

**AN ENGINEERED GENETIC CIRCUIT FOR PROTEIN
GLYCOSYLATION IN *ESCHERICHIA COLI***

**A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF
MASTER OF SCIENCE
IN
MATERIALS SCIENCE AND NANOTECHNOLOGY**

By

ERAY ULAŞ BOZKURT

MAY 2021

**AN ENGINEERED GENETIC CIRCUIT FOR PROTEIN GLYCOSYLATION
IN *ESCHERICHIA COLI***

By Eray Ulaş Bozkurt

May 2021

We certify that we have read this dissertation and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Urtu Özgür Şafak Şeker (Advisor)

Fatih İnci

Ceren Özkul Koçak

Approved for the Graduate School of Engineering and Science:

Ezhan Karaşan

Director of the Graduate School

ABSTRACT

AN ENGINEERED GENETIC CIRCUIT FOR PROTEIN GLYCOSYLATION IN *ESCHERICHIA COLI*

Eray Ulaş Bozkurt

M.Sc. in Materials Science and Nanotechnology

Advisor: Urartu Özgür Şafak Şeker

May 2021

Protein glycosylation is one of the most crucial and common post-translational modifications. Glycosylation provides certain advantages to host organisms and extend the proteome beyond genetic material. After the discovery of bacterial glycosylation mechanisms and especially after its transfer into laboratory work-horse *E. coli*, studies utilizing this mechanism increased exponentially. It has been previously showed that utilizing *N-Linked Glycosylation*, certain recombinant proteins have been furnished with improved features, such as stability and solubility. In this study, we utilized *N-linked Glycosylation* to glycosylate alkaline phosphatase (ALP) enzyme in *E. coli* and investigate the effects of glycosylation on an enzyme.

Considering the glycosylation mechanism is highly dependent on the acceptor protein, ALP constructs carrying glycosylation tag at different locations of the gene has been created and glycosylation rates have been calculated. The most glycosylated construct has been selected for comparison with the native enzyme. Studies showed that glycosylated ALP performed better at optimal conditions. In order to extend the knowledge on the differences due to glycosylation, several conditions were applied.

Both enzymes were tested at elevated temperatures for different incubation times, different pH conditions, protease treatment and under denaturing conditions. Also, secondary structure analysis was performed for each condition to elaborate on these differences. Experiments showed that glycosylated ALP performs remarkably better at all conditions tested. Therefore, *N-linked Glycosylation* mechanism can be employed for enzyme engineering purposes and is a useful tool for industrial applications that require enzymatic activity.

Keywords: Protein Engineering, Enzyme Kinetics, *N-linked Glycosylation*

ÖZET

ESCHERICHIA COLI BAKTERİSİNDE PROTEİN

GLİKOLİZLENMESİ İÇİN DEVRE TASARIMI

Malzeme Bilimi ve Nanoteknoloji, Yüksek Lisans

Tez Danışmanı: Urartu Özgür Şafak Şeker

Mayıs, 2021

Protein glikozilasyonu, proteinlerin translasyon sonrası değişimleri arasında en önemli ve en çok görülen değişimlerden biridir. Glikozilasyon, hücre proteome'una büyük bir çeşitlilik ve pek çok açıdan avantaj sağlar. Bakteriyel glikozilasyon mekanizmalarının keşfedilmesi ve laboratuvarlarda en çok kullanılan *E. coli* bakterisine aktarılmasından sonra, bu konu üzerine yapılan çalışmalar hız kazanmıştır. Bugüne kadar yapılan çalışmalarda, bazı proteinlerin glikolizlenme mekanizması kullanılmasıyla özelliklerinin geliştirildiği not edilmiştir. Bu özelliklere proteinlerin stabilitesi ve solübilitesi örnek verilebilir. Bu çalışmada, glikolizasyon mekanizması kullanılarak glikolizlenmiş *E. coli* alkalın fosfataz enzimi üretilmiş ve üretim sonucu oluşan rekombinant protein, doğal hali ile kıyaslanarak incelenmiştir.

Bakteriyel glikozilasyon mekanizmasının verimliliği, şeker grubunu kabul eden proteine göre yüksek değişkenlik göstermesi göz önüne alınarak, alkalın fosfataz enziminin birden fazla türevi oluşturulmuştur. Bu türevlerde alkalın fosfataz enzimi, şeker grubunu üzerinde bulunduracak olan glikolizasyon bölgesinin proteinin farklı bölgelerine konulmuştur. En fazla glikozilasyon verimliliğini gösteren türev, doğal alkalın fosfataz enzimi ile karşılaştırılmak için seçilmiştir. Yapılan çalışmalar, ideal şartlar altında, glikozile alkaline fosfataz enziminin, glikozile olmamış alkalın

fosfataz enzimine göre daha iyi performans gösterdiği gözlenmiştir. Aradaki farkı daha iyi gözlemlemek için çeşitli şartlar altında enzim performansları incelenmiştir.(farklı sürelerde ve farklı sıcaklıklarda inkübasyon, farklı pH şartları, proteaz muamelesi ve denaturasyon şartları). Her deney grubu için, proteinlerin ikincil yapılarında meydana gelen değişimler Dairesel Dikroizm (Circular Dichroism) ile takip edilmiştir. Test edilen her durumda, glikozile alkalın fosfataz enzimi, glikozile olmayan enzime göre daha iyi çalışmıştır. Sonuç olarak, glikozilasyon mekanizmalarının enzim mühendisliği çalışmalarında ve özellikle enzim aktivitesi gerektiren endüstriyel süreçlerde kullanılmasının avantajlı olabileceği gösterilmiştir.

Anahtar Kelimeler: Protein Mühendisliği, Enzim Kinetiği, *N-bağlantılı Glikozilasyon*,

ACKNOWLEDGEMENTS

First and foremost, I would like to express my gratitude to Dr. Urartu Özgür Şafak Şeker. I started my research career in his lab as an undergraduate 4 years ago, knowing nothing. During this period, I was honored to work as a group member of his, benefitting his knowledge, experience and guidance in both research and life. I would like to thank him for giving me the opportunity to work with him and providing an environment which I can improve myself and build my career. I would like to thank the jury members for reviewing my work and valuable comments to take my research to the next step.

During my undergraduate years and as a freshman in M.Sc., I received great help and support from Dr. Behide Saltepe and Dr. Ebru Şahin Kehribar. They were my mentors in the lab and I am happy that I achieved the honor of having publications with both of them. I was great pleasure working alongside them and I would like to thank you for their guidance and patience. I would like to thank İrem Niran Çağıl for contributing my thesis work. She was a quick-learner and she achieved to learn and contributed a lot to my thesis. I also would like to thank to Musa Efe Işılak, Murat Alp Güngen and Nedim Hacıosmanoğlu who spent time on our work and contributed the publications. I also would like to thank to Recep Erdem Ahan for his constructive comments and suggestions on every project I was working on. He was there all the time to answer my questions. Lastly, Gizem Makas, our lab manager was working closely with us to keep things running smoothly, I would like to thank her for her cooperation and support.

During this journey, I was fortunate to build great relationships as well. Firstly, I would like to thank Behide Saltepe for her friendship. We spent lots of time together.

She made tiring and long-lasting experiments fun. She was there for me when I was down and it would be much harder, if she was not around. I cannot thank her enough for her support and help. I also would like to thank Cemile Elif Özçelik for our nice conversations during breaks. Especially, I really enjoyed our coffee routines during the day which helped me to keep my mind out of work. I also would like to thank Nilsu Turay for her good companionship. We always had great conversations and had lots of fun during the little time we could spare. Çisil Köksaldı and Gökçe Özkul was very nice to me all the time and I feel lucky to have them as friends.

I also would like to thank the past and current SBL group members (Ahmet Hınçer, Anooshay Khan, Büşre Merve Kırpat, Dr. Ebru Aras, Goozel Binte Shadid, Julian Ostaku, Merve Erden Tüçer, Merve Yavuz, Nedim Kurt, Sıla Köse, Suat Tüçer) for their cooperation. I also would like to thank my iGEM Team members (Artun Bülbül, Azra Atabay, Berk Ürgen, Büşra Nur Ata, and Ezgi Dikici). We started together to learn synthetic biology and made many contributions to each other during this process. We also worked as a team for two years and accomplished a lot. I really enjoyed being a part of such team.

Lastly, I would like to express my deepest gratitude to my family for their endless support. I would not be able to be who I am, if it was not for them. I would like to thank my mother, Nuran Bozkurt, my father, Yaşar Bozkurt and my sister, Seray Başak Bozkurt.

This project was partially supported by TÜBİTAK Project Number 216M127.

Table of Contents

CHAPTER 1	1
1. Introduction	1
1.1. Glycosylation as a Post Translational Modification	1
1.2. Bacterial Glycosylation	3
1.3. <i>Campylobacter jejuni</i> N-Linked Glycosylation	6
1.4. Utilization of Glycosylation Systems in Synthetic Biology	9
1.5. The Aim of the Study	10
CHAPTER 2	12
2. Materials and Methods	12
2.1. Buffers, Media and Strains	12
2.2. Cloning and Transformation	18
2.3. Protein Expression	19
2.4. SDS-PAGE, Western Blot and Lectin Blot	20
2.5. Glycosylation Rate Calculations	22
2.6. Protein Purification and Buffer Exchange	22
2.7. Unit Enzyme Calculations	24
2.8. Alkaline Phosphatase Activity Assays	24
2.8.1. Enzyme Activity at Optimal Conditions	24
2.8.2. Preparation of Enzymes for Assessment of Enzyme Activity at Elevated Temperatures	25
2.8.3. Preparation of Enzymes for Assessment of Enzyme Activity at Different pH Values	25
2.8.4. Preparation of Enzymes for Assessment of Enzyme Activity upon Protease Treatment	25
2.8.5. Preparation of Enzymes for Assessment of Enzyme Activity under Denaturant Conditions	26
2.9. Circular Dichroism	26
2.9.1. Acquisition of CD Data for ALP at Elevated Temperatures	27
CHAPTER 3	28
3. Results and Discussion	28

3.1.	Design and Cloning of the Constructs	28
3.2.	Expression and Purification of ALP Variants	30
3.3.	Glycosylation Rate Calculation	32
3.4.	Overexpression of PglB Enzyme	33
3.5.	Enzyme Activity Comparison of Nonglycosylated and Glycosylated ALP	41
3.5.1.	Enzyme Activity and Secondary Structure Assessment at Elevated Temperatures	45
3.5.2.	Enzyme Activity and Secondary Structure Assessment at Different pH	52
3.5.3.	Enzyme Activity Assessment upon Protease Treatment	55
3.5.4.	Enzyme Activity and Secondary Structure Assessment under Denaturant Conditions	57
CHAPTER 4		58
4.	Conclusion	58
4.1.	References	60
4.2.	Appendix A	63
4.2.1.	DNA Sequences of Genes	63
4.3.	Appendix B	84
4.3.1.	List of Primers	84
4.4.	Appendix C	85
4.4.1.	Plasmid Maps	85
4.5.	Appendix D	91
4.5.1.	Sanger Sequencing Results	91
4.6.	Appendix E	92
4.6.1.	Amino acid sequences of Proteins	92

LIST OF FIGURES

Figure 1: Post-Translational Modifications shared by all life forms. Created with BioRender.com. Adapted from Wang et al.[1]	1
Figure 2: N-linked glycosylation mechanism in eukaryotes. Created with Biorender.com	3
Figure 3: <i>C. jejuni pgl</i> glycosylation pathway. Created with Biorender.com. Adapted from Tan et al.	7
Figure 4: PCR result of pglB amplification.	28
Figure 5: Digest Result of ALP DQNAT pET22B.	29
Figure 6: Colony PCR results of T7 pglB ALP-DQNAT pET22b Amp.	29
Figure 7: Created ALP Variants was created to assess Glycosylation Rate.	30
Figure 8: Western Blot Results of Created ALP Variants.	31
Figure 9: Lectin Blot Results of created ALP Variants.	31
Figure 10: Calculated glycosylation rate by Vilber Evolution Edge software. (n=3)	32
Figure 11: The representation of pglB overexpressing ALP variant, T7 pglB ALP-DQNAT + pgl-pACYC.	33
Figure 12: pglB Overexpressing version of ALP variant performs better in terms of glycosylation.	34
Figure 13: PglB overexpression enhanced the glycosylation rate.	35
Figure 14: Lectin Blot of pglB ALP-DQNAT indicates that ALP-DQNAT is glycosylated.	35

Figure 15: Putative <i>N-linked</i> glycosylation analysis results performed by GlycoPP software.	37
Figure 16: SDS-PAGE result of the purified ALP-DQNAT and Glycosylated ALP-DQNAT	39
Figure 17: Western Blot result of ALP-DQNAT and Glycosylated ALP-DQNAT	40
Figure 18: Lectin Blot results of purified ALP-DQNAT and Glycosylated ALP-DQNAT	40
Figure 19: Comparison of ALP and Glycosylated ALP. Glycosylated ALP obtained higher catalysis activity.	42
Figure 20: Normalized enzyme activities of ALP-DQNAT and Glycosylated ALP-DQNAT at optimal conditions.	43
Figure 21: Secondary structure analysis performed to assess the differences between non-glycosylated and glycosylated ALP and predicted secondary structures.	44
Figure 22: ALP-DQNAT and Glycosylated ALP-DQNAT are treated with varying temperatures.	45
Figure 23: Enzyme Kinetics of ALP-DQNAT and Glycosylated ALP-DQNAT at 55°C for different treatment times.	47
Figure 24: Enzyme Kinetics of ALP-DQNAT and Glycosylated ALP-DQNAT at 75°C for different treatment times.	48
Figure 25: Enzyme Kinetics of ALP-DQNAT and Glycosylated ALP-DQNAT at 95°C for different treatment times.	49
Figure 26: Secondary structure analysis of ALP-DQNAT and Glycosylated ALP-DQNAT at elevated temperatures.	51
Figure 27: Enzyme activity results of ALP-DQNAT and Glycosylated ALP-DQNAT at different pH conditions.	52

Figure 28: Enzyme Kinetics of ALP-DQNAT and Glycosylated ALP-DQNAT at different pH values.	54
Figure 29: Proteinase K treatments of ALP enzymes for different incubation times.	55
Figure 30: Effect of protease treatment on the enzyme activity of ALP-DQNAT and Glycosylated ALP-DQNAT.	56
Figure 31: The plasmid map of pgl-pACYC	86
Figure 32: The plasmid map of DQNAT-ALP pET22b	87
Figure 33: The plasmid map of ALP-DQNAT pET22b	88
Figure 34: The plasmid map of ALP-DQNAT (x2) pET22b	89
Figure 35: The plasmid map of T7 pglB ALP-DQNAT pET22b	90
Figure 36: Sanger sequencing result of DQNAT-ALP pET22b	91
Figure 37: Sanger sequencing result of ALP-DQNAT pET22b	91
Figure 38: Sanger sequencing result of T7 pglB ALP-DQNAT pET22b	92

LIST OF TABLES

Table 1: Most studied bacterial glycosylation mechanisms. Adapted from <i>Lu et al.</i>	4
Table 2: LB media ingredients required for 1L.	12
Table 3: ZYM5052 medium ingredients required for 1L.	12
Table 4: 1L of 1M TAE (pH 8.6) buffer ingredients	13
Table 5: 50 ml of TSS buffer ingredients and preparation	13
Table 6: 10 mM Imidazole (pH 7.4) binding buffer ingredients	14
Table 7: 150 mM Imidazole (pH 7.4) elution buffer ingredients	14
Table 8: 1L of 0.5M MES buffer (pH 6) ingredients and preparation	14
Table 9: 100 ml of 1M Tris-HCl (pH 7.6) ingredients and preparation	15
Table 10: 50 ml of 0.1 M Sodium Phosphate buffer (pH 7.4) ingredients and preparation	15
Table 11: 6x Loading Dye preparation	16
Table 12: 1L of 1X Tris Buffered Saline Tween (TBS-T) ingredients	16
Table 13: pNPP reaction buffer ingredients	17
Table 14: 100 ml of 8M Urea buffer ingredients	17
Table 15: Experimental groups utilized in the study	Error! Bookmark not defined.
Table 16: Experimental groups for the assesment of enzyme activity upon protease treatment.	Error! Bookmark not defined.
Table 17: The enzyme behaviours calculated for ALP-DQNAT and Glycosylated ALP-DQNAT from enzyme activity assays.	42
Table 18: The secondary structure analysis results at different pH conditions.	54

Table 19: The DNA sequences of genes that was used in this work.	63
Table 20: Primers that were used in this study.	84

CHAPTER 1

1.Introduction

1.1. Glycosylation as a Post Translational Modification

The majority of the proteins in eukaryotes is processed after they are formed by ribosome. This process is called Post-Translational Modifications(PTMs). PTMs allow a diversification and create a variety of options in terms of what has been coded by the genome.[1] These modifications play an important role on the biochemical properties of the proteins that are newly formed. They regulate the localization, activity and the association with other molecules.

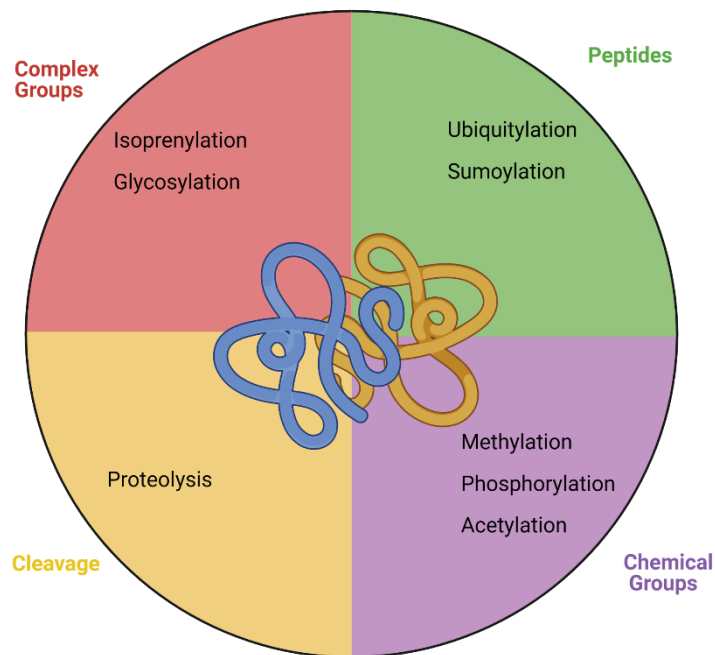


Figure 1: Post-Translational Modifications shared by all life forms. Created with BioRender.com. Adapted from Wang et al.[1]

One of the important modifications that cells use to regulate their proteins is glycosylation. Glycosylation is a so common post-translational modification that 70% of all eukaryotic proteins are known to be glycosylated.[2] Protein glycosylation refers to the attachment of sugar residues to proteins has been thought to be the most complex post-translational modification. Glycosylation is a common phenomenon for all domains of life. Glycosylation was first discovered by Neuberger and colleagues in ovalbumin by defining GlcNAc- β -Asn linkage in 1961 in hen's egg. After that, in 1989 Wieland and colleagues have discovered that glycosylation in *Halobacterium halobium*. Furthermore, with the discovery of glycosylation mechanisms in *archaea* in 1997, Glycosylation is now considered a common phenomenon for all domains of life. [3]

Today, it is a well-known fact that protein glycosylation is crucial in formation of protein structure, signal transduction, immune responses, embryonic development, interaction between cell and its environment. [4-9] Glycotransferases and glycosidase enzymes which performs the reaction that attaches carbohydrate groups to proteins are found in endoplasmic reticulum (ER) and Golgi apparatus. [1] Glycosylation can be classified into four main categories in mammalian cells: *N*-linked glycosylation, *O*-linked glycosylation *C*-linked glycosylation and glypiation.[10] However, the most common one is the *N*-linked glycosylation and shared by three domains of life.

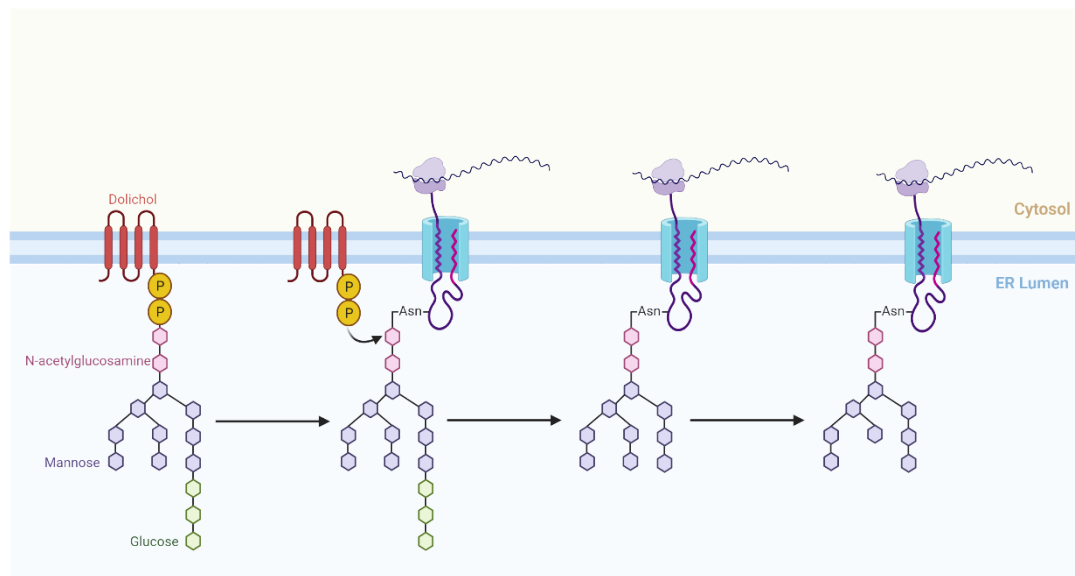


Figure 2: N-linked glycosylation mechanism in eukaryotes. Created with Biorender.com

In eukaryote, dolichol, a polymer of isoprene units that consists the lipid-linked oligosaccharide (LLO) is built on the cytoplasmic side of the cell to the ER lumen. ER lumen contains multi subunit oligosaccharyltransferase enzymes and utilize the

glycans are transferred *en bloc* to the Asparagine residue on the nascent protein during translation.

1.2. Bacterial Glycosylation

In the past decade, protein glycosylation in bacteria has become a hot-spot for glycobiology. With the increasing evidence of protein glycosylation in bacteria, including *Campylobacter jejuni*, *Pseudomonas aeruginosa*, *Neisseria meningitidis*, *Streptococcus parasanguis*, *Mycobacterium tuberculosis* *Haemophilus influenzae*. [11-17], it became obvious that pathogenicity was common characteristics.

Therefore, glycosylation in bacteria was associated with virulence factors.

Glycosylation mechanisms in bacteria contributes to adhesiveness of the outer surface. Therefore, pathogenic bacteria can hold on host surfaces.[18]

Like eukaryotes, glycosylation occurs in variety of ways in bacteria as well:

N-Linked Glycosylation, *O-Linked Glycosylation* and *S-Linked Glycosylation*.

N-Linked Glycosylation refers to attachment of sugar moieties to the nitrogen atoms of asparagine or arginine amino acids. *O-linked Glycosylation* means that sugar groups are added to the hydroxy groups of serine or threonine residues and *S-Linked Glycosylation* refers to attachment of sugars onto thiol groups on cysteines.[19]

Table 1: Most studied bacterial glycosylation mechanisms. Adapted from *Lu et al.*

Organism	Glycosyltransferase	Linkage	Residues Modified	Glycan
<i>Haemophilus influenzae</i>	HMW1C and HMW2C	N	Asparagine	Hex and di-Hex (Hex: Glc or Gal)
<i>Campylobacter jejuni</i>	PglB	N	Asparagine	Heptasaccharide
EPEC	NleB	N	Arginine	GlcNAc
<i>Clostridium</i> spp.	TcdA/B and the large <i>Clostridial</i> cytot oxin family	O	Threonine	Glc or GlcNAc

<i>Photorhabdus asymbiotica</i>	PaTox	O	Tyrosine	GlcNAc
<i>Legionella pneumophila</i>	Lgts	O	Serine	Glc
DAEC/ETEC/<i>Citrobacter rodentium</i>	AAH/TibC/BAHT _{cr} (the BAHT family)	O	Serine	Hep
<i>Streptococcus parasanguinis</i>	Gtf1/Gtf2 complex	O	Serine	GlcNAc
	Gtf3	O	GlcNAc	Glc
	dGT1	O	Glc	Glc

N-Linked Glycosylation and *O-Linked Glycosylation* are the most common pathways in bacteria kingdom. One of the greatest examples of utilization of *N-linked glycosylation* mechanism in bacteria is *Haemophilus influenza* (*H. influenzae*). *H. influenzae*, the cause of the disease called nontypeable *H. influenza* initiates the infection by colonizing the epithelial cells of the host. *H. influenzae* utilizes two adhesin proteins called HMW1 and HMW2. HMW1 and HMW2 proteins are secreted to outer membrane and play significant role for attaching to the host surface.[20] Another outstanding example is how *C. jejuni* utilize this mechanism. On the contrary what has been taught previously, the main purpose of the *C. jejuni* glycosylation is not adhesiveness. *C. jejuni* has over 60 proteins that has been observed to be glycosylated. There are reports suggesting that these proteins may contribute to the fitness of the bacterium in the gut by providing protection from proteases. Furthermore, free oligosaccharides may play role in osmotic regulation of the bacterium.[18]

O-Linked Glycosylation is also being employed by *Clostridium difficile* as an effective virulence factor. *C. difficile*, the cause of infectious diarrhea encodes for two toxin proteins called TcdA and TcdB. These proteins have N-terminal glucosyltransferase activity. Once these toxins were internalized by endocytosis, they confer their toxicity by inhibiting Rho GTPase activation. Therefore, they interfere with host's membrane cycling process, leading the death of the host cell.[21]

1.3. *Campylobacter jejuni* N-Linked Glycosylation

C. jejuni was the first discovered bacterium that is found to *N-glycosylate* proteins.

C. jejuni utilizes a heptasaccharide glycan

(GalNAc- α 1,4-GalNAc- α 1,4-[Glc β 1,3]-GalNAc- α 1,4-GalNAc- α 1,3-diNacBac- β 1; where diNacBac is N',N'-diacetyl bacillosamine [2,4-diacetamido-2,4,6-trideoxyglucopyranose]). The operon that performs *N-glycosylation* is called *pgl*

(protein glycosylation). *pgl* pathway is composed of 12 genes that completes sequential addition of glycan onto its acceptor protein. It is initiated by formation of uridine diphosphate-di-N-acetyl-bacillosamine (UDP-diNacBac). PglF enzyme converts UDP-GlcNAc to UDP-2-acetamido-2,6-dideoxy- α -D-xyl-4-hexulose. Then, pglE forms UDP-4-amino-4,6-dideoxy- α -D-GlcNAc. PglD enzyme transfers the acetyl group onto product of pglE, forming UDP-diNacBac.

