T.C.

ÇANAKKALE ONSEKIZ MART UNIVERSITY

SCHOOL OF GRADUATE STUDIES

DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS

RECOMBINANT PRODUCTION AND CHARACTERIZATION OF A NOVEL *N*-GLYCOSIDASE (EndoBI-2) FROM *BIFIDOBACTERIUM LONGUM SUBSP. INFANTIS* 157F

MASTER OF SCIENCE THESIS

Hatice DUMAN

Thesis supervisor

Assoc. Prof. Dr. SERCAN KARAV

ÇANAKKALE – 2022

T.C.

ÇANAKKALE ONSEKIZ MART UNIVERSITY

SCHOOL OF GRADUATE STUDIES

DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS

**RECOMBINANT PRODUCTION AND CHARACTERIZATION OF A NOVEL
N-GLYCOSIDASE (EndoBI-2) FROM *BIFIDOBACTERIUM LONGUM SUBSP.
INFANTIS* 157F**

MASTER OF SCIENCE THESIS

Hatice DUMAN

Thesis supervisor
Assoc. Prof. Dr. SERCAN KARAV

T.C.
ÇANAKKALE ONSEKİZ MART ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

Hatice Duman tarafından Doç. Dr. Sercan Karav yönetiminde hazırlanan ve **26/12/2022** tarihinde aşağıdaki jüri karşısında sunulan "**Recombinant Production and Characterization of a Novel *N*-glycosidase (EndoBI-2) from *Bifidobacterium longum subsp. infantis* 157F**" başlıklı çalışma, Çanakkale Onsekiz Mart Üniversitesi Lisansüstü Eğitim Enstitüsü **Moleküler Biyoloji ve Genetik Anabilim Dalı**'nda **YÜKSEK LİSANS TEZİ** olarak oy birliği ile kabul edilmiştir.

Jüri Üyeleri	İmza
Doç. Dr. Sercan KARAV (Danışman)	…………………………
Prof. Dr. Kemal Melih Taşkın (Üye)	…………………………
Doç. Dr. Hacı Mehmet Kayılı (Üye)	………………………….

Tez No : 10513290

Tez Savunma Tarihi : 26/12/2022

Doç. Dr. YENER PAZARCIK
Enstitü Müdürü
26/12/2022

ETİK BEYAN

Çanakkale Onsekiz Mart Üniversitesi Lisansüstü Eğitim Enstitüsü Tez Yazım Kuralları'na uygun olarak hazırladığım bu tez çalışmasında; tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi, tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu, tez çalışmasında yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi, kullanılan verilerde herhangi bir değişiklik yapmadığımı, bu tezde sunduğum çalışmanın özgün olduğunu, bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi taahhüt ve beyan ederim.

In this thesis, I wrote in accordance with the School of Graduate Studies of Çanakkkale Onsekiz Mart University's thesis writing guidelines. I hereby affirm that I received the information, facts, and materials I utilized for my thesis study ethically and within the parameters of academic standards. I also affirm that I properly cited all the works I used for my thesis study. Otherwise, I now agree and declare that I will accept any loss of rights that may be asserted against me.

Hatice DUMAN

26/12/2022

ACKNOWLEDGEMENT

Hatice Duman

Çanakkale, December 2022

ÖZET

BİFİDOBACTERİUM LONGUM SUBSP. İNFANTİS 157F KÖKENLİ ÖZGÜN *N*-GLİKOSİDAZ (ENDOBI-2) ENZİMİNİN REKOMBİNAT ÜRETİMİ VE KARAKTERİZASYONU

Hatice DUMAN

Çanakkale Onsekiz Mart Üniversitesi

Lisansüstü Eğitim Enstitüsü

Moleküler Biyoloji ve Genetik Anabilim Dalı Yüksek Lisans Tezi

Danışman: Doç. Dr. Sercan KARAV

26/12/2022, 58

Gastrointestinal (GI) sistem, sağlığın korunmasına ve hastalıkların ilerlemesine önemli ölçüde katkıda bulunan çeşitli mikroorganizmaları barındırır. GI sisteminde yaşayan mikroorganizmaların hayatta kalması, temel olarak konakçı organizma tarafından sindirilemeyen karmaşık karbonhidratları (oligosakkaritler) kullanma yeteneklerine bağlıdır. İnsan sütü, yeni doğmuş bir bebek için birincil beslenme kaynağıdır ve içeriğinde laktoferrin, immünoglobülinler vb. önemli biyoaktif proteinlerin büyük çoğunluğu, glikozile formda bulunmaktadır. Bu proteinlere bağlı formda bulunan biyoaktif glikanlar bebekler tarafından sindirilemez ve yalnızca belirli bağırsak mikroflorasında yaşayan yararlı mikroorganizmalar, sahip oldukları enzimler sayesinde, glikozile proteinlerden bu yapıları ayırarak bu bileşikleri tek karbon kaynağı olarak kullanmaktadır. Bu sayede bağırsak mikroflorasının sağlıklı bir şekilde oluşumu sağlanmaktadır. Glikosidazlar, glikoproteinlerden *N*-glikan salınımı için kullanılan enzim türüdür. Ticari olarak temin edilebilen glikosidazlar, bu glikanları proteinlerden doğal hallerinde ayırmada yetersizdir. Bu eksikliğin üstesinden gelmek için, *N*-glikanların biyolojik rollerini ve bunların büyük ölçekli üretimini daha iyi anlamak için daha etkili deglikozilasyon yaklaşımları gereklidir.

Bu çalışma kapsamında, *N*-glikanların benzersiz özelliklerini anlamak için *Bifidobacterium longum subsp. infantis* kaynaklı yeni ve özgün endo-*ß-N*-asetilglukozaminidaz 2 (EndoBI-2) rekombinant olarak üretilmiş ve karakterize edilmiştir. EndoBI-2'nin yapısal özelliklerini ve substrat seçiciliğini araştırmak için karşılaştırmalı modelleme yaklaşımları kullanılmıştır. Ayrıca, EndoBI-2'nin optimum reaksiyon/kinetik

parametreleri glikoproteinler kullanılarak araştırılmış ve salınan biyoaktif glikanlar MALDI-(TOF)/TOF-MS aracılığıyla karakterize edilmiştir.

Elde edilen veriler ile EndoBI-2'nin kinetik ve yapısal karakterizasyon parametreleri, bu yeni enzimin, glikoprotein kaynaklarından biyoaktif glikanların büyük ölçekli salınımını kolaylaştırabilen önemli bir enzim olduğunu göstermektedir. Bu yeni enzimin ve salınan biyoaktif glikanların gıda endüstrisinde başarılı bir şekilde uygulanması, fonksiyonel gıdaların geliştirilmesini sağlayacaktır.

Anahtar Kelimeler: Endoglikosidaz, Glikoprotein, *N*-Glikan, *B. infantis*, Rekombinant DNA teknolojisi

ABSTRACT

RECOMBINANT PRODUCTION AND CHARACTERIZATION OF A NOVEL *N*-GLYCOSIDASE (EndoBI-2) FROM *BIFIDOBACTERIUM LONGUM SUBSP. INFANTIS* 157F

Hatice DUMAN

Çanakkale Onsekiz Mart University

School of Graduate Studies

Master of Science Thesis in Molecular Biology and Genetics

Advisor: Assoc. Prof. Dr. Sercan KARAV

26/12/2022, 58

The gastrointestinal (GI) tract harbors several microorganisms, which significantly support both the advancement off diseases and the maintenance of health. The survival of microorganisms that live in the GI tract is mainly dependent on their ability to utilize complex oligosaccharides that are not digestible by the host organism. Human milk is the primary nutrition for a newborn. Glycosylated milk components including lactoferrin, immunoglobulins, etc., are unsuitable for infant consumption, and certain gut microbes have the enzymes to liberate glycans from glycoproteins and utilize these molecules as the only carbon source. The enzyme class known as glycosidases is utilized to liberate *N*-glycan from glycoproteins. Commercially available glycosidases are inefficient at cleaving these glycans from proteins in their native state. So, efficient deglycosylation approaches are required to comprehend the biological functions of *N*-glycans and their widespread production.

Within this study, to comprehend the special characteristics of *N*-glycans, the novel endo-*ß*-*N*-acetylglucosaminidase 2 (EndoBI-2) was isolated from *Bifidobacterium longum subsp. infantis* was recombinantly produced and characterized. Comparative modeling methods were employed to investigate the enzyme structural characteristics and substrate selectivity. Furthermore, the optimum reaction/kinetic parameters of EndoBI-2 were investigated using glycoproteins, and released bioactive glycans were characterized with MALDI-(TOF)/TOF-MS.

The research outcomes from characterization analysis of EndoBI-2 point to its potential to enhance the widespread release of bioactive glycans from glycoprotein source. It will be possible to develop functional foods for health and wellness if this unique enzyme and its released bioactive glycans are successfully applied in the food sector.

CONTENT

CHAPTER I

INTRODUCTION

CHAPTER II

THEORETICAL FRAMEWORK/PREVIOUS STUDIES

CHAPTER III

MATERIAL METHOD

CHAPTER IV

RESEARCH FINDINGS

CHAPTER V
RESULTS AND RECOMMANDATIONS

SYMBOLS AND ABBREVIATIONS

2-AA	2-Aminobenzoic acid
ACN	Acetonitrile
Asn	Asparagine
BLOSUM	Blocks Substitution Matrix
Cα	Carbon Alpha
DHB	2,5-Dihydroxybenzoic acid
ENGase	endo-β-N-acetylglucosaminidases
dH$_2$O	Distilled Water
DMSO	Dimethyl sulfoxide
HexNAc	N-acetylglucosamine
HMO	Human Milk Oligosaccharides
h	hour
kDa	Kilodalton
LF	Lactoferrin
LPO	Lactoperoxidase
MD	Molecular Dynamics
N-glycan	N-linked glycan
M	Molar
MALDI-MS	Matrix Assisted Laser Desorption/Ionization- Mass spectrometry
min	Minute
mL	Milliliter
MS	Mass Spectrometry
Mw	Molecular Weight
NaCNBH$_3$	Sodium cyanoborohydride
nm	Nanomater
OD	Optical Density
O-glycan	O-linked glycan
PNGase F	Peptidyl N-glycosidase F
Thr	Threonine
Ser	Serine
SDS-PAGE	Sodium Dodecyl-Sulfate Polyacrylamide Gel Electrophoresis
pH	Power of Hydrogen
RPM	Revolutions per minute

Kg	Kilograms
g	Grams
%	Percent
°C	Degrees Celcius

LIST OF TABLES

CHAPTER 1
INTRODUCTION

Mammalian milk has evolved as uniquely over the course of 200 million years to be nourishing and protective. It is a highly specialized fluid that shapes all aspects of neonatal development (Hamosh, 2001). Human milk, which has a dynamic and rich content, has macro- and micronutrients that are necessary for infant nutrition. Proteins (9-12 g/L), lipids (32-36 g/L), and lactose (67-78 g/L) are the main macronutrients present in human milk, whereas vitamins, minerals, etc., are considered micronutrients found in it and these components can differ depending on the conditions such as diet. Importantly, numerous bioactive components including immunological and growth factors, metabolic hormones, oligosaccharides & glycans, non-coding RNAs, cytokines and cytokine inhibitors, etc., can be found in human milk composition (Ballard & Morrow, 2013).

All the biomolecules present in human milk, notably oligosaccharides & glycans, have not been thoroughly studied by scientists (Hamosh, 2001). Milk contains abundant and highly specialized oligosaccharides that can be found in free form (human milk oligosaccharides, HMOs) and glycosylated form (*N*- and *O*- glycosylation). It is now known how HMOs affect the gut flora during a crucial period of neonatal development. Further, glycans have emerged as the major primary substrates for the gut microbiome, and this community enables the conversion of glycans inaccessible to the host into substrates that are actively taken up by the gut epithelium (Pokusaeva et al., 2011).

In the gut microbiome, the importance of these glycans has been sparsely examined (Wang et al., 2017) and their importance in shaping the gut microbiome is not known. *N*-glycans have been shown to bear similar prebiotic specificities to HMOs, which are exquisitely specific in selecting for the growth of taxa that are important to neonatal health (Karav et al., 2016). Additionally, highly adapted gut symbionts have developed efficient mechanisms to access *N*-glycans, which they deploy *in-vivo* to utilize these glycans, despite the relative abundance of HMOs in the neonatal colon (Karav et al., 2019). Given their structural and compositional complexity, the large degree of conservation of *N*-glycan structures in milk glycoproteins, and recent findings demonstrating the utility of these glycans, it is highly likely that they also play a role in shaping the developing microbiome.

Extensive research has characterized the impact of free glycans, especially abundant in human milk relative to other mammals, on the gut microbiome, but the impact of protein-conjugated glycans is lacking. We have previously demonstrated that specialized mechanisms are required to access conjugated glycans, which are largely inaccessible to bacteria without either specialized enzymes to release them or enzymes to remove terminal sugars from the *N*-glycan structure (Karav et al., 2016).

