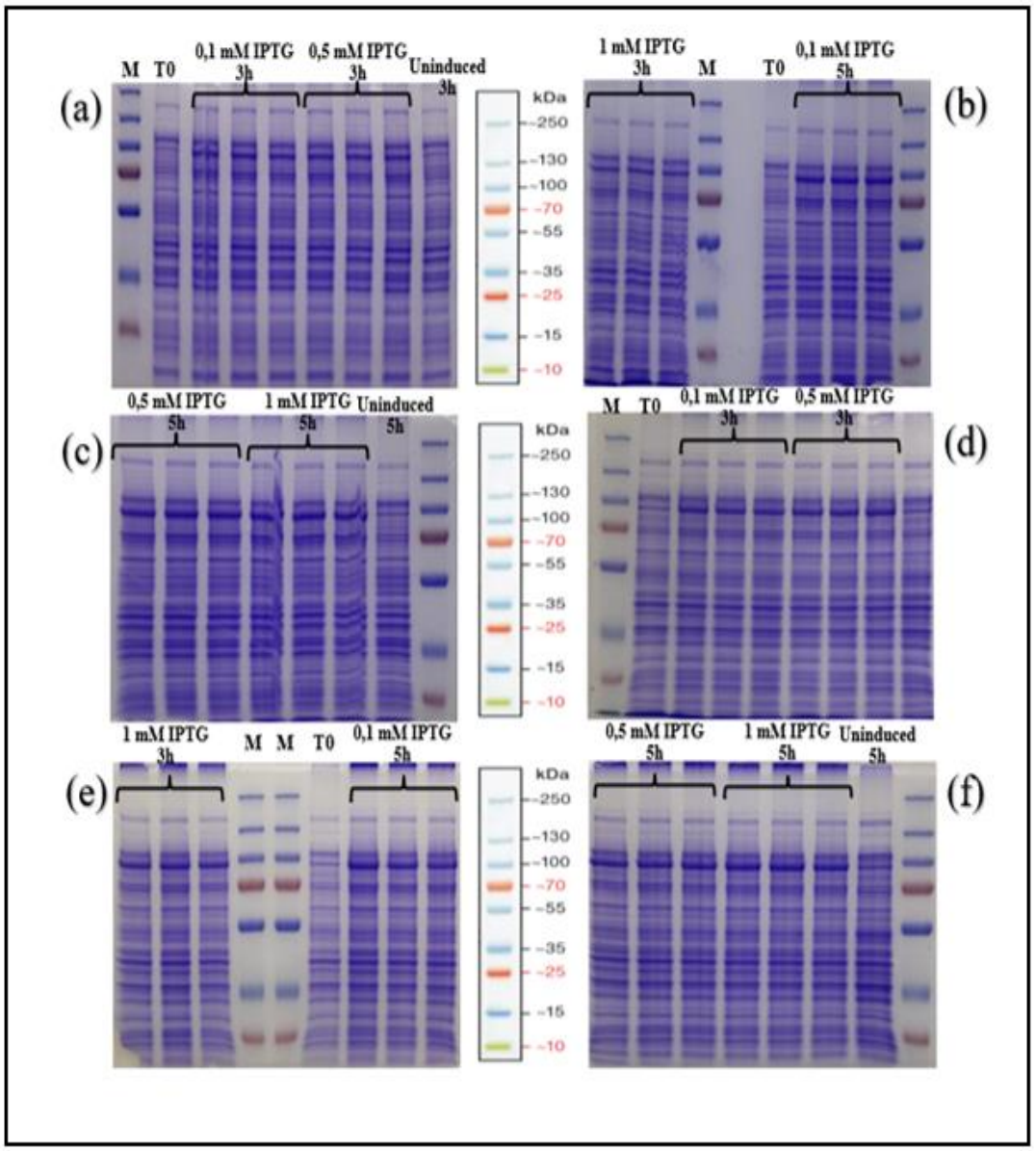


Figure 4.10: S1 protein expression with SHuffle and BL21 (DE3) strains at 25 ℃, different IPTG concentration and time intervals. Protein expression in the SHuffle strain was induced with 0.1 mM, 0.5 mM, and 1 mM IPTG, and cells were harvested after 3, and 5 hours (a), (b), (c). The BL21 (DE3) strain was induced to produce proteins with 0.1 mM, 0.5 mM, and 1 mM IPTG, and the cells were harvested after 3 and 5 hours (d), (e), (f).