The formed UDP-diNacBac is linked to Und-P by phosphoglycosyltransferase PglC by a pyrophosphate bond. PglA and PglJ perform sequential reactions and catalyze the addition of α 1,3- and α 1,4-linked N-acetylgalactosamine. Then, PglH comes into play and adds three α 1,4-linked GalNAc residues on the glycan structure. PglI ends the formation of heptasaccharide that is built on Und-PP lipid carrier. The completed glycan is then flipped as an LLO from the cytoplasmic side of the cell to the periplasmic site by the help of ABC transporter enzyme, PglK. In the final step, pglB enzyme, an oligosaccharyltransferase, the heptasaccharide is transferred to the glycosylation sequon (D/E-X1-N-X2-(S/T)).[22]

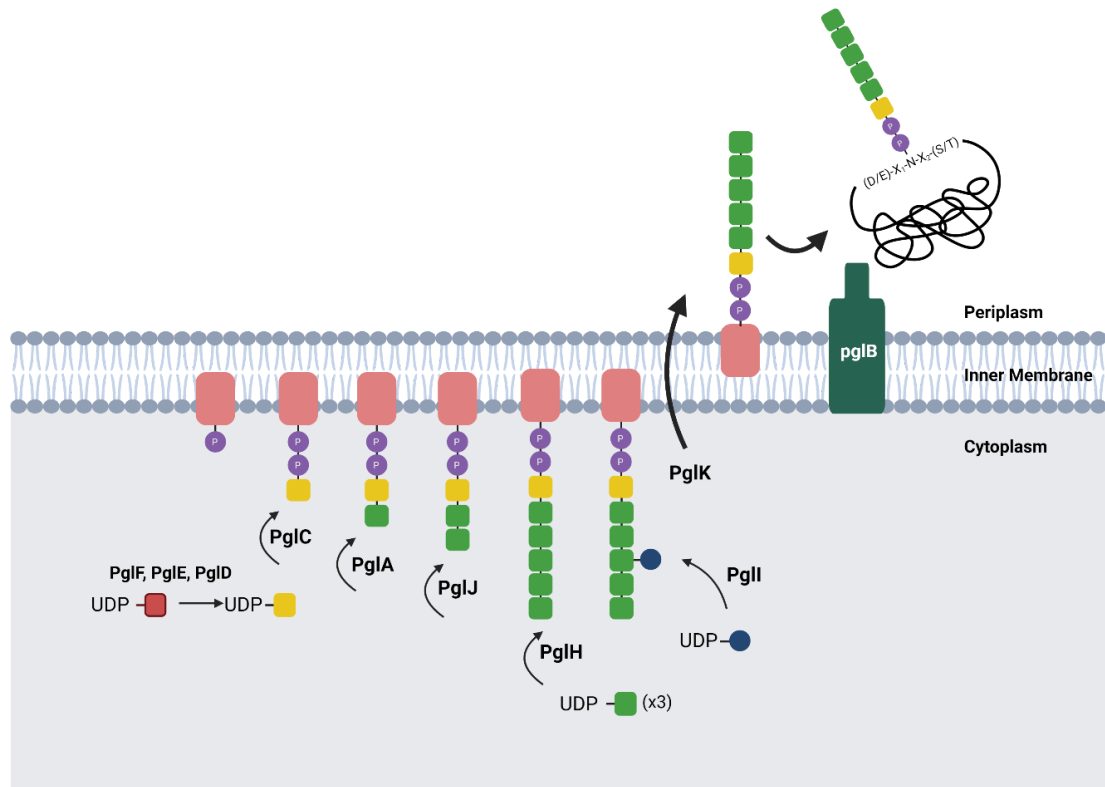


Figure 3: *C. jejuni* *pgl* glycosylation pathway. Created with Biorender.com. Adapted from Tan *et al.*

pgl pathway is a highly conserved machinery, given that all *Campylobacter* species, except one, demonstrates high similarity. This conservation applies also the reducing end sugars among *Campylobacter* species. On the other end, through time, other end of the sugars has undergone changes with evolutionary pressure due to immunogenicity.[23]

After gaining sufficient insights, researchers started to investigate ways to improve and adapt the mechanism to desired applications in various ways. Crystal structure and computational analysis of PglB enzyme revealed important information about the acceptor sequences and required structural features. [24] Directed evolution

approach has been employed to create variants of PglB enzyme in order to create better-performing oligosaccharyltransferase.

Work done on PglB uncovered that PglB enzyme has relaxed specificity over *N-glycan* residues. This showed that a variety of glycan can be utilized to glycosylate proteins. In 2005, after only three years the mechanism is transferred to *E. coli*, it has been reported that PglB enzyme can transfer different O-antigens from *E. coli* and *Pseudomonas aeruginosa*. [25] This showed that different *N-glycans* can be utilized and novel glycans can be engineered for new applications. Although novel glycan engineering work started before bacterial glycosylation systems discovered, it has been rapidly adopted and employed for bacterial systems as well. In *H. pylori*, native glycans were modified to include azide-containing groups.[26]

At last, but not least, high-throughput methods were also developed to screen the recombinant glycoproteins. Lectin proteins have been widely used to detect glycoproteins since they are naturally carbohydrate-binding abilities. Lectin microarrays have been developed to satisfy the need. However, they showed increased non-specificity, when they are subject to a test with bacterial glycans. [6]

An array, composed of a large library of oligosaccharides can also be used to test possible glycosylation of native or recombinant proteins. However, these assay are often lack standardization.[27]

A more recent system, called Glycophage Display, have been developed to meet the demand for high-throughput screening. As in the phage display method, phage coat protein g3p is fused to glycoprotein of interest. When phages infect *E.coli* carrying the *N-linked Glycosylation* mechanism, the end product becomes a glycosylated

protein, and it is displayed on the surface of the phage. Such method is extremely useful for screening, optimization and engineering of large libraries.[28]

1.4. Utilization of Glycosylation Systems in Synthetic

Biology

Shortly after the discovery of *N-Linked Glycosylation* pathways in bacteria and its successful transfer into *E.coli*, the power horse of laboratory[29], gave rise to a new discipline known as bacterial glycoengineering. As mentioned before, glycans have many fate-determining effects on proteins that they are attached to. They can regulate folding state and stability of the proteins. They can serve as a quality control check before they are sent to their final destination and have control over the protein's molecular interactions with environment and other proteins. They also play role in cell-cell interactions and immune response of host cells. [30] Therefore, studies to understand the mechanism better also led to engineering approaches have been developed to control and tune such effects for desired applications.

One of the vaccine subtypes, called conjugate vaccines, are produced by linking polysaccharide antigens with a carrier protein. Conjugation event provides a longer T-cell memory, and such vaccines became highly effective in the battle against pathogenic bacteria. Since chemical conjugation has its drawbacks in terms of cost and time, glycoprotein engineering has been adopted. [23] *C. jejuni* AcrA protein was glycosylated utilizing *N-Linked Glycosylation* machinery and employed to fight against *Shigella dysenteriae*. [31] Another groups, utilized *C. jejuni* CmeA protein and *N-Linked Glycosylation* machinery to produce a conjugate vaccine against *Brucella abortus*. The proposed vaccine showed an IgG response but unable to perform well in challenge experiments.

N-Linked Glycosylation is also a good approach to produce therapeutic proteins to tune biophysical and pharmacokinetic properties. *Lizak et al.* utilized *N-linked glycosylation* in *E. coli* and showed the successful glycosylation of single-chain antibody (scFv 3D5). It has been reported that *N-glycan* addition to the antibody increased the proteolytic stability and solubility of the antibody.[24]

1.5. The Aim of the Study

Protein glycosylation is one of the most common post-translational modifications in nature and shared by all three kingdoms of life, archaea, bacteria, and eukaryotes. It plays a fate-determining role for the protein and can alter its behavior in terms of stability, folding, activity and interaction with other molecules as well as with the environment. The discovery of bacterial glycosylation systems and its transfer into the *E. coli*, gave rise to a new concept in biotechnology: microbial glycobiology. One of the best studied mechanisms, *C. jejuni N-Linked Glycosylation* offers recombinant production of glycosylated proteins and therefore, allows protein of interest to be altered in terms of its behavior and tuned for desired applications. So far, *N-Linked Glycosylation* has been employed to produce glycan-conjugated vaccines and therapeutic proteins in order to increase their stability and solubility. In this study, we adopted the same machinery to investigate the effects of glycans on enzymes behavior.

To show our work, we utilized alkaline phosphatase (ALP) enzyme. Since *N-linked Glycosylation* is not very efficient system and efficiency changes according to protein of interest, I first tested different variants of ALP with glycosylation tag (DQNAT) at different locations of ALP in order to investigate the most suitable structure for efficient glycosylation. After finding the best performing one, I worked

on improving the glycosylation rate by overexpressing pglB enzyme. After achieving significant glycosylation rate, both unglycosylated and glycosylated ALP enzymes were purified to assess their differences due to glycosylation. Both versions were subject to enzyme kinetics assays, and treated with heat, protease and buffers with different pH in order to analyze the effect of glycosylation on the enzyme and its different behavior, compared to native ALP. Nonglycosylated and glycosylated ALP were also analyzed by circular dichroism (CD) after each treatment in order to obtain a relationship between enzyme kinetics and secondary structure changes.

Overall, it has been known that glycosylation had positive impacts for the protein in terms of its stability and solubility. However, this work uncovers the effect of glycosylation on an enzyme and how glycosylation will affect the activity of the enzyme and many other factors, including but not limited to, their working conditions in terms of pH and temperature, their resistance to denaturant conditions and protection against proteases.

CHAPTER 2

2. Materials and Methods

2.1. Buffers, Media and Strains

Cells were grown in Lysogeny Broth (LB) growth medium for cloning experiments.

ZYM5052 autoinduction media was used to inoculate bacteria for protein production.

Table 2: LB media ingredients required for 1L.

Tryptone	10 gr
Yeast Extract	5 gr
NaCl	10gr

Table 3: ZYM5052 medium ingredients required for 1L.

Tryptone	10 g
Yeast Extract	5 g
Disodium Phosphate (Na₂HPO₄)	3.55 g
Monopotassium Phosphate (KH₂PO₄)	3.4 g
Ammonium Chloride (NH₄Cl)	2.67 g

Sodium Sulfate (Na_2SO_4)	0.71 g
Magnesium Sulfate Heptahydrate ($\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$)	0.49 g
Trace Metals (50.000x)	20 μl
Glycerol	5 ml
Glucose Monohydrate	0.5 g
Lactose	2 g

TAE buffer was used both for agarose gel preparation and as running buffer in agarose gel electrophoresis.

Table 4: 1L of 1M TAE (pH 8.6) buffer ingredients.

2M Tris Base	242 g
1M Disodium EDTA	18.61 g
1M Acetic Acid	59.955 g

Chemically competent cells were prepared by TSS buffer for cloning experiments.

Table 5: 50 ml of TSS buffer ingredients and preparation.

10% Polyethylene glycol	5 g
--------------------------------	-----

5% Dimethyl Sulfoxide	2.5 ml
20mM MgCl₂	1 ml
Autoclaved LB	to 50 ml

Different concentrations of imidazole buffer were used for protein purification in binding and elution steps.

Table 6: 10 mM Imidazole (pH 7.4) binding buffer ingredients.

20mM Sodium Phosphate
0.5M NaCl
10mM Imidazole

Table 7: 150 mM Imidazole (pH 7.4) elution buffer ingredients.

20mM Sodium Phosphate
0.5M NaCl
150mM Imidazole

Cobalt Resin which was used for protein purification was regenerated by MES buffer.

Table 8: 1L of 0.5M MES buffer (pH 6) ingredients and preparation.

MES (free acid)	97.6 g (dissolve in 800 ml ddH ₂ O)
NaOH (to adjust pH to 6)	
ddH₂O	to 1L

Proteins were buffer exchanged to Tris-HCl to determine the activity depending on the temperature and incubation time.

Table 9: 100 ml of 1M Tris-HCl (pH 7.6) ingredients and preparation.

Tris base	12.11 g
Deionized H₂O (diH₂O)	80 ml
HCl (to adjust pH to 7.6)	
dH₂O	to 100 ml

Sodium Phosphate buffer was used to exchange the buffer of proteins to examine the activity change depending on pH to eliminate the effect of chloride ion in CD experiments.

Table 10: 50 ml of 0.1 M Sodium Phosphate buffer (pH 7.4) ingredients and preparation.

1 M Na₂HPO₄	3.87 ml
1 M NaH₂PO₄	1.13 ml
NaOH or H₃PO₄ (to adjust the pH to 7.4)	
diH₂O	to 50 ml

6X Loading Dye was used to visualize the samples during the run on SDS gel.

Table 11: 6x Loading Dye preparation.

375 mM Tris-HCl (pH 6.8)	
SDS	9% (w/v)
Glycerol	50% (v/v)

Bromophenol blue	0.03% (v/v)
-------------------------	--------------------

Washing and antibody incubation for immunoblotting experiments was performed with Tris Buffered Saline (TBS) containing two different amounts of Tween: 0.1% and 0.05%.

Table 12: 1L of 1X Tris Buffered Saline Tween (TBS-T) ingredients.

	TBS-T (0.1%)	TBS-T (0.05%)
10X TBS	100 ml	100 ml
Tween 20 Detergent	1 ml	0.5 ml
diH₂O	900 ml	900 ml

Enzyme activity is measured by catalysis of one of the substrates of Alkaline Phosphatase which is pNPP. Different concentrations of pNPP are prepared with pNPP reaction buffer.

Table 13: pNPP reaction buffer ingredients.

0.1 M Glycine
1 mM MgCl₂

1 mM ZnCl₂

Urea buffer was used to test the activity and structural change in proteins.

Table 14: 100 ml of 8M Urea buffer ingredients.

Urea	48.048 g (dissolved in 50 ml ddH ₂ O)
ddH₂O	to 100 ml

E. coli DH5α strain (New England Biolabs, Inc) was used in cloning experiments.

Bacterial Growth Conditions

E. coli cell stocks were prepared with 50% sterile glycerol diluted at 1:2 in LB media containing the cells and stored at -80 °C. They were inoculated from the stock in LB media into falcons overnight at 200 rpm at 37 °C for cloning. They were also inoculated from the stock in ZYM5052 into either falcons or Erlenmeyer's for 20-24 hours at 200 rpm at 37 °C for protein purification and whole-cell immunoblotting. Corresponding antibiotics were diluted at 1:1000 in those media.

2.2. Cloning and Transformation

Plasmids and primers that were constructed were designed using Benchling. PCR reactions were performed using Q5 High-Fidelity DNA Polymerase (New England Biolabs Inc.) by following the manufacturer's instructions. Primers that were used in PCR reactions were ordered from Oligomer. Annealing temperatures for PCR reactions were calculated using NEB Tm Calculator tool.

At the end of each PCR reactions, 20 μ L of the amplicon was mixed 6X purple loading dye (New England Biolabs Inc) and run on 1% agarose gel. Agarose gels were prepared by dissolving 0.6g of agarose in 60 mL 1X TAE buffer. Agarose gel electrophoresis was performed at 140V for 30 min. 1 kb+ Ladder (New England Biolabs Inc) was used to track DNA fragment lengths. After electrophoresis, agarose gels were visualized using Blue light transilluminator. Desired bands were excised, and gel extracted using Macherey-Nagel GmbH & Co. kit following manufacturer's instructions. The yield was measured using Nanodrop 2000 spectrophotometer (Thermo Fisher).

Restriction Digestion was performed using New England Biolabs enzymes and by following the instructions. At the end of each reaction, digested fragments were run on 1% agarose gel and extracted using Macherey-Nagel GmbH & Co. PCR Clean-Up kit. Ligation reactions were set up according to manufacturer's protocols.

To prepare competent cells, glycerol stocks were incubated overnight at 37°C 200 rpm shaking incubator. Growing cells were diluted at 1:100 and incubated at 37°C until the cells reached OD₆₀₀ 0.4 and 0.6. Cells were centrifuged at 10.000 g for 10

min at 4°C supernatants were removed and resuspended in 1:10 TSS buffer. The cells were aliquoted as 100 µL and stored at -80°C.

Definitive version of the constructed plasmids was transformed into competent cells. Competent cells were first thawed on ice and plasmid was added onto cells. After 20 minutes of incubation, cells were applied 42°C heat for 30 seconds and placed on ice for 2 min to complete the heat-shock procedure. 1 mL LB was added to each tube and incubated at 37°C for 1 hour. Cells were centrifuged at 8000 g for 5 minutes, supernatant was removed. Cells were resuspended in residual supernatant and overlaid on agar plates with appropriate antibiotics. Agar plates were incubated at 37°C overnight. Next day, colonies were picked, and colony PCR was performed. Positive colonies were selected, inoculated into LB media, and incubated at 37°C overnight in a shaking incubator. Plasmids were isolated from grown cultures using Gene-JET Miniprep Kit (Thermo Scientific) and plasmid concentrations were measured using Nanodrop 2000. Plasmids were sent for sequencing (Genewiz Inc) for verification.

2.3. Protein Expression

200-400 ml of cells containing ALP DQNAT and 400-600 ml of cells containing Glycosylated ALP DQNAT were inoculated from stock in ZYM5052 autoinduction media into Erlenmeyer flasks with corresponding antibodies. After 20-24 hours, cells were transferred to ultracentrifuge tubes and centrifuged at 3500 rcf for nearly 30 min. Supernatant was discarded and the pellet was resuspended by 10 mM imidazole buffer (pH 7.4). 1 mM of phenylmethanesulfonylfluoride (PMSF) (AMRESCO Inc.) was added. Lysing procedure was done via sonication. 30 seconds of pulse-on and 59 seconds of pulse-off were applied 5 times to bacteria with 0.35 amplitude. Then, they

were centrifuged at 12000 rpm for 1 hour. Supernatant was collected as total protein and stored at +4 °C.

Protein Extraction from Periplasmic Space

100 ml of cells were inoculated from the stock in ZYM5052 autoinduction media into an Erlenmeyer flask. After induction was completed in nearly 20 hours, the pellet of the cells was resuspended with 5mL of 30mM Tris (pH 7.5) at 4°C. Then, they were centrifuged at 6,000g for 15 minutes at 4°C. Supernatant was discarded and pellet was resuspended again with 30mM Tris (pH 7.5) containing 20% sucrose. The solution was incubated at room temperature for 15 minutes and centrifuged at 6,000g for 20 minutes at 4°C. Supernatant was collected as the extract from periplasmic space.

2.4. SDS-PAGE, Western Blot and Lectin Blot

Whole-cell western blotting and lectin blotting were performed for verifications of recombinant protein production. Cells were inoculated from stock in ZYM5052 autoinduction media with appropriate antibiotics into falcons at 200 rpm at 37 °C. After 20-24 hours, cells were precipitated at 3500 rpm for 5 min and the supernatant was discarded. The pellet was dissolved with different amounts of water to reach different concentrations. Heat inactivation was applied at 95°C for 5 min. 20 ul of cell extracts were run on SDS gel prepared via BioRad SDS Gel casting system by using 6x Loading Dye. 120 V was applied to the gel when proteins were in separating gel and the voltage was increased to 150 V for the rest of the run.

For SDS-PAGE, the gel was put into the Coomassie blue staining dye and heated but not boiled by microwave. Then, the gel was incubated on a shaker for 1 hour. After

washing the excess dye with water, destaining solution was added. The gel was washed again and visualized by Image Lab Software (Bio-Rad).

For the immunoblotting part, proteins were transferred into polyvinylidene difluoride (PVDF) membrane (Thermo Fisher Scientific). Filters were put into Turbo Transfer buffer (Bio-Rad) and placed on the cassette. The PVDF membrane was activated by methanol in nearly 5 min and put in buffer. By putting the membrane between filters, the sandwich was generated. Semi-dry transferring via Transblot Turbo Transfer System (Bio-Rad) was applied in 7 min.

Western Blotting

Blocking of the membrane was done for 2 hours in 1x Tris Buffer Saline Tween (TBS-T (0.1%)) and 5% skimmed milk. Then, the membrane was incubated for 1 hour with primary antibody (His-Tag Mouse McAb) (Proteintech Europe) diluted at 1:10000 in blocking solution. Wash steps were performed in 1x TBS-T (0.1%) three times for 5-, 10-, and 10-min. Incubation with secondary antibody (HRP conjugated goat anti-mouse) (Abcam) diluted at 1:10000 in blocking solution was completed in 1 hour. Finally, wash steps are performed again. Visualization of the membrane was done by HRP-conjugate secondary detection reagent. Enhanced chemiluminescence (ECL) (Bio-Rad) solution was prepared by mixing 500 µl of Clarity Western Peroxide Reagent with 500 µl of Clarity Western Luminol/Enhancer Reagent. After pouring the ECL solution onto the membrane, it was incubated for nearly 2 min in dark. Finally, the membrane was visualized via Image Lab Software (Bio-Rad).

Lectin Blotting

Blocking of the membrane was done for 1 hour in filtered 1x TBS-T (0.05%) and 3% BSA. Then, the membrane was incubated for 2 hours with primary antibody diluted at 1:5000 in blocking solution. Wash steps were performed in filtered 1x TBS-T (0.05%) five times for 5 min each. Visualization of the membrane was done by ECL (Bio-Rad).

2.5. Glycosylation Rate Calculations

Western Blotting analysis was used to calculate glycosylation rate of the constructs. Cells were grown in ZYM5052 as either 3 mL exclusively western blot or as a large batch for protein purification following western blotting. Grown cultures were centrifuged, resuspended in water and 6X SDS loading dye was added. Prepared samples were heated at 95°C for 5 minutes. Western Blotting was performed according to the Immunoblotting section. After visualization of western blot with Vilber Lourmat FUSION SOLO 6 imaging system, the analysis of bands was performed using Vilber Evolution Capt Edge software. Band intensity analysis was performed to the bands in each lane separately. By calculating each band intensity in one lane, slower migrating band intensity, which corresponds to the glycosylated ALP, was proportioned to the total intensity coming from the lane. Two replicates were used for consistency.

2.6. Protein Purification and Buffer Exchange

Cobalt resin stored at +4 °C with EtOH was regenerated by washing with ddH₂O and MES buffer respectively after discarding EtOH. Binding buffer which is 10 mM imidazole was added to the resin and then discarded. Lysis supernatant was

transferred on top of the resin and the mixture was rotated for 1 hour to make the His-tagged ALP bind to the resin. Unbound proteins were collected separately, and the cobalt resin was washed three times with binding buffer. Then, elution was performed three times by elution buffer which is 150 mM imidazole buffer (pH 7.4). After each use, the cobalt resin was regenerated again by washing with ddH₂O and MES buffer, and EtOH was added on top of the resin for storage.

Protein Quantification

Buffers of the proteins were exchanged with three different buffers for further experiments via Hi-Trap Desalting column (Sigma Aldrich). Proteins in 25 mM Tris (pH 8) were used to test the activity depending on temperature, incubation time, and pH. Sodium Phosphate buffer was used to observe the changes in structures of the proteins in solutions with different pH via Circular Dichroism (CD). The buffer also exchanged with 8 M urea solution to indicate both the activity and structural change of proteins. After equilibrating the column with 10 ml ddH₂O and 5 ml of buffer, 1.5 ml of proteins were collected with 2 ml of buffer. Then, BCA Assay (Thermo Fisher Scientific) quantification was performed. BCA standards were prepared by diluting from 2 mg/ml of stock solution. BCA Working Reagent was prepared by mixing 50 parts of Reagent A and 1 part of Reagent B. 200 μ l of working reagent was added to 25 μ l of standards and proteins. After gently shaking on shaker for 30 sec, the reaction was incubated at 37°C for 30 min. Absorbance values of standards and samples were measured at 540 nm. Finally, concentrations of the proteins were calculated by using the standard curve created by standards.

2.7. Unit Enzyme Calculations

Both unglycosylated and glycosylated ALP was purified, and buffer exchanged to 25 mM Tris (pH 8.0). pNP dilutions, which is the final product after ALP enzyme function, were prepared. 7 serial dilutions of both ALP enzymes were prepared and mixed with 0.5 mM pNPP. Reactions were placed into a spectrophotometer (SpectraMax M5, Molecular Devices) which was pre-set to 37°C and kinetic measurement was taken for 20 minutes for 30 seconds intervals. Reaction velocities were calculated and the interpolated enzyme concentration which catalyzes 1 μ M pNP / min was selected as 1 unit of enzyme.

2.8. Alkaline Phosphatase Activity Assays

Proteins were diluted to 2 μ g/ml to get 1 U of each. Enzymatic activity was calculated by measuring the catalysis of pNPP substrate. pNPP substrate solutions were prepared with pNPP reaction buffer. Different pNPP concentrations (0 mM, 1 mM, 2mM, 3 mM, 4 mM, and 5 mM) were used to create the *Michaelis-Menten* kinetics graph.

2.8.1. Enzyme Activity at Optimal Conditions

50 μ l of 2 μ g/ml proteins were put in each well and incubated at 37°C in the incubator for 10 min to make them reach optimal temperature. The M5 microplate reader was also adjusted to 37°C. 50 μ l of pNPP substrate concentrations were added on top of the proteins with a multichannel pipette rapidly. The absorbance of well at 405 nm was measured at 37°C and reaction rate was calculated using the pNP standard curve.