1.1. Human Milk Oligosaccharides (HMOs) and Their Characteristics

HMOs are structurally diverse, non-conjugated glycans that range in size from two to over twenty units. Numerous free oligosaccharides found in human milk vary from person to person and based on the lactation stage, the concentration of oligosaccharides as well as their composition can vary. HMOs are mostly made of a lactose core that has been altered by *N*-acetlylactosamine blocks. No matter if it is linear or branching, additional elongation is made possible using *N*-acetylneuraminic and/or acid fucose as additives. Five monosaccharides serve as the building blocks for the neutral and anionic species that make up HMOs. D-galactose (Gal), D-glucose (Glc), *N*-acetylglucosamine (GlcNAc), L-fucose (Fuc), and acetylneuraminic acid (also known as sialic acid; NeuAc) are considered as the building blocks. Neutral *N*-containing (non-fucosylated) HMOs (42–55%), neutral (fucosylated) HMOs (35–50%), and acidic (sialylated) HMOs (12–14%) are the three primary groups. To produce 2′-3-Fucosyllactose (FL) or 3′-6-Sialyllactose, the lactose core is conjugated with Fuc or NeuAc (SL). The repeats of lacto-*N*-biose (Galβ1- 3GlcNAc; LNB), which are connected to the lactose core, are referred to as type 1 chains. Lacto-N-tetraose (LNT), a type 1 HMO, is the most prevalent. The *N*-acetyllactosamine unit (LacNAc; Galβ1-4GlcNAc) is formed when the lactose core is joined with it, creating the type 2 chain. In HMOs, lacto-N-neotetraose (LNnT) is an example of the type 2 chain (Duman et al., 2021) (Figure 1).
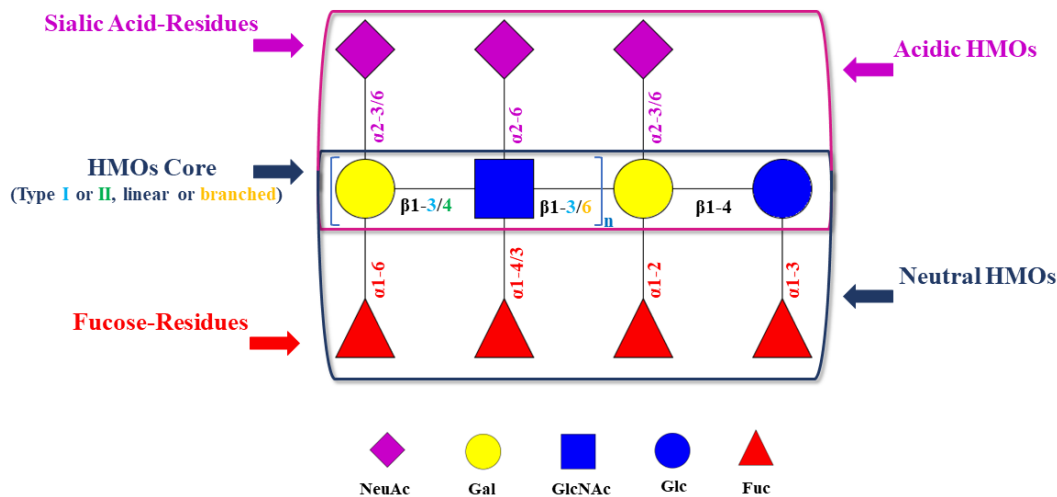
Figure 1. Human milk oligosaccharides composition and structures (Duman et. al., 2021)

HMOs are the third most prevalent substance in human milk, yet newborns cannot digest these bioactive molecules because they lack glycosidases in their gastrointestinal tracts. But LoCascio et al. (2009) demonstrated that different *Bifidobacteria* strains preferentially use these oligosaccharides (LoCascio et al., 2009). Being a prebiotic source and influencing the microbial ecology of the developing GI tract are two of HMOs' most well-known uses. HMOs are used by specialist gut microorganisms in the colon, which have the crucial molecular machinery for the transportation and metabolization of these complicated compounds. Some species of infant-adapted bacteria may break down and consume oligosaccharides using special genomic capabilities. Additionally, short-chain fatty acids (SCFA) are formed once HMO has been digested in the colon. These substances contribute to an acidic environment (low pH), which promotes the development of some microbes while simultaneously preventing the growth of pathogens (Frese et al., 2017) (Figure 2).

The antipathogenic activity of HMOs, which includes HIV, enteropathogenic *Escherichia coli* (*E. coli*), *Streptococcus pneumonia*, and *Vibrio cholera*, is another crucial function of these molecules. These small-chain carbohydrates have recently been revealed to have bifidogenic action and antiadhesive antibacterial properties. They are also crucial for the development of the brain (Duman et. al., 2021) (Figure 2).
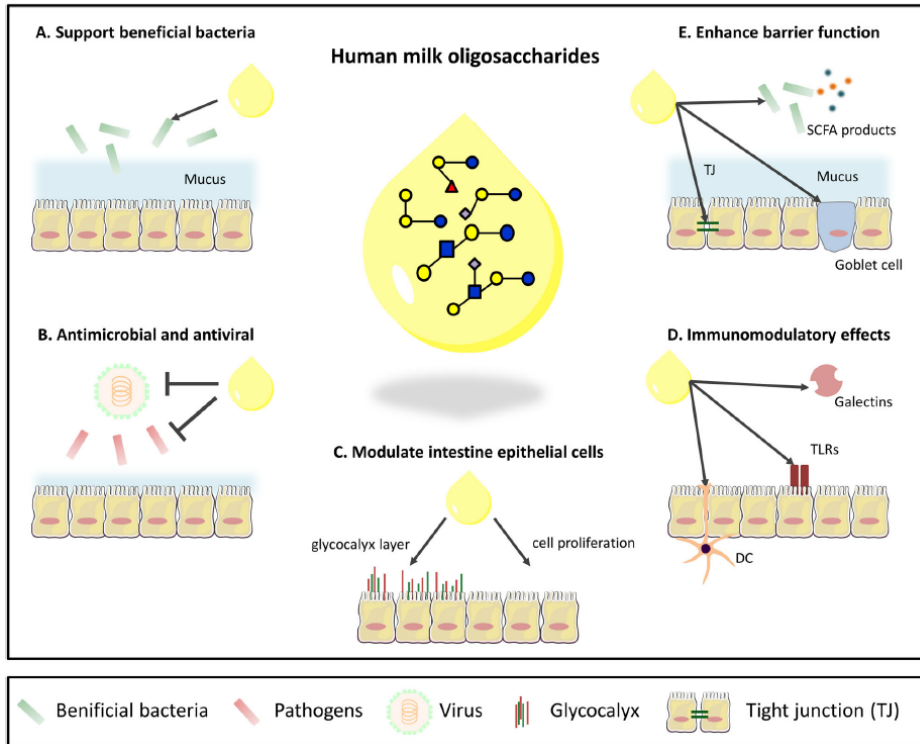
Figure 2. Effects on early immunological development and HMO diversity (Cheng et al., 2021).

1.2. Glycans and Their Importance as Prebiotics

Among other modifications, the post-translational glycosylation of proteins in the endoplasmic reticulum is critical to proper protein folding, extracellular transport, and protein function (Figure 3). The relative importance of protein glycosylation is evidenced by the high degree of conservation across evolutionarily distant phyla (Varki et al., 2017). Additionally, proteins' biological functions depend on glycans, and changes to glycosylation can have a variety of effects.
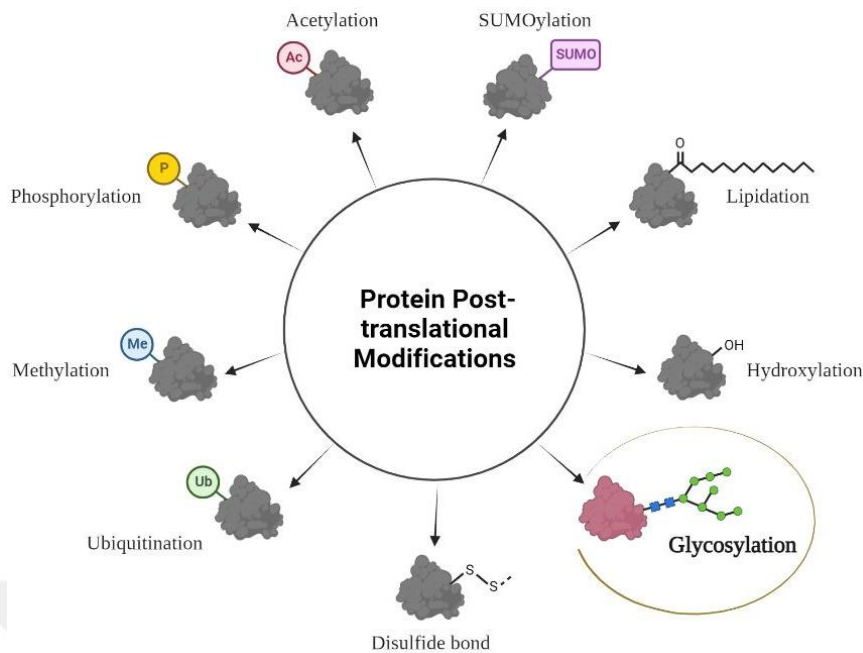
Figure 3. An illustration of various types of protein post-translational modifications.

N- and O-glycosylation, which account for over 70% of all proteins in milk, are the two most common kinds of glycosylation (Varki et. al., 2017). A protein's hydroxyl group is linked to *O*-linked glycans (*O*-glycans) by the serine (Ser) or threonine (Thr) residues via *N*-acetylgalactosamine, and their common structure can be changed into distinct core classes, (Figure 4) In contrast, *N*-linked glycans also refers to *N*-glycans are conjugated to proteins' asparagine (Asn) residues' amide groups via *N*-acetylglucosamine (HexNAc) in the Asn-X-Ser/Thr or Asn-X-Cys (cysteine) (where X can be any amino acid, with the exception of proline) (Lafite & Daniellou, 2012). Basically, two *N*-acetylglucosamine (GlcNAc) residues form the sole core of an *N*-glycan, which is then followed by three mannose groups. Further glycosylation establishes three primary categories of *N*-glycans depending on composition: high mannose (HM), hybrid, and complex. HM glycans typically have 5–9 residues of mannose connected to the chitobiose (GlcNAc2) core, along with unsubstituted terminal mannose sugars. Hybrid glycans differ from HM glycans in that they replace the mannose residues with a *N*-acetylglucosamine bond. GlcNAc residues are found in complex type glycans at both the α-3 and α-6 mannose sites. Additionally, aside from the core structure, they don't contain mannose residues as high mannose glycans do. Additional monosaccharides may be present in complex glycans, which end with sialic acid residues and repeated lactosamine (GlcNAc-β(1→4)Gal) units (Varki et. al, 2017).

The process modification of proteins by *N*-glycosylation is well characterized, and many roles for *N*-glycosylation have been described, and span from fundamental alterations to protein chemistry (e.g. solubility) to highly specific modifications for the function of the resulting glycosylated protein in protein trafficking, recognition by receptors, and protein activity. Overwhelmingly, these modifications have been studied for their properties within cells their role in disease (e.g. cancer or congenital diseases), or their impact on protein function (Krainer et al., 2013).
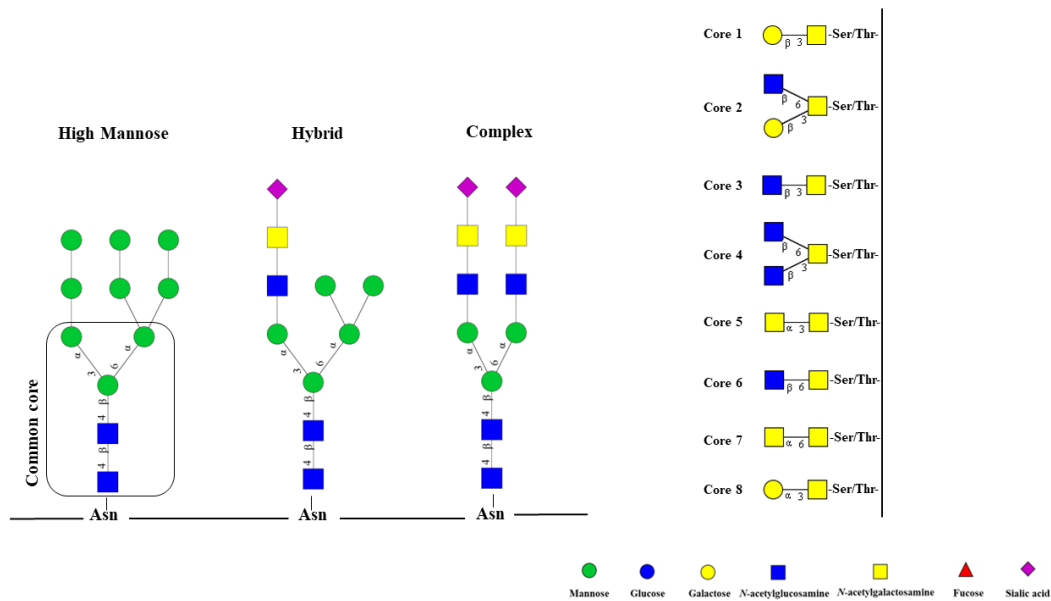
Figure 4. Schematic representation of *N*- and *O*-glycans (GlycoWorkbench).

In numerous cellular processes that affect both health and, glycans play significant roles. These are cell adhesion and mobility, receptor activation, protein folding and degradation, the calnexin-calreticulin cycle, immune signaling, defense against various microbial and viral attacks, and modulation and development of the gut microbiome of infant. (Barboza et al., 2012). A bacterial strain is known as *B. infantis* was found in infant feces and has been demonstrated to have a broad spectrum of glycan consumption capacities. It has already been shown that *B. infantis* can consume a number of oligosaccharides found in human milk (HMOs) (Sela et al., 2012). Due to their similarity to HMOs in terms of monosaccharide content and bond types, *N*-glycans can also exhibit strong bifidogenic activity by encouraging the development of beneficial microbes and exhibiting antimicrobial activity (Mussatto & Mancilha, 2007). Recent research has shown that *N*-glycans, which are

6

produced from bovine milk glycoproteins by a unique enzyme called Endo-*β-N*-acetylglucosaminidase (EndoBI-1), preferentially promote the growth of *B. infantis* while *Bifidobacterium animalis* (*B. animalis*) is unable to utilize *N*-glycan structures (Karav et al., 2016).