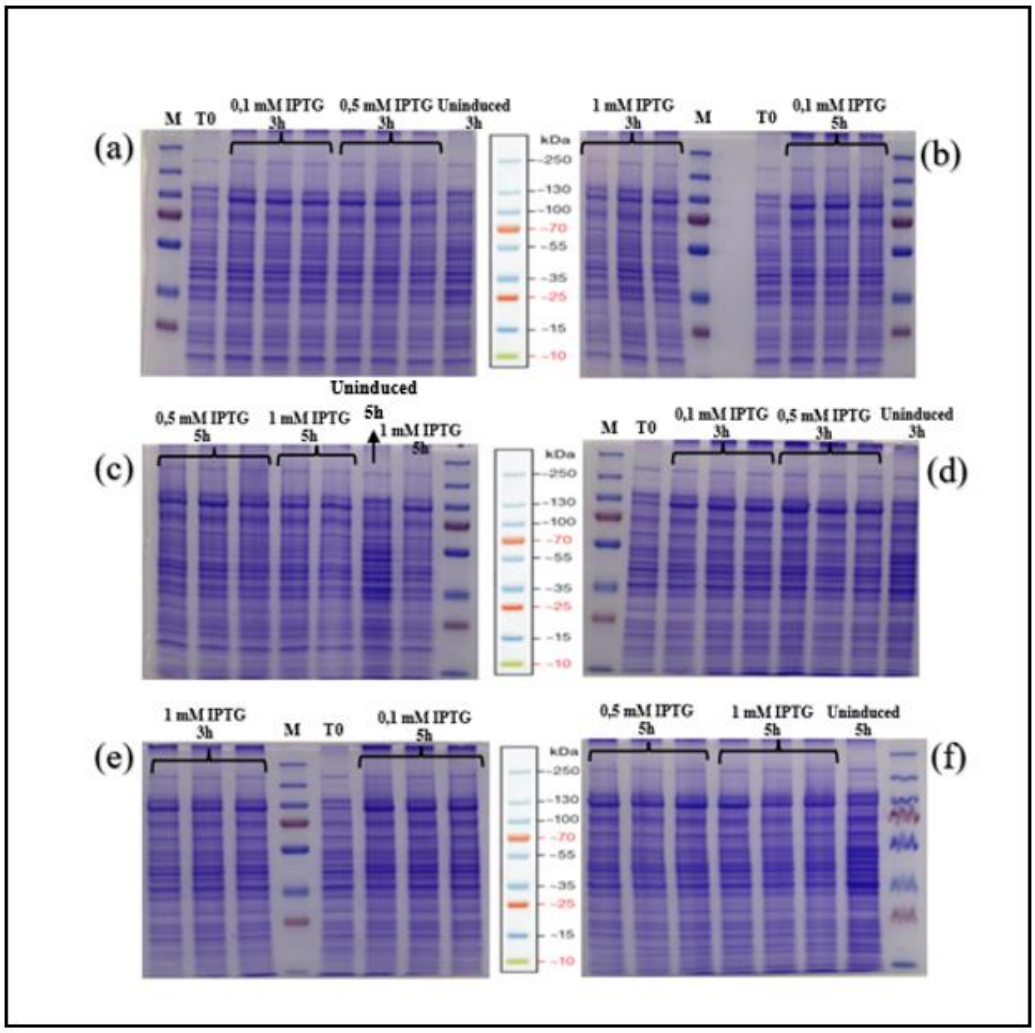


Figure 4.11: S1 expression with SHuffle and BL21 (DE3) strains at 30 ℃, various IPTG concentrations, and different time intervals. The SHuffle strain was treated with 0.1 mM, 0.5 mM, and 1 mM IPTG, and the cells were harvested after 3 and 5 hours (a), (b), (c). With 0.1 mM, 0.5 mM, and 1 mM IPTG, the BL21 (DE3) strain was induced to produce proteins, and the cells were harvested after 3 and 5 hours (d), (e), (f).

When the harvesting time was analyzed at 25 ℃, it was found that, depending on the time, there was an increase in the expression of the S1 protein in both strains. When the concentrations of IPTG were compared, it was observed that a concentration of 0.1 mM IPTG increased the expression in SHuffle cells relatively more than the other concentrations (Figure 4.10 a,b,c). However, in BL21 (DE3) cells, an administration of IPTG at a high concentration was found to be more effective on protein expression (Figure 4.10 d,e,f).

Expressions at temperature 30 ℃ revealed an expression profile similar to that of temperature 25 ℃. The expression of total proteins was enhanced by extending

the harvesting period. It was seen that low IPTG concentrations resulted in comperatively higher protein expression in SHuffle cells, but changing IPTG concentrations in BL21 (DE3) cells didn't make a big difference in expression (Figure 4.11).