2.8.2.Preparation of Enzymes for Assessment of Enzyme Activity at Elevated Temperatures

Enzyme activity was measured in elevated temperatures with different incubation times to observe the resistance of the protein. 2 ug/ml proteins incubated at 3 different temperatures: 55, 75, 95°C, for 4 different incubation times: 10, 30, 60, 120 min. Temperatures were adjusted by using thermal cycler (Bio-Rad). As indicated previously, 50 µl of 2 ug/ml proteins and 50 µl of pNPP substrate concentrations were put in each well and the absorbance of well at 405 nm was measured at 37°C and reaction rate was calculated using the pNP standard curve.

2.8.3.Preparation of Enzymes for Assessment of Enzyme Activity at Different pH Values

Effect of pH on the enzymatic activity was measured by diluting proteins to 2 ug/ml with 25 mM Tris having pH 5, 6, 7, 8, 9, 10, 11. The same procedure was applied to measure the absorbance and calculate the reaction rate. Initial assays only include a rough enzyme activity measurement utilizing only 1 pNPP concentration. After determining the working range for both nonglycosylated and glycosylated ALP, *Michealis-Menten* enzyme kinetics experiments were performed.

2.8.4.Preparation of Enzymes for Assessment of Enzyme Activity upon Protease Treatment

Enzymatic activities of 2 ug/ml proteins were calculated and compared in the presence of different Proteinase K concentrations (100 uM, 20 uM, 10 uM, 5 uM, 0.5 uM, 0.05 uM, 0.005 uM, 0.0005 uM) with different incubation time (60, 120, 240

min). 50 μ l of 2 μ g/ml proteins and 5 μ l of Proteinase K was added to each well and the plate was incubated at 37°C in the incubator for different time periods. This step was repeated for each Proteinase K concentration and incubation time. The same procedure was applied to measure the absorbance and calculate the reaction rate.

2.8.5.Preparation of Enzymes for Assessment of Enzyme

Activity under Denaturant Conditions

Enzymatic activities of 2 μ g/ml proteins were calculated and compared in 8 M urea solution. After the protein purification step, buffer exchange procedure was applied to proteins with 8 M urea. The proteins were diluted to 2 μ g/ml to get 2 μ g/ml of proteins by 25 mM Tris-HCl. 50 μ l protein and 50 μ l pNPP substrate concentrations were added to each well. The same procedure was applied to measure the absorbance and calculate the reaction rate.

2.9. Circular Dichroism

Circular Dichroism (CD) is a commonly used to analyze secondary structure of the protein of interest. In this work, I have utilized CD (JASCO J-815) to analyze unglycosylated ALP and glycosylated ALP structures. For optimal conditions, both enzymes were prepared as 1 μ M in Tris buffer (pH 8.0). Three channels were selected as circular dichroism, voltage, and absorbance. Digital integration time (D.I.T) was selected as 4 seconds. Band width was 1 nanometer. Wavelength range was selected as 190 nm – 250 nm. Sensitivity is selected as standard. Scanning speed was set 100 nm/min. Accumulation mode was set to 3 repeats. A blank measurement was first taken as Tris (pH 8.0) at the beginning of each experiment. For optimal

conditions, in order to assess the secondary structure of unglycosylated ALP and glycosylated ALP, both enzymes were arranged to 1 μ M in pH 8.0 Tris.

2.9.1.Acquisition of CD Data for ALP at Elevated Temperatures

In order to track changes due to elevated temperatures, both unglycosylated ALP and glycosylated ALP were subjected to heat. Both enzymes were treated with heat at 55°C, 75°C, and 95°C for 2 hours. After that, samples were immediately measured with CD.

CHAPTER 3

3. Results and Discussion

3.1. Design and Cloning of the Constructs

The ALP variant constructs used in this study were designed by former lab member Musa Efe Işılak. *pgl*-pACYC containing DH10B cells were gifted by Markus Aebi. After selecting the construct with the highest glycosylation, *pglB* enzyme in the *pgl* pathway was amplified by primers that contain appropriate sites for Gibson Assembly.

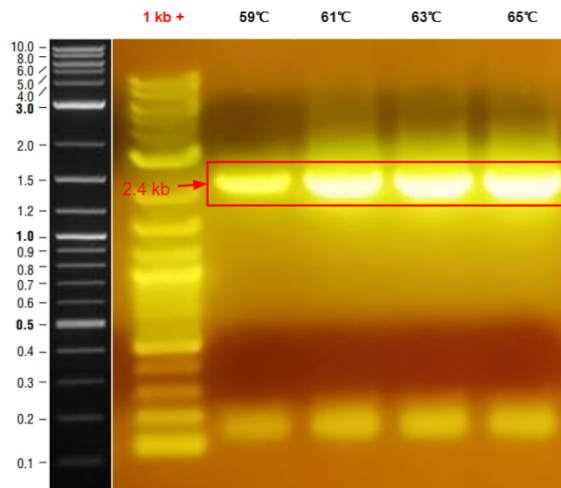


Figure 4: PCR result of *pglB* amplification.

According to the gel result, *pglB* was successfully amplified at expected size. Next, the circuit with the highest glycosylation, ALP-DQNAT w/ *pgl* were digested with BamHI to insert *pglB* enzyme.

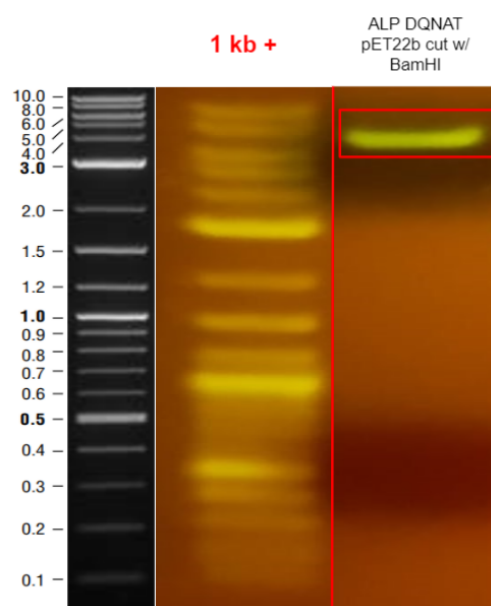


Figure 5: Digest Result of ALP DQNAT pET22B.

Restriction digest result confirms the single digestion of ALP DQNAT pET22b plasmid. The PCR and Digest products were extracted from agarose gels. Gibson Assembly was performed according to manufacturer's instructions. The Gibson Assembly product was transformed to Dh5 α cells and overlaid to agar plate with appropriate antibiotics. The plate was incubated overnight at 37°C. After colonies were observed, 8 colonies were picked for colony PCR.

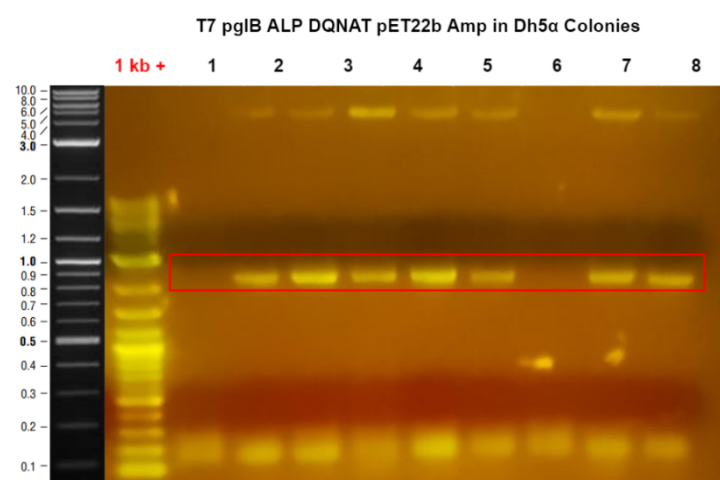


Figure 6: Colony PCR results of T7 pgIB ALP-DQNAT pET22b Amp.

According to colony PCR results, Colony #2 and #4 were sent for Sanger Sequencing and verified. Glycerol stocks were prepared from overnight cultures.

Electro transformation was performed for pgl-pACYC plasmid due to its size.

Pgl-pACYC containing BL21 DE3 cells were made competent, and the constructed plasmids were chemically transferred to BL21 DE3 cells.

3.2. Expression and Purification of ALP Variants

Three different ALP Variants; one with N-terminus glycosylation tag, one with C-terminus glycosylation tag and one with two glycosylation tag at the C-terminus was created.

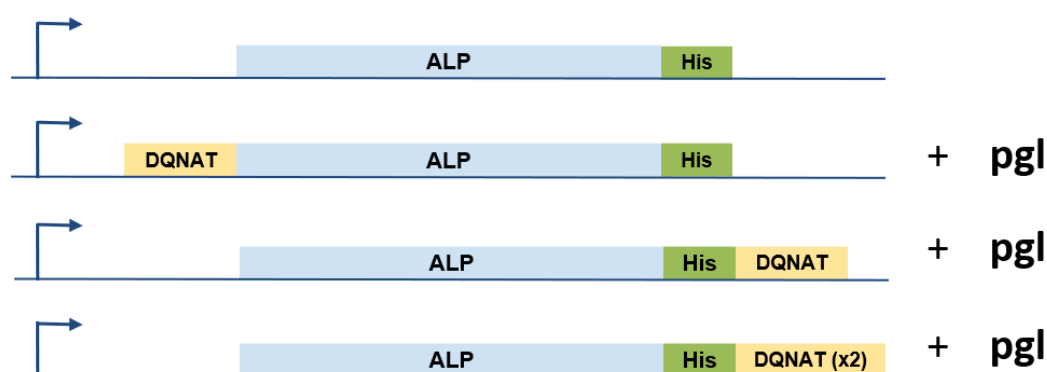


Figure 7: Created ALP Variants was created to assess Glycosylation Rate.

BL21 DE3 cells containing ALP-DQNAT pET22b, DQNAT-ALP pET22b + pgl-pACYC, ALP-DQNAT pET22b+pgl-pACYC, pglB ALP DQNAT(x2) pET22b + pgl-pACYC plasmids were grown and whole-cell western blot was performed to check expression of ALP variants.

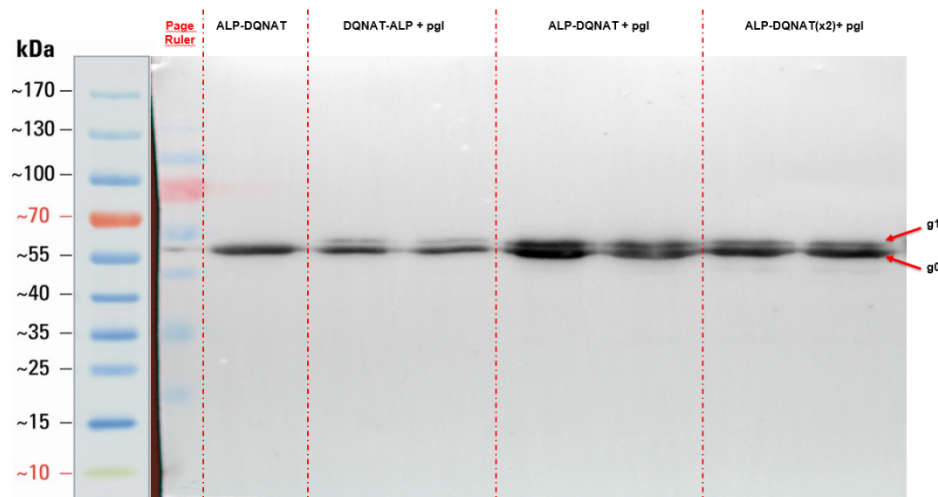


Figure 8: Western Blot Results of Created ALP Variants.

According to Western Blot results, it was observed that ALP-DQNAT has one band, while other constructs have two bands in each lane. The higher band in each lane corresponds to glycosylated version of ALP enzyme produced in the cells. The higher band is denoted as “g1”. The lower band corresponds to remaining ALP which are not glycosylated and denoted as “g0”. Glycosylation of this constructs were confirmed by Lectin Blot.

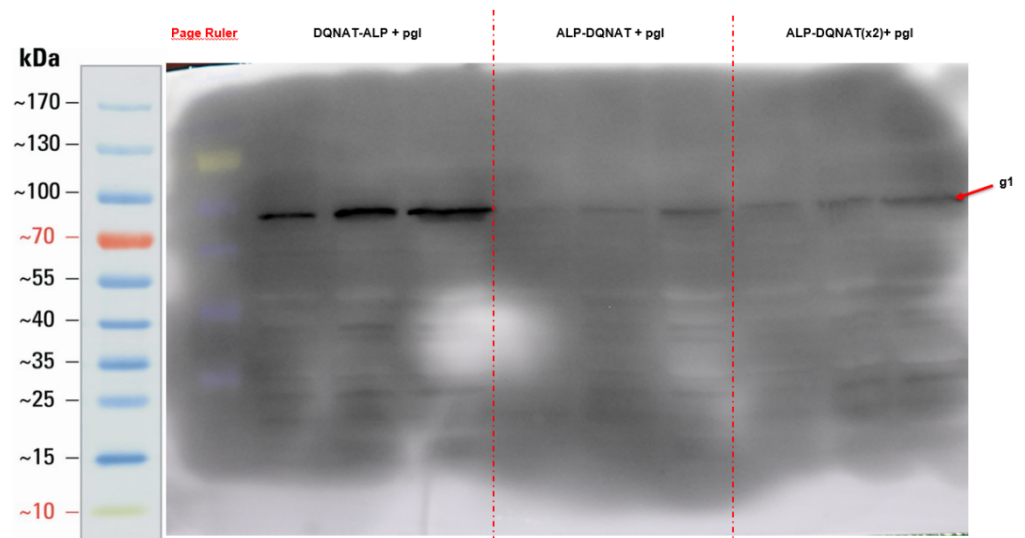


Figure 9: Lectin Blot Results of created ALP Variants.

After showing the successful glycosylation, to select the best performing construct, a glycosylation rate analysis must be performed.

3.3. Glycosylation Rate Calculation

According to the results, there are differences in glycosylation amounts of each ALP variant. To calculate these differences, Bands in each lane in western blot images were compared to each other (as $g1/(g0+g1)$) and a ratio was determined for each construct.

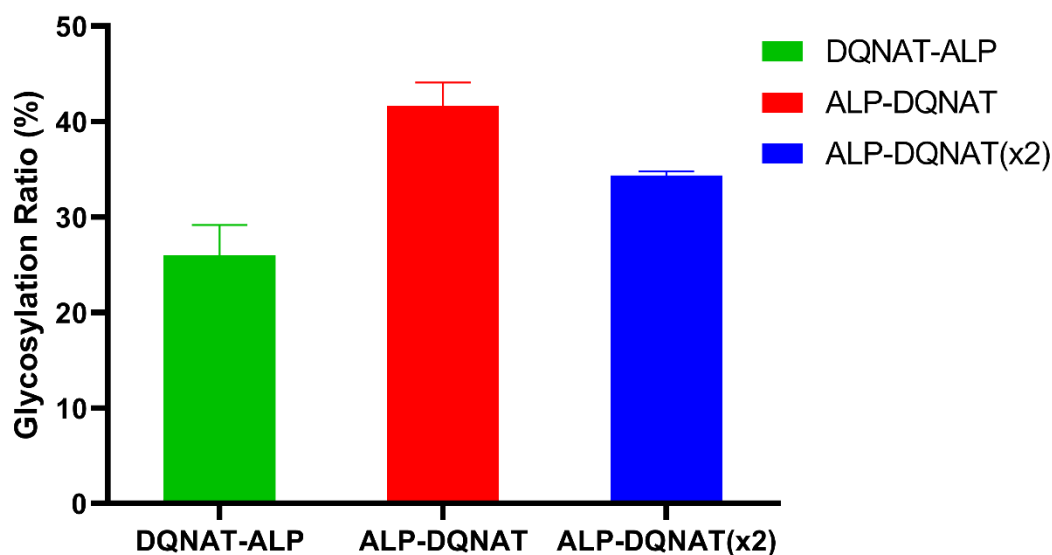


Figure 10: Calculated glycosylation rate by Vilber Evolution Edge software. (n=3)

Calculation of glycosylation rates indicated that the highest glycosylation rate can be achieved, when DQNAT motif was placed at the C terminus of ALP enzyme.

Furthermore, increasing the number of DQNAT motifs did not enhance the glycosylation rate. On the contrary, a slight decrease has been observed. According to the results, ALP-DQNAT pET22b + pgl-pACYC was selected as the best performing ALP variant. However, the glycosylation rate achieved (45%) was not

adequate to make a meaningful comparison between Nonglycosylated (g0) and glycosylated ALP enzymes (g1). Therefore, new strategies had to be adopted to further increase the glycosylation rate.

3.4. Overexpression of PglB Enzyme

To increase the glycosylation rate, the key enzyme in the glycosylation pathway, pglB can be overexpressed. To achieve this, the best performing ALP variant construct, ALP-DQNAT + pgl, which has 45% glycosylation rate, was selected and pglB enzyme was integrated into ALP-DQNAT pET22b plasmid. Figure shows the representation of the construct.

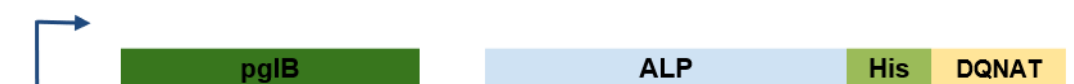


Figure 11: The representation of pglB overexpressing ALP variant, T7 pglB ALP-DQNAT + pgl-pACYC.

After successfully cloning the T7 pglB ALP-DQNAT + pgl-pACYC, the BL21 DE3 cells containing the construct was grown in ZYM5052 medium and whole cell western blot was performed to assess the glycosylation rate.

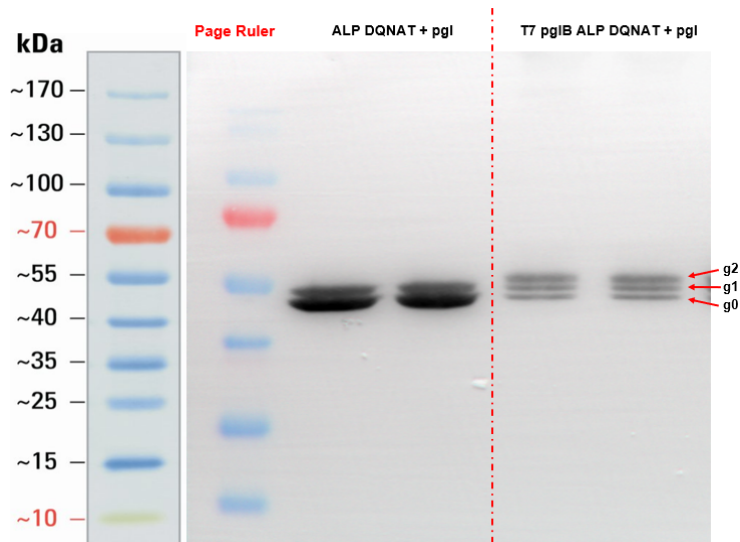


Figure 12: pglB Overexpressing version of ALP variant performs better in terms of glycosylation.

According to the results, T7 pglB ALP-DQNAT + pgl-pACYC yields three band in each lane. In the literature, there are examples of such results in Western Blot, when using the same glycosylation pathway and it indicates double glycosylation on the same protein. Considering that, proteins get heavier when glycosylated, the most upper band corresponds to the di-glycosylated version of ALP, denoted as “g2”, the middle band corresponds to mono-glycosylated ALP, denoted as “g1” and the lower band corresponds to non-glycosylated ALP, denoted as “g0”. It can be said that pglB overexpression enhanced the glycosylation. Therefore, a new glycosylation rate by comparing glycosylated bands to non-glycosylated band as $(g1+g2)/(g0+g1+g2)$, has been calculated.

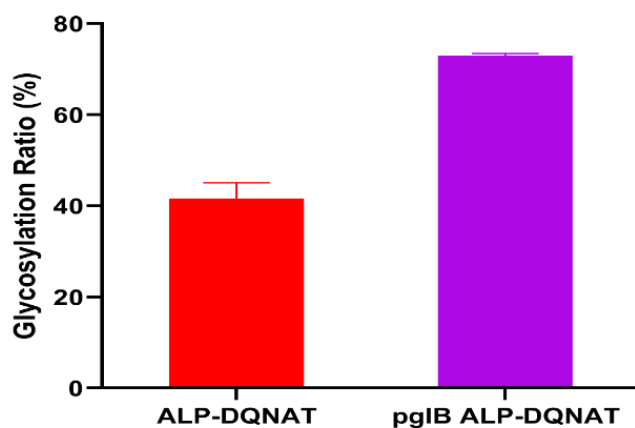


Figure 13: PglB overexpression enhanced the glycosylation rate.

According to glycosylation rate results, overexpression of pglB, enhanced the glycosylation rate from 45% to 73%. Glycosylation of pglB ALP-DQNAT + pgl-pACYC was confirmed by Lectin Blot.

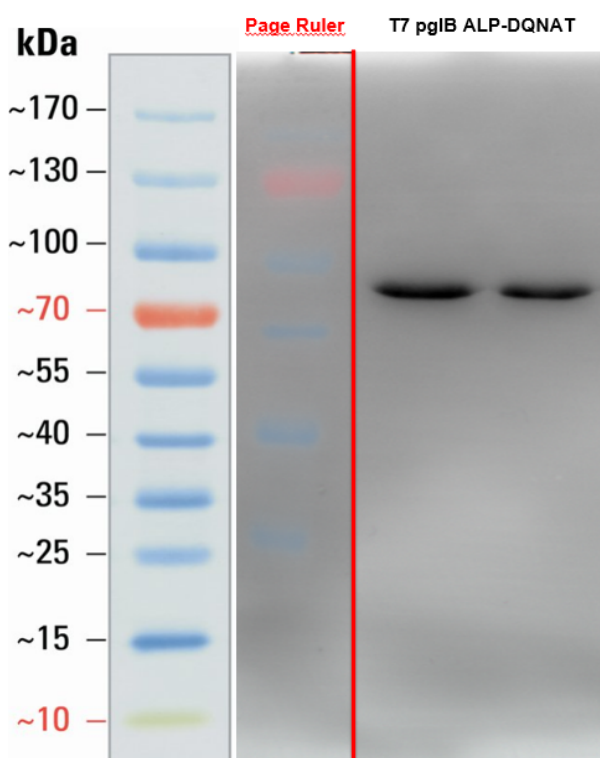


Figure 14: Lectin Blot of pglB ALP-DQNAT indicates that ALP-DQNAT is glycosylated.

When a comparison is made between Western Blot and Lectin Blot, it is observed that Western Blot suggests that two glycosylation states of ALP-DQNAT (mono-glycosylated and di-glycosylated). However, Lectin Blot results indicates only one band, meaning there is only one state of glycosylation, in this case, mono-glycosylation. One possible explanation is that the affinity of antibodies that are used in Lectin Blot to sugar moieties is lower than the affinity of antibodies used in Western Blot against His tag residues. This may clarify the difference in Western and Lectin Blot results. Another explanation, since glycan residues are placed on DQNAT motif on the protein of interest by PglB enzyme, placing of two glycan residues on only five amino acid residues could cause a steric hinderance effect. Therefore, di-glycosylated state cannot be recognized by the Lectin Blot antibodies. Another possible explanation of observing two different bands can be utilization of two different glycan residues. It has been previously stated that the last glucose residue is not a constraint to complete glycosylation process.[32] It is known that *C. jejuni* utilizes only one type of sugar moiety, glycan residue, the lack of the last glucose on glycan structure may result in three different sized bands. Furthermore, utilization of pgl pathway may interfere with *E.coli*'s native pathways. It is also known that pglB enzyme has a relaxed specificity, which refers to possible glycosylation sites on the acceptor protein. [33] Glycosylation can be occurred at another site on ALP enzyme which has resemblance to the recognition motif of pglB enzyme, along with the current glycan addition to the glycosylation tag, DQNAT. [34] To test this, GlycoPP, an online tool to define putative glycosylation sites were used. The tool allows identification of N-linked and O-linked glycosylation sites prokaryotes. [35] Analysis of putative N-linked glycosylation sites revealed that six more potential glycosylation sites. Among them, three sites were scored significantly

better and near to the actual glycosylation sequon, DQNAT. On the other hand, a potential glycosylation contradicts the lectin blot result which suggest only one glycosylated version. However, lectin blot may not be sensitive or the site may not be accessible for the antibodies.

Position	Residue	Score	Prediction
10	<u>N</u> RA	-0.63298651	Non-glycosylated
45	<u>N</u> II	-0.46195507	Non-glycosylated
64	<u>N</u> YA	-0.41202729	Non-glycosylated
91	<u>N</u> KK	-0.41927758	Non-glycosylated
118	<u>N</u> GA	0.20435362	Potential Glycosylated
146	<u>N</u> VS	0.21998587	Potential Glycosylated
182	<u>N</u> AL	-0.39246534	Non-glycosylated
198	<u>N</u> AR	0.012481975	Potential Glycosylated
245	<u>N</u> SV	0.21858822	Potential Glycosylated
251	<u>N</u> QQ	-0.23148619	Non-glycosylated
264	<u>N</u> MP	-0.40149138	Non-glycosylated
279	<u>N</u> ID	-0.36998229	Non-glycosylated
290	<u>N</u> PQ	-0.76691008	Non-glycosylated
294	<u>N</u> DS	-0.21727658	Non-glycosylated
314	<u>N</u> EK	0.073995766	Potential Glycosylated
335	<u>N</u> PC	-0.25937741	Non-glycosylated
362	<u>N</u> TL	0.042110277	Potential Glycosylated
392	<u>N</u> TK	-0.12752397	Non-glycosylated
405	<u>N</u> SE	-0.27856817	Non-glycosylated
429	<u>N</u> VV	-0.14465687	Non-glycosylated
453	<u>N</u> AT	0.37399455	Potential Glycosylated

Figure 15: Putative *N-linked* glycosylation analysis results performed by GlycoPP software.

It has been also reported that due to its relaxed specificity, pglB is also able to transfer O7- antigens. *Feldman et al.*, showed that pgl pathway, when transferred to *E.coli*, glycosylates AcrA proteins in both *O-linked* and *N-linked* manner.

Furthermore, both glycosylation types were transferred on the the same recognition motif, the *N-linked glycosylation* tag.[25] Therefore, *O-linked* glycosylation may contribute to the third band, given that excess pglB is present which will facilitate the non-canonical glycosylation pathways. When *O-linked* glycosylation putative site analysis were performed using GlycoPP, there were 14 putative sites found.