The bifidogenic potential of *N*-glycans released from human and bovine whole skim milk was studied by Wang et al. in 2017. After being quantified, the *N*-glycans were supplied to the growth medium in varying quantities. After 12 hours, the growth of the Bifidobacteria was assessed, and the addition of *N*-glycans from human milk caused bacterial growth. *N*-glycans from bovine milk also showed a growth-promoting impact when compared to those present in human milk (Wang et al., 2017).

The growth of *B. infantis* is specifically stimulated by *N*-glycan structures from immunoglobulins and lactoferrin (LF), according to research by Karav et al. (2019). Additionally, this research revealed that EndoBI-1 from *B. infantis*, which has a distinct structural specificity and function for the cleavage of *N*-glycans from glycoprotein present in milk, is functional *in-vivo* (Karav et al., 2019). These *in-vitro* and *in-vivo* research have revealed that various glycan types affect the gut microbiota in various ways.

As the technology has improved, there has been a better resolution of glycosylation sites and glycan structures using mass spectrometry and the field has gained important insight as to how these glycans, their relative location on proteins, and their composition play a role in disease and normal cellular function. A variety of approaches have been employed to not only measure but compare glycan structures in systems by including genetic alterations to glycosylation pathways, chemical inhibition of glycan modification, and enzymatic removal of *N*-glycans from glycosylation sites (e.g. endoglycosidases). As the relative importance of these glycans has become appreciated, it has also become apparent that intrinsic selective pressures govern the proper construction and function of these glycans (Stanley, 2016).

1.3. Glycan Release Strategies from Glycoproteins

With the growing demand for the mass manufacturing of glycans acquires worthwhile deglycosylation methods to better comprehend their crucial functions. Deglycosylation methods essentially fall into two categories: enzymatic and chemical (Figure 5). These methods are preferred to isolate these biologically active glycans with great efficacy, broad diversity, affordable costs, and simple applications. Another important point is that deglycosylation processes are the ability to release glycans without altering their structural integrity. So, researchers may choose to employ a single approach to obtain glycans, or they can combine methods to achieve a deglycosylation process effectively (Karav, 2019).

Deglycosylation Strategies for Glycan Release

Chemical Deglycosylation			Enzymatic Deglycosylation		
β-elimination	Hydrazination	Trifluoroacetic acid Treatment	Endoglycosidases EndoF, EndoH, etc.	PNGases	O-glycosidases
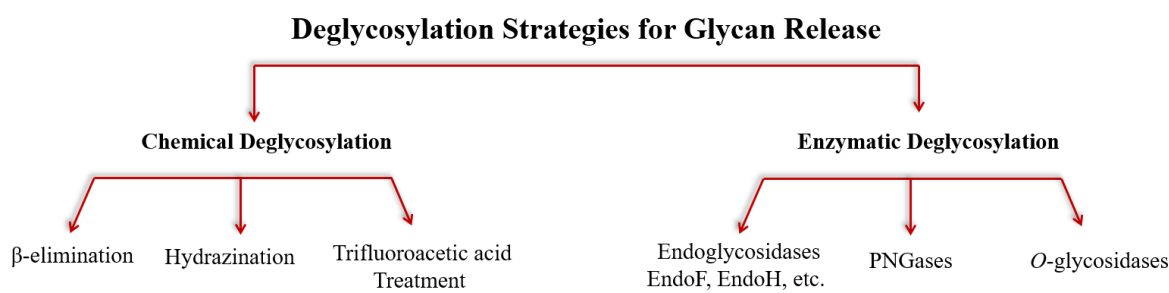
Figure 5. Classification of deglycosylation methods.

1.3.1. Glycan Release Strategies from Glycoproteins Using the Chemical Methods

When it comes to the deglycosylation of glycoproteins in glycobiology, chemical approaches are frequently employed. The two most popular deglycosylation with chemical processes are hydrazination and β-elimination (Dwek et al., 1993). While chemical methods are often preferred because of their quick and simple application, affordable cost, and excellent substrate specificity, these methods have many drawbacks that complicate further glycan research. For example; the liberated glycans and the residual polypeptide chain may be damaged during the chemical reaction in the β-elimination process achieved under alkaline conditions, even if reducing agents such as sodium borohydride is used to prevent any deterioration of the chemical structure. The drawback of this approach is that the reductive agent converts the single labeling group on oligosaccharides into alditols, which

restricts the capacity to attach a fluorophore or a chromophore. The monitoring of the deglycosylation process is hence difficult. With this method, sample losses can be very high during the purification of the samples from excessive saltiness (Carlson, 1968; Turyan et al., 2014).

Another chemical deglycosylation technique is hydrazine treatment, which relies on the hydrolysis reaction that occurs when the glycoprotein is treated with anhydrous hydrazine. By altering the reaction parameters, like temperature, this approach has the advantage of allowing both *N*- and *O*-glycans to be liberated under regulated terms. Turyan et al. demonstrated that hydrazinolysis is more influential than β-elimination to liberate *O*-glycans in regard to activity, indicating the higher potential of hydrazine therapy (Turyan et al., 2014). Anhydrous versions of HF and TMFS have also been used for chemical deglycosylation (Figure 6) (Sojar & Bahl, 1987).

Chemical Deglycosylation Methods

β-elimination	Hydrazination	Trifluoroacetic acid Treatment
✓ Easy and Fast application	✓ Low cost	✓ High substrate specificity
↓ Sample losses	✗ Degradation of released glycans and remaining part	✗ Challenging to monitor glycan release
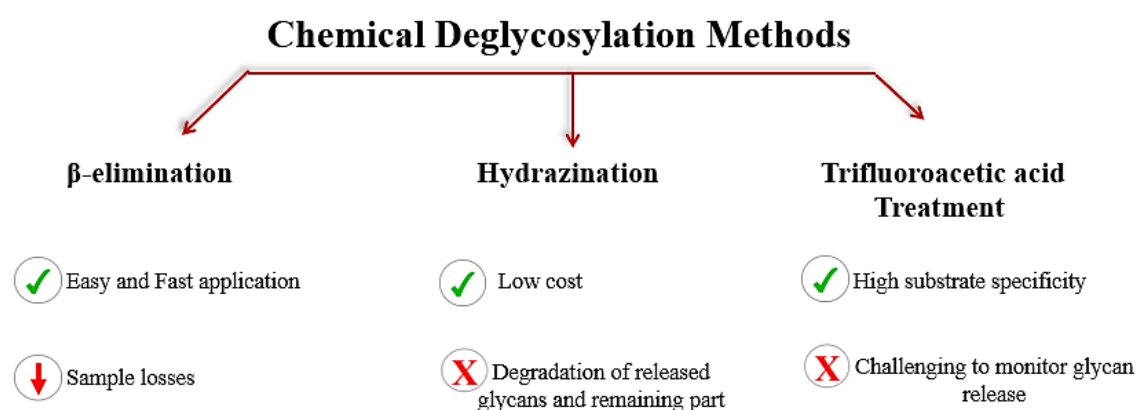
Figure 6. Summary of chemical deglycosylation methods and their advantages/disadvantages

1.3.2. Glycan Release Strategies from Glycoproteins Using the Enzymatic Methods

Peptidyl-*N*-glycosidases (PNGases) in particular PNGase F are the common enzymes that release glycans from glycoproteins. Regardless of their charge or size, commercially available PNGases hydrolyze amide side chains glycoproteins by releasing *N*-glycans (Szabo et al., 2010). These enzymes, however, are unable to release these glycans if *N*-acetylglucosamine is coupled to fucose. The activity of PNGases on *N*-glycans is another

limitation. PNGase activity necessitates harsh treatment, such as the denaturation of glycoproteins with high heat and detergent, to make the glycans more accessible to the enzyme. The liberated glycan and residual polypeptide structures may be disrupted by this severe treatment, much like chemical deglycosylation techniques. Compared to PNGases, a number of endoglycosidases, including F1, F2, F3, etc., exhibit greater activity on native forms of glycoproteins. On multiple-antennary glycans, nevertheless, their activity is incredibly modest (Figure 7) (Trimble & Tarentino, 1991). Due to complicated different cores of *O*-glycans and a lack of commercial *O*-glycosidases that are effective on these cores, the amount of *O*-glycan release from glycoproteins that can be accomplished through enzymes is relatively limited (Karav, 2019).
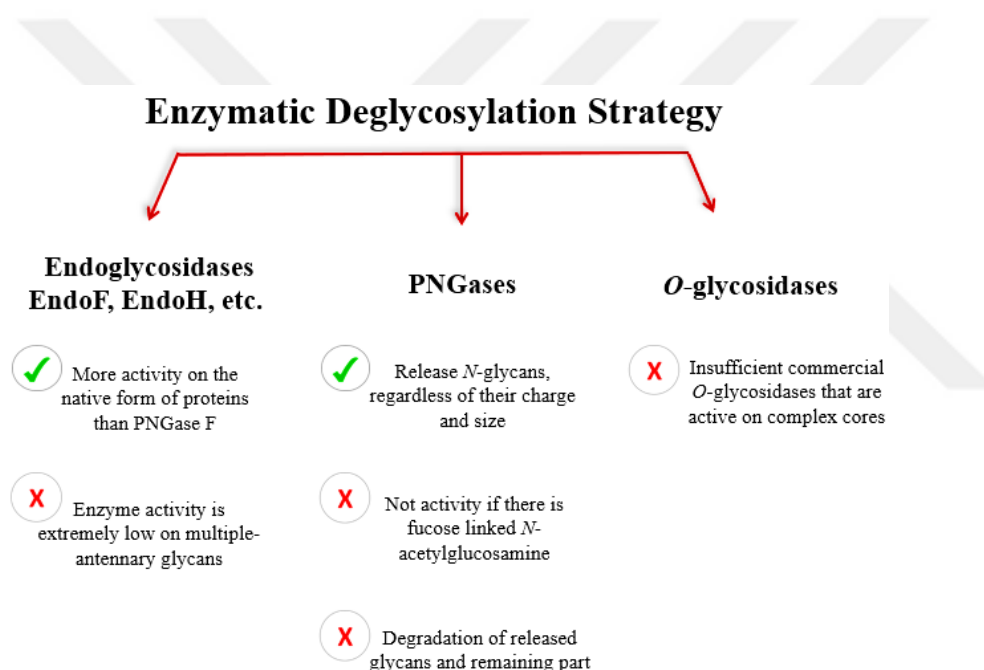
Figure 7. Summary of enzymatic deglycosylation strategies.

1.3.3. A novel Endo-*β*-*N*-acetylglucosaminidases for Better Glycan Release

For the purpose of removing glycans from glycoproteins, a variety of enzymes (glycosidases) and various chemical techniques are frequently employed. The exploration of the biological functions of glycans is hampered by their numerous drawbacks, which include low selectivity of glycan release, high expense, disruption of glycans, and the remaining polypeptide structure. In conclusion that, unique methods for releasing these glycans from dairy streams should be taken into consideration (Karav, 2019).

EndoBI-1, from *B. infantis* ATCC 15697, was demonstrated by Karav et al. (2015) to have the capacity to separate the *N-N′*-diacetylchitobiose partition present in the *N*-glycan core of whole classes of *N*-glycans. The *N*-glycan core's fucosylation has no impact on the activity of EndoBI-1. When compared to currently available *N*-acetylglucosaminidases like PNGase F, it was also demonstrated to be heat resistant, enabling its usage in heat treatments at 95°C heat treatments and the ability to release glycans without any substrate denaturation (Karav et al., 2015).

With this thesis, a novel enzyme, endo-*β-N*-acetylglucosaminidase 2 (EndoBI-2) was recombinantly produced and characterized. EndoBI-2 was first discovered from *B. infantis* 157F in 2012 is a unique and very effective enzyme for the glycan release from glycoproteins but this enzyme is not commercially available. It was also proven to be heat resistant, which is critical for heat-required processes. Compared to conventional chemical and enzymatical methods, this enzyme is highly active on the glycoproteins and releases different types of *N*-glycans including high mannose, complex (Garrido et al., 2012). This unique characteristic of EndoBI-2 offers a potentially useful method for generating *N*-glycans from glycoprotein-rich proteins like immunoglobulin, LF, etc.

1.4. Structural Characterization and Comparative Modelling Approaches

Comparative modeling and bioinformatic research could yield priceless structural data on this enzyme. The distinctive characteristics of the EndoBI-2 enzyme can be understood using the structural data that was collected. The correct structure/function connection information might also be used to find additional enzymes from either a healthy human microbiome or other organisms with similar characteristics and distinctive specificities. (Sliwoski et al., 2014).

One of the most crucial cases for comprehending the biological activities of living things at the molecular level is the existence of a protein's structural model. The identification and generation of protein structures that have several error in experimental and theoretical models is a significant issue. Large-scale structure determination pipelines, as

well as experimental structure validation, structure prediction, and modeling, have all been made possible by recent advancements in computational approaches. However, whereas new DNA sequences are being produced by DNA sequencing techniques at an ever-increasing rate, the rate at which new protein structures are being identified experimentally has not increased at the same rate. Therefore, using computational approaches has become a key strategy for producing more reliable and stable protein structures (Bienert et al., 2017).

In this thesis, comparative modeling approaches were used to generate 3D models for the EndoBI-2 enzyme together with EndoBI-1 enzyme. According to data acquired from homologous proteins with known structures, comparative modeling (homology modeling) is a method for creating 3D models of proteins for which experimental structures (template) are not yet accessible (Bienert et al., 2017).