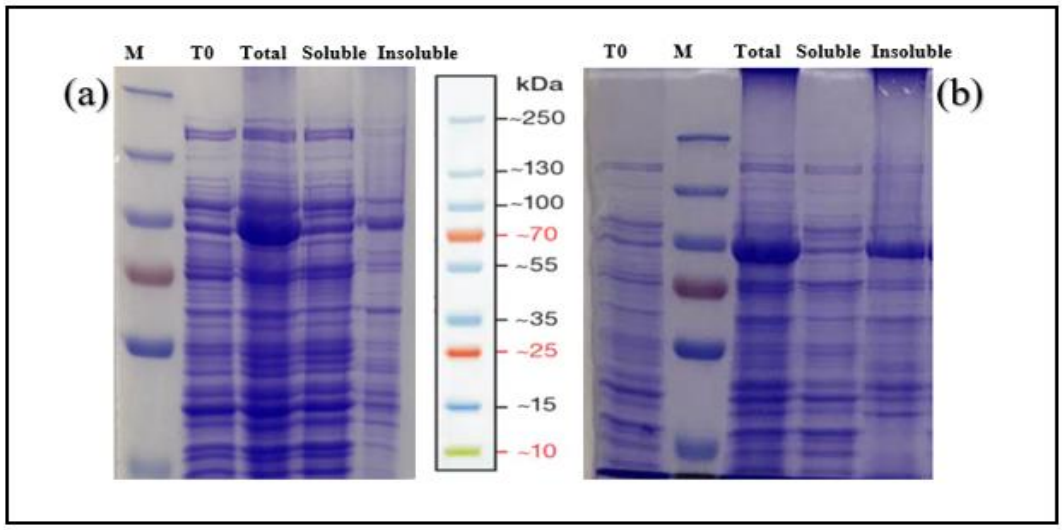


Figure 4.12: Total, soluble and insoluble proteins in S1 protein expression. The SDS-PAGE image belongs to the expression of SHuffle and BL21 (DE3) cells (a) and (b), respectively.

Figure 4.12 shows the SDS-PAGE image of proteins separated as total, soluble and insoluble at 25 ℃. 5 hours following the induction of expression with 0.5 mM IPTG, cells were harvested. The expression of S1 protein is more intense in SHuffle cells, and as it is seen in the gel the amount of soluble protein is also higher (Figure 4.12).

Low temperature and slow shaking speed are known to be effective for the production of soluble proteins in *E. coli* expression systems. For this reason, SHuffle and BL21 (DE3) strains were subjected to a low-temperature test for S1 protein expression. After induction with 0.4 mM IPTG, cells were incubated at 16 ℃ with 140 rpm shaking for 24 hours, and then they were harvested. After separating total, soluble, and insoluble proteins, SDS-PAGE and Western Blot analyses were carried out (Figure 4.13).

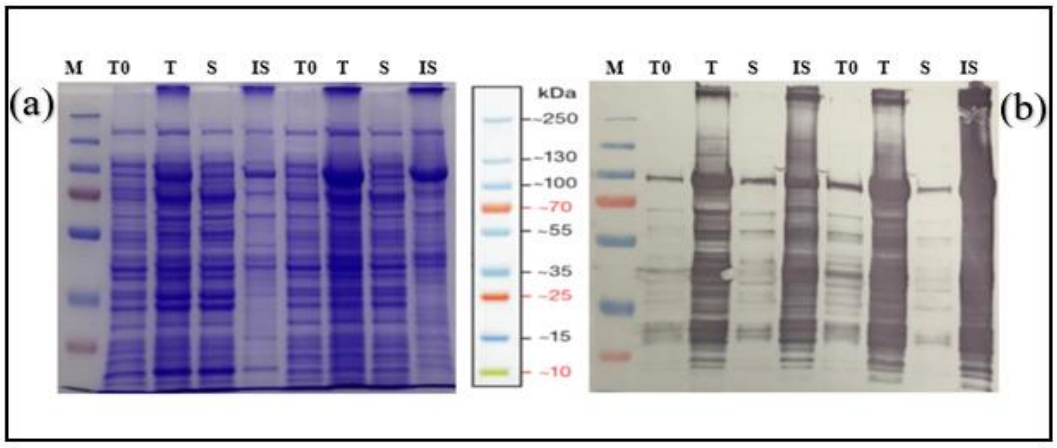


Figure 4.13: Total, soluble and insoluble proteins in S1 protein expression. The SDS-PAGE image belongs to the expression of SHuffle and BL21 (DE3) cells (a). Western Blot image demonstrates the expression of SHuffle and BL21 (DE3) (b).

When total protein expression in cells is compared, both SDS-PAGE image and western blot image show that the expression of S1 protein is higher in BL21 (DE3) cells at 16°C. However, when the amount of soluble protein is examined, it was seen that the S1 protein expressed as soluble by SHuffle cells is higher than BL21 (DE3) cells. Among all the temperatures, IPTG concentration and time trials, the conditions determined that the S1 protein was produced more soluble than the other trials were accepted as optimum for the production of S1 protein (SHuffle strain, 16 °C, 0.4 mM IPTG, 24 hours). Protein production was carried out using these conditions for further purification experiments.

4.1.5. Purification of SARS-CoV- 2 S1 Protein

S1 protein samples obtained as soluble were purified using the Akta Avant Fast Pressure Liquid Chromatography (FPLC) system. Figures 4.14 and 4.15 show the device data collected during purification and the chromatogram of the purification result, respectively.

Sample Application	: Sample Flow Rate	: 0.5 ml/min	
	Conductivity	: 20.90 mS/cm	
	UV	: 3196 mAU	
	Fractionation type	: Fixed Volume Fractionation	
Wash	: System Flow Rate	: 1 ml/min	
	Cond	: 30.84 mS/cm	
	UV	: 45.93 mAU	
	Fractionation type	: Fixed Volume Fractionation	
Elution	: Flow Rate	: 0.4 ml/min	
	Fractionation Type	: Fixed Volume Fractionation.	
	Eluted protein was collected at a volume of 0.5 ml.		