Therefore, it can be concluded that ALP-DQNAT can also carry O7 antigens natively produced in *E.coli*. Construction of mutated putative sites will be helpful to understand the causative of the third band to elaborate further.

After assessing the glycosylation rate from whole-cell proteins, proteins were purified utilizing His-tag affinity. Purified proteins were analyzed in SDS-PAGE for purity, and Western and Lectin Blot to compare with whole-cell results.

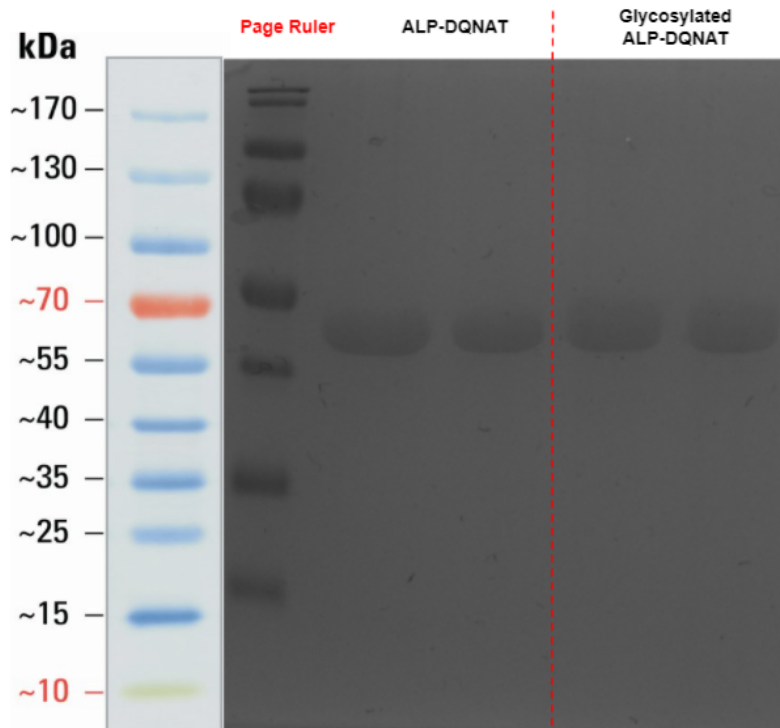


Figure 16: SDS-PAGE result of the purified ALP-DQNAT and Glycosylated ALP-DQNAT

SDS-PAGE analysis showed that proteins were purified without any significant purity. Since ALP-DQNAT is overexpressed, the band separation was not possible.

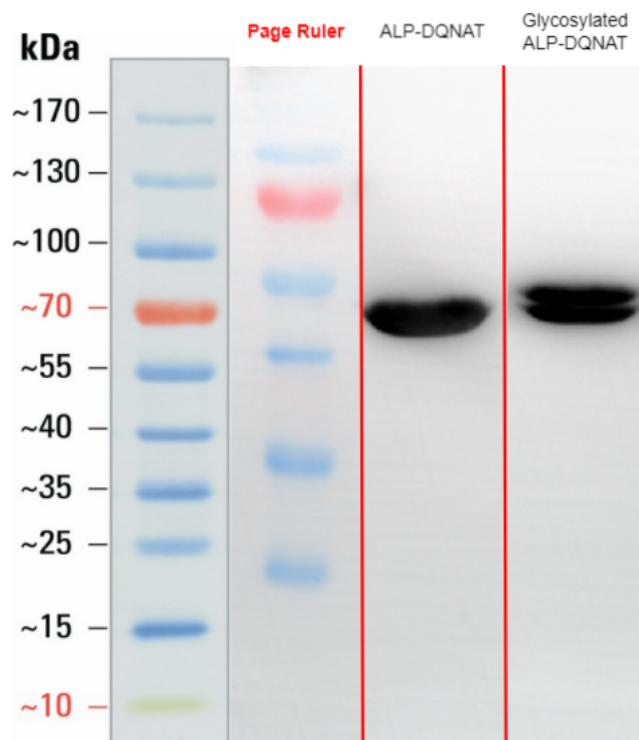


Figure 17: Western Blot result of ALP-DQNAT and Glycosylated ALP-DQNAT

According to western blot results, the miscellaneous band was eliminated during the purification step which simplifies the analysis of the protein characteristics.

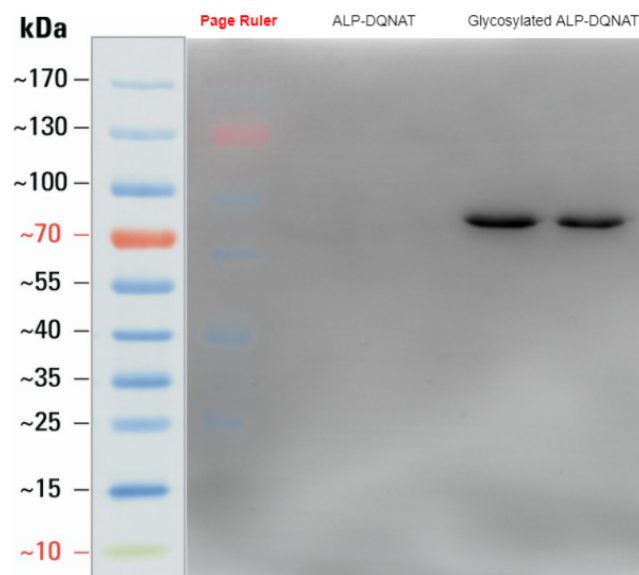


Figure 18: Lectin Blot results of purified ALP-DQNAT and Glycosylated ALP-DQNAT

Lectin blot results for the purified enzymes confirm the western blot results, suggesting that there is only one glycosylated state. The glycosylation rate is

calculated to be approximately 55% from the western blot and experiments were performed using the same batch with 55% glycosylation rate.

3.5. Enzyme Activity Comparison of Nonglycosylated and Glycosylated ALP

After achieving high glycosylation rate and observing successful expression of each ALP variants, the next step is to assess the enzymatic behaviors of non-glycosylated and glycosylated ALP. Para-Nitrophenylphosphate (pNPP) is the substrate of alkaline phosphatase enzyme. pNPP is cleaved into two parts: orthophosphate and NPP which is a yellow-colored product. The alkaline phosphatase activity can be assessed by the rate of the color change in a transparent medium.[36]

To compare two enzymes catalyzing the same substrates, using enzyme concentrations is useful. Therefore, enzymes were diluted to 2 ug/mL and incubated with different substrate concentrations, 0 μ M, 1000 μ M, 2000 μ M, 3000 μ M, 4000 μ M and 5000 μ M. Michealis - Menten analysis was performed to investigate the differences between two constructs.

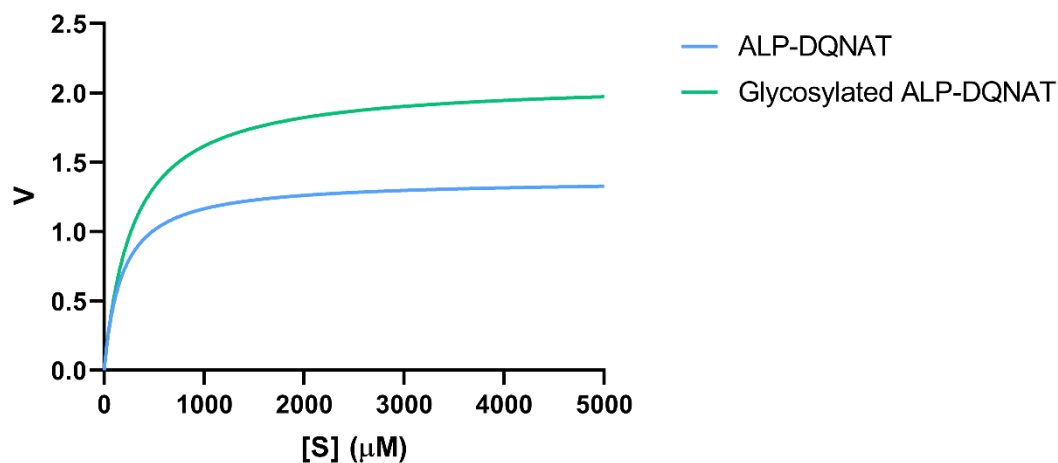


Figure 19: Comparison of ALP and Glycosylated ALP. Glycosylated ALP obtained higher catalysis activity.

Table 15: The enzyme behaviours calculated for ALP-DQNAT and Glycosylated ALP-DQNAT from enzyme activity assays.

	ALP-DQNAT	Glycosylated ALP-DQNAT
$v_{\max}$	1.376	2.088
K_M	180.2	290.8

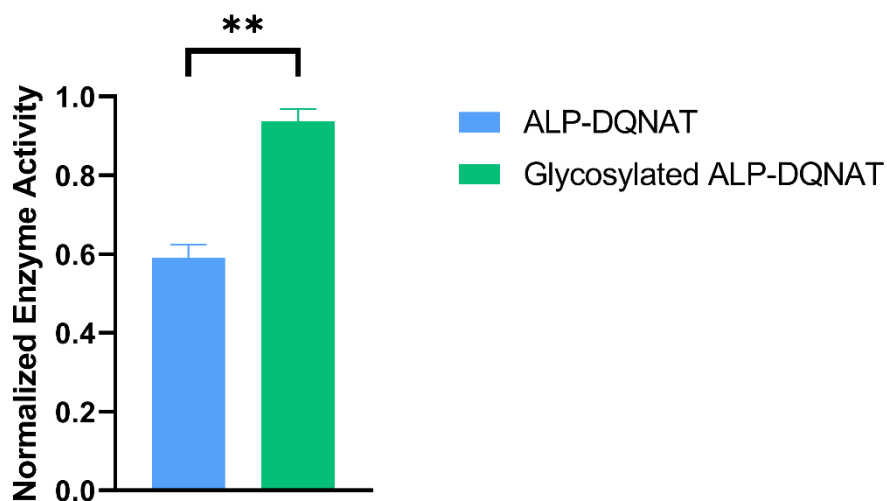


Figure 20: Normalized enzyme activities of ALP-DQNAT and Glycosylated ALP-DQNAT at optimal conditions.

When enzyme activities are compared, glycosylated ALP-DQNAT functions significantly better, almost 2 times faster than the native enzyme. Michealis – Menten analysis cleared it out that $v_{\max}$ has significantly increased from 1.376 to 2.088, when the ALP is glycosylated. This means that glycosylation at C-terminus provided ALP better catalyzing properties. On the other hand, enzyme's affinity to its substrate has decreased, upon glycosylation.

In order to have a better understanding of the changes in enzymatic activity, secondary structure analysis has been performed. Circular Dichroism (CD) is a widely utilized method to study secondary structure of the protein of interest.. the technique is based on the measurements taken of different absorption values between the left and right circularly polarized light. According to peptide bond conformation of the peptide bonds and their H-bonding pattern which are the features of the secondary structure, proteins yield different spectral profiles. Most proteins give a differential spectra between 170 nm and 250 nm. Therefore, analyzing the spectra between this range will give valuable information about the secondary structure of the protein analyzed. [37]

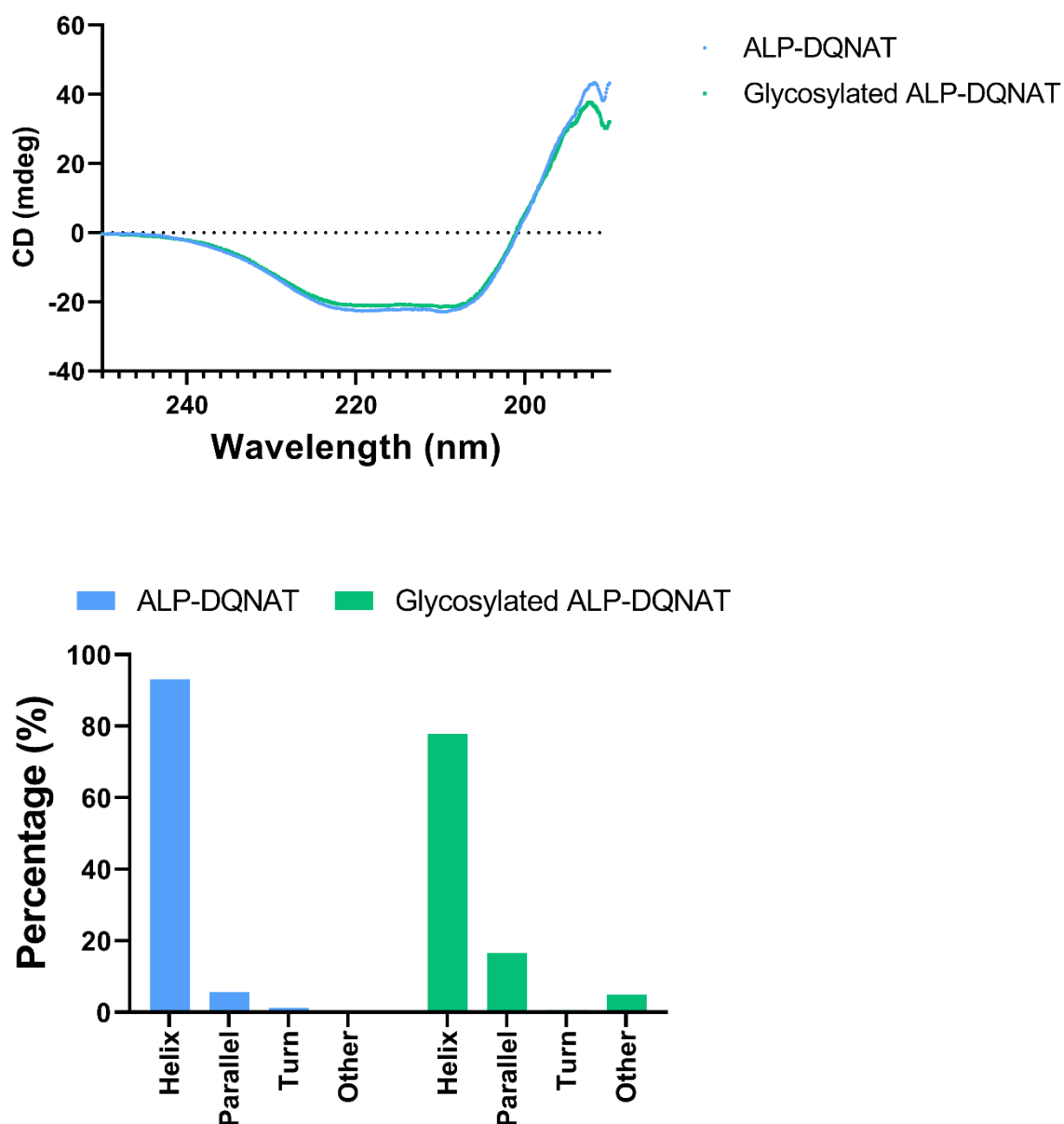


Figure 21: Secondary structure analysis performed to assess the differences between non-glycosylated and glycosylated ALP and predicted secondary structures.

According to the secondary structure analysis, there is a decrease in total helical structure of ALP-DQNAT upon glycosylation. This decrease is compensated by the increase in parallel structure. These differences may play a significant role in the difference observed in the enzymatic activity.

3.5.1.Enzyme Activity and Secondary Structure Assessment at Elevated Temperatures

To test and the compare the stability of non-glycosylated and glycosylated ALP, the enzymes were subjected to varying elevated temperatures (55°C, 75°C, 95°C) for varying incubation times (15 min, 30 min, 60 min, 120 min). Enzyme activities are normalized, including the control groups (optimal conditions) in **Figure 13**.

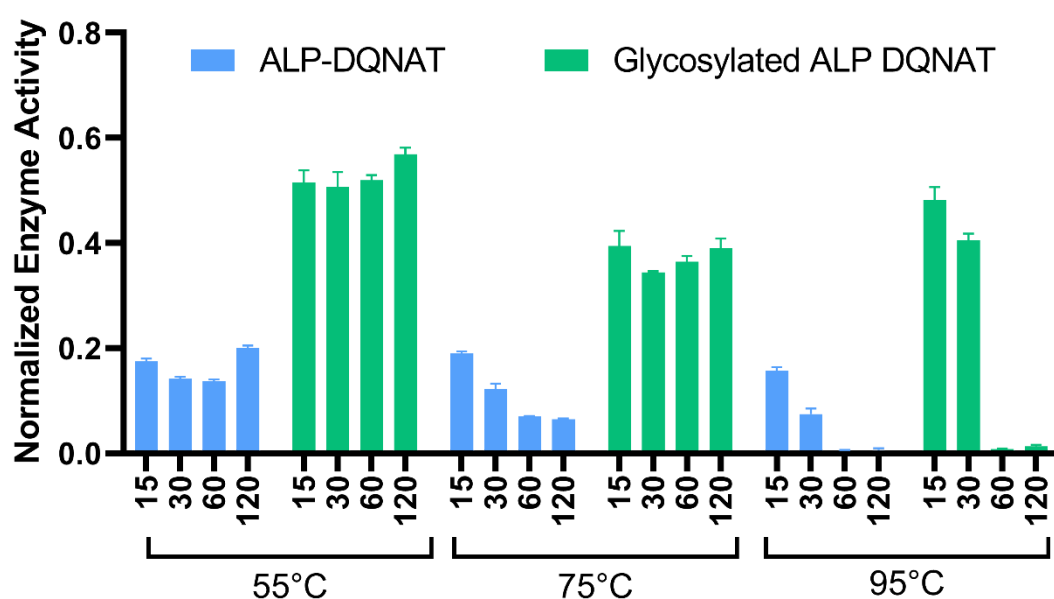


Figure 22: ALP-DQNAT and Glycosylated ALP-DQNAT are treated with varying temperatures.

Treatment to different temperatures indicated that glycosylated ALP-DQNAT showed enhanced enzyme activity, when a comparison is made. At 55°C, Both groups activities were deviated from the control groups for all treatment times. However, while ALP-DQNAT activity dropped from 0.59 to 0.17 which corresponds to 3.4 times decrease, glycosylated ALP-DQNAT activity dropped from 0.93 to 0.51 which corresponds to 1.8 times decrease. At 75°C, glycosylated ALP-DQNAT

activity is retained for all treatment times, whereas, ALP-DQNAT activities are dropped, as the treatment time increases. The difference between 15 min treatment and 120 min treatment was 3 fold. However, the same does not apply to glycosylated ALP-DQNAT which does not show a significant difference. At 95°C, even at 30 minutes, ALP-DQNAT activity is almost ceased, 90% less active compared the control group. Glycosylated ALP-DQNAT works at half of its activity and approximately the same with the native enzyme at optimal conditions. After 30 minutes of treatment, both enzymes did not exhibit meaningful enzyme activities. From this point, it can be suggested that glycosylation significantly improved the stability of ALP enzyme.

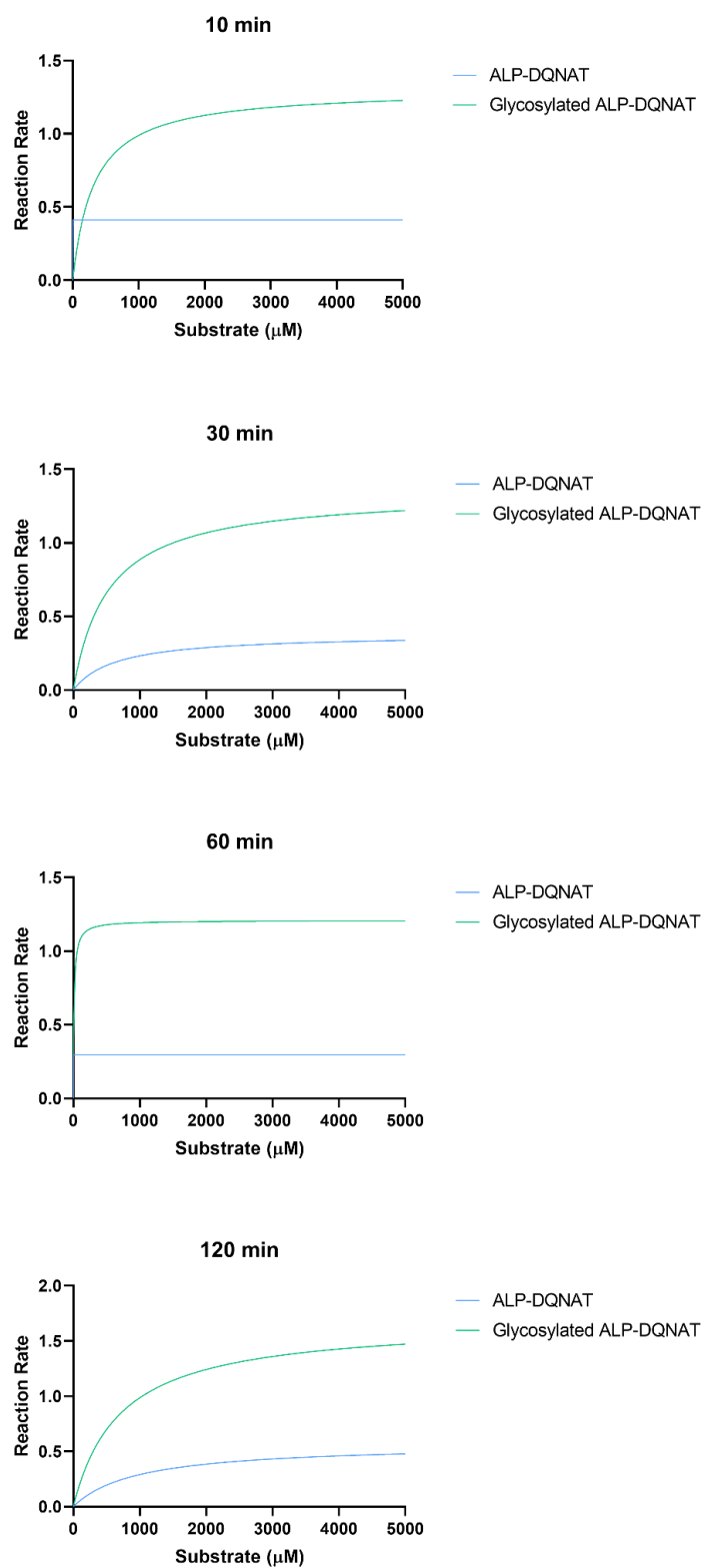


Figure 23: Enzyme Kinetics of ALP-DQNAT and Glycosylated ALP-DQNAT at 55°C for different treatment times.

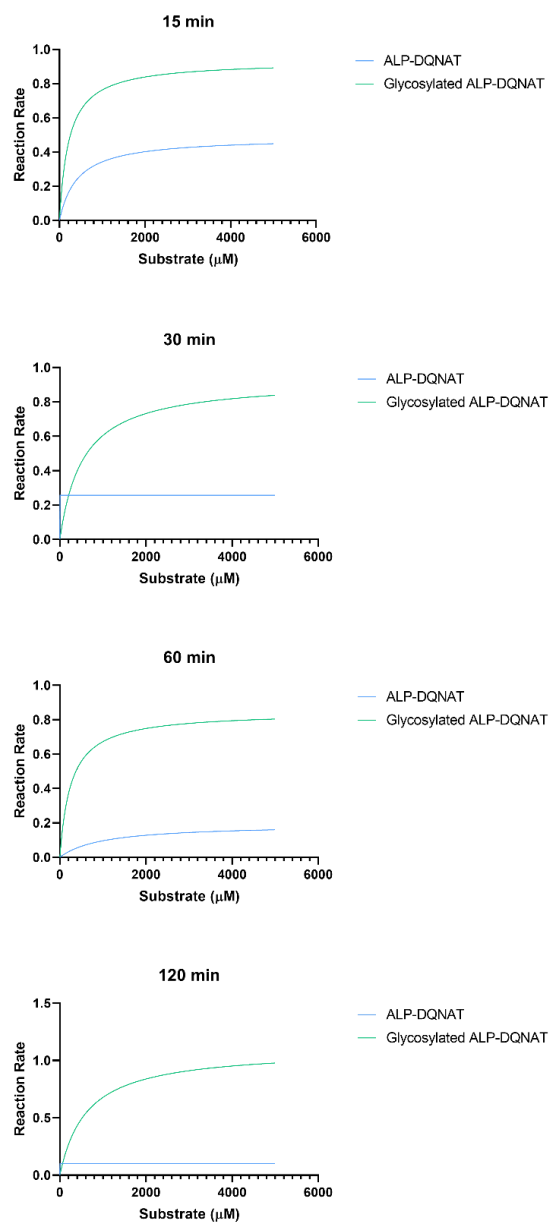


Figure 24: Enzyme Kinetics of ALP-DQNAT and Glycosylated ALP-DQNAT at 75°C for different treatment times.