Comparative modeling has developed into a common technique to create 3D models in recent years, making these analyses crucial to molecular biology research. In fact, a number of new tools and specialized web servers are developed and released each month. example: the SWISS-MODEL, the I-TASSER, and the ROBETTA are examples of different comparative modeling web-based servers and databases.

1.5. The Overall Objectives of Presented Thesis

N-glycans obtained from glycoproteins, selectively promote the development of good microbes in the gut microbiome. This unique characteristic of *N*-glycans is considered a new approach for microbiome treatments and wellness since the microbial community can be manipulated by specific glycan structures. A better understanding of important biological roles and commercialization of *N*-glycans requires effective deglycosylation methods. Recombinantly produced enzyme, EndoBI-2 represents a great opportunity to produce *N*-glycans from different glycoproteins. Within the context of this thesis, the general purpose is to assess the function of EndoBI-2 and determine specificities, kinetics, and activity on various glycoproteins. The novel EndoBI-2 enzyme which is not commercially available was produced and characterized recombinantly to understand the unique features of *N*-glycans and large-scale production. The aims and objectives achieved under the thesis were detailed below.

In this thesis, the aims are to;

- Characterize the 3D structure of EndoBI-2 to assess unique features of the enzyme
- Recombinantly produce EndoBI-2 enzyme by using an *in-vivo* cloning system
- Characterize and investigate EndoBI-2 activity on model glycoprotein
- Release *N*-glycans from glycoprotein by EndoBI-2
- Characterize the released *N*-glycans by MALDI-(TOF)/TOF-MS

In this thesis, the objectives are to;

- 3D structure of unique EndoBI-2 enzyme was generated using comparative modeling approaches.
- Novel EndoBI-2 enzyme was produced recombinantly for further analysis including enzyme characterization, optimization, and immobilization studies.
- EndoBI-2 enzyme activity was tested on different glycoproteins and potential bioactive glycans were characterized by advanced mass spectrometry techniques.
- Literature related to efficient deglycosylation methods and glycobiology has been contributed.
- A contribution has been made to international potential collaborations in various research areas, such as human health, related to the proposed technology.

CHAPTER 2
THEORETICAL FRAMEWORK/PREVIOUS STUDIES

Within the purview of this thesis, novel EndoBI-2 from an infant's gut–associated *B. infantis* was recombinantly produced and characterized in extensive studies. The literature regarding EndoBI-2 has very limited information and analysis. The research on this enzyme that has been done indirectly or directly is outlined below.

2.1.   Theoretical Framework/Previous Studies

One of the key variables influencing the content of the infant's gut microbiota from the beginning of life is breastfeeding. The substantial amount of oligosaccharides present in human milk, which are used as a sole carbon source by saccharolytic beneficial microbes such as the Bifidobacterium species, influence this process. For using these oligosaccharides as a sole carbon source, infant-borne Bifidobacteria have a number of genomic capabilities. From this perspective, it was anticipated that these species would also interact with host glycoprotein *N*-glycans, which share structural similarities with HMOs Garrido et. al., 2012. In this research, *B. breve*, *B. infantis*, and some isolates of *B. longum* were shown to include endo-*ß-N*-acetylglucosaminidases, and their presence was connected with the capacity of these bacteria to deglycosylate glycoproteins. EndoBI-2 was isolated from *B. infantis* 157F using molecular techniques without signal and transmembrane sequences to promote protein expression by Garrido et. al., 2012. Together with gene cloning experiments, general properties such as thermostability, and optimum enzyme activity conditions were investigated. Hence, EndoBI-2 was considered a heat-resistant enzyme and they have activity on various types of *N*-glycans at 30-45°C pH 5 based on the first study results (Garrido et al., 2012).

CHAPTER 3
MATERIAL & METHOD

3.1. Material

All of the materials required to complete the experiments for this thesis, including chemicals, substrates, and used laboratory equipment, was listed under this subsection.

3.1.1. Chemicals, Kits, and Necessary Items

The necessary chemicals, kits, and other supplies were mentioned in Table 1 for the specific investigation.

Table 1

List of chemicals, kits, and necessary items

Name of Materials	Catalog Number	Necessary for
Acetic acid	137130	SDS-PAGE analysis
Methanol	106012	SDS-PAGE analysis
Coomassie Brilliant Blue	0472-25G	SDS-PAGE analysis
SureCast™ Acrylamide Solution (40%)	HC2040	SDS-PAGE analysis
SureCast™ TEMED	HC2006	SDS-PAGE analysis
Ammonium Persulfate	17874	SDS-PAGE analysis
Isopropanol	961.023.2500	SDS-PAGE analysis
Sodium dodecyl sulfate (SDS)	SDS001.500	SDS-PAGE analysis
10X Running Buffer (Tris-Glycine SDS)	TGS10	SDS-PAGE analysis
Laemmli Sample Buffer (2X)	LSB-2x	SDS-PAGE analysis
Tris base	TRS001.1	SDS-PAGE analysis
Qubit™ Protein Assay Kit	Q33211	Determination of protein concentration
Glycobuffer 2 Glycoprotein Denaturing Buffer, NP-40	P0704S	Buffers for deglycosylation process
L-Rhamnose monohydrate	83650	Induction of protein expression
HisPur™ Ni-NTA Resin	88222	Protein purification
Imidazole	56750	Cell lysis process
Sodium Chloride	31434-5Kg-R	Cell lysis process

Table 1 (continue)

Sodium Phosphate Monobasic Anhydrous	0571-1Kg	Enzymatic deglycosylation buffer
Sodium Phosphate Dibasic Dihydrate	04272-1Kg	Enzymatic degylcosylation buffer
10-kDa-cut-off centrifugal filter	UFC9010	Enzyme concentration

3.1.2. Substrates

The enzymatic deglycosylation studies were performed using LF from bovine colostrum (Code, L4765-50MG), lactoperoxidase (LPO) from bovine milk (Code, L2005-10MG), all of which were provided by Sigma-Aldrich. In addition to this, RNase B protein (Code, P07817S) and PNGase F enzyme (P0704S) from New England Biolabs was used to confirm enzyme activity.

3.1.3. Laboratory Equipment Used for Molecular Analysis

The Molecular Biology and Genetics Department's research laboratory at Çanakkale Onsekiz Mart University (COMU) was used for this thesis studies. The following is a list of the equipment used in this research's research laboratory (Table 2)

Table 2

List of Research Laboratory Resources and Equipments

Device name	Brand name
Cooling centrifuge	Hettich Mikro 200 R
Centrifuge	Beckman Allegra X-15R
Speedvac evaporator	Labconco
Orbital shaker	STUART
Analytical balance	Shimadzu
Dry bath	Benchmark Scientific BSH1002-E
Thermometer-Portable	IsoLab
pH meter	IsoLab
Ice generator	Izmak
Autoclave	NÜVE
-20°C freezer	Arçelik
Qubit™ 3 Fluorometer	Invitrogen
Vortex	Vortex Genie 2
Incubator	Indem Nüve EN 400

Table 2 continued	
Pure water system	Millipore
Power supply	Bio Rad
Bruker rapifleX™ MALDI Tissuetyper™	Bruker Daltonik GmbH, Bremen, Germany

3.2. Methods
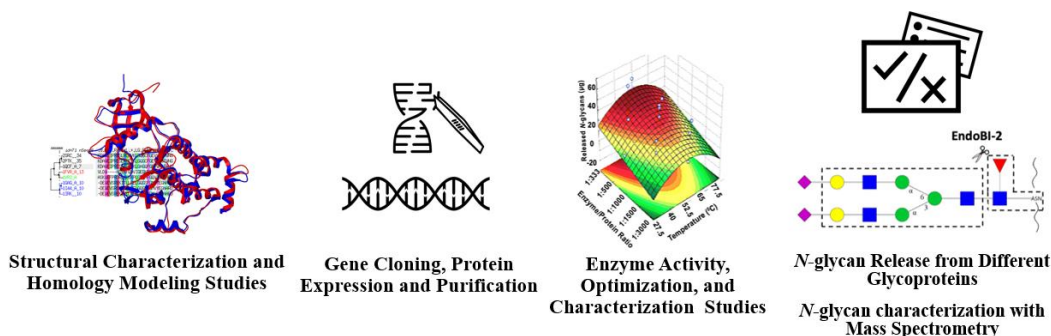
Figure 8. Overview of the thesis's general methodology.

3.2.1. Structural Characterization of EndoBI-2
3.2.2. Sequence Alignment and Database Search

BLAST was used to investigate EndoBI-1 and EndoBI-2 homologs. BLASTp was preferred for the initial searches. BLOSUM62 scoring matrix was chosen for this sequence alignment. EMBL webserver-based Clustal-Omega software was used for the multiple sequence alignments process. Sequence similarity and secondary structure data from the aligned sequences were rendered using the ESPript-3.0 program. With the use of the MEGA software, phylogenetic trees were created and displayed to understand the relationships among the enzymes. WebLogo 3 was used to prepare the sequence logos from the alignment results to better visualization.

3.2.3. Comparative Modelling Studies

Cordyceps militaris' Endo-COM crystal structures (PDB: 6KPL) were selected as the model template for the comparative modeling research. The full-chain protein structure prediction server ROBETTA, which is based on Rosetta software, was used to generate the enzymes' entire homology models, whereas the I-TASSER webserver was preferred to model the glycosidase domains of both enzymes. Verification of the reliabilities of the generated models as well as templates was performed SAVES web server. The structures

were superimposed using the MultiSeq module and STAMP algorithm of VMD. With the use of PDBsum, the structures of generated protein models, and templates were examined and summarized.

3.2.4. Molecular Docking Studies

SwissDock molecular docking server was used for the molecular docking of ligand analysis. Fuc-GlcNAc as a Fuc-containing "plus" subsite ligand was used. Marvin 17.6, 2017, ChemAxon was used to prepare the ligand, which included drawing, showing, and describing its structures and substructures. The interactive analysis and visualization program UCSF-Chimera was chosen for the preparation of protein models prior to docking analysis as well as the final docking results evaluation.

3.2.5. Molecular Dynamic Simulation Studies

NAMD was used to do all molecular dynamics (MD) simulations using CHARMM 36 together with an all-atom force field. Using VMD, the docked complexes and partial models of the enzymes were solvated and ionized. The TIP3P model was used to characterize the water molecules. The simulations of the NPT system were performed with 2-fs time steps and Langevin pressure coupling at 1.0 atm and 310 K. Using the PME approach, electrostatic interactions were calculated. VMD was used for the investigation and display of MD trajectories. 2 ps minimization and a greedy approach were used to run MD simulations for 10 nanoseconds (ns) in order to minimize and equilibrate the models and complexes. The average outcomes for all cases were revealed after each molecular simulation was performed three times with varied velocity seeds.

3.2.6. Molecular Cloning of EndoBI-2

The Expresso Rhamnose Sumo cloning and expression system (Lucigen cat no: 49013) was used for recombinant production of the EndoBI-2 enzyme (BLIF_1310 gene) without any enzymatic treatment. This approach allows for the 6xHis fusion protein tag to be used to tag the EndoBI-2 enzyme, enhancing its expression as well as the purity, stability, and solubility of recombinant protein. The Expresso system vector and acquired PCR products are combined after the target gene has been amplified using primers. The vector

mixture is then immediately transformed into *E. coli* 10G competent cells. Following the transfer, a colony PCR experiment was run to confirm that the target gene had successfully transformed (Figure 9) (Steinmetz, 2011).
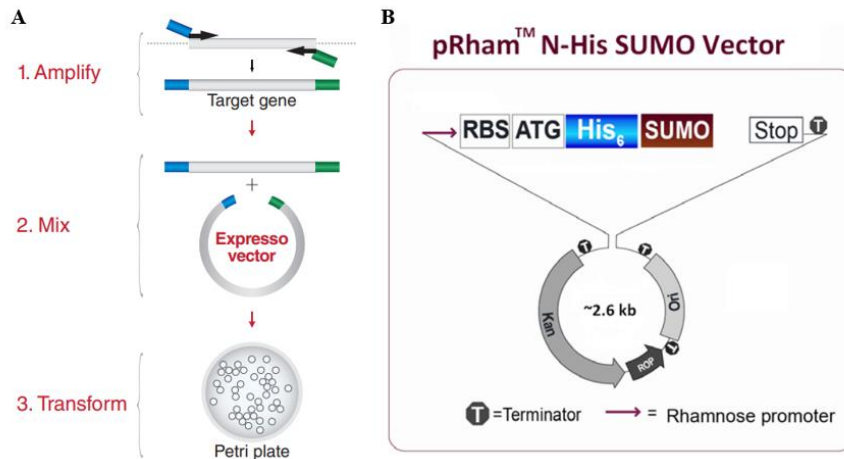
Figure 9. *In-vivo* cloning system strategy and pRham N-His SUMO vector design.

3.2.7. Design of the BLIF_1310 Gene with N-His Sumo Tag

To amplification of the desired region of the gene, particular forward and reverse primers were designed according to the manufacturer's recommendations. In addition, primers were designed without signal and transmembrane region sequences to promote further protein expression experiments. Using sterile water, the primer concentration was made to be 100 μM, while the fresh stock preparation for the PCR amplification phase was made to be 10 μM (Figure 10).