Figure 4.14: Datas collected from AKTA Avant FPLC system.

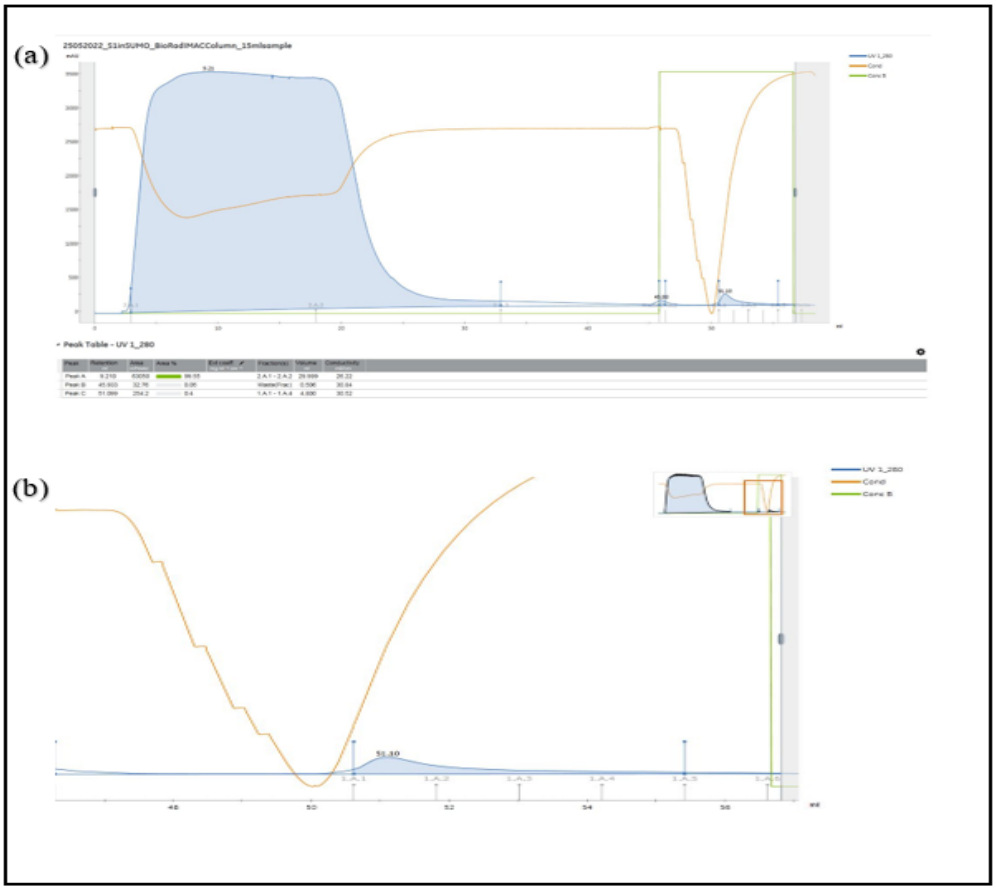


Figure 4.15: Chromatogram of purification of S1 protein (a) and (b).

Analyzing the chromatogram showed that a low-level UV peak was obtained during the elution. This low-level peak indicates a low concentration of protein. Therefore, S1 protein was collected in 3 fractions (A1, A2 and A3) in a 0.5 ml volume to increase the elution concentration.

SDS-PAGE and Western Blot analysis was used to examine purification samples (Figure 4.16). The concentration of elution samples was measured using the BCA assay (Figure 4.17).