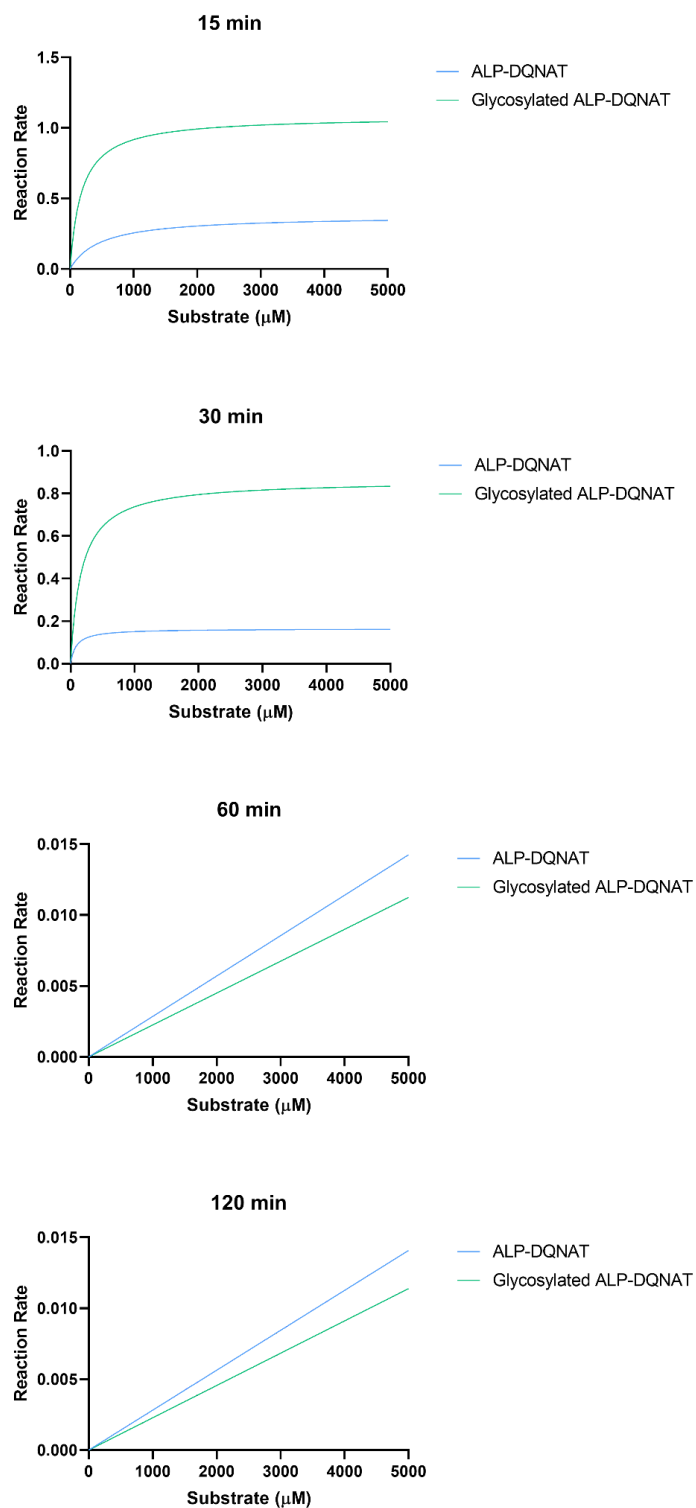


Figure 25: Enzyme Kinetics of ALP-DQNAT and Glycosylated ALP-DQNAT at 95°C for different treatment times.

In order to have a clearer understanding of these differences, secondary structure analysis has been performed.

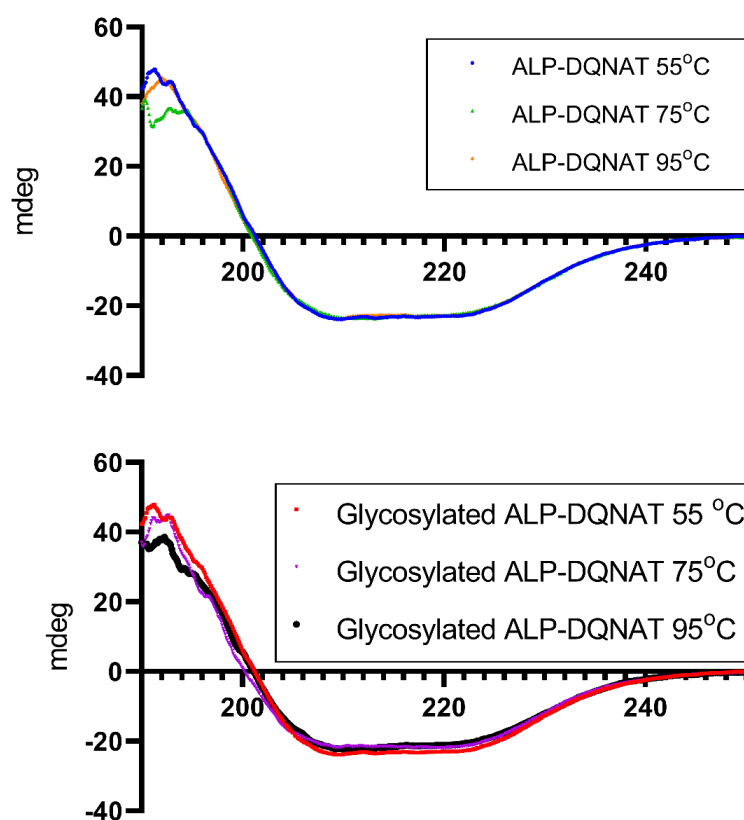


Figure 26: CD spectra of ALP-DQNAT and Glycosylated ALP-DQNAT at elevated temperatures.

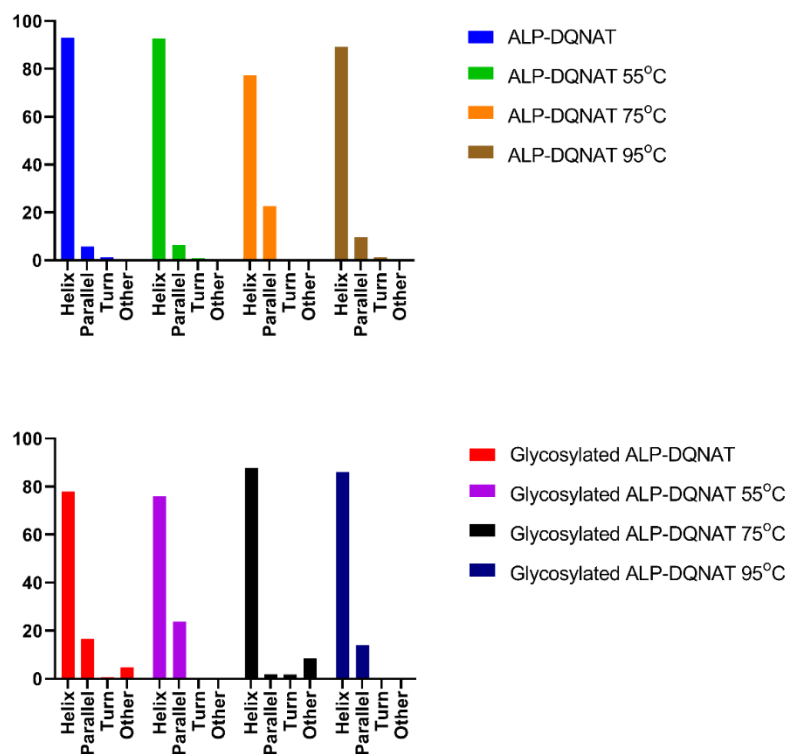


Figure 27: Secondary structure analysis of ALP-DQNAT and Glycosylated ALP-DQNAT at elevated temperatures.

CD spectra of ALP-DQNAT displayed a minimum between 205 nm and 225 nm and maximum at 195 nm. The only significant difference between experiment groups is between 190 nm and 200 nm. The signal obtained from this interval is associated with helical secondary structures.

When the secondary structure prediction data is taken into consideration, ALP-DQNAT did not show a significant difference in secondary structure, as the temperature increases. Glycosylated ALP-DQNAT has a different pattern for all treatments, compared to ALP-DQNAT. As the temperature is elevated, parallel structure abundance fluctuates. An increase is seen at 55°C, and completely disappears at 75°C and increases again at 95°C. Furthermore, when a comparison

made between 75°C treated samples, ALP-DQNAT possess 1.9% parallel structure, whereas Glycosylated ALP-DQNAT possess 22.6%. considering that there is a 7-fold enzyme activity difference between these groups, these parallel structures may account for the enhanced enzyme activity.

3.5.2. Enzyme Activity and Secondary Structure Assessment at Different pH

One of the biggest factors in enzyme activity is the pH of the buffer the enzyme working. Enzymes are greatly affected by the pH change and optimum pH is enzyme dependent. Although enzymes generally lost their activity at extreme low and extreme high pH values, this also needs to be investigated for every enzyme.

E. coli Alkaline phosphatase enzyme shows the maximum activity at pH 8.0. In order to investigate the effect of glycosylation on working range of the enzyme, Tris-HCl buffers were prepared with different pH values, between pH 5.0 and pH 13.0 and ALP-DQNAT and glycosylated ALP-DQNAT activity were measured in these conditions.

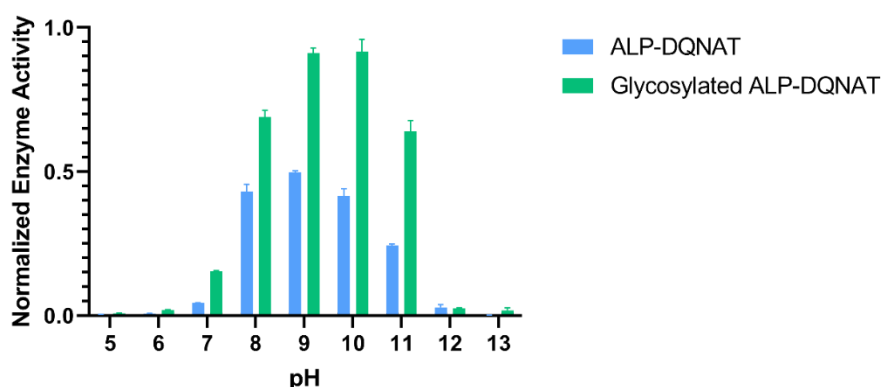
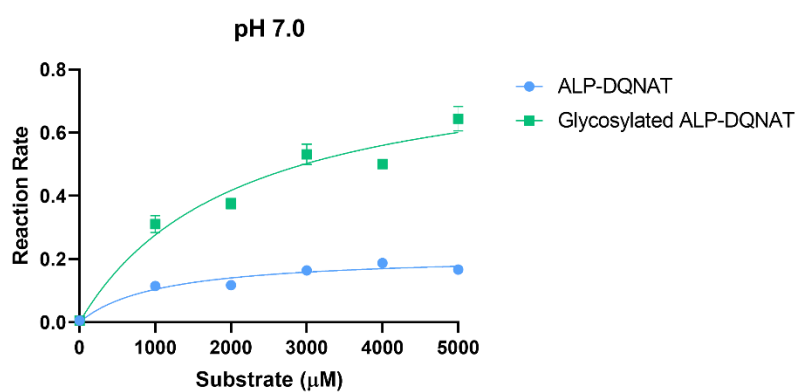
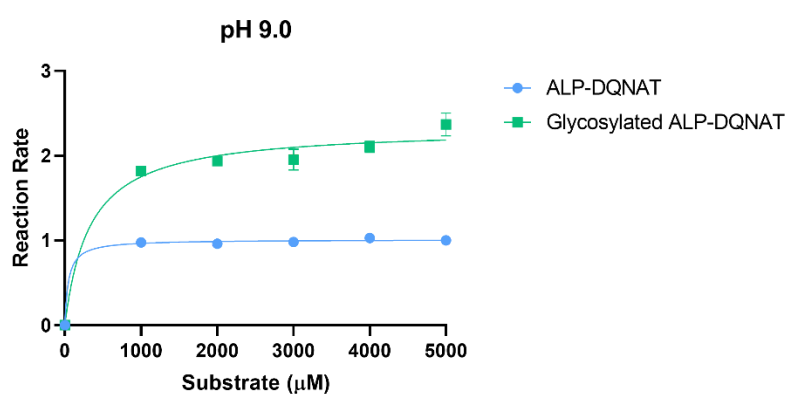


Figure 28: Enzyme activity results of ALP-DQNAT and Glycosylated ALP-DQNAT at different pH conditions.

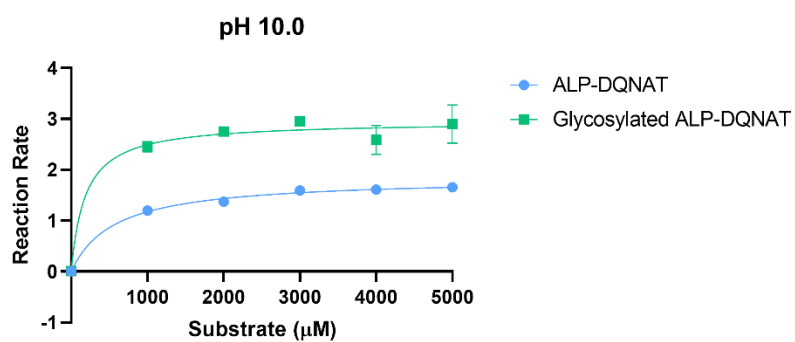
a.



b.



c.



d.

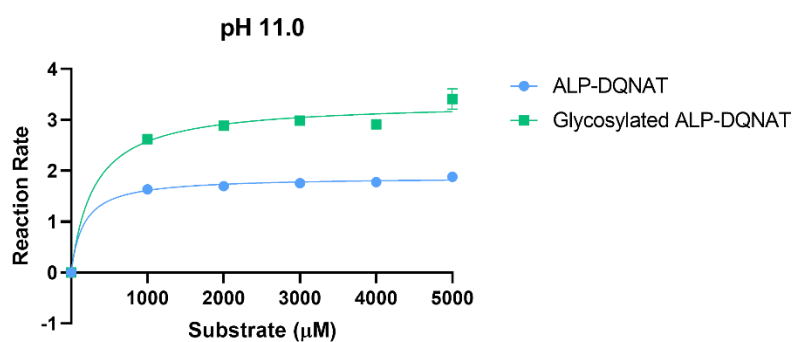


Figure 29: Enzyme Kinetics of ALP-DQNAT and Glycosylated ALP-DQNAT at different pH values.

According to the results, ALP-DQNAT has an optimum pH range between 8 and 10, as expected. When pH is lowered, ALP-DQNAT loses its activity considerably at pH 7.0. ALP-DQNAT does not show any activity at pH values lower than 7 and higher than 12.

Compared to ALP-DQNAT, glycosylated ALP-DQNAT works significantly better at pH 8.0, approximately 1.6-fold. Surprisingly, at pH 9.0 and pH 10.0 glycosylated ALP-DQNAT activity increases, whereas ALP-DQNAT activity stays at the same levels. Glycosylated ALP-DQNAT activity is 1.85 and 2.25 times higher than ALP-DQNAT, at pH 9.0 and pH 10.0, respectively. Furthermore, the highest fold change is achieved at pH 7.0. Glycosylated ALP-DQNAT performed nearly 4 times better than ALP-DQNAT.

Table 16: The secondary structure analysis results at different pH conditions.

	Helix	Antiparallel	Parallel	Turn	Others
ALP-DQNAT pH 5	76,7	4,3	1,3	0	17,7
ALP-DQNAT pH 7	71	18,4	7,7	0	3
ALP-DQNAT pH 8	82,3	4,8	0	0	12,9
ALP-DQNAT pH 9	75,4	3,2	1	0	20,5
ALP-DQNAT pH 10	74,5	8,5	0	0	16,9
ALP-DQNAT pH 11	76,6	6,1	3,8	0	13,5
Glycosylated ALP-DQNAT pH 5	61,6	3,3	0	5	30
Glycosylated ALP-DQNAT pH 7	55,5	0	7,9	4,6	32
Glycosylated ALP-DQNAT pH 8	58	4,7	1,6	6,6	29,2
Glycosylated ALP-DQNAT pH 9	66,6	6,8	0,1	6,1	26,4
Glycosylated ALP-DQNAT pH 10	60,1	0	0	5,4	34,5
Glycosylated ALP-DQNAT pH 11	46,5	5,2	5,8	4,2	38,5

Overall, the results indicated that glycosylation increases the efficiency of the ALP enzyme at all pH values. Moreover, enhancement of glycosylated ALP-DQNAT

activity at pH 9.0 and pH 10.0 suggest that glycosylation shifts the optimum pH value of the enzyme.

3.5.3. Enzyme Activity Assessment upon Protease Treatment

Proteases are enzymes that performs breaking down of proteins into peptides by cleaving the peptide bonds. We used a widely utilized serine protease, Proteinase K, in order to assess the stability of ALP against proteases. Proteinase K has an optimum working pH of 8.0 as ALP and it is suitable to see the action of both enzymes in the same medium.

We first tried to optimize the incubation time and ALP enzymes were subjected to 5 μ L of 20 μ g/mL Proteinase K for 60 minutes, 120 minutes and 240 minutes.

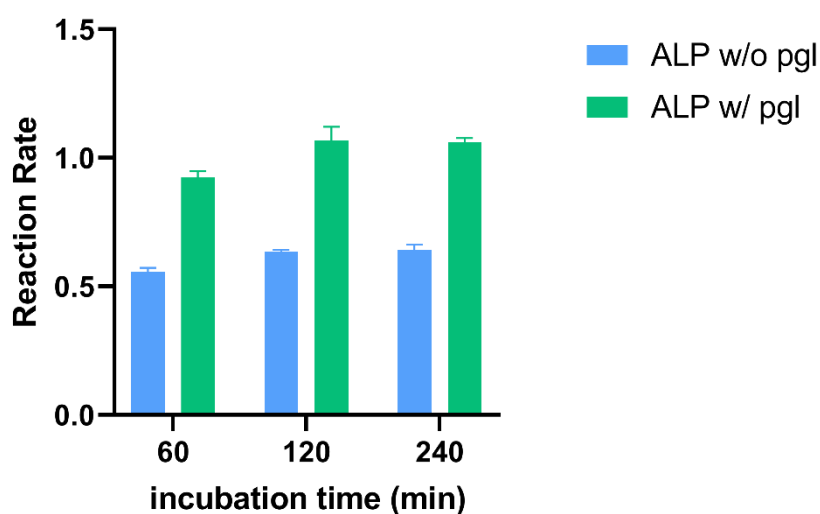


Figure 30: Proteinase K treatments of ALP enzymes for different incubation times.

According to the results, although overall enzyme activity decreased for both ALP enzymes, elongated incubation times did not increase the effect of Proteinase K. Therefore, 60 minutes of incubation was selected for Proteinase K concentration optimization.

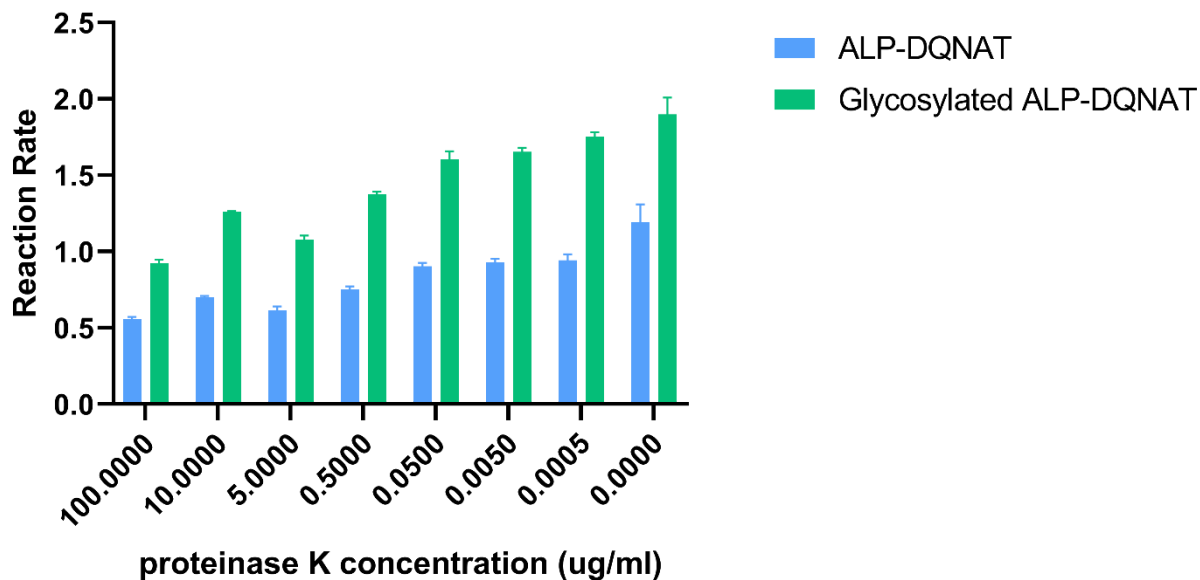


Figure 31: Effect of protease treatment on the enzyme activity of ALP-DQNAT and Glycosylated ALP-DQNAT.

Dilutions of Proteinase K were prepared and 5 μ L of each dilution were added to 1 U of ALP-DQNAT and Glycosylated ALP-DQNAT. According to the results, increasing concentrations caused further decrease in the enzyme activity. For each protease concentration, Glycosylated ALP-DQNAT worked better, compared to ALP-DQNAT.

3.5.4. Enzyme Activity and Secondary Structure Assessment under Denaturant Conditions

According to the results so far, ALP-DQNAT and Glycosylated ALP-DQNAT showed a variety of enzyme activity levels. However, according to CD results, they did not lose their secondary structure completely.

In order to investigate this possibility as well, we decided to apply denaturants to break the secondary structure of ALP. Both ALP-DQNAT and Glycosylated ALP-DQNAT was buffer exchanged to 8 M Urea solution. Both enzymes were diluted to 2 $\mu\text{g/mL}$ utilizing different concentrations of urea solution to understand the effect of urea on enzyme activity and secondary structure.

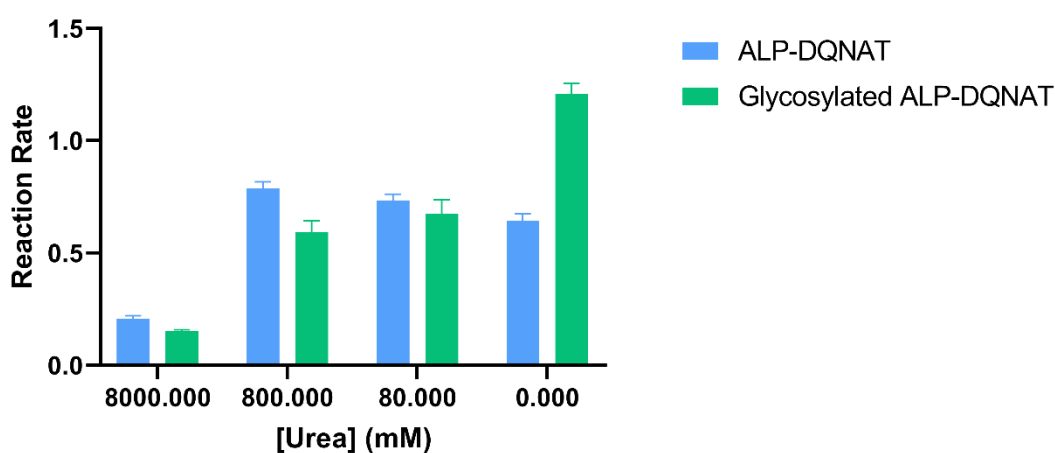


Figure 32: Enzyme activity of ALP-DQNAT and Glycosylated ALP-DQNAT under denaturing conditions.

CHAPTER 4

4. Conclusion

Protein glycosylation is an essential post-translational modification that can determine the fate of the protein that undergoes the modification. It regulates the protein's stability, activity, and interactions with its environment. After the discovery of bacterial glycosylation mechanisms, it became a hot-topic for both studying and utilizing. Successful transfer of *C. jejuni N-linked glycosylation mechanism* into *E.coli* has greatly expanded the utilization of the mechanism and generate a new discipline called Microbial Glycobiology. *N-linked Glycosylation* mechanism has been previously used many times to glycosylate recombinant proteins and it has been reported that glycosylation improves certain features of the protein, such as stability and solubility. In this study, it is aimed to utilize this mechanism in order to glycosylate an enzyme, alkaline phosphatase, to develop a better performing enzyme.

Considering that success of *N-linked Glycosylation* is a highly acceptor protein dependent, 3 alkaline phosphatase variants have been created. In these variants, one has the glycosylation tag at N-terminus, one has at C-terminus and also two glycosylation tag was added to C-terminus to understand if there is a further increase in glycosylation. It was found that the variant that has the glycosylation tag at C-terminus is the most glycosylated variant, approximately 48%. In order to further increase this rate, PglB enzyme was overexpressed. As the result of overexpression, glycosylation rate increased up to 70%. This variant was selected as it has the highest glycosylation rate for the next steps.

Both unglycosylated ALP and glycosylated ALP were analyzed using *Michealis-Menten* enzyme kinetics. It has been observed that glycosylated ALP worked 1.5-fold better than nonglycosylated ALP at optimal conditions.

Furthermore, different temperature treatments were applied to both enzymes for varying of times. The difference between unglycosylated ALP and glycosylated ALP became clearer at non-optimal temperatures. Both enzymes were affected by the temperature. However, glycosylated ALP activity was significantly less hampered, when compared with unglycosylated ALP. Glycosylated ALP worked between 3-to-7 times better than unglycosylated ALP, when all temperatures and incubation times were taken into account. Working pH range analysis, showed that the range did not change, however, glycosylated ALP still performed better at all pH values. The results where similar ALP is tested upon protease treatment and under denaturing conditions. Secondary structure analysis by circular dichroism was performed to assess the changes in protein's structure at each condition. Overall, the results shows that glycosylation of ALP improved the enzyme activity, due to increased stability.

The results suggest that *N-linked Glycosylation* of enzymes is a promising approach to achieve better enzymes in terms of activity and stability. Glycosylated enzymes could solve the problem of inefficiency in enzymatic reactions due to suboptimal conditions. The approach utilized in this study could be useful and speed up industrial applications that require enzymatic reactions.