BLIF1310 Forward Primer

5'- CGC GAA CAG ATT GGA GGT GTT GCG AAC GCC CAG GAG

Oligo No	211206-1-4	GC	%61	Tm (Basic)	71.28°C	Total nmol	40.91 nmol
Skala	50 nmol	MW	11200.31	Conc	1636.61 ng/µl	Total ng	458251.96 ng
Saflastirma	DSLT	A260	51.9	OD	14.5	100 µM stok - µl TE	409.1

BLIF1310 Reverse Primer
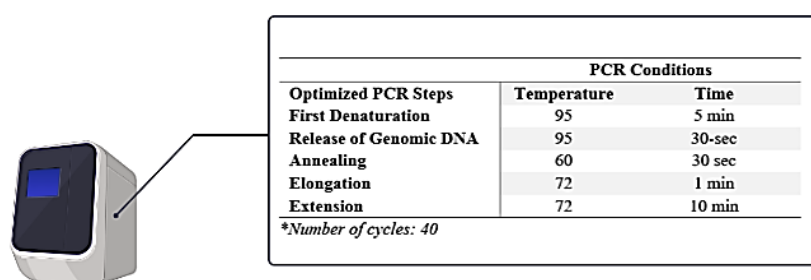
5'- GTG GCG GCC GCT CTA TTA CGC CGC GTT TCT GGC CGT

Oligo No	211206-1-5	GC	%67	Tm (Basic)	73.56°C	Total nmol	42.22 nmol
Skala	50 nmol	MW	11027.14	Conc	1662.82 ng/µl	Total ng	465590.36 ng
Saflastirma	DSLT	A260	47.6	OD	13.3	100 µM stok - µl TE	422.2

Figure 10. Designed specific PCR primers and their information.

3.2.8. PCR Amplification of BLIF_1310 Gene Using N-His Sumo Tag Primer Set

PCR was used to amplify the desired genes. The 50 µL volume used for the amplification step contained 2 µL of template DNA, 1 µL of each primer, 0.2 mM, 25 µL of master mix, and 21 µL of DNase/RNase-free water. PCR tubes were put into a thermal cycler after the mixture had reached the desired consistency. As follows are the PCR steps:

PCR Conditions		
Optimized PCR Steps	Temperature	Time
First Denaturation	95	5 min
Release of Genomic DNA	95	30-sec
Annealing	60	30 sec
Elongation	72	1 min
Extension	72	10 min
*Number of cycles: 40		

Figure 11. PCR steps information for gene of interest amplification.

Using agarose gel electrophoresis, the size of DNA amplified were investigated. Initially, the samples and DNA ladder were combined with the SafeRed loading dye (5:1 (v/v). 1% agarose gel electrophoresis system using 1X TBE running buffer was performed

at 100 V for 1 hour (h). PCR results were viewed using the ST4 1100 gel documenting system.

3.2.9. Transformation of N-His Sumo Tagged PCR Products and pRham Vector into *E. cloni* 10G Cells Using Heat Shock Method

For this process, 1 µL PCR product and 0.5 µL pRhamTM Vector DNA were transferred to 40 µL of *E. cloni* 10G cell gently. The obtained mixture found in falcon tube was incubated on ice for 30 min. Following the incubation on ice, the mixture containing was placed in the water bath (42°C) for 45 seconds (sec) and then placed onto the ice for 2 min to complete the heat shock process. A tube was filled with 960 mL of recovery media, which enables quick cell healing. The falcon tube was centrifuged for 10 minutes (min) at 4000 rpm after being incubated at 37°C for 1 h at 250 rpm. The pellet part was dissolved and homogenized in 100 mL of recovery media, and the supernatant part was discarded. Then obtained culture was spread to the LB agar (30 µg/mL kanamycin). Finally, plates were placed at 37°C overnight (Steinmetz, 2011).

3.2.10. Verification of Successful Transformants Using Colony PCR Technique

Colony PCR was used to determine if the transformants had recombinant genes or not. For this confirmation process, colonies from the incubated LB agar plates were specified randomly. After that, one portion of each colony was inoculated into a 15 mL sterile falcon tube which contains 5 mL of LB liquid media (30 µg/mL kanamycin). One portion of each colony was transferred to a PCR tube for PCR amplification. In the colony PCR experiment, primers included with the kit were used. The PCR procedure was carried out exactly as in the PCR amplification of the gene of interest step. Results were visualized using a 1% agarose gel electrophoresis system. Hence, successful transformants can be identified in this manner. On the other hand, inoculated cultures were incubated at 37°C, 200 rpm overnight. Based on the colony PCR results, successful transformants were stored at -80°C using a 15% glycerol stock solution (Figure 12) (Steinmetz, 2011).
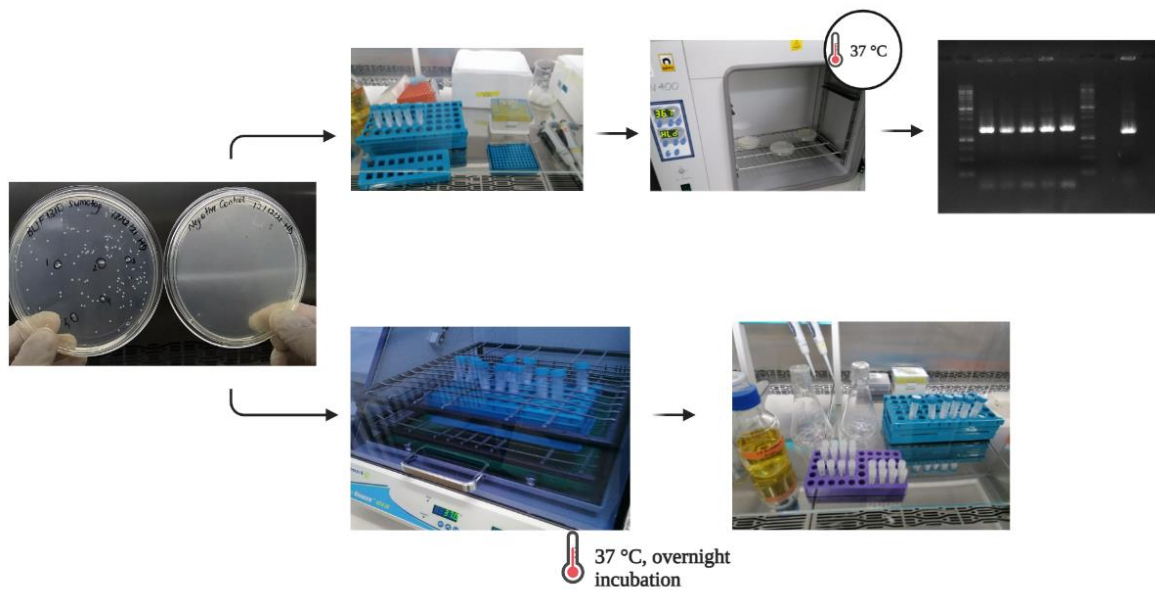
Figure 12. Summarization of colony PCR technique.

3.2.11. Protein Expression and Purification Studies

Protein Expression Using L-rhamnose Induction

Before scaling up for purification, a small-scale expression trial (50 mL) was carried out to assess the target protein's expression and solubility. Lysogeny Broth (LB) medium (30 µg/mL kanamycin) was inoculated with a colony of *E.cloni* 10G cells containing a pRham expression construct. The obtained culture was incubated at $37\,^{o}C$ 160-180 rpm (OD_{600}=0.5). To induce expression, L-rhamnose solution was added to the culture (0.2%). Then, the culture was waited at $24^{o}C$, 160-180 rpm overnight after the addition of L-rhamnose solution. Following the incubation process, the induced culture was centrifugated under optimized conditions (15 min., +4°C, and 4000 rpm). The pellet part was stored for the protein production step at -80°C. In addition, 1 mL of aliquots from the induced culture and prior to the addition of L-rhamnose solution to the culture were separated for the confirmation of expression of the protein on SDS-PAGE.

Protein Purification of N-His Tagged EndoBI-2 Using Batch Method

The initial step in the extraction and purification of proteins is cell lysis. To rupture of the *E.coli* cell membrane and release the content of the cell, lysis buffer and sonication were performed. The cell lysis process was started by dissolving the induced pellet part with

5 mL dH$_2$O. Following centrifugation under optimized parameters, the supernatant was discarded, and 6300 µL of lysis buffer was added to the remaining portion and incubated for 30 min. on ice. To complete cell lysis, sonication was applied. A sonicator's pulse mode was configured to cycle on for 10 s and cycle off for 59 s, with a 37% amplitude. The centrifugation method was used to remove the cell debris after sonication.

Using the batch approach, the purification process was accomplished. Two resin-bed volumes of equilibration buffer were initially introduced to and equilibrated in the Ni-NTA resin. Following centrifugation, the supernatant fraction was collected in a single, sterile falcon, mixed 1:1 with equilibration buffer, and placed in a falcon tube with resin. The resulting mixture was agitated for 30 min. at room temperature (rt), and 200 rpm to help the proteins adhere to the resin more firmly. The sample was then centrifuged for two min. at a speed of 700 x g at 4°C to remove the supernatant. To the remaining resin, ~20 mL of wash buffer was added and was centrifuged at 700 x g, 4°C, for 2 min. to eliminate any potential contaminants. At least three times this process was carried out. Following the wash stage, 10 mL of elution buffer was added to release the proteins bound to the resin, and the mixture was centrifuged at 700 x g, 4°C, for two min. An amicon tube was afterward employed to concentrate the eluted protein (10-kDa cut-off). Prior to usage, aliquots of the concentrated enzyme were kept at -80°C.

3.2.12. Confirmation of Purity of EndoBI-2 Enzyme Using SDS-PAGE Analysis

For confirmation of the purity of the EndoBI-2 enzyme, SDS-PAGE analysis was preferred. For this gel electrophoresis system, a 4-12% discontinuous gel system was prepared based on the stacking and separating gel contents and their concentration. The recombinant protein was mixed with 2X Laemmli buffer (1:1, v/v). Then, resulted mixture contained Laemmli buffer, and the protein was placed at 95°C for 5 min. After that, the sample loading onto an SDS-PAGE gel, this system was run at 200 V. By using the Coomassie Brilliant Blue solution, the gel was stained to completely visualize protein bands. To ensure the bands' clarity, a destaining solution (50% dH$_2$O, 10% Glacial Acetic Acid, and 40% Methanol) was administered for three hours.

3.2.13. Investigation of EndoBI-2 Enzyme Activity Using Model Glycoprotein

Ribonuclease B (RNase B) (5 mg/mL) was reconstituted in 0.2 M Na_2HPO_4 (pH 5), and EndoBI-2 was mixed with the sample. PNGase F was used as a positive control. The resulted two mixtures were incubated for 1 h at 37°C 200 rpm. The results were visualized using SDS-PAGE on 4–12% polyacrylamide gels. Under denaturing conditions, this experiment was conducted (Karav et al., 2015).

3.2.14. Kinetic Characterization and Optimization Studies of EndoBI-2

To optimize the enzymatic reaction conditions of the EndoBI-2 enzyme on RNase B (5 mg/mL) model glycoprotein, different pH/temperature conditions were determined and combined. The enzymatic reactions using EndoBI-2 under constant pH: 5 together with increased temperature (24°C, 30°C, and 37°C) and constant temperature (37°C) together with increased pH were performed. The reactions were terminated after 1 h.

To calculate the Vmax value of EndoBI-1, EndoBI-2, and their combination on RNase B, LF, whey proteins, soy, and pea proteins, released glycan concentrations were determined at different substrate concentrations (0.1-0.8 mg/mL) with a constant enzyme concentration (0.025 mg/mL). Then Vmax value for each substrate was calculated by using the linearized Lineweaver-Burk plotting technique.

3.2.15. Enzymatic Deglycostion of Different Glycoproteins Using EndoBI-2 Enzyme and Characterization of Released *N*-glycans by Advanced Mass Spectrometry

Enzymatic Deglycosylation of Different Glycoproteins Using EndoBI-2 Enzyme

Releasing of glycans from various glycoproteins including LF from bovine colostrum, LPO from bovine milk, and RNAseB model glycoprotein (5 mg/mL) was carried out by enzymatic deglycosylation with EndoBI-1, EndoBI-2, and enzyme cocktail which contains EndoBI-1 and EndoBI-2 enzymes to characterize different glycan profiles obtained from various glycoproteins and compare the enzyme activity. 5 mg of LF and LPO was

dissolved in 1 mL dH$_2$O, and 1 μL of EndoBI-2. The reaction was performed at 37 °C for 1 h (Karav et al., 2015).

2-Aminobenzoic acid (2-AA) was preferred to tag the released *N*-glycans. The obtained glycan was combined with 25 μL of 2-picoline borane and 20 μL of 2-AA tag. Afterward, the resulting mixtures were incubated for 2 h at 65°C. The pipette tip had a little piece of cotton placed at the bottom. Before loading, the 2-AA tagged *N*-glycans were combined with 85 μL of ACN. The pipette tip was initially cleaned three times with MQ and 85% ACN. To remove contaminants, the pipette tip was washed three times with 85/14/1, ACN/MQ/TFA, (v/v/v), and 85% ACN. Using 20 μL MQ, elution was accomplished by repeatedly (Pekdemir et al., 2022).

MALDI/MS Analysis

The characterization of the released glycan analysis was performed using a Bruker rapifeX^{TM} MALDI Tissuetyper^{TM} outfitted with a Smartbeam 3D laser. The equipment was initially calibrated using a peptide standard. On the AnchorChip MALDI-target plate, 1 μL of the purified *N*-glycans were spotted and given time to air dry. The next step was adding 1 μL of DHB matrix. In the investigation, a 25 kV acceleration voltage with a 160 ns extraction delay and 10,000 shots at 5000 Hz were added together for each spectrum. The software ProteinScape, which uses the GlycoQuest algorithm, was used to insert the obtained 2-AA tagged *N*-glycan peak data. By adjusting the MS tolerance of 50 ppm, the identified peaks were compared to a glycan database (Carbbank). Profiling of the characterized *N*-glycans based on the mass spectrum peaks was performed using the GlycoWorkBench software (Pekdemir et al., 2022).