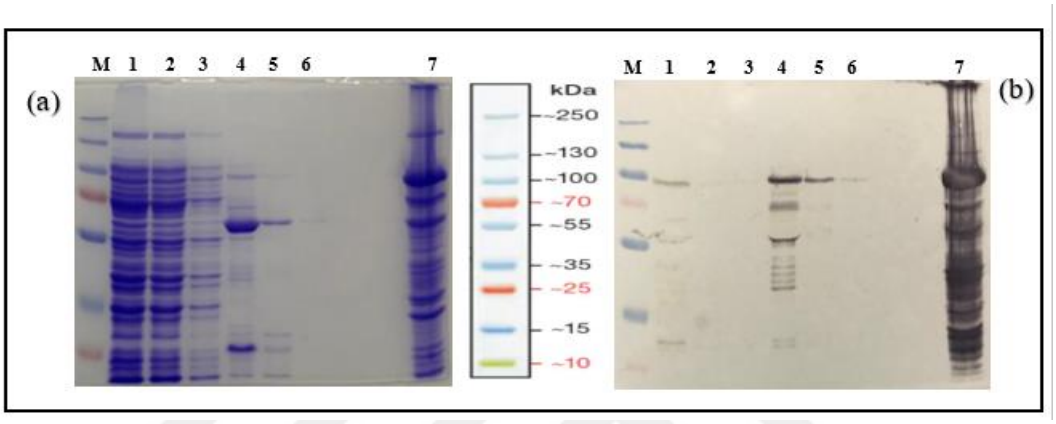
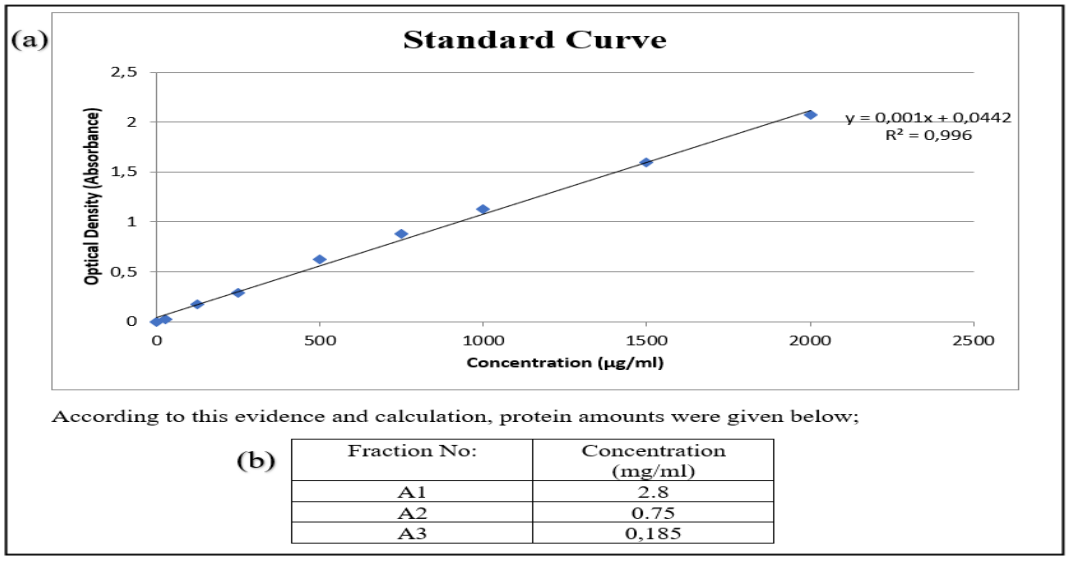


Figure 4.16: SDS-PAGE and Western Blot images of purification samples. 1- Sample load, 2- Unbound samples, 3- Wash samples, 4, 5, 6- Elution samples (A1, A2 and A3), 7- Insoluble samples (a) and (b).



Fraction No:	Concentration (mg/ml)
A1	2.8
A2	0.75
A3	0,185

Figure 4.17: BCA results of purified S1 protein. The standard curve was prepared using BSA standards (a). Protein concentration was calculated with an equation obtained from the standard curve (b).

In the purification of S1 protein, SDS-PAGE and Western blot analyses showed that protein loss during purification was minimal. Although the S1 protein was obtained in the expected size (83 kDa) in the elution samples, it was also observed that there were cellular proteins rather than the target protein in the elution samples, especially in A1 (Figure 4.16).

Concentration of all 3 fractions that belongs to elution samples were measured by BCA technique. The decrease in protein amount in elution samples was visually demonstrated by SDS-PAGE and Western blot. In addition, the BCA method was used to determine the concentration, and the decreasing protein concentration in the fractions was recorded numerically (Figure 4.17).

4.1.6. Cleavage of SUMO from SUMO Fused S1 Protein

To obtain the native S1 protein, SUMO protease was used to cleave the N-terminal peptide containing the 6XHis tag and SUMO. After removal of the N-terminal peptide, the S1 protein, known to have a size of 70 kDa, and the 6XHis tag and SUMO fragment (13 kDa) are expected to be cleaved.

When the cleavage experiment was conducted, purified S1 protein samples which is stored at -20 ℃ were found to be degraded. Therefore, the cleavage of SUMO experiments could not be completed.

4.2. Expression of S1 Protein with DsbC Signal Peptide

In this thesis, the effect of DsbC signal peptide on the soluble production of SARS-CoV-2 S1 protein was investigated. For this purpose, SHuffle cells containing DsbC signal peptide in their genotype were used for expression control.

4.2.1. Transformation of S1 gene Ligated pET30a (+) Plasmid

100 ng of isolated plasmid was transformed into *E. coli*. SHuffle strain. The cells were spread on 2xYT agar media, and transformants were obtained after overnight incubation at 37 °C. Colonies were selected from the agar plate and confirmed by colony PCR (Figure 4.18).