4.1. References

1. Wang, Y.C., S.E. Peterson, and J.F. Loring, *Protein post-translational modifications and regulation of pluripotency in human stem cells*. Cell Res, 2014. **24**(2): p. 143-60.
2. Dell, A., et al., *Similarities and differences in the glycosylation mechanisms in prokaryotes and eukaryotes*. Int J Microbiol, 2010. **2010**: p. 148178.
3. Spiro, R.G., *Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds*. Glycobiology, 2002. **12**(4): p. 43R-56R.
4. Shental-Bechor, D. and Y. Levy, *Folding of glycoproteins: toward understanding the biophysics of the glycosylation code*. Curr Opin Struct Biol, 2009. **19**(5): p. 524-33.
5. Rudd, P.M., et al., *Glycosylation and the immune system*. Science, 2001. **291**(5512): p. 2370-6.
6. Reid, C.W., K.M. Fulton, and S.M. Twine, *Never take candy from a stranger: the role of the bacterial glycome in host-pathogen interactions*. Future Microbiol, 2010. **5**(2): p. 267-88.
7. Haltiwanger, R.S., *Regulation of signal transduction pathways in development by glycosylation*. Curr Opin Struct Biol, 2002. **12**(5): p. 593-8.
8. Balog, C.I., et al., *Mass spectrometric identification of aberrantly glycosylated human apolipoprotein C-III peptides in urine from Schistosoma mansoni-infected individuals*. Mol Cell Proteomics, 2010. **9**(4): p. 667-81.
9. Marth, J.D. and P.K. Grewal, *Mammalian glycosylation in immunity*. Nat Rev Immunol, 2008. **8**(11): p. 874-87.
10. Moremen, K.W., M. Tiemeyer, and A.V. Nairn, *Vertebrate protein glycosylation: diversity, synthesis and function*. Nat Rev Mol Cell Biol, 2012. **13**(7): p. 448-62.
11. Brimer, C.D. and T.C. Montie, *Cloning and comparison of fliC genes and identification of glycosylation in the flagellin of Pseudomonas aeruginosa a-type strains*. J Bacteriol, 1998. **180**(12): p. 3209-17.
12. Grass, S., et al., *The Haemophilus influenzae HMW1 adhesin is glycosylated in a process that requires HMWIC and phosphoglucomutase, an enzyme involved in lipooligosaccharide biosynthesis*. Mol Microbiol, 2003. **48**(3): p. 737-51.
13. Wu, H., et al., *Isolation and characterization of Fap1, a fimbriae-associated adhesin of Streptococcus parasanguis FW213*. Mol Microbiol, 1998. **28**(3): p. 487-500.
14. Stimson, E., et al., *Meningococcal pilin: a glycoprotein substituted with digalactosyl 2,4-diacetamido-2,4,6-trideoxyhexose*. Mol Microbiol, 1995. **17**(6): p. 1201-14.
15. Castric, P., *pilO, a gene required for glycosylation of Pseudomonas aeruginosa 1244 pilin*. Microbiology (Reading), 1995. **141** (Pt 5): p. 1247-1254.
16. Dobos, K.M., et al., *Definition of the full extent of glycosylation of the 45-kilodalton glycoprotein of Mycobacterium tuberculosis*. J Bacteriol, 1996. **178**(9): p. 2498-506.

17. Szymanski, C.M., et al., *Evidence for a system of general protein glycosylation in Campylobacter jejuni*. Mol Microbiol, 1999. **32**(5): p. 1022-30.
18. Lu, Q., S. Li, and F. Shao, *Sweet Talk: Protein Glycosylation in Bacterial Interaction With the Host*. Trends Microbiol, 2015. **23**(10): p. 630-641.
19. Nothaft, H. and C.M. Szymanski, *Protein glycosylation in bacteria: sweeter than ever*. Nat Rev Microbiol, 2010. **8**(11): p. 765-78.
20. St Geme, J.W., 3rd, S. Falkow, and S.J. Barenkamp, *High-molecular-weight proteins of nontypable Haemophilus influenzae mediate attachment to human epithelial cells*. Proc Natl Acad Sci U S A, 1993. **90**(7): p. 2875-9.
21. Voth, D.E. and J.D. Ballard, *Clostridium difficile toxins: mechanism of action and role in disease*. Clin Microbiol Rev, 2005. **18**(2): p. 247-63.
22. Tan, F.Y., C.M. Tang, and R.M. Exley, *Sugar coating: bacterial protein glycosylation and host-microbe interactions*. Trends Biochem Sci, 2015. **40**(7): p. 342-50.
23. Nothaft, H. and C.M. Szymanski, *Bacterial protein N-glycosylation: new perspectives and applications*. J Biol Chem, 2013. **288**(10): p. 6912-20.
24. Lizak, C., et al., *X-ray structure of a bacterial oligosaccharyltransferase*. Nature, 2011. **474**(7351): p. 350-5.
25. Feldman, M.F., et al., *Engineering N-linked protein glycosylation with diverse O antigen lipopolysaccharide structures in Escherichia coli*. Proc Natl Acad Sci U S A, 2005. **102**(8): p. 3016-21.
26. Koenigs, M.B., E.A. Richardson, and D.H. Dube, *Metabolic profiling of Helicobacter pylori glycosylation*. Mol Biosyst, 2009. **5**(9): p. 909-12.
27. Oyelaran, O. and J.C. Gildersleeve, *Glycan arrays: recent advances and future challenges*. Curr Opin Chem Biol, 2009. **13**(4): p. 406-13.
28. Celik, E., et al., *A filamentous phage display system for N-linked glycoproteins*. Protein Sci, 2010. **19**(10): p. 2006-13.
29. Wacker, M., et al., *N-linked glycosylation in Campylobacter jejuni and its functional transfer into E. coli*. Science, 2002. **298**(5599): p. 1790-3.
30. Baker, J.L., E. Celik, and M.P. DeLisa, *Expanding the glycoengineering toolbox: the rise of bacterial N-linked protein glycosylation*. Trends Biotechnol, 2013. **31**(5): p. 313-23.
31. Ihssen, J., et al., *Production of glycoprotein vaccines in Escherichia coli*. Microb Cell Fact, 2010. **9**: p. 61.
32. Cain, J.A., et al., *Identifying the targets and functions of N-linked protein glycosylation in Campylobacter jejuni*. Mol Omics, 2020. **16**(4): p. 287-304.
33. Wacker, M., et al., *Substrate specificity of bacterial oligosaccharyltransferase suggests a common transfer mechanism for the bacterial and eukaryotic systems*. Proc Natl Acad Sci U S A, 2006. **103**(18): p. 7088-93.
34. Ollis, A.A., et al., *Substitute sweeteners: diverse bacterial oligosaccharyltransferases with unique N-glycosylation site preferences*. Sci Rep, 2015. **5**: p. 15237.
35. Chauhan, J.S., et al., *GlycoPP: a webserver for prediction of N- and O-glycosites in prokaryotic protein sequences*. PLoS One, 2012. **7**(7): p. e40155.
36. Sayler, G.S., M. Puziss, and M. Silver, *Alkaline phosphatase assay for freshwater sediments: application to perturbed sediment systems*. Appl Environ Microbiol, 1979. **38**(5): p. 922-7.

37. Micsonai, A., et al., *BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra*. Nucleic Acids Res, 2018. **46**(W1): p. W315-W322.

4.2. Appendix A

4.2.1.DNA Sequences of Genes

Table 17: The DNA sequences of genes that was used in this work.

Plasmid Name	Gene Name	DNA Sequence
Pgl-pAC YC	<i>pglK</i>	TCATTTCTCCTCTTTAAGCTTACCGTGTTCTAAAC GATAAACCTTATCACATTGCGTTATCGTAGAAAGG CGATGTGCGATAATAATCATGGTTTTATCTTTAGA AATTTTATAAATTTTCATCCATGATTTTTGCTTCACT TTGAGTATCAAGCGCAGAAGTTGCTTCATCAAGC ACTAAAATTTCAAGGCTCTAAATACAAAGCCCTTG CTATAGCTATGCGTTGTTTTGTCCCCCGCTTAAA TTACTCCCCCATCGCCTACTTTTGTCTGAACTCC TTGAGGTAAATTTTTTATAAAATGCTCTAAATTTG CTTGTTTGATAACCTTATTAAGTTTTTCTTCATCAA CCGCATCTCCAAAAGTGATATTTTAGCTATGCTA TCATTAAAAAGATAGATATTTTGCGGGATATAGCC TATTTTTTGGCGATAATTTTTTGCATTACTTGCATT TAATTCTTGCTTATCAATTAATAAATTTGCCCTTCTTT TGGTTTTAAAGTCCTATGATAAGATCTACTAAGG TACTTTTTCCACAACCACTCTCCCCTATAAAAGCG ATTTTTTCACCTTTTTTAATGTTTAAATTAAGATTT TTAAATAAATATTTTTTTCCCTCATAACCAAAGCT AAGATTGCAAATTTTAAGCTCTTGATTAAAACTTA ATTTCCCTCGCCCAAATTTTCTTCTTCTTGTCTTA AATTTTGATAAATAATATTCAAAGAAGAATGATAA TAAAGCAAATCATGATAACTTGTGATAATACGATT TGCGCTTGGCATTAAAGCGATAAAGCGCTAAAACA AAAATAGAAATCGTGGATAAAAATACCTGAAATAT CACTTTCATTTTTTCAAACCAAAAATACCACGAT AAAAACAAGCACACAAAAGCCTATTCTTCAAG ATAAATTCTAGGCACCGCAGCTACGCTTTCATTGG TAATATTTGCTTTAGAAAAAGCTTCACTTTGTGCT TTAAAAAGACTTAATACTCCATCTTCTTTGGTTTT AAGCTTGATAAATTTAAAATTATTTAAATTTGTATT TAAAATTTTGAAGAAATTTTTTCATCGCTTCTTCGC GTCTTAGTCCTGCTTTTTTTAATGATAGGGCTTAAA ATTTTCACTAAAATAAAAGCATTTAACACCATAAA AATACTTAAAAATAAAGTGATTTTATAATTAATCA AAAGCATTAAAGCATAAAGCAAAAGCACTACAA AAATTTCACTCATCAAAAGTAAAAATGATGAAAT CATAGTGCTTAGATTATAAACTTCCCCTGTAATGG ACTTTAAAATTTTCAAGATTGATTTTTTTGAGTAAAT TTTTCATAATTAATATTTAAAAATTTAGAAAAAAC CTTATAAGCGATCGCATGATAACGTCCTTTAGAAA ATCTTGCCAAAAGATGAAAATAATACGCATTTAA CAAAGCTCTAAATACATAAAAAACAATAAGCCCC ACTCCAAAATAAACAATGATTTCAAAAACAGGG ATATTAAGATATTCTTTTAGGCTGATTAAATATTTA TTACGATCAAAATAAGAAAAATCACTAGCCAAAG TGATAAAAGGCATTACCAAAGAAATCGCAAAAGT TTCTATAAAAGAAACAAATACTGAAAAACAAG

		CAAGAAAAATAAAAAATTTTTATCTTCCTTGCTTA AAATAAAAAAAGTTTTTTTAGCAC
	<i>pglH</i>	TTAGGCATTTTTAACCTCGACTATAAGCTTAAGCC ATTCTTCTTTGATGTTAGAAATTTCAAAGTCTTTA CACCTTTGTTTGGCATTATTTACAAGTTCTTTTCT AAAATTTTCATCATTTTAAAACAAGTTCAAGTTTTT TAGCAAGTGCTATTTTCATCATCACAACTACCAA AAGCCCATCATGATTATCTTTGATTAAATCCTTAG CACCATTATAATAAGAACTTGAAATTCTACAAACC TCAAATACAACTTTCAATTAAAACCGTTGGCA AACCCTCTACAAAAGAACAAGGCAAAGCACTT TTGCTTTTTCATAAAGAGCCTTGACATTTTCAACG CGTCCTAAAAAATCAACTTTTATTCCTAAAGATTT TACCTTGTATTCAAGTTCTTGCCTTAACTCTCCAT CTCCTGCTATAACAAATTTATAATTTTCTTGCAAAT TTTTATCCAAATGTGCTATAGCTTTTAAAAACATT ACAGGGTTTTTGTGTGATCTAAACGCCCTATAA AAAGAACCAAATTTTCTTTTCAAACCTAGAAATC AAAAGGAATTCATCGCTAAAATGACAAGGGTTT AATAAAAGCTTAACCTTTTACAAATCTTTCATA ATACACCTTATCACTGCTTCCAAGCACACTTAAA GCATCACAAAAGGATAGCTTACCCTTCTTAAAA AACGCCAAGTTTTAGGTTTTAAATACGCTTCATTA CTATGCTCGCTAATGATGAGTGGAGTTTTAAGTCC TATATTAGCTAGAATACAAGCAATATTAGTTGTATC CAAAAAAGAAATAAAAACATCAGCCTTACTTTCT TTTAAAGCCTTTCTTAAGGCAAAAAATTTCTTAA AACGACTTGCGATTTTATGATAAAGCGTGTCAAA TCTAAATTGTTCTAAACTTGTAACTTTAACTTCAT TTTCAAGCTTATAAAAAGATTCTCCTGTATGAAAT TTAATAATACTTACTTCATGCTCTTTGCAAAGCGC ATTAGCTAAGGTTACTAAAACACGCTCAGCGCCT CCTGAATTTAAAGTTGCGATAATAAAGCTTATTTT CATCAT
	<i>pglI</i>	TTAGGCATTTTTAACCTCGACTATAAGCTTAAGCC ATTCTTCTTTGATGTTAGAAATTTCAAAGTCTTTA CACCTTTGTTTGGCATTATTTACAAGTTCTTTTCT AAAATTTTCATCATTTTAAAACAAGTTCAAGTTTTT TAGCAAGTGCTATTTTCATCATCACAACTACCAA AAGCCCATCATGATTATCTTTGATTAAATCCTTAG CACCATTATAATAAGAACTTGAAATTCTACAAACC TCAAATACAACTTTCAATTAAAACCGTTGGCA AACCCTCTACAAAAGAACAAGGCAAAGCACTT TTGCTTTTTCATAAAGAGCCTTGACATTTTCAACG CGTCCTAAAAAATCAACTTTTATTCCTAAAGATTT TACCTTGTATTCAAGTTCTTGCCTTAACTCTCCAT CTCCTGCTATAACAAATTTATAATTTTCTTGCAAAT TTTTATCCAAATGTGCTATAGCTTTTAAAAACATT ACAGGGTTTTTGTGTGATCTAAACGCCCTATAA AAAGAACCAAATTTTCTTTTCAAACCTAGAAATC

		AAAAGGAATTTTCATCGCTAAAATGACAAGGGTTT AATAAAAGCTTAACCCTTTTTACAAATCTTTCATA ATACACCTTATCACTGCTTCCAAGCACACTTAAA GCATCACAAAAAGGATAGCTTACCCTTCTTAAAA AACGCCAAGTTTTAGGTTTTAAATACGCTTCATTA CTATGCTCGCTAATGATGAGTGGAGTTTTAAGTCC TATATTAGCTAGAATACAAGCAATATTAGTTGTATC CAAAAAAGAAATAAAAAACATCAGCCTTACTTTCT TTTAAAGCCTTTCTTAAGGCAAAAAATTTCTTAA AACGACTTGCGATTTTATGATAAAGCGTGTCAAA TCTAAATTGTTCTAAACTTGTAACCTTAACTTCAT TTTCAAGCTTATAAAAAGATTCTCCTGTATGAAAT TTAATAATACTTACTTCATGCTCTTTGCAAAGCGC ATTAGCTAAGGTTACTAAAACACGCTCAGCGCCT CCTGAATTTAAAGTTGCGATAATAAAGCTTATTTT CATCAT
	pglJ	TTATCCTAATAAATATTTCAAAGCATCGCGTGCAA TTTTTACTTTTATCAAAAGCTTTAGCCCTAGTTTTA GCCTTATTTTTATACGCTTTTCTTAATTTATCGTCT TCAAGCATAGTCTTTAAACCCTGAAACATGGAGT TTTCATTATCTACTTCTACTAAAAGTCCAAATTCAT CATCGCCAAAAAGCTCTCTTGCGCCACTTTTATG ATCAGTGCAAACCACCGCACAGAACAAGCTAA GCTTTCAATTAAAACATTTGAAAAGCCTTCAAAT ACAGAAGCAAAAGCAAAAAATTCGCATTTAGCC ATGTATTTATAAGGATTATTATCAAATCCTAAAAGC AAAACCTTTTCTTCTAAATTTAACTCTTTAATCAA AGCTAAAAGCTCATCTTTAAGCACACCTTCACCT AAAATCACAAGCTTTAAATCTGTTTTCAATCTTGC ATAAGCACGAATTAATAAAGCATGATTTTTACCTT TATCAAGCCTGCCTACACTTAAAATAAATTTATCT TTTAAAGCTACATCTTCAAGAGCTTTTTGCTCTAT ATTTTCTAAATCGATGGCATTGTATAAAATTTACAC ATTTTTTTGGATTTATACTAAAATTTTGCCTAAAT CTTCTAAATTTCCCTTAGAATTAGGCAAGATTAAA TCAGCTTTTGGGTAAAGCAAAGAAATTAAAAAATT TATTTGCCAAAGAATTAAAATTATTTTTCGCATAC ATCACACTTGGCGTAGTGCATTCATTGATCACAA GGCGAGTTTTGTTTCCAAACATTCTTGCCATTAA GGCTATATAATTAGGTGATTTAAAAAAACAAATT CTGTATCAATGCCTAAATTTCTGCAAAGTTTTTTA TATTTTAAAGCCAAAAAAGGTAGTTTTAAAAAATT TCAAAATAGGATTTTCACTAGGTTTTGAACATTCT AAAAAATGAATTTGACACTCTGGAATTTTATAAG AAATTTTATCATTCAATTAATAAATCAAATGCACTTCA AATTTCAAACCTTAAAATAGGCAATAAAGTCGCCA CAACTCTTTCAGCACCACTTCTTAAAGAATA AATAAAAAATGCCTAATTTTTTGCAT
	pglB	TTAAATTTTAAAGTTTAAAAACCTTAGCATCTCTAG AATTAATCACCAAGTCAAATAAATTCTTATCATAA

	TTTCCTAAAAAAAACATTTGCACATAAGCACTATT AAACATGGTTTTATCCATTAAAATAAATTGTGCGT AAGGAATAGCACTATCCTTTAAATAAAAAATATAA AACTGAGCCTTATCATCAATTGGAGTGATTTTGTA TTCACCTTGTTTAATAGAATTAATCTCTACGATACT ATTTACAGAAACCACATTATCACCTATTTTAAAC TTCTAAAATCATCGCTTAAAACCACTCCGTTGCTA AGATAAATTTCTCCATTTTAAACATCAAGTGGATA AGCTGTGCTAAAGGTAAAAGGTTTATCCAAAAC CCTGTATCTAAATTAATAAAAGAAAACTAGCCA CCGTAGAAAAAATCAAAGACATTCTAGCGGGCAT ATAAAGATAAATATCACGAGTTTTTGGCGTATCGA TTTTAAAATCAGGTTTTTGATAATGAAGCTAGAAA CAAATCCACATTGCTTTGATTATAATCTTTCATCAT GGCTTGCAAAATGTCTGTTTTTAAAATATCATTTT GCGGAGCATAAAAGCTTTTTTCTGTATATTCTACA CTAAGTCTTGCCATATTAGCTGCAGCTTGTTTCATC TTTGCTTAAAGCAAAAGAAGGGAAAAAATTATCC TTACCTAAATGCTTTCCACCATCTACTAAAGTTTT CACATCGCTATAATAACGCACAGGATAACCATAAT CCCACCAAGTTACCACATAATCTTCTCTATTGGCT ATATTTTTTAATTGATTTAATAATGATGCTTCATTTT GAGAAAAAACTGTTGGCGCTTTATAGTTGTAAAT ATGGATAAATACTGGAGCTAAAGTCAAAATAGTT GCAAAAACAATACAAACATTTGAAGTTAATTGGC TATATTTTTTAACCATATAGCCTTAAACTCGCTCA ATAAAAAACCAAATCCTAAGGCCATTACAGGTAC AGAATAAATGGTAAATCTAAGCCCCCTTTTAAG GCTAAAAACCCAAGCACCAATATAGGTAAAGCCA TAATCATACTTTTATGTTTTTCTCAAAAGCCATACA AAACCAAACAAAGAAAAACAAAAAACAATTTCA CTACCACTAATTCTTCGCATAAATTTCGCTAAGATC TACATTTTCAACTTCTTGTATGGTTTGATTGACATT AAAATACATAAAGCCCTGCGTTAAATTTCGCACTT TCATCACTTCTAAAAATATAAAATTTAAGCTGATA AAGTATAGGATCAACCCCACTTAAAATCAAA AATATCAAAGTTGCACTACCTAAAATTCCTATAAT CATAAAATTTAAGCGTTTTTGCTCTAAGGCGAATA AAGCAAAAAGTATTACTATAATGGCACTTTGATAA AACCATGCTATATTTGAAAGAGTAAGAGAAGACA AAATCACAGCTATATAAAAAATCTTTTCTTTTCTAT GAAAAATAAGTGTATAAATTAAAAAAAGTCCAAT TAAAGCTACATTTAAAGTATAACTTGAAGGATACC ACCAAAGATAAATTCCTATAAATAACGGCAAGGC AATCAATGAAAAAAAGTCTTTTTTTTAAATCATTC TTACCATAAAAAATAAAATAAACATAGGCAAAAC AATTACCAGCATATCCGTATCATAATACCCACTCAT AGTGCGATTATAAATACTGTTTGCTATACTTGCTA AAAGAGCAGCTACAAAGCCCATTAAGGACGTT TGTATTCGTTAGCTAGCAAAATAGTAGGAATCACC
--	---

		ACCAAAGAAGATAAAAAAGTACTCATATATAAAA TGATACTTTCAAAAAGAAAAAGGTGTGATTTTATA AAGCCAATAAGTAAGCGCGGATAAAGAAGATCC ATAATAACTCAAATCATTAGGCTGATGAAAACCT GCTATCATATCTCTTGCGCCCTCAGCAAAAAGCATA GCCATCATTTGAAATGATCATTAAGTATTATTGA AAAAATACTCATTAAGTCACTTGCCCAACCAAC CCAATAAACCTGCAAAATACACTAAAAACATAA GCTAATATAATCATCGCAAAACAAACTAAATAAG GGTTTTTTAAATACTCTTTTTTCAACAT
	pglA	TCATACATTCTTAATTACCCTATCATAAAGTTTTAA ATAACGCTGTGCGATATTATTTTCATCGTATTGTAA AGCATCTTGGGCAGCATTTTTAGCTAAATTTAATC TTAATTTTTTCATCTTCTAATAAAAGTGAAATTTTT CGCTTAAATCCTTAGCATTTTTTGTTTTTGCCCAA AGTCCATCATAAGCATTAGAAATAGCCTCTACACA ACCTTCACAATCACTCACCCTATAGCCTTGCCA CAAGCTTTTGCCTCCAAAACACTTACAGGAAAG CCTTCTTTATAGCTTGGTAAAACAAAAATATCACA ATTTTGCAAAAGCTCGACTATATCACTTCTAGCAC CCAAATAATGCACCACACCCGAGTTTAAAAACTC CAAACCTCGCACAAGAAGGATTTTCATCTCTTCCA CCAATAAAACAAAATTTGCTTTGTCTTTTAGCAT AGTAGCACTTTCATAAAATTCTTTTACACCTTTAT GCCATAAAGCTCTTGCTATCATAAGAACAATAGGT TTTTTATCTATATTTAAATTTCTCCAAAAAAGCTCT TTTTTTTCCGATTCTATATAAATAGGAAAAAATTTT TTTAAATTGATCCCTACGGATTTTATCACACAAAT TTTATTTTCCTTAAGTCCTAAATTCCGCATAAACT CGGCATTACTTTCATTGACAAAAATAAATTGGTGT GCAAATTTAAACTTAATTTATAAAGATTATTAATA ACAAAACGTACTAAATTTGCCTTAAACCTTGAT CTATATAAAAAGATCCCAAGCCTTCTACCAAAGC AAAACGATAAGGAATTTTTGCCCATTTTGCCGCT AAAATTCCAAAGGTATTGCTTTTGTGTGCCGCAC TTTGAATAAGATCAAGATTAAATTTTTTAAAACC TTAGCAAGATAAAAAAAATTCTTTAAACTACAA AAGGATTTAACTAGCTCTTGAAAACCTACATAAC AATTACTTTTAAACCAAGATCTCTAAGTTTTTGCG TGTATTCATCTTGCGGCACTATAACAAAAACTTCA TCTTTTCTATCTTTTAATGCTTTTATAATAGGCATT CTAAAATGATAAATACTTGCTCCTGCATGTGATAA AAATCCTATTCTCAT
	pglC	TCAGTTCTTGCCATTAAATTTCTCTGTTGTAACAT GGCCTTCTTTGCTTACCCCGCTTCGTTTTAAAACC TTTAAAGCTGTAAAAAACATGATTTTTTAAATCAAG CAAAAAAGAAATATTTTTTACATAATACACATCA AGTTCGAATTTTTTCTGCCAAGAAATAGCATTCT GCCATTACCTGCGCCCATCCTGTTATACCTGGAC GCACCTTATGGCGCAATTTTTGCTCTTCATTATAA

		AGGGATAAATACTCAACCAAAAGAGGTCTAGGC CCCACAAAACATCATATCCCCTTTTAAAACATTAAA AAGTTGCAAAAGCTCATCCAACTTAAGCTTCTA ACAATTTTTCCAAAGGCTTTCAAACGCAATTCAT CGCTTAATAACTCACCTTCTCATCTCTTTCATCG CTCATGGTTTTAAATTTATAAATTTTAAAAATTTT TCATCTAACCCAGGGCGATTTTGAGTGAAAATCA CACTTCCTTGAGTGATTTTAAAAGTAAAGCAGT GATTAAAATCACCGGAGAAAAAAGTACTAAAAG CACTAAAGCTAAAATAAAATCAAAAATTCTTTTA AAAACTTTTTCATACAT
	pglD	TCATATTTTTCTTTTTGCAGGCACTCCCACAAAAA CACCTTTTTTCATTTTGGCTTTTAACCAAAGTTGCT CCACCACCCAAAATACTATCATCTGCTAAGCTTAA ATTAGGCAAAACACAAGAATTAATCCCTAAAAAA CAATTTTTACCGATTTTACATTACCCGCACATTTA GCCCCTACGCTTACATGAGAAAATTCCCCTATCAC GCATTCATGCTCAATCACACTTGAAGTATTTAAAA TCACGCCTTTTTCTATTTTGGCTTTAGCGTTAATC ACTACATAGGGCATAATTAAAATCCCTGCATTTTC TTCCACGCTTGCACTAGGACTTATAAGTGCACCTT TATGAATAAGATTAACCTATTTTAAAGCCATTTTCT GAAATTTTTTGATAAATCTTTTTTCGAATTCATTG TTTCCTATAGCTATAAAAAAATCATATTTAGGCAA GGTGTTTTCAAATTCATTCCTTTAAATCATCTA AAAAAATACATTCTTTATAGCCCATATTTTTAGCC ACATCTTCACAAACAAGCCCATGACCACTAGCAC CATAAATATAAATTTTTTTCAGTTCTTGCCATTAAAT TTCTCTGTTGTAACAT
	pglE	TTAAGCCTTTATGCTCTTTAAGATCAGTTTTGAAA TTTCATAAACATCATCTTTGCTCATCGCCGTACCA CTTGGCAAACAAATTCCTTTTTTGAAAAATAACT CACTATTGCCATTAAGATAAGCCTTAGTTCCTTTA AATACTTCTTGAGTGTGCATAGCTTTCATAATGG TCTTGTTTCTATTTGTTCAATTTTCAAATCTTCTAT GAGTTTTGAAATTTTTGGATGCAAAGTAATATTTT TTTGACTGATATTTATATCTTTTTGACAAGCATTA GTTCAATTTTTATCAAAATCAATCAAAGCTGTACTT AACCAGCGATTACTTCTTGAATTTTCTAATTCATC TAAAAAGCTAAAATACTCTCCTAAAAAATTCTTTAT ACCACTCATAAATTTCTCTTTTTTTAAGCACTCTT TGTTCTAAAACCTCCATTTGTGCTACACCAATAGC TCCTAAAACATTGCTTAAACGGTAATTATAACCAT AGTCTAAATGTTTATAATGCAAACAATTTCCCTA GCTTGAGTGCTATAAAATCTTGCTTTTTCAATCTT TTCTTTATTTTTTCCTATAAGCATACCTCCACCTGA AGTGGTGATAATTTTATTGCCATTATAAGAATAAG CTCCAAATTCTCCAAAAGTTTCTAAGGCTTTATTC TTATAAAAACCTTCCTAAAGCTTCAGCAGCATCTTC GATTAAAACAATTTCAATTTTCTTGCAAATTTCAA