CHAPTER 4
RESEARCH FINDINGS

This thesis describes the recombinant production and comprehensive characterization of unique EndoBI-2 from gut-associated *B. infantis*. The findings related to the experiments and analysis planned within the scope of the thesis are explained in detail in the subsections.

4.1. Structural Characterization of EndoBI-2

Sequence Alignment and Database Search

On the basis of homologs' similarity in amino acid sequence identified inside the Carbohydrate-Active enZymes (CAZy) Database, endo-*β-N*-acetylglucosaminidases (ENGase) enzymes are divided into two categories, glycoside hydrolase (GH) families 18 and 85. Blon_2468 is a gene that produces EndoBI-1, which is found in the bacteria *B. infantis* ATCC 15697. The enzyme is categorized as a GH20 member in the Genetic Sequence Database maintained by the NCBI (GenBank: ACJ53522.1) and the EBI (ENA: CP001095.1), but as a GH18 member in the Universal Protein Resource Knowledgebase (UniProtKB: B7GPC7). A product of the BLIF_1310 gene is EndoBI-2 from *B. infantis* 157F. In NCBI GenBank, the enzyme is categorized as a GH18 member (BAJ71450.1). Bifidobacterium such as *B. longum* SC 116, SC 630, and SC 706, *B. infantis* SC 142 and SC 143, and *B. breve* SC 559 are also capable of producing EndoBI-2 (Garrido et al., 2012). Together with its isoenzyme EndoBI-2, the EndoBI-1 entry is located under the GH18 members' page in the CAZy as an endo-*β-N*-acetylglucosaminidase (EC 3.2.1.96).

There are four *β-N*-acetylhexosaminidase (HEX, EC 3.2.1.52) enzymes among the thirteen GH20 members of *B. infantis* listed on the GH20 members' page in the CAZy (GenBank: ACJ51575.1/ACJ51836.1/ACJ53413.1/BAJ71575.1 and UniProtKB: B7GN69/B7GPV6/B7GNQ7/E8MUY1). On the other hand, there are only five GH18 members of *B. infantis* including EndoBI-1 and EndoBI-2. In contrast to the GH20 member enzymes of the same organism, EndoBI-1 shares just 14–19% of its primary structure with GH20, but EndoBI-2 and its isoenzyme EndoBI-1 share roughly 60% of their main structures (Table 3).

Table 3

Multiple Sequence Alignment Results of GH18 and GH20 Member Enzyme Sequences that Belong to *B. infantis* in CAZy.

	ACJ53522.1	BAJ71450.1	BAJ70116.1	VEG44861.1	ALE08664.1	ALE09291.1	BAJ71575.1	ACJ51836.1	BAJ68336.1	VEG40981.1	ALE09497.1	ACJ51575.1	BAJ68060.1	VEG40447.1	ALE08560.1	ACJ53413.1	BAJ70005.1	VEG44538.1
EndoBI-1	100	58.18	100	100	99.63	14.75	17.9	18.39	18.39	18.39	17.82	17.98	17.98	17.98	17.85	17.85	17.85	17.85
EndoBI-2	58.18	100	58.18	58.18	58.18	14.25	19.23	18.06	18.06	18.06	17.78	15.49	15.49	15.49	16.44	16.99	16.99	16.99

Although the GenBank entry of EndoBI-1 shows contradiction with other databases, GH18 and GH20 families are both consists of retaining glycosidase enzymes which acts with neighboring group participation mechanism and uses same amino acid residues as catalytic sites. The GH18 enzymes are mostly endo-type enzymes whereas GH20 exo-type. However, the main difference between these two families is their conserved active patterns which GH18 family has the D-X-E motif which is part of a diagnostic D-X-X-D-X-D-X-E motif whereas GH20 family has a D-E amino acid pair which is preceded in the primary sequence by the consensus H-X-G-G motif. The EndoBI-1 primary sequence contains a GH18 active site D177-G178-L179-D180-I181-D182-M183-E184 motif together with a GH20 active site like D316-R317-D318-G319-R320-N321-Y322-D323-E324-D325 motif preceded by the H305-N306-L307-G308-G309 motif which all three are in close proximity in overall 3D structure (Figure 13).
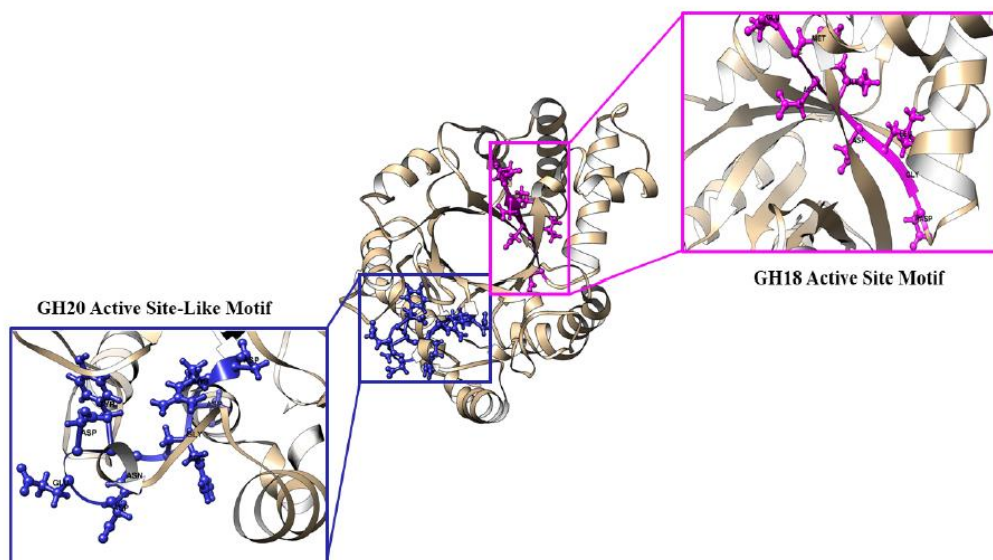
Figure 13. GH18 Active Site Motif and GH20 Active Site-Like Motif of EndoBI-1.

A BLOSUM62 matrix from the PDB data set was used for the BLAST alignments. Table 2 presents the BLASTp findings for EndoBI-1 and EndoBI-2, respectively. It was

discovered that there was a 32.41% and 34.11% percent homology between the query sequences and the best hits, respectively. Figure 14 shows the GH18 EndoS-like subfamily (cd06542) of the GH18 chitinase-like cluster from NCBI's Conserved Domain Database (CDD) hits (cd00598). Despite the fact that both enzymes have transmembrane C-terminal helixes, EndoBI-2 has a gram-positive LPXTG cell wall anchor motif (pfam00746), which suggests cell wall anchoring rather than cell membrane anchoring, based on the CDD search (Table 4).

Table 4.

Results for EndoBI-1 (A) and EndoBI-2 (B) alignment in PDB database with BLOSUM62 Scoring Matrix

Query account verifier	Subject account verifier	% Identity	E value	Bit score	% Positives
ACJ53522.1	6KPL_A	32.414	1.05×10^{-30}	122	47.93
	6KPN_A	31.724	9.26×10^{-30}	119	47.93
	6E58_A	27.986	5.64×10^{-16}	82.4	44.03
	4NUY_A	26.557	4.15×10^{-12}	70.1	42.62
	4NUZ_A	26.230	2.42×10^{-11}	67.4	42.3
	6EN3_A	25.902	6.48×10^{-11}	66.2	42.3
	1EDT_A	25.397	1.4	32.3	40.21
	1C90_A	24.868	2.9	31.2	40.21
	1C8X_A	25.000	3.8	30.8	40.43
	1C91_A	24.868	3.9	30.8	40.21
	1C3F_A	24.868	4.0	30.8	40.21
	1C92_A	24.868	4.4	30.8	39.68
	4AXN_A	27.907	9.9	29.6	45.35

B

Query account verifier	Subject account verifier	% Identity	E value	Bit score	% Positives
BAJ71450.1	6KPL_A	34.114	1.90×10^{-36}	138	49.16
	6KPN_A	33.445	2.04×10^{-35}	135	49.16
	6E58_A	25.949	4.06×10^{-15}	79.7	44.62
	4NUY_A	24.262	2.94×10^{-10}	64.3	42.95
	4NUZ_A	23.934	1.49×10^{-9}	62.0	42.62
	6EN3_A	23.607	5.13×10^{-9}	60.5	42.62
	3IAN_A	25.581	0.88	33.1	43.02
	4W5Z_A	33.663	3.7	31.2	47.52
	4HMC_A	33.663	6.1	30.8	47.52

A. List of Domain Hits for EndoBI-1

B. List of Domain Hits for EndoBI-2

Signal Helix Glycoside Hydrolase Domain Active Site Anchor Helix
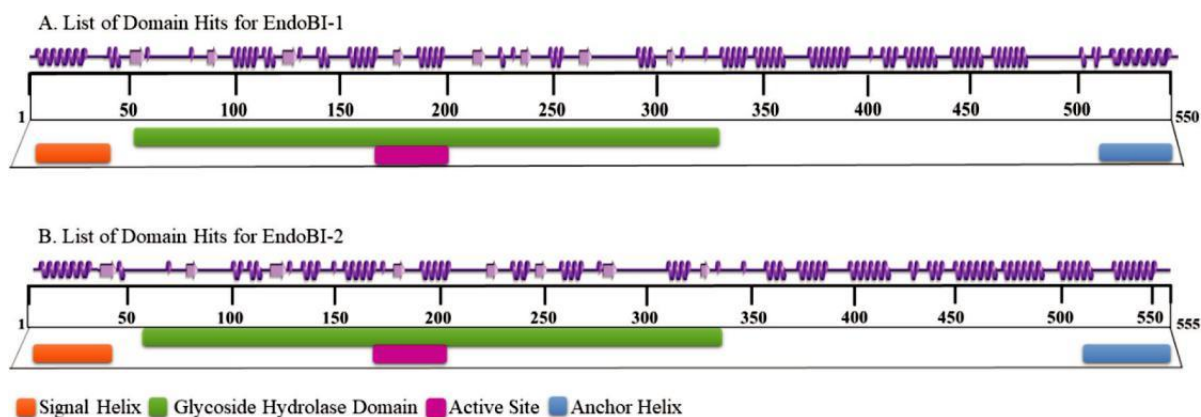
Figure 14. Secondary structural information from the EndoBI-1 and EndoBI-2 homology models is summarized graphically (PDBsum) along with information on conserved domains (CDD).

The Overall 3D Structure of the EndoBI-1 and EndoBI-2 Enzymes

A signal helix (1-36), the active ENGase (37-517), and a transmembrane helix (518-545) are all parts of the EndoBI-1 enzyme, whereas a signal peptide (1–60), the active ENGase (61–515), and anchor helix (520 555) are all components of the EndoBI-2 enzyme (Figure 14). Using the SIB ExPASy Bioinformatics Resource Portal's Compute pI/Mw tool, the masses of the entire enzymes were determined to be 56.1 and 59.6 kDa, respectively. The N-terminal glycosidase domain is present in both enzymes. EndoBI-1's GH domain is made up of amino acids 51-366, while EndoBI-2's GH domain is made up of amino acids 61-360. The enzymes' second domain is a substrate binding domain that may comprise a carbohydrate-binding module (CBM) and a 4-helix up-down bundle domain (Figure 15).
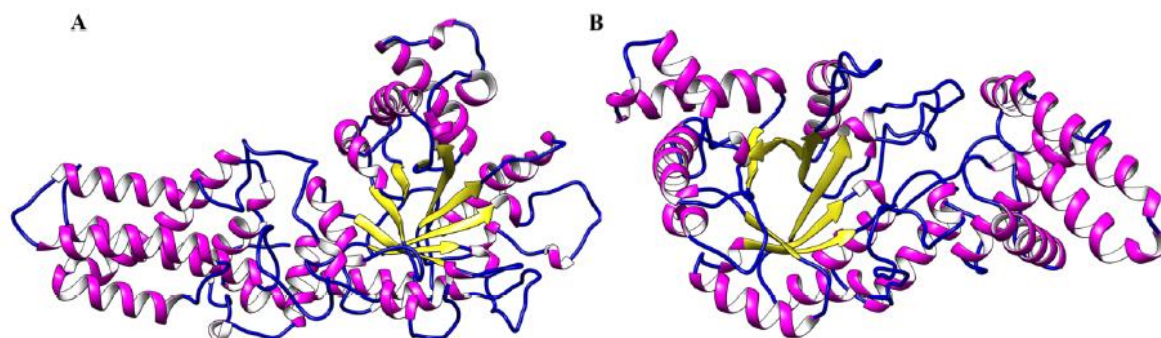
Figure 15. Overall 3D structures of EndoBI-1 (A) and EndoBI-2 (B) models.

Using the SAVES server, the complete models were validated. Results from VERIFY show that for EndoBI-1 and EndoBI-2, 95.3% and 98.0% of the residues, respectively, have average 3D-1D scores of ≥0.2. The total quality factors were determined to be 90.2% and 95.4%, respectively, using the ERRAT program.