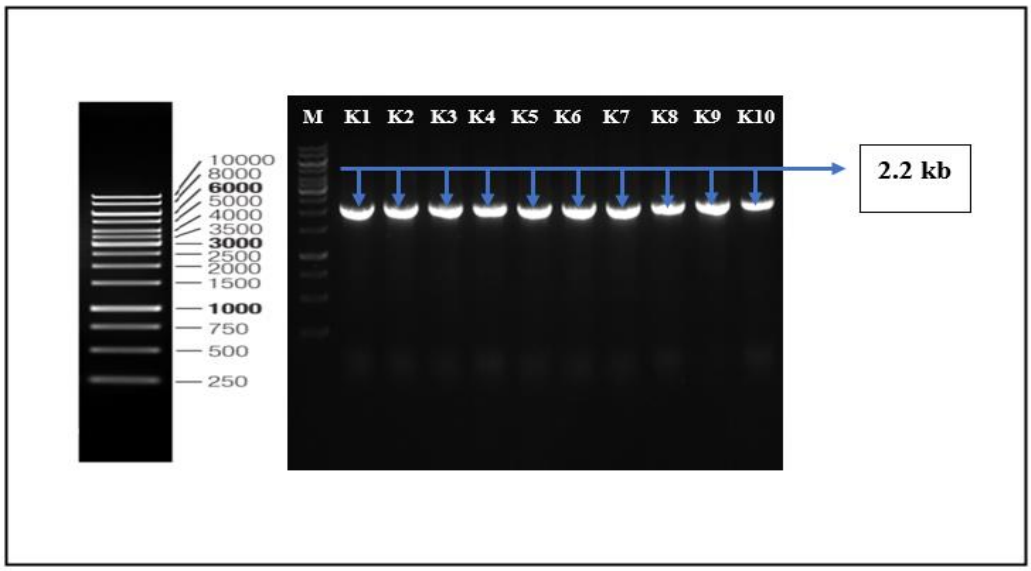


Figure 4.18: Result of colony PCR for the SHuffle strain. (M) GeneRuler 1 kb DNA Ladder, (1, 2, 3, 4, 5, 6, 7, 8, 9, 10) PCR products.

According to colony PCR results, all selected colonies bear gene of interest (2.2 kb). In the next step, all of the selected colonies were evaluated to determine SARS-CoV-2 S1 protein expression.

4.2.2. Expression of S1 Protein with DsbC Signal Peptide

Utilizing PCR-selected transformants, S1 expression was evaluated. For that purpose, transformants were grown overnight at 37 °C in 5 ml of TB medium, and expression cultures were grown in 20 ml of TB medium. After cultures were induced with 0.4 mM IPTG, they were incubated at 30 °C or 16 °C. After five hours, cells incubated at 30 °C were harvested for cell lysis to obtain total protein. Cells incubated at 16 °C were harvested after 24 hours. With centrifugation, soluble and insoluble proteins were separated from one another. The total protein samples were analyzed with SDS-PAGE (Figure 4.19).

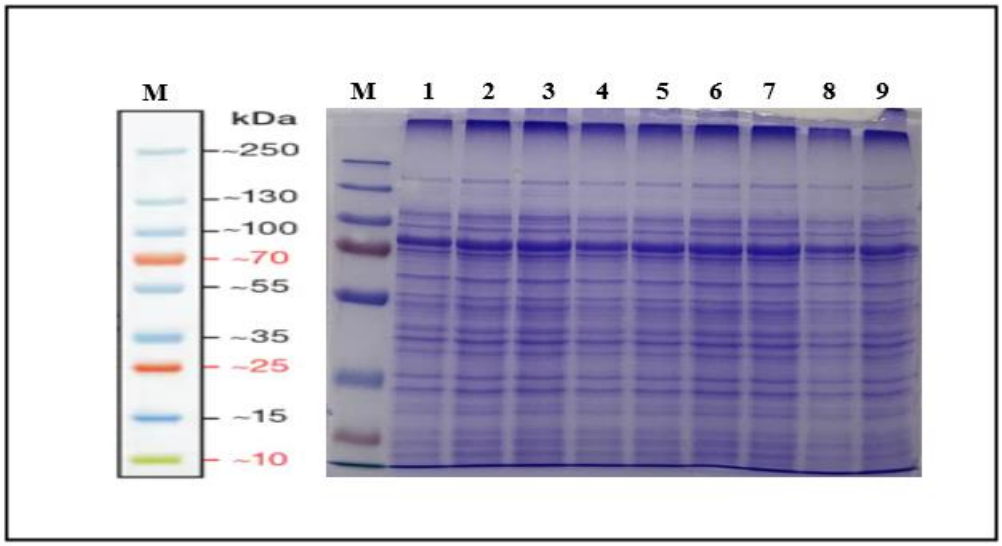


Figure 4.19: Total protein samples expressed in SHuffle cell.

When the SDS-PAGE results were examined, it was observed that the target protein was obtained at the expected size (~70 kDa). There was no significant difference in expression level between colonies when the expression was performed at 30 ℃ in SHuffle cells.

SDS-PAGE and Western Blot analyses of the total, soluble, and insoluble protein samples from randomly selected colonies were performed (Figure 4.20).