		CAATTTTCATCCATTTTGTAGCCGCATTGCCATAAAGA TGAGTTAAAATCAATGCTTTTGGTTTTTTTTTTCACA TTCCTTAATAGCAAGTTTTAACAAATCTACATCGA TATTATAAGTTTCATCACAATCTATAAATACAGGTT TTGCTTTAAGATAACAAATAGGCGCCACTGAAGC GATAAAAGTAAAAGAAGAAGCCAAAACAATATC ATCTTGTTTTTACCCCTGCCACCCTTAAAGCTAAAT GCAAAGCTGCTGTAGCTGAATTTAAAGCTAAGGC ATTTTCACTTTTGCTATAATCTTTTACACTTTGCTC AAAGCGATTTACAAATTCACCCAAAGGTGCTATA TAATTGCTTTTGAAAACCTTCTTCTATATATTTTAAT TCATTACCGCCCATATGCGGAGGAGAAAGAAAAA ATCTCAT
	pglF	TTATATGCCTTCTTTATTGTGTTTAAATTCAGGTAC AATCTTTAATAAAGCCTCTGAAATATCTTCACATA TTTGTAATTTTCTATCTCTTTATTAAGCCAATCAA GATCAACCTTCTCATTCTTTGCTACAAAAATACTC TCATATTGGGTTTTAGCATCATTTTCATCAATCAA AAGCTCTTCATAAAGCTTCTCACCTTTTCTTAAAC CTGTGATTTTAATTTCTAAATCATTGCGATTAGAA AGTAAAAGCATTTTTTTTAGCTAAATCTATGATTTT CACAGGCTTACCCATATCCAAAACAAAAAGTTCT CCCCCTTTTGCGATAGCTCCAGCTTGCAAAACAA GTTGCACTGCCTCAGCCACAAGCATAAAATAACG CACTATATCAGGGTGTGTTAAAGTTAAAGGCTCAT TATTGGCAATTTGTGCTTTAAATTTTCGGTATCACA CTACCACTAGAACCTAAAACATTACCAAAACGCA CACAAGCAACTTCAAAATTTTCATCACTCATACTT AAAGTATAAAGCTCACAACTCTCTTAGTGCAAC CCATAATATTTGTTGGTCGTACTGCTTTATCTGTAC TTATCATCACAAATTTAGCTACTTTATTTTCTTTAG CACTGTTCGCATAAAATTTTGGTTCCTAAAATATTA TTGATTACTGCTGAATGTGGATTTTGTTTCGCAAAG AGGCACATGTTTATAAGCGGCTGCATGTAAAATA AGCTCAGGTTTATAAGTTTTTAATACCTCATCTAA ACTTTGCTTATCTAAAATACTTAACAAAATAGGAG TAATTTTTTCTTTATATAAATTTAAATCATCATTGAT CTTATAAAGATTATACTCACTATGATCAACCATAAT AAGATGCTTAGCACCAAATTTAATACATTGCTTAC AAAGTTCACTGCCTATAGTTCCACCTGCTCCACTT ACCAAAACTACCTTGTCTTTTAAAAAAGCCGCCA CAGCACTATCGTCTAAATCTTTTGGTTTTCTAGCA AGCAAGTCTTCTATACTAATATCTTTGCTTCATT CTTGTAAGAAAGAAAATATTTTACATCACAAATACC ATAAGCTACAAGTTCTTCAAAAAGTTTTTTAAGC TCTTCTTGTTCAAGTCTTAAAGCAATAATGGCAGT TTTTACCCCTTGTTCTACATAAGATTTTATTTTTTC TTTTTCTTCTACAACAAATTTATCACATAAAGTCC CTATAAGCTCTTTTCTCGCATCAACTACGCCACA GGAAAAAGCCCTAAAGAACCTTCTTTTGCGCCTT

		TTAACAAATGTAAAGCCTTAGAAGTTGCCCTAC TACAATACAAGGAGTTTCTTCTTCTTTTATTAG AAGGTTTAAAATCCACAAGCATTCTTTTGCTAATT CTTAAAGTACCTATAAACATATAAGAAAGAACAA AATCTATCACAATAGCACTTCTTGGAAGGATT AAAAAATCACTAAAAAATAAAAAATAAGAAA AAAACAAAACCTCAGCTAAAAGCAAAGCGATAAA AATTTTTCTTGCTTCATTGAGGGAAAAAATCTC CAAGCTACTTTATAAATTCTAAAAACAAACAAA ATGAAAGTTTTAAAACAAGCAAATAATAGCAGA AACCATCATGCCATGATAAAAAATACTCGGAATAT CTCCACTAAATCTCAAAGAAAAAGCCAGATAAAC ACTAAGTAAAATTAAAACAATATCTGAAGTTAAA AAAAATGCTAATCTTTTGCTTTTATAAAAAATCAT
	pglG	TCATGGCGACCACACCCGTCCTGTGGATCCCCCG GGCTGCAGGAATTCGATCATGGTTTTTTGAATTCC ACCTTCTTCTTTTTTTGTCCTTCCTAATACTTCTTG TTCGATCAAAGTTTTTGCCAAATCAAAAATAGCT AAAGCTAAAGTTATGGCAATGGTACTTTCAAACA TATGTTCTATGCCTATAGCTTGATTATCAAAAAGA CTAAAAAAACCTTTTTGAAATAAAAAACAAGTAA TCACAAATAAAACCAGTGCGAAACAAAATACA CCAAACGACTAAATTGAGTAAAAACAACTCAA ATTTTCCTGCTTGAATAATCTTTAAATATCTTCAA GCTTGATATCAATACAAACAACAAAAGCAAGTT ATTTTATCATCGTAAATTGGTATAGAAGCTGTTA CGCAAAGTTCATTGTTTAAAACCGAAGGATAAGG ATCGCTCAAACACATCGTTTTAATTTTACAGCTC TATAAAAATAAGAACGAGTACTTCTATCCTCACCC TGCCCTATTTTCATGAGCCTTATTTAAGGAAATCGC ATTTTCTATTTGCATGCCTGAAATATCTAAAATATA CAAAGCATCAAAAACAACAACCTTCATGAGCTATT TTATCCAAAGCTGTTTTAATCGTTTCAAGGTGAAT ATCAGGAAGTTTATTGGGTAGATTTCTTGTTAAAA TATAACTCATATAAGCTCTAGCACGATAGCGATTA TCTTCAAATCTTTGTATATCTTTTATATACAT
DQNAT- ALP pET22b	DQNAT- ALP	ATGAAACAAAGCACTATTGCACTGGCACTCTTAC CGTTACTGTTTACCCCTGTGACAAAAGCCGATCA GAACGCGACCCGGACACCAGAAATGCCTGTTCT GGAAAACCGGGCTGCTCAGGGCGATATTACTGCA CCCGGCGGTGCTCGCCGTTTAAACGGGTGATCAGA CTGCCGCTCTGCGTGATTCTCTTAGCGATAAACCT GCAAAAAATATTATTTTGCTGATTGGCGATGGGAT GGGGGACTCGGAAATTACTGCCGCACGTAATTAT GCCGAAGGTGCGGGCGGCTTTTTTAAAGGTATAG ATGCCTTACCGCTTACCGGGCAATACACTCACTAT GCGCTGAATAAAAAAACC GGCAAACCGGACTAC GTCACCGACTCGGCTGCATCAGCAACCGCCTGGT CAACCGGTGTCAAACCTATAACGGCGCGCTGG GCGTCGATATTCACGAAAAAGATCACCCAACGAT

		TCTGGAAATGGCAAAAGCCGCAGGTCTGGCGAC CGGTAACGTTTCTACCGCAGAGTTGCAGGATGCC ACGCCCCGCTGCGCTGGTGGCACATGTGACCTCGC GCAAATGCTACGGTCCGAGCGCGACCAGTGAAA AATGTCCGGGTAAACGCTCTGGAAAAAGGCGGAA AAGGATCGATTACCGAACAGCTGCTTAACGCTCG TGCCGACGTTACGCTTGGCGGCGGCGCAAAAAC CTTTGCTGAAACGGCAACCGCTGGTGAATGGCA GGGAAAAACGCTGCGTGAACAGGCACAGGCGC GTGGTTATCAGTTGGTGAGCGATGCTGCCTCACT GAATTCGGTGACGGAAGCGAATCAGCAAAAACC CCTGCTTGGCCTGTTTGTGCTGACGGCAATATGCCA GTGCGCTGGCTAGGACCGAAAGCAACGTACCAT GGCAATATCGATAAGCCCGCAGTCACCTGTACGC CAAATCCGCAACGTAATGACAGTGTACCAACCCT GGCGCAGATGACCGACAAAGCCATTGAATTGTTG AGTAAAAATGAGAAAGGCTTTTTCCTGCAAGTTG AAGGTGCGTCAATCGATAAACAGGATCATGCTGC GAATCCTTGTGGGCAAATTGGCGAGACGGTCGAT CTCGATGAAGCCGTACAACGGGCGCTGGAATTCG CTAAAAAGGAGGGTAACACGCTGGTCATAGTCA CCGCTGATCACGCCACGCCAGCCAGATTGTTGC GCCGGATACCAAAGCTCCGGGCCTCACCCAGGC GCTAAATACCAAAGATGGCGCAGTGATGGTGATG AGTTACGGGAACTCCGAAGAGGATTCAACAAGAA CATACCGGCAGTCAGTTGCGTATTGCGGCGTATG GCCCCGATGCCGCCAATGTTGTTGGACTGACCGA CCAGACCGATCTCTTCTACACCATGAAAGCCGCT CTGGGGCTGAAAGGCGGCGGCAGCCTCGAGCAC CACCACCACCACCACTGA
ALP-DQ NAT pET22b and T7 pglB ALP DQNAT	ALP-DQ NAT	ATGAAACAAAGCACTATTGCACTGGCACTCTTAC CGTTACTGTTTACCCCTGTGACAAAAGCCCGGAC ACCAGAAATGCCTGTTCTGGAAAACCGGGCTGC TCAGGGCGATATTACTGCACCCGGCGGTGCTCGC CGTTTAAACGGGTGATCAGACTGCCGCTCTGCGTG ATTCTCTTAGCGATAAACCTGCAAAAAATATTATT TTGCTGATTGGCGATGGGATGGGGGACTCGGAAA TTACTGCCGCACGTAATTATGCCGAAGGTGCGGG CGGCTTTTTTAAAGGTATAGATGCCTTACCGCTTA CCGGGCAATACACTCACTATGCGCTGAATAAAAA AACCGGCAAACCGGACTACGTCACCGACTCGGC TGCATCAGCAACCGCCTGGTCAACCGGTGTCAA AACCTATAACGGCGCGCTGGGCGTCGATATTCAC GAAAAAGATCACCCAACGATTCTGGAAATGGCA AAAGCCGCAGGTCTGGCGACCGGTAACGTTTCTA CCGCAGAGTTGCAGGATGCCACGCCCGCTGCGC TGGTGGCACATGTGACCTCGCGCAAATGCTACGG TCCGAGCGCGACCAAGTGAAAAATGTCCGGGTAA CGCTCTGGAAAAAGGCGGAAAAGGATCGATTAC CGAACAGCTGCTTAACGCTCGTGCCGACGTTACG

		CTTGGCGGCGGCGCAAAAACCTTTGCTGAAACG GCAACCGCTGGTGAATGGCAGGGAAAAACGCTG CGTGAACAGGCACAGGCGCGTGGTTATCAGTTG GTGAGCGATGCTGCCTCACTGAATTCGGTGACGG AAGCGAATCAGCAAAAACCCCTGCTTGGCCTGTT TGCTGACGGCAATATGCCAGTGCGCTGGCTAGGA CCGAAAGCAACGTACCATGGCAATATCGATAAGC CCGCAGTCACCTGTACGCCAAATCCGCAACGTAA TGACAGTGTACCAACCCTGGCGCAGATGACCGA CAAAGCCATTGAATTGTTGAGTAAAAATGAGAAA GGCTTTTTCCTGCAAGTTGAAGGTGCGTCAATCG ATAAACAGGATCATGCTGCGAATCCTTGTGGGCA AATTGGCGAGACGGTTCGATCTCGATGAAGCCGTA CAACGGGCGCTGGAATTCGCTAAAAAGGAGGGT AACACGCTGGTCATAGTCACCGCTGATCACGCC ACGCCAGCCAGATTGTTGCGCCGGATACCAAAGC TCCGGGCCTCACCCAGGCGCTAAATACCAAAGAT GGCGCAGTGATGGTGATGAGTTACGGGAACCTCC GAAGAGGATTCACAAGAACATACCGGCAGTCAG TTGCGTATTGCGGCGTATGGCCCGCATGCCGCCA ATGTTGTTGGACTGACCGACCAGACCGATCTCTT CTACACCATGAAAGCCGCTCTGGGGCTGAAAGAT CAGAACGCGACCGGCGGCGGCAGCCTCGAGCAC CACCACCACCACCACTGA
ALP-DQ NAT(x2) pET22b	ALP-DQ NAT(x2)	ATGAAACAAAGCACTATTGCACTGGCACTCTTAC CGTTACTGTTTACCCCTGTGACAAAAGCCCGGAC ACCAGAAATGCCTGTTCTGGAAAACCGGGCTGC TCAGGGCGATATTACTGCACCCGGCGGTGCTCGC CGTTTAACGGGTGATCAGACTGCCGCTCTGCGTG ATTCTCTTAGCGATAAACCTGCAAAAAATATTATT TTGCTGATTGGCGATGGGATGGGGGACTCGGAAA TTACTGCCGCACGTAATTATGCCGAAGGTGCGGG CGGCTTTTTTAAAGGTATAGATGCCTTACCGCTTA CCGGGCAATACACTCACTATGCGCTGAATAAAAA AACCGGCAAACCGGACTACGTCACCGACTCGGC TGCATCAGCAACCGCCTGGTCAACCGGTGTCAA AACCTATAACGGCGCGCTGGGCGTCGATATTCAC GAAAAAGATCACCCAACGATTCTGGAAATGGCA AAAGCCGCAGGTCTGGCGACCGGTAACGTTTCTA CCGCAGAGTTGCAGGATGCCACGCCCGCTGCGC TGGTGGCACATGTGACCTCGCGCAAATGCTACGG TCCGAGCGCGACCAAGTGAAAAATGTCCGGGTAA CGCTCTGGAAAAAGGCGGAAAAGGATCGATTAC CGAACAGCTGCTTAACGCTCGTGCCGACGTTACG CTTGGCGGCGGCGCAAAAACCTTTGCTGAAACG GCAACCGCTGGTGAATGGCAGGGAAAAACGCTG CGTGAACAGGCACAGGCGCGTGGTTATCAGTTG GTGAGCGATGCTGCCTCACTGAATTCGGTGACGG AAGCGAATCAGCAAAAACCCCTGCTTGGCCTGTT TGCTGACGGCAATATGCCAGTGCGCTGGCTAGGA

	CCGAAAGCAACGTACCATGGCAATATCGATAAGC CCGCAGTCACCTGTACGCCAAATCCGCAACGTAA TGACAGTGTACCAACCCTGGCGCAGATGACCGA CAAAGCCATTGAATTGTTGAGTAAAAATGAGAAA GGCTTTTTCTGCAAGTTGAAGGTGCGTCAATCG ATAAACAGGATCATGCTGCGAATCCTTGTGGGCA AATTGGCGAGACGGTCGATCTCGATGAAGCCGTA CAACGGGCGCTGGAATTCGCTAAAAAGGAGGGT AACACGCTGGTCATAGTCACCGCTGATCACGCCC ACGCCAGCCAGATTGTTGCGCCGGATACCAAAGC TCCGGGCCTCACCCAGGCGCTAAATACCAAAGAT GGCGCAGTGATGGTGATGAGTTACGGGAACTCC GAAGAGGATTCACAAGAACATACCGGCAGTCAG TTGCGTATTGCGGCGTATGGCCCGCATGCCGCCA ATGTTGTTGGACTGACCGACCAGACCGATCTCTT CTACACCATGAAAGCCGCTCTGGGGCTGAAAGT CGACGATCAGAACGCGACCGGCGGGCGGCAGCCT CGACGATCAGAACGCGACCGGCGGGCGGCAGCCT CGAGCACCACCACCACCACCACTGA
--	---

4.3. Appendix B

4.3.1. List of Primers

Table 18: Primers that were used in this study.

Primer Name	Sequence
EUB1	5'CTAGAATTAAAGAGGAGAAAGGTACGCGGATCCATGTTGAA AAAAGAGTATTTAAAAA 3'
EUB2	5' TAGCAGCCGGATCTCAGT 3'
EUB3	5' TAGCAGCCGGATCTCAGT 3'

EUB4	5' TAGCAGCCGGATCTCAGT 3'
SEQFWD	5' TAGCAGCCGGATCTCAGT 3'
MatBio2	5' TAGCAGCCGGATCTCAGT 3'