PROCHECK web-based server was used to produce the Ramachandran plots for the enzymes. EndoBI-1's Ramachandran plot displays 95.2% of the amino acids (456) in the regions that are most favored, 4.2% amino acids (20) in the additional favored regions and only 0.6% amino acid (3) found in the disallowed regions. These findings indicated a generally respectable structure. Similar to this, EndoBI-2's Ramachandran plot displays 95.7% of the amino acids (442) in the regions that are most favored, 3.7% amino acids (17) in the additional favored regions and only 0.6% amino acid (3) found in the disallowed regions. These results showed a generally good structure (Figure 16).
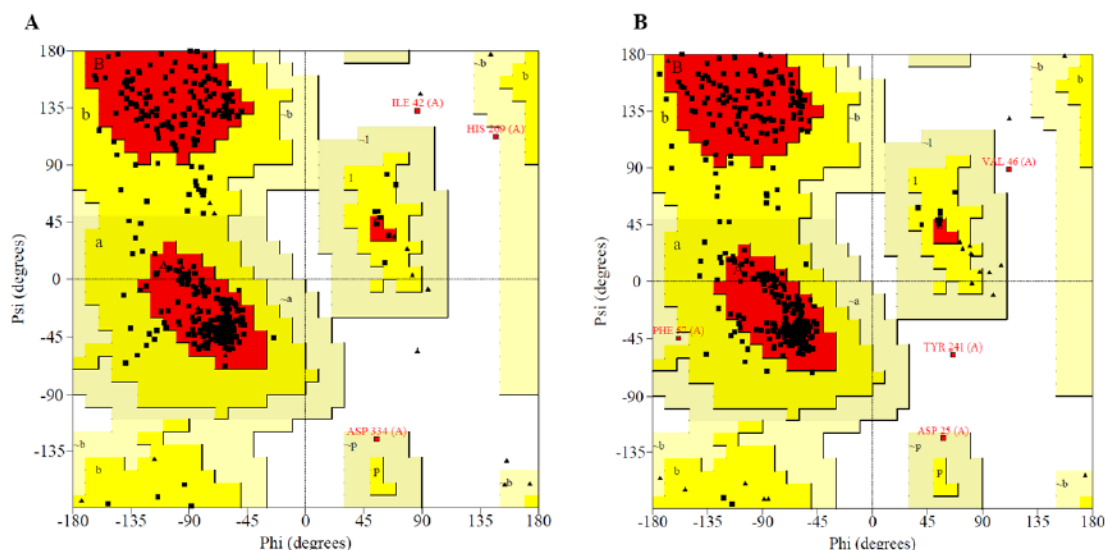
Figure 16. Ramachandran plots of EndoBI-1 (A) and EndoBI-2 (B).

EndoS and EndoS2 from *Streptococcus pyogenes* were shown to have higher structural alignment findings for enzymes than *Cordyceps militaris* chitinase, Endo-COM, despite the fact that sequence alignment outcomes were higher with this enzyme. EndoBI-1 was discovered to be structurally most comparable to the EndoS (Dali Z-Score = 30.2, RMSD = 3.5, %ID = 25), while Dali Z-Score indicates that EndoBI-2 is more comparable to EndoS2 (Dali Z-Score = 33.7, RMSD = 2.2, %ID = 24) (Figure 17).
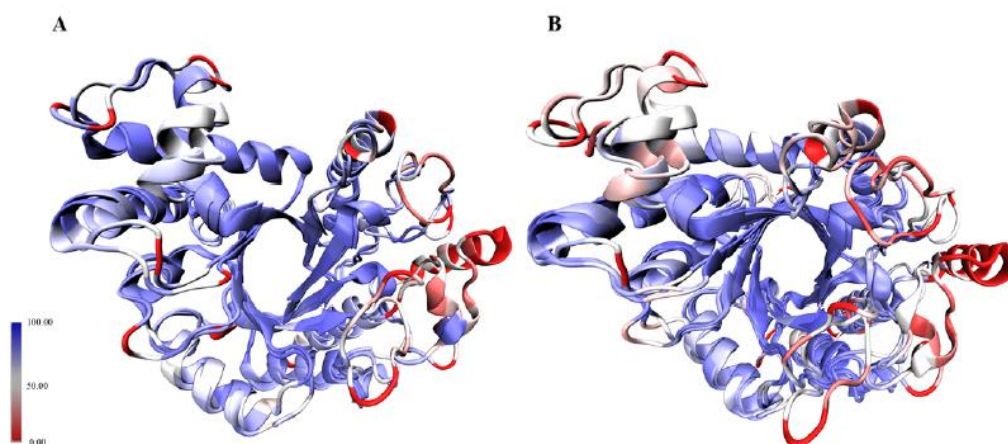
Figure 17. GH domain structural alignment findings belong to A) EndoBI-1 and EndoBI-2; and B) EndoBI-1, EndoBI-2, and template (Endo-COM).

EndoBI-1 and EndoBI-2 Contains N-terminal Signal Sequences

The enzyme EndoBI-1 has a signal helix (1-36). The transit peptide of Sus scrofa GTP-specific succinyl-CoA synthetase (UniProtKB: O19069) and isoform II of Sus scrofa GTP-specific succinyl-CoA synthetase (69.2% identical, E value: 9.7) (UniProtKB: O19069-1) are comparable to the signal helix sequence of the EndoBI-1 (69.2% identity, E value: 9.6). The signal peptide 1-60 is also present in the EndoBI-2 enzyme. The signal peptide of one of the most important extracellular proteinases, *Streptomyces griseus* trypsin (UniProtKB: P00775), shares 52.0% identity with the signal helix sequences of EndoBI-2 (E value: 2.2).

Both enzymes' signal peptides have the distinctive tripartite structure, which consists of a hydrophobic helical core, a polar N-terminal area, and a hydrophilic C-terminal portion. The GH domains of both enzymes are intimately connected to their C-terminal regions.

EndoBI-1 and EndoBI-2 Contains C-terminal Anchor Helixes

Both EndoBI-1 and EndoBI-2 have C-terminal helixes for anchoring in addition to the signal peptides. The anchor helix sequence of the EndoBI-1 is most similar (91.7% identity, E value: 7.6×10^{-1}) to the C-terminal transmembrane helix of *Raphidocelis subcapitata* amino acid permease, a twelve-pass transmembrane amino acid transporter

(UniProtKB: A0A2V0PHW6) and C-terminal transmembrane helix of the LRAT domain-containing protein of *Sarcophilus harrisii* (UniProtKB: G3VKG8), another integral membrane protein (73.7% identity, E value: 2.7×10^{-1}). So that EndoBI-1 is considered an outer membrane bound enzyme that acts on the exterior matrix of the cells, these findings are compatible with the literature. On the other hand, EndoBI-2 C-terminal sequence has both cell wall/membrane associated (K4R7Q6) and cell membrane associated (A0A0N0HIK6) homolog sequences in UniProtKB. The anchor helix sequence of the EndoBI-2 is most similar to the uncharacterized proteins S9PCE9 (76.5% identity, E value: 2.7×10^{-1}) and A0A1C4N4E0 (63.2% identity, E value: 1×10^{-4}). Therefore, although the existence of a anchor helix from gram positive LPXTG cell wall suggests a cell wall interaction for EndoBI-2, this finding should be validated due to the existence of similar enzymes with cell membrane association.

Glycosidase Domain Analysis and Its Unique Structural Features

The TIM barrel conformation $((\beta/\alpha)8)$, which is typical of bacterial endoglycosidases, is adopted by the EndoBl glycosidase domains. GH domain of EndoBI-1 shares the greatest structural similarities with the corresponding domain of EndoS (Dali Z-Score = 41.1, RMSD = 1.6, %ID = 24), while GH domain of EndoBI-2 shares the most structural similarities with the comparable to Endo-COM domain (Dali Z-Score = 39.6, RMSD = 0.9,%ID = 32) (Figure 18).

Figure 18. Structural alignment of GH domains of (A) EndoBI-1 with EndoS: and (B) EndoBI-2 with EndoS2.

The EndoBI-1 and EndoBI-2 Enzymes Use a Double Displacement Reaction Mechanism with Neighboring Group Participation

The catalytic residues part of both enzymes are situated in a conserved (LIVMFY)-(DN)-G-(LIMFY)-(DN)-(LIMFY)-DN)-X-E motif, just like other GH18 enzymes (Figure 20). The highly conserved motif shared by all ENGases with known protein structures is L-D-G-L-D-I-D-M-E. EndoBI-1 and EndoBI-2's catalytic residues are similar (Figure 19). E184 and E193 residues for EndoBI-1 and EndoBI-2 respectively may function as a proton donor general acid to protonate the glycosidic oxygen to aid in the aglycone's removal. During the hydrolysis of the intermediate oxazolinium ion, they function as a universal base to deprotonate the nucleophilic water molecule (Figure 21).

Figure 19. A) Hydrolytic cleave site for EndoBI-1 and EndoBI-2, B) Phylogenetic tree, and C) Active site multiple sequence alignment of ENGases in PDB database together with EndoBI-1 and EndoBI-2.

Figure 20. Sequence logo for the EndoBI-1 and EndoBI-2 active sites of ENGases in the PDB database.

Figure 21. Proposed catalytic mechanism for EndoBI-1 and EndoBI-2.

The Docking Studies

The position of the catalytic residues and the 3D form of the TIM barrel are fundamentally preserved in both EndoBI-1 and EndoBI-2. But unlike other enzymes, the loops that connect the barrel helixes are very different. The structural differences between EndoBI-1 and EndoBI-2 were demonstrated using the loop displaying system. The loop system indicates that the EndoBI-1 and EndoBI-2 plus subsites are in loop 6, while the EndoBI-1 and EndoBI-2 minus subsites are close to loops 1 and 2. To test whether the protein side of N-glycan substrates can be recognized by the GH domain of EndoBI-1 and EndoBI-2, these places received Fuc-GlcNAc docking. The E184 and E193 OE atoms were used as the focus of the docking studies, with a 10 $Å^3$ grid box and 3 Å flexibility.

Figure 22. The graphical represantatios of loop organization belongs to EndoBI-1 (A) and EndoBI-2 (B).

The Endo-COM dock results were consistent with the crystal structure data (6KPN) (Figure 23). While GlnNAc is located in the +1-side cleft in loop 6, Fuc is located in the +1' side cleft. The substrate and enzyme interact via H bonds. With the R groups of the N193 and Y216 as well as the backbone of R218 the Fuc molecule forms many H bonds. The GlcNAc molecule makes numerous H bonds with the R groups of the Q156 and perhaps the W243. The minor variations in the +1 and +1' recognition sites may control the variances in affinities between various ENGase enzymes. The alteration of these sequences (Figure 24) may have a significant impact on the EndoBI enzymes' ability to bind substrate, which has important implications for molecular genetic research.

35
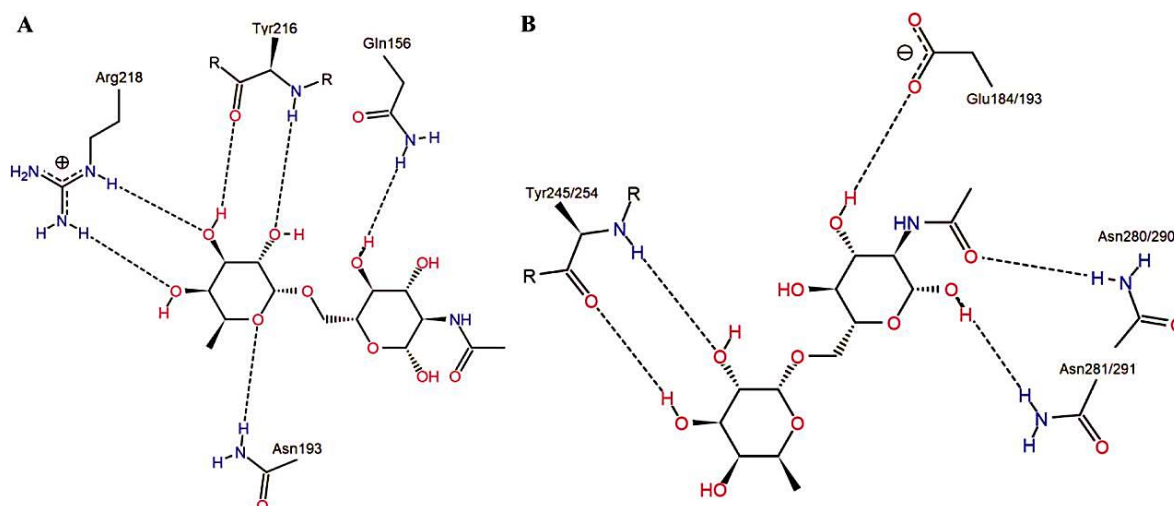
Figure 23. PoseView images show the mutant template (D154N/E156Q) crystal structure (6KPN) (A) and the docking outcome using EndoBI-1 and EndoBI-2 enzymes (B).
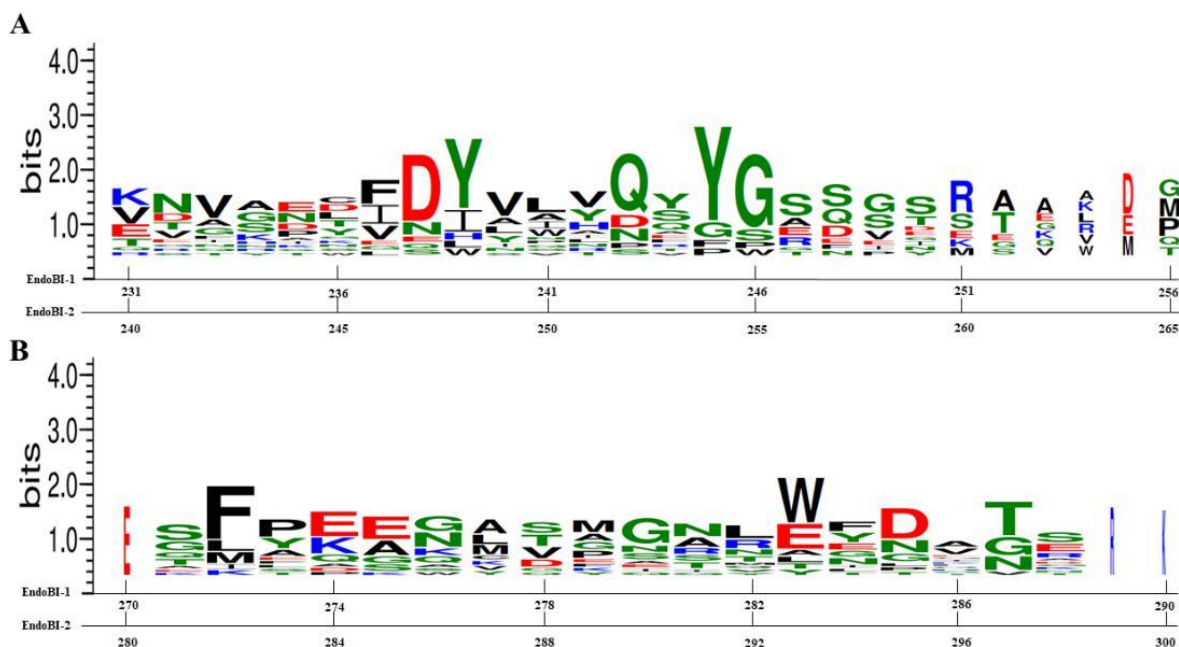
Figure 24. Sequence logos of the amino acid residues that enable ENGase enzymes' +1' (A) and +1 (B) subsites to recognize substrates.

Molecular Dynamic Simulation

The clusters with the highest full-fitness values were chosen after the docking studies, and MD simulations were used to assess the stability of the GH domain model with ligand. The complexes' RMSD, RMSF, and Rg patterns, as well as the ligands' RMSD patterns, were supplied in the figures 25-27.
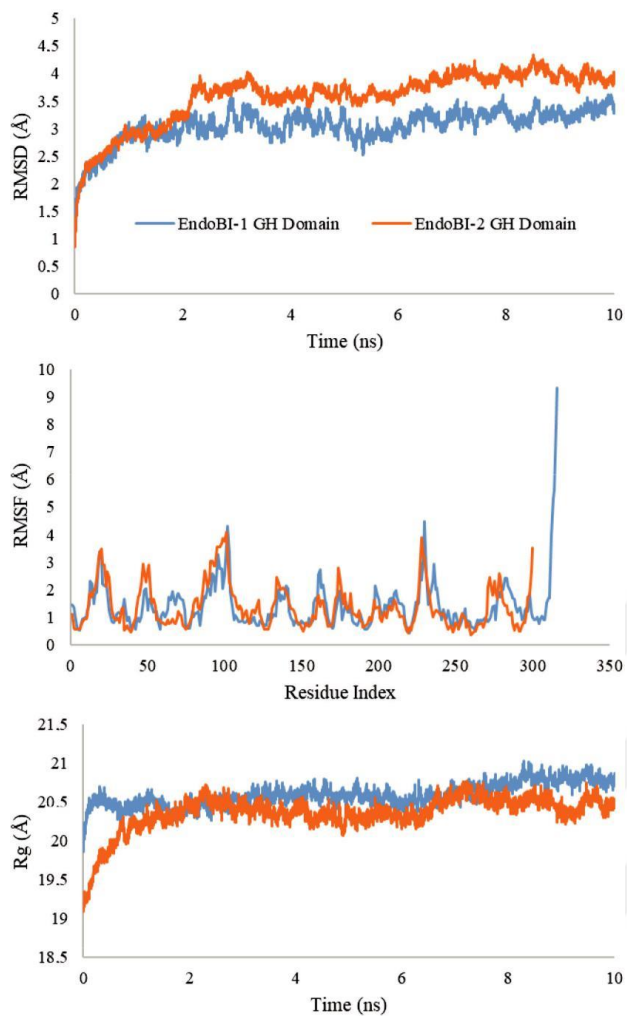
Figure 25. RMSF, RMSD, and Rg patterns of the GH domain of EndoBI-1 and EndoBI-2 enzyme models over MD trajectories
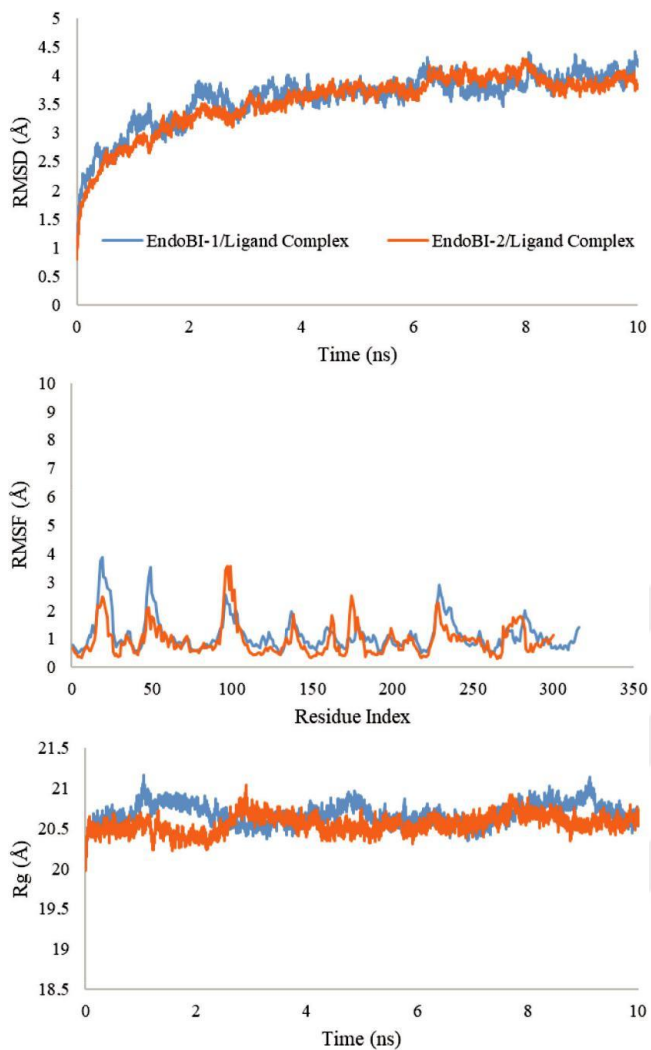
Figure 26. RMSF, RMSD and Rg patterns of the GH domain – ligand complexes of EndoBI-1 and EndoBI-2 enzymes over MD trajectories.
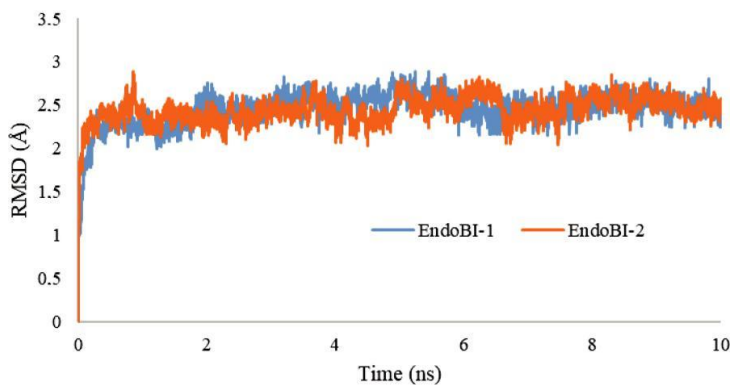
Figure 27. MD trajectories of the bound ligands' RMSD patterns.

Figure 28. MD trajectories of the mean distance between ASP182/191 OD atoms and LIGH21 atoms.

4.2. Molecular Cloning of EndoBI-2

Primer Design of the BLIF_1310 Gene with N-His Sumo Tag

Figure 29. Graphical view of designed amplification primers and their detailed information from NCBI Primer-Blast

PCR Amplification of BLIF_1310 Gene Using N-His Sumo Tag Primer Set

Figure 30. PCR amplification results of BLIF_1310 gene using N-His sumo tag primer set.

In this PCR amplification of desired gene from bacterial genome, 1266 base pair (bp) was expected for BLIF_1310 gene.

Transformation of N-His Sumo Tagged PCR Products and pRham Vector into *E. cloni 10G* Cells Using Heat Shock Method
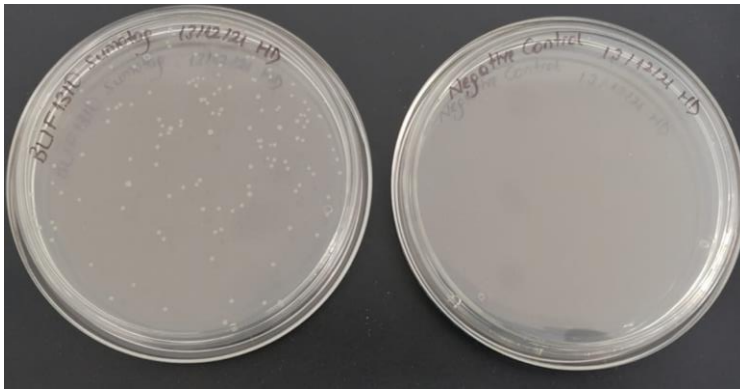
Figure 31. Transformation results and obtained colonies on LB agar plate with negative control.

Verification of Successful Transformants Using Colony PCR Technique
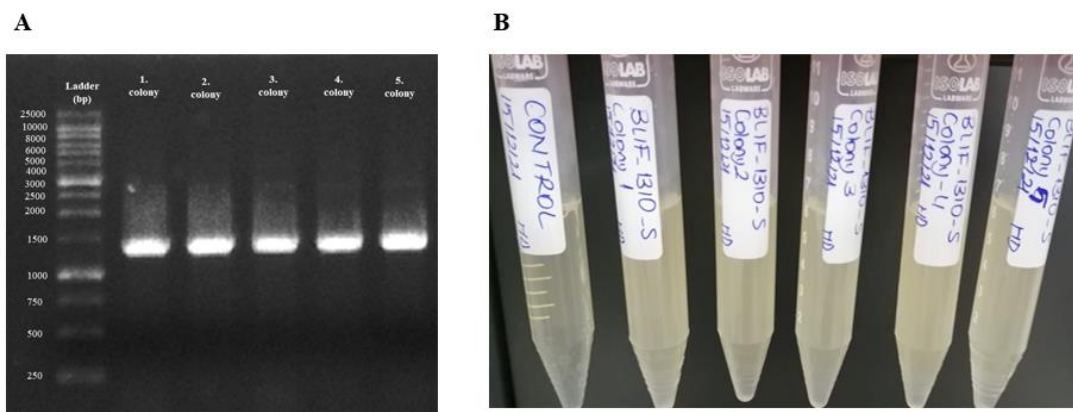
Figure 32. Colony PCR experiment results and the visualization of successful transformants using agarose gel electrophoresis system (A), and inoculated cultures with transformants.

According to the colony PCR experiment findings, vector and desired gene were successfully inserted into *E.coli* 10G competent cells.

4.3. Protein Expression and Purification Studies
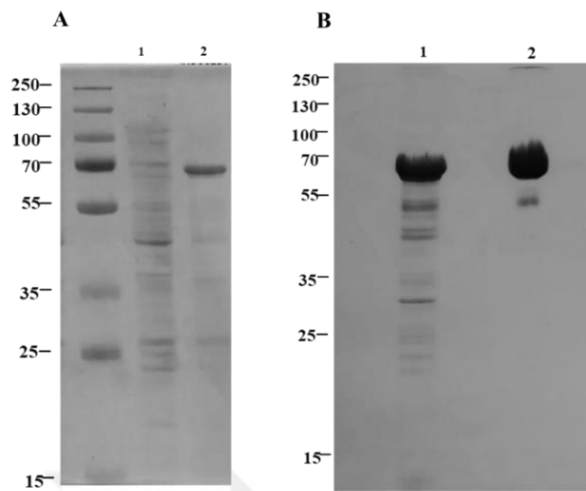
Figure 33. Protein expression experiment results on SDS-PAGE Gel (4-12%) of protein expression experiment results. Lane 1: Uninduced sample (10 µL); Lane 2: induced sample (10 µL) (A). SDS-PAGE (4-12%) gel of recombinant EndoBI-2. Lane 1: Lysate (5 µl); Lane 2: EndoBI-2 (5µL) (~58.45 kDa) (B).
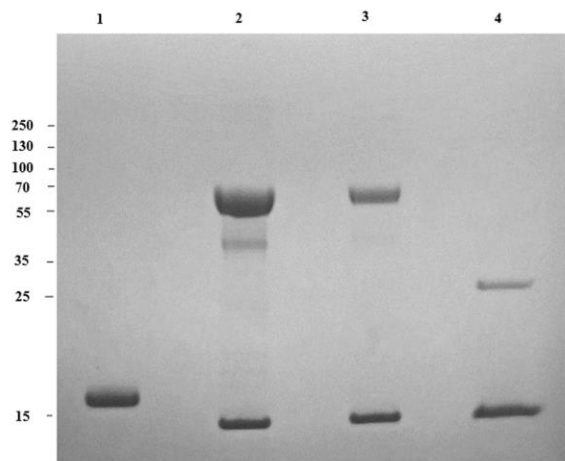
Figure 34. Deglycosylation of denatured RNase B by EndoBI-2 and PNGase F on 4-12% SDS-PAGE gel. Lane 1: Glycosylated RNase B (17kDa). Lane 2: Denatured RNase B deglycosylated by EndoBI-2 (2 µL) Lane 3: Denatured RNase B deglycosylated by EndoBI-2 (0.5 µL) (14kDa). Lane 4: Denatured RNase B deglycosylated by PNGase F. RNase B denaturation step for PNGase F deglycosylation also includes detergents as the developer suggest.

4.4. Enzyme Kinetic