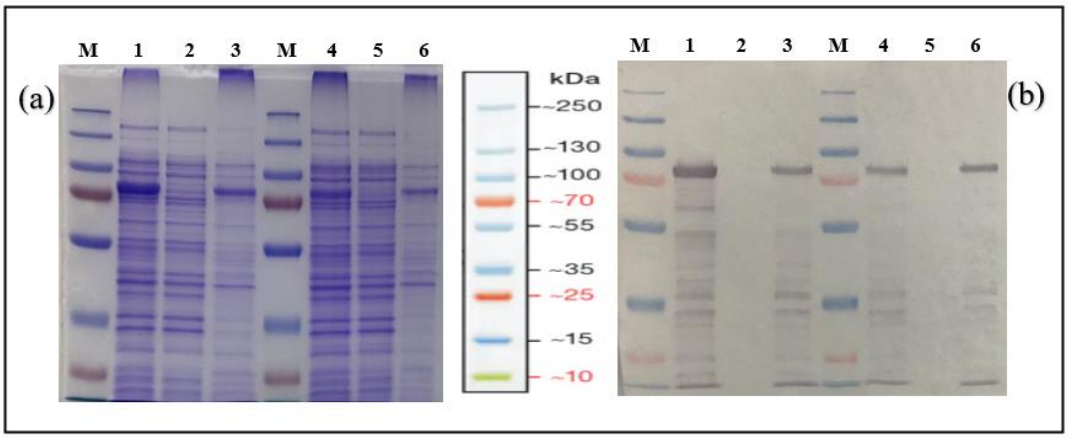


Figure 4.20: Samples of total, soluble and insoluble proteins belonging to 16 ℃ and 30 ℃ expressions. Examples 1,2 ,3 show the expression made at 16 ℃, including total, soluble and insoluble proteins, respectively (a) and (b). Examples 4,5 and 6 show the expression at 30 ℃ and include total, soluble and insoluble proteins, respectively.

When the SDS-PAGE and Western blot analyzes were evaluated, it was observed that the total protein amount in the culture incubated overnight at 16 °C was higher than that obtained after incubation at 30 °C for 5 hours. When considering soluble and insoluble protein samples, it was observed that the S1 protein was produced as insoluble in both incubation conditions.

5. DISCUSSION AND CONCLUSION

SARS-CoV-2 is the newest member of the *Coronaviridae* family that causes the Covid-19 disease. Of the four main protein subunits of SARS CoV 2, the Spike protein, which helps the coronavirus enter the cell, has subunits S1 and S2 [Shang., 2020].

Recombinant production of proteins, including SARS-CoV-2 proteins, using mammalian, bacterial, yeast, and baculovirus expression systems are commonly used systems [Tai et al., 2020]. Bacterial cells are chosen as a recombinant protein expression because they offer inexpensive growth conditions, short life cycle, well-known genetics, and easy genetic manipulation [Sahdev et al., 2008]. There are studies involving the production of SARS-CoV-2 proteins using the bacterial expression system [Prahlad et al., 2021].

The most significant disadvantage of producing recombinant proteins using the bacterial expression system is that the resulting protein is insoluble [Butt et al., 2005]. In order to solve this problem, signal peptides and fusion tags have been widely used [Lee et al., 1996]. In this study, we aimed to produce SARS-CoV-2 S1 protein as soluble with a SUMO fusion tag and a DsbC signal peptide.

This study aims to produce SARS-CoV-2 S1 protein recombinantly in the *E.coli* expression system. In order to obtain the protein as soluble, firstly, the S1 gene was cloned into the pET SUMO vector containing the SUMO fusion tag. Plasmids containing the S1 gene, whose sequence information was confirmed by the sequencing method, were transformed into SHuffle, BL21 (DE3), and BL21 Star(DE3) pLysS strains of *E.coli* for expression control. Expression control was performed at different temperatures, IPTG concentrations, and time intervals. Although the expression level of S1 protein was high in the BL21 (DE3) strain, S1 protein was produced as insoluble. It was observed that the S1 protein expressed on the pET SUMO vector was partially soluble as a result of incubation in TB medium at 16 ℃ with 140 rpm shaking for 24 hours with *E.coli* SHuffle strain.

In small-scale expression studies, although the S1 protein was partially soluble, it was observed that the solubility of the S1 protein was lost when large-scale production was made.

Elutions at low concentrations were obtained from purification by affinity chromatography of the S1 protein, which was produced as soluble when expressed with the SUMO fusion tag. It was determined that non-specific proteins were intensely present in these elution samples. In addition to affinity chromatography, ion exchange chromatography and size exclusion chromatography can be used to remove non-specific proteins.

Previous studies with the SUMO fusion tag have shown it to increase soluble protein production. However, when we look at the studies, it is seen that the proteins desired to be produced as soluble are of small sizes [Butt et al., 2005], [Pratheesh et al., 2019], [Lu et al., 2021]. In contrast to the studies that have been published previously, we used the S1 protein, which is a protein that is significantly larger (70 kDa). Because of this, it is possible that the size of the protein is what causes it to be produced in a partially soluble form for the S1 protein.

As another approach for the soluble production of S1 protein, the expression of the S1 gene with the DsbC signal peptide was examined. To achieve this, SHuffle cells containing the DsbC signal peptide were transformed with the pET30a (+) plasmid containing the SARS-CoV-2 S1 gene. Expression control was performed at different temperatures, IPTG concentrations, and time intervals. Low level production of SARS-CoV-2 S1 protein was achieved with different expression conditions. In addition to this, S1 protein was produced as insoluble in all tested conditions.

There are a variety of published strategies for the production of recombinant proteins using Dsb systems. The combined use of DsbA and DsbC signal peptides causes Dsb systems to be more effective in protein expression. DsbA is responsible for the formation of disulfide bonds, while DsbC is responsible for their isomerization [Gleiter et al., 2008]. The combination of two signal peptides can provide soluble expression of the S1 protein.

REFERENCES

Arechaga I., Miroux B., Runswick M. J., Walker J. E., (2003), "Over-expression of Escherichia coli F1Fo–ATPase subunit a is inhibited by instability of the uncB gene transcript", FEBS letters, 547 (1-3), 97-100.

Baneyx F., (1999), "Recombinant protein expression in Escherichia coli", Current opinion in biotechnology, 10 (5), 411-421.

Butt T. R., Edavettal S. C., Hall J. P., Mattern M. R., (2005), "SUMO fusion technology for difficult-to-express proteins", Protein expression and purification, 43 (1), 1-9.

Correa A., Oppezzo P., (2015), "Overcoming the solubility problem in *E. coli*: available approaches for recombinant protein production", In Insoluble proteins 27-44.

Cui Z., Chang H., Wang H., Lim B., Hsu C. C., Yu Y., Huang W. E., (2020), "Development of a rapid test kit for SARS-CoV-2: an example of product design", Bio-design and Manufacturing, 3 (2), 83-86.

Gleiter S., Bardwell J. C., (2008), "Disulfide bond isomerization in prokaryotes", Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 1783 (4), 530-534.

Guerrero F., Ciragan A., Iwaï H., (2015), "Tandem SUMO fusion vectors for improving soluble protein expression and purification", Protein expression and purification, 116, 42-49.

Hu B., Guo H., Zhou P., Shi Z. L., (2021), "Characteristics of SARS-CoV-2 and COVID-19", Nature Reviews Microbiology, 19 (3), 141-154.

Johnson E. S., (2004), "Protein modification by SUMO", Annual review of biochemistry, 73 (1), 355-382.

Kuo D., Nie M., Courey A. J. (2014), "SUMO as a solubility tag and in vivo cleavage of SUMO fusion proteins with Ulp1", In *Protein Affinity Tags* 71-80.

Lan J., Ge J., Yu J., Shan S., Zhou H., Fan S., Wang X., (2020), "Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor", Nature, 581 (7807), 215-220.

Lee S. Y., (1996), "High cell-density culture of Escherichia coli", Trends in biotechnology, 14 (3), 98-105.

Lu X., Hu S., Li Q., Song X., Zhou L., Wang Y., (2021), "Expression, purification and antimicrobial activity analysis of recombination peptide subtilosin A in Escherichia coli using SUMO fusion technology".

Marblestone J. G., Edavettal S. C., Lim Y., Lim P., Zuo X. U. N., Butt T. R., (2006), "Comparison of SUMO fusion technology with traditional gene fusion systems: enhanced expression and solubility with SUMO", Protein science, 15 (1), 182-189.

Pratheesh P. T., Nimisha S., Jess V., Asha K., Agarwal R. K., (2019), "Expression and purification of an immunogenic SUMO-OmpC fusion protein of Salmonella Typhimurium in Escherichia coli", Biologicals, 62, 22-26.

Prahlad J., Struble L. R., Lutz W. E., Wallin S. A., Khurana S., Schnaubelt A., Borgstahl G. E., (2021), "CyDisCo production of functional recombinant SARS-CoV-2 spike receptor binding domain", Protein Science, 30 (9), 1983-1990.

Sahdev S., Khattar S. K., Saini K. S., (2008), "Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies", Molecular and cellular biochemistry, *307* (1), 249-264.

Shang J., Wan Y., Luo C., Ye G., Geng Q., Auerbach A., Li F., (2020), "Cell entry mechanisms of SARS-CoV-2", Proceedings of the National Academy of Sciences, 117 (21), 11727-11734.

Tai W., He L., Zhang X., Pu J., Voronin D., Jiang S., Du L., (2020), "Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine", Cellular & molecular immunology, 17 (6), 613-620.

Zhao R., Li M., Song H., Chen J., Ren W., Feng Y., Sun L., (2020), "Early detection of SARS-CoV-2 antibodies in COVID-19 patients as a serologic marker of infection", Clinical infectious diseases.

BIOGRAPHY

Melike Yağmur ÜNAL was graduated from Gebze Technical University, Department of Molecular Biology and Genetics. She started her MSc. education in Gebze Technical University, Graduate School of Natural and Applied Sciences, Department of Molecular Biology and Genetics. She performed her studies under the supervision of Prof. Dr. Tamer YAĞCI and co-supervisor Dr. Hasan Ümit ÖZTÜRK.