4.4. Appendix C

4.4.1. Plasmid Maps

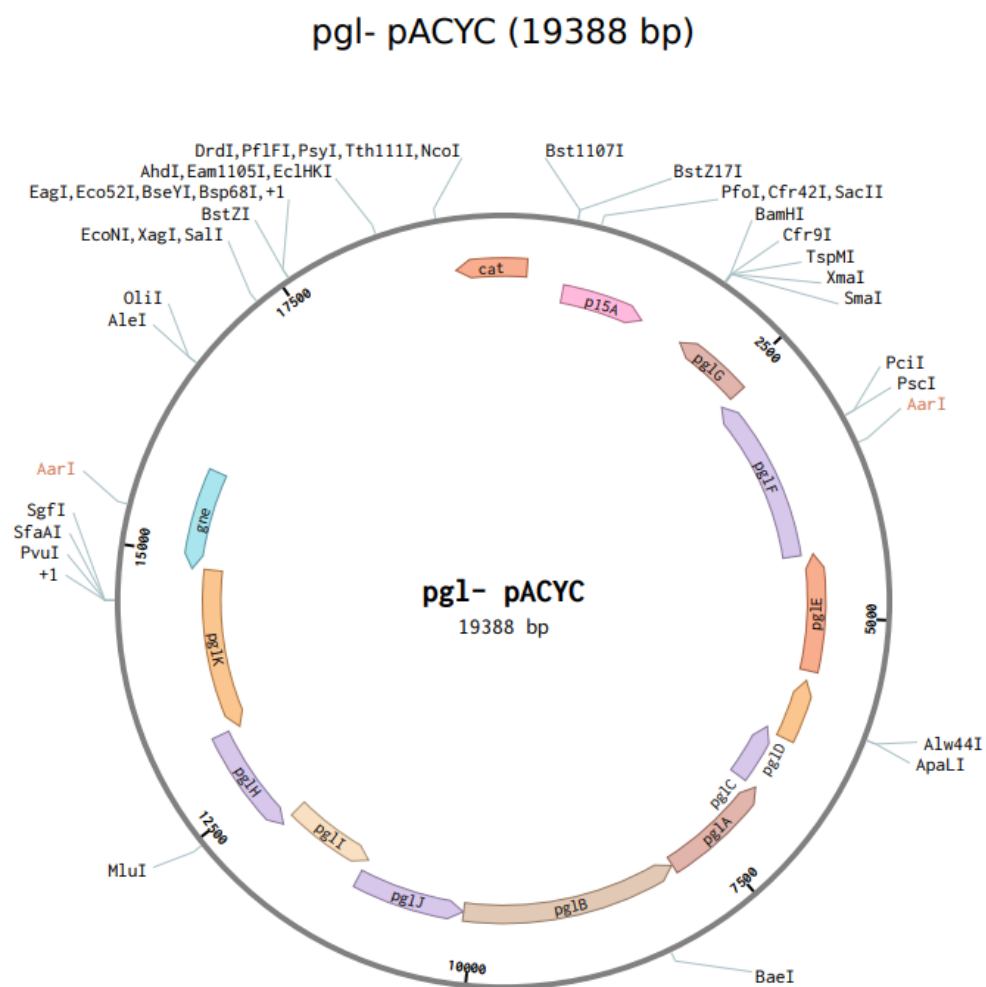


Figure 33: The plasmid map of pgl-pACYC

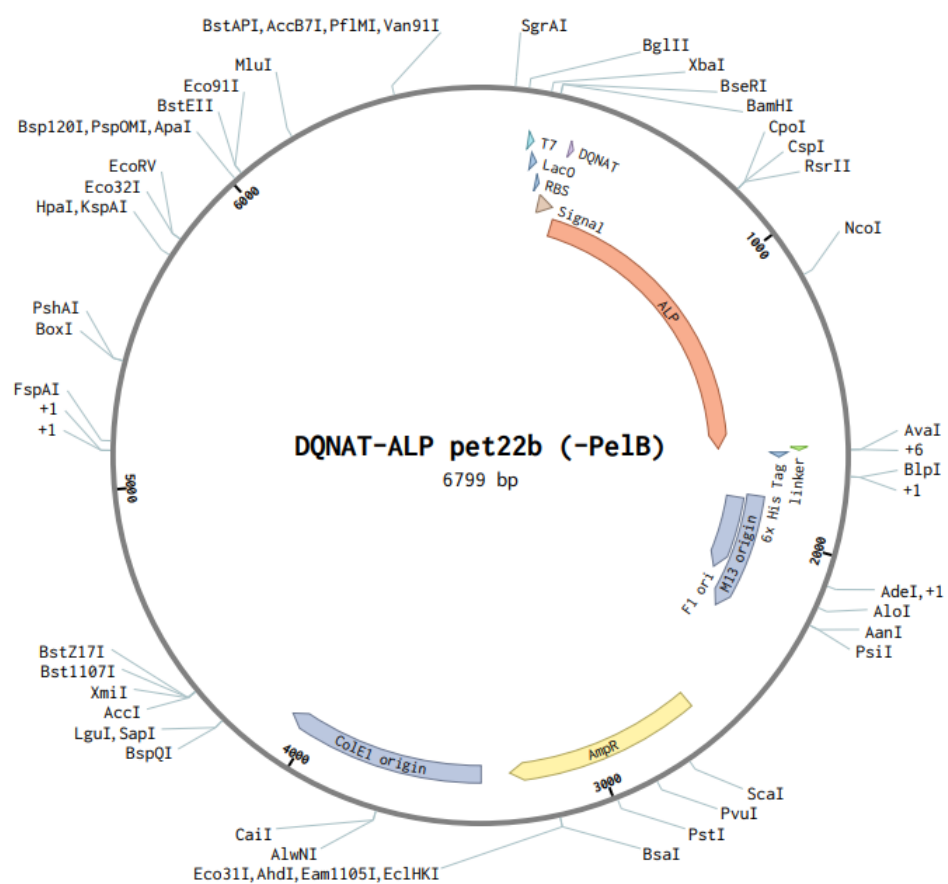


Figure 34: The plasmid map of DQNAT-ALP pET22b

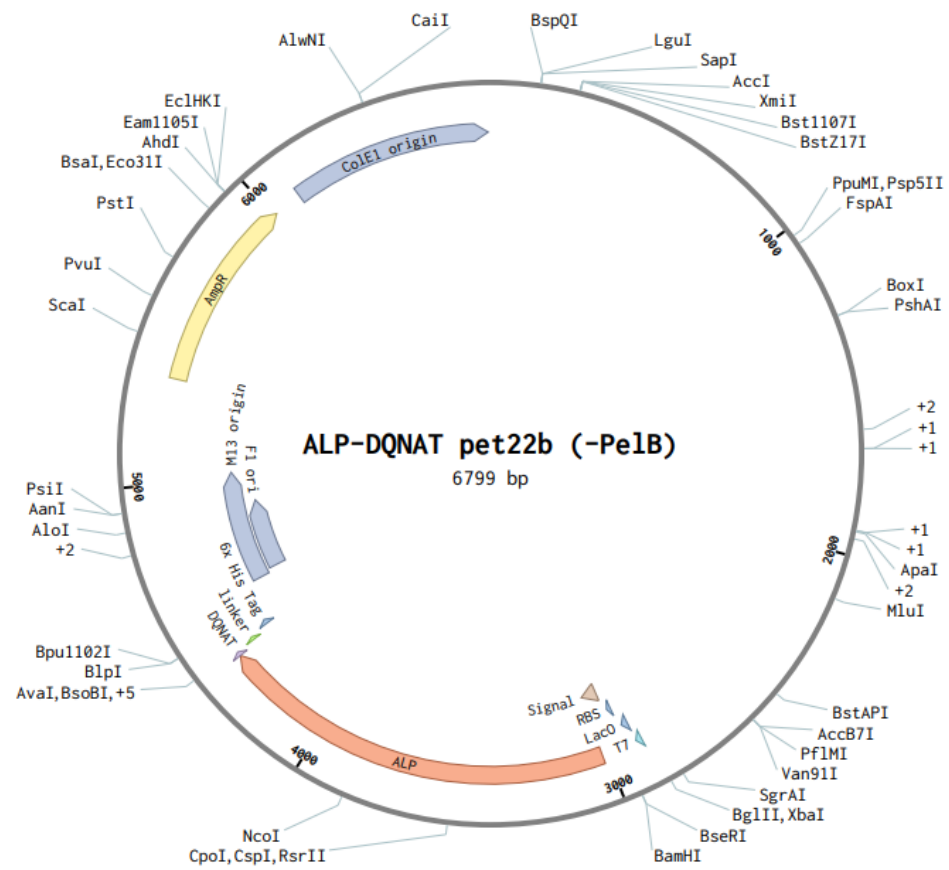


Figure 35: The plasmid map of ALP-DQNAT pET22b

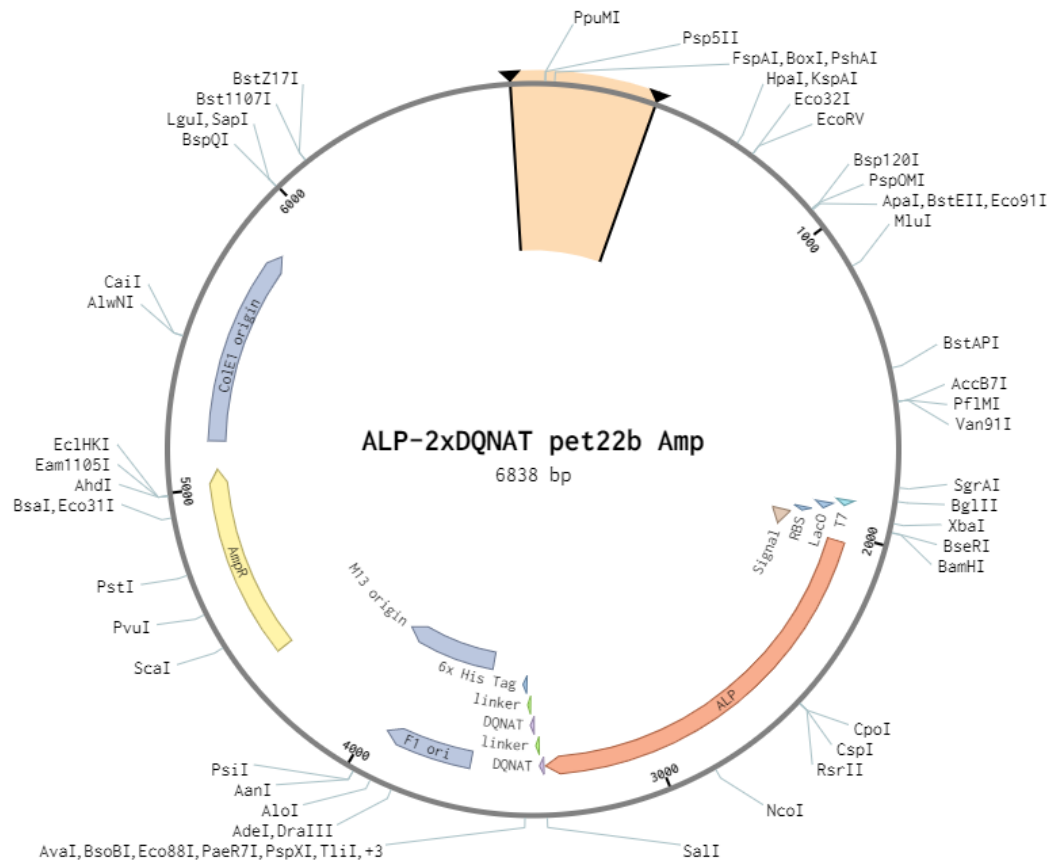


Figure 36: The plasmid map of ALP-DQNAT (x2) pET22b

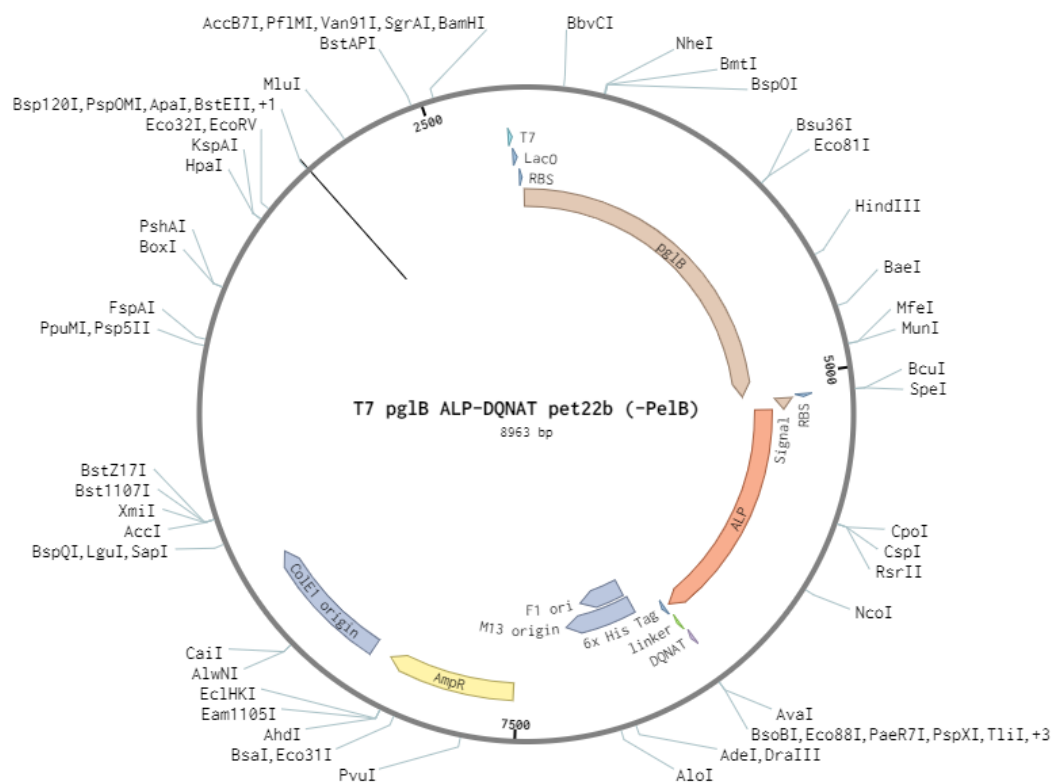


Figure 37: The plasmid map of T7 pglB ALP-DQNAT pET22b

4.5. Appendix D

4.5.1.Sanger Sequencing Results

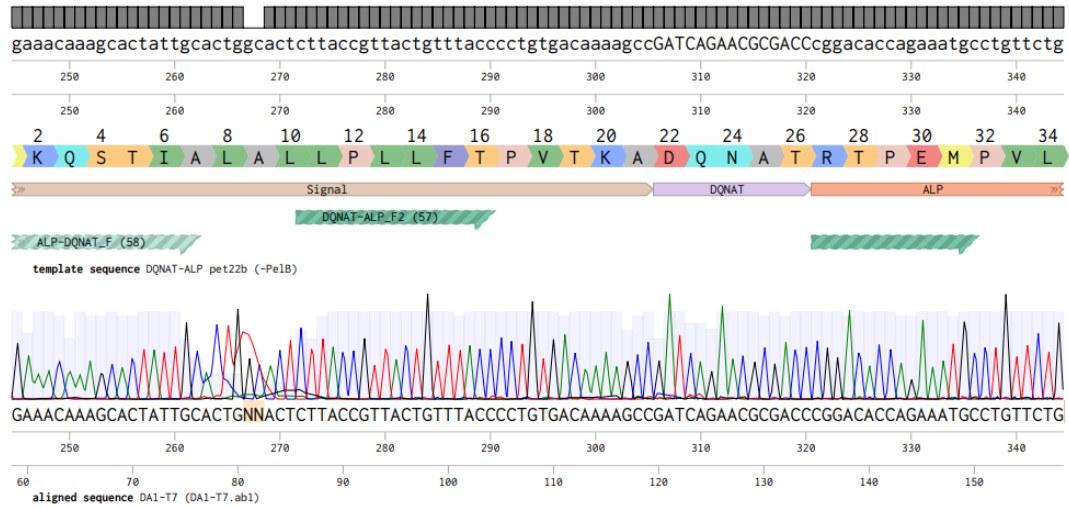


Figure 38: Sanger sequencing result of DQNAT-ALP pET22b

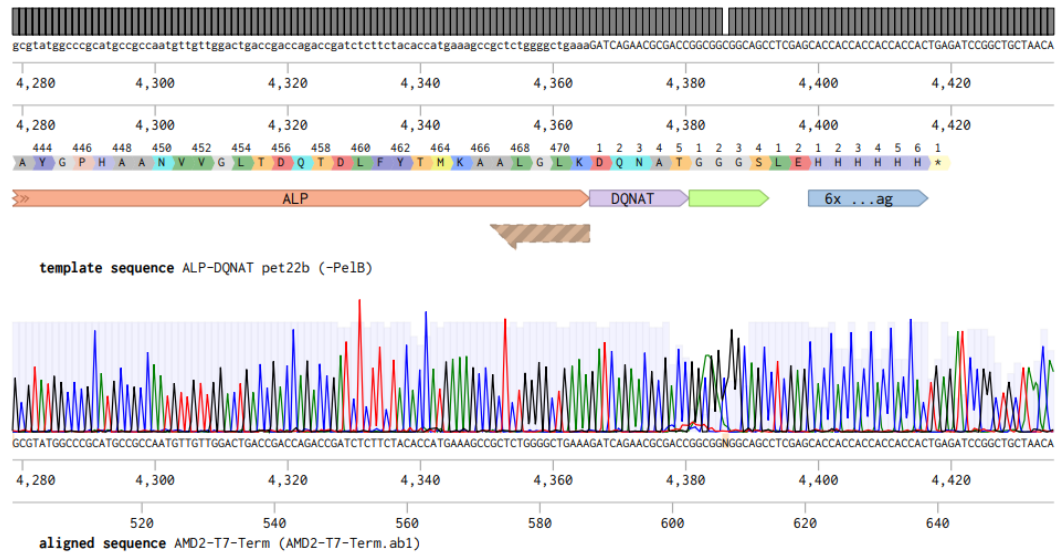


Figure 39: Sanger sequencing result of ALP-DQNAT pET22b

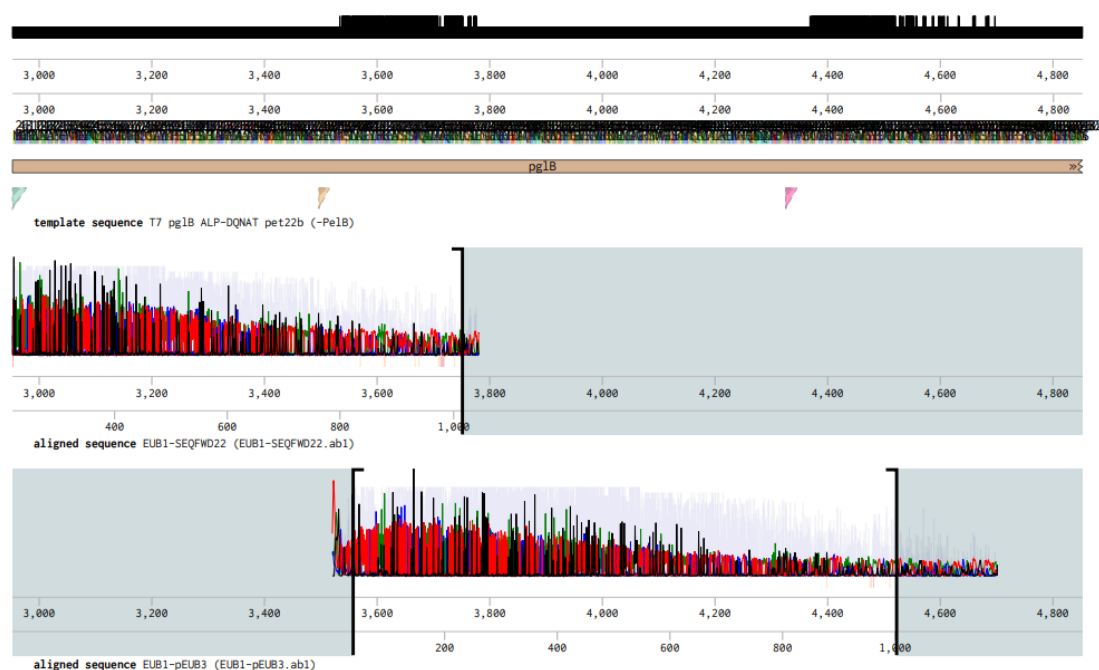


Figure 40: Sanger sequencing result of T7 pglB ALP-DQNAT pET22b

4.6. Appendix E

4.6.1. Amino acid sequences of Proteins

Plasmid Name	Gene Name	Amino acid Sequence
Pgl-pACYC	<i>pglK</i>	VLKKLFFILSKEDKNFLFFLLVFSVFVSFIETFAIS LVMPFITLASDFSDFDRNKYLISLKEYLNIPVFEI IVYFGVGLIVFYVFRALLNAYYFHLLARFSKGR YHAIAYKVFSKFLNINYEKFTQKNQSEILKSITG EVYNLSTMISSFLLLMSEIFVLLLYALMMLINY KITLFLSIFMVLNAFILVKILSPIKKAGLRREEA MKNFFEILNTNLNNFKFIKLKTKEDGVLSLFA QSEAFSKANITNESVAAPRIYLEGIGFCVLVFIV VFLVLKNESDISGILSTISIFVLALYRLMPSANRII TSYHDLLEYHSSLNIIYQNLQEEENLGEGKLSF NQELKICNLSFGYEGKKYLFKNLNLNIKKGEKI AFIGESGCGKSTLVDLIIGLLKPKEGQILIDKQEL NASNAKNYRQKIGYIPQNIYLFNDSIAKNITFGD AVDEEKLNVIKQANLEHFIKNLPQGVQTKVG DGGSNLSGGQKQRIAIARALYLEPEILVLDEATS ALDTQSEAKIMDEIYKISKDKTMIIAHRLSTITQ CDKVYRLEHGKLKEEK*
	<i>pglH</i>	MMKISFIIATLNSGGAERVLVTLANALCKEHEV SIIKFHTGESFYKLENEVKVTSLEQFRFDLTLYHK IASRFKKFFALRKALKESKADVFI SFLDTTNIACI LANIGLKTPLIIEHSNEAYLKPKTWRFLLRRVSY PFCDALSVLGSSDKVYYERFVKRVKLLNPNCHF SDEIPFDSSFEKENLVLFGRDLHNKNPVMFLKA

		IAHLDKNLQENYKFVIAGDGELRQELEYKVKS LGIKVDLGRVENVKALYEKAKVLCLCSFVEGL PTVLIESLYFEVCRISSSYYNGAKDLIKDNHDGL LVGCDDEIALAKKLELVLNDENFRKELVNNAK QRCKDFEISNIKEEWLKLIVEVKNA*
	pgII	MPKLSVIVPTFNRQVLLEKAIKSIQNQDFKDLEI IVSDDNSSDDTKSVVQNLQKDDDRIKYFLNQN YKQGPNGNKNNGLDQASGEFVTFLDDELLS GALSTLMQKANEGYAHVFGNCLIEKEGNLSKE FSGKGLEKDSEISKKDFLMAKFSGEFFSVFKKS LLENKRFNEEFYGNEATLWVNLYKEKSFYIHKA FRIYRIFRQDSVTLGASKNAHRVYLGYLELAKI LENELRMSKDKDYKKTCASYKMAAYYAKLA KNYKALYKCLFKSLSIKINAPALILLISIIPNNMI EKLSKIRVALCKN*
	pgIJ	MQKLGIFIYSLGSGGAERVVATLLPILSLKFEVH LILMNDKISYEIPECQIHFLCSKPSENPIKFLK LPFLALKYKKLCRNLGIDTEFVFLNRPNYIALM ARMFGNKTRLVINECTTPSVMYAKNNFNLAN KFLISLLYPKADLILPNSKGNLEDLVQNFSINPK KCEILYNAIDLENIEQKALEDVALKDKFILSVGR LDKGKNHALLIRAYARLKTDLKLVILGEGVLKD ELLALIKELNLEEKVLLLGFDNNPYKYMAKCEF FAFASVFEGFSNVLIESLACSCAVVCTDHKSGAR ELFGDDEFGLLVEVDNENSMFQGLKTMLEDDK LRKAYKNKAKTRAKAFDKVKIARDALKYLLG*
	pgIB	MLKKEYLKNPYLVLFAMIILAYVFSVFCRFYVW WWASEFNEYFFNNQLMIISNDGYAFAEGARDM IAGFHQPNDLSYYGSSLSALTYWLYKITPFSFESI ILYMSTFLSSLVVIPTILLANEYKRPLMGFVAAL LASIANSYYNRTMSGYYDTDMLVIVLPMFILFF MVRMILKKDFFSLIALPLFIGIYLWWYPSSYTIN VALIGLFLIYTLIFHRKEKIFYIAVILSSLTLSNIAW FYQSAIIVILFALFALEQKRLNFMIIIGILGSATLIF LILSGGVDPILYQLKFYIFRSDSANLTQGFMYF NVNQTIQEVENVDLSEFMRRISGSEIVFLFSLFG FVWLLRKHKSMIMALPILVLGFLALKGGLRFTI YSVPVMALGFGFLLSEFKAIMVKKYSQLTSNV CIVFATILTAPVFIHIYNYKAPT VFSQNEASLLN QLKNIANREDYVVTWWDYGYPVRYYSVDVKT VDGGKHLGKDNFFPSFALSKDEQAAANMARLS VEYTEKSFYAPQNDILKTDILQAMMKDYNQSN VDLFLASLSKPDFKIDTPKTRDIYLYMPARMSLI FSTVASFSFINLDTGVLDKPFTFSTAYPLDVKN EIYLSNGVVLSDDFRSFKIGDNVSVNSIVEINSI KQGEYKITPIDDKAQFYIFYLKDSAIPYAQFILM DKTMFNSAYVQMFFLGNYDKNLFDLVINSRDA KVFKLKI*
	pgIA	MRIGFLSHAGASIYHFRMPIIKALKDRKDEVFVI VPQDEYTQKLRLDLGLKVIVYEFSRASLNPFVVL

		KNFFYLAKVLKLNLDLIQSAAHKSNTFGILAA KWAKIPYRFALVEGLGSFYIDQGFKANLVRFVI NNLYKLSFKFAHQFIFVNESNAEFMRNLGLKEN KICVIKSVGINLKKFFPIYIESEKKELFWRNLD KKPIVLMIARALWHKGVKEFYESATMLKDKAN FVLVGGRDENPSCASLEFLNSGVVHYLGARSDI VELLQNCDFVLPSYKEGFPVSVLEAKACGKAI VVSDCEGCVEAISNAYDGLWAKTKNAKDLSEK ISLLEDEKLRLNLAKNAAQDALQYDENNIAQ RYLKLYDRVIKNV*
	pglC	MYEKVFKRIFDFILALVLLVLFSPVILITALLKIT QGSVIFTQNRPGLEKIFKIYKFKTMSDERDEK GELLSDELRLKAFGKIVRSLDELQLFNVLK GDMSFVGPRPLLVEYLSLYNEEQKLRHKVRPGI TGWAQVNGRNAISWQKKFELDVYYVKNISFL DLKIMFLTALKVLKRSGVSKEGHVTTEKFNGK N*
	pglD	MLQQRNLMARTEKIYIYGASGHGLVCEDVAKN MGYKECIFLDDFKGMKFENTLPKYDFFIAIGNN EIRKKIYQKISENGFKIVNLIHKSALISPSASVEE NAGILIMPYVVINAKAKIEKGVILNTSSVIEHEC VIGEFSHVSVGAKCAGNVKIGKNCFLGINSCL PNLSLADDSILGGGATLVKSQNEKGVFVGVPK RKI*
	pglE	MRFFLSPPHMGGNELKYIEEVFKSNYIAPLGEF VNRFEQSVKDYSKSENALALNSATAALHLALR VAGVKQDDIVLASSFTFIASVAPICYLKAKPVFI DCDETYNIDVDLLKLAIKECEKKPKALILTHLY GNAAKMDEIVEICKENEIVLIEDAAEALGSFYK NKALGTFGEFGAYSNGNKIITTSGGGMLIGKN KEKIEKARFYSTQARENCLHYEHLDYGYNYRL SNVLGAIGVAQMEVLEQRVLKKREIYEWYKEF LGEYFSFLDELENSRSNRWLSTALIDFDKNELN ACQKDINISQKNITLHPKISKLIEDLKNEQIETRP LWKAMHTQEVEFKGTKAYLNGNSELFFQKGICL PSGTAMSKDDVYEISKILKSIKA*
	pglF	MIFYKSKRLAFFLTSDIVLILLSVYLAFSLRFSGD IPSIFYHGMMVSAIILLVLKLSFLFVFRIYKVAWR FFSLNEARKIFIALLLAEFCFFLIFYFFSDFNPPF RSAIVIDFVLSYMFIGHTLRISKRMVDFKPSKMK EEETPCIVVGATSKALHLLKGAKESLGLFPVG VVDARKEGITYCDKFVVEEKEKIKSYVEQGV KTAIIALRLEQEELKKLFEELVAYGICDVKIFSFT RNEARDISIEDLLARKPKDLDDSAVAFLKDKV VLVSGAGGTIGSELCKQCIKFGAKHLIMVDHSE YNLYKINDDLNLYKEKITPILLSILDKQSLDEVL KTYKPELILHAAAYKHVPLCEQNPHSAVINIL GTKILCDSAKENKVAKFVMISTDKAVRPTNIMG CTKRVCELYTLMSDENFEVACVRFGNVLGSSG SVIPKFKAQIANNEPLTLTHPDIVRYFMLVAEAV

		QLVLQAGAIKGGELFVLDMGKPVKIIDLAKK MLLLSNRNDLEIKITGLRKGEKLYEELLIDENDA KTQYESIFVAKNEKVDLDWLNKEIENLQICEDIS EALLKIVPEFKHNKEGI*
	pglG	MYIKDIQRFEDNRYRARAYMSYILTRNLPNKLP DIHLETIKTALDKIAHEVVVFDALYILDISGMQIE NAISLNKAHEIGQGEDRSTRSYFYRAVKLRKRCV LSDPYPSVLNNELCVTASIPYDDKNNLLFVVC DIKLEDILKIIQAGKFEFVFTQFSRLVYFCFALVL FVITCFLFQKGFFSLFDNQAIGIEHMFESTIAITL ALAIFDLAKTLIEQEVLGRTKKEEGGIQKTMIEF LQPGGSTGRVWSP*
DQNAT- ALP pET22b	DQNAT- ALP	MKQSTIALALLPLLFTPVTKADQNATRTPEMPV LENRAAQGDITAPGGARRLTGDQTAALRDSLSD KPAKNIILLIGDGMGDSEITAARNYAEGAGGFFK GIDALPLTGQYTHYALNKKTGKPDYVTDSAAS ATAWSTGVKTYNGALGVDIHEKDHP TILEMAK AAGLATGNVSTAELQDATPAALVAHVTSRKCY GPSATSEKCPGNALEKGGKGSITEQLLNARADV TLGGGAKTFAETATAGEWQGKTLREQAQARGY QLVSDAASLNSVTEANQQKPLLGLFADGNMPV RWLGPKATYHGNIDKPAVTCTPNPQRNDSVPTL AQMTDKAIELLSKNEKGFFLQVEGASIDKQDH AANPCGQIGETVDLDEAVQRALEFAKKEGNTL VIVTADHAHASQIVAPDTKAPGLTQALNTKDGA VMVMSYGNSEEDSQEHTGSQLRIAAYGPHAAN VVGLTDQTDLFYTMKAALGLKGGGSLEHHHH HH*
ALP-DQ NAT pET22b and T7 pglB ALP DQNAT	ALP-DQ NAT	MKQSTIALALLPLLFTPVTKARTPEMPVLENRA AQGDITAPGGARRLTGDQTAALRDSLSDKPAKN IILLIGDGMGDSEITAARNYAEGAGGFFKGIDAL PLTGQYTHYALNKKTGKPDYVTDSAASATAWS TGVKTYNGALGVDIHEKDHP TILEMAKAAGLA TGNVSTAELQDATPAALVAHVTSRKCYGPSATS EKCPGNALEKGGKGSITEQLLNARADVTLGGG AKTFAETATAGEWQGKTLREQAQARGYQLVSD AASLNSVTEANQQKPLLGLFADGNMPVRWLG PKATYHGNIDKPAVTCTPNPQRNDSVPTLAQMT DKAIELLSKNEKGFFLQVEGASIDKQDHAANPC GQIGETVDLDEAVQRALEFAKKEGNTLVIVTAD HAHASQIVAPDTKAPGLTQALNTKDGA VMVMS YGNSEEDSQEHTGSQLRIAAYGPHAANVVGLT DQTDLFYTMKAALGLKDQ NATGGGSLEHHHH HH*
ALP-DQ NAT(x2) pET22b	ALP-DQ NAT(x2)	MKQSTIALALLPLLFTPVTKARTPEMPVLENRA AQGDITAPGGARRLTGDQTAALRDSLSDKPAKN IILLIGDGMGDSEITAARNYAEGAGGFFKGIDAL PLTGQYTHYALNKKTGKPDYVTDSAASATAWS TGVKTYNGALGVDIHEKDHP TILEMAKAAGLA TGNVSTAELQDATPAALVAHVTSRKCYGPSATS

		EKCPGNALEKGGKGSITEQLLNARADVTLGGG AKTFAETATAGEWQGKTLREQAQARGYQLVSD AASLNSVTEANQQKPLLGLFADGNMPVRWLGP KATYHGNIDKPAVTCTPNPQRNDSVPTLAQMT DKAIELLSKNEKGFFLQVEGASIDKQDHAANPC GQIGETVDLDEAVQRALEFAKKEGNTLVIVTAD HAHASQIVAPDTKAPGLTQALNTKDGAVMVMS YGNSEEDSQEHTGSQLRIAAYGPHAANVVGLT DQTDLFYTMKAALGLKVDDQNATGGGSLDDQ NATGGGSLEHHHHHH*
--	--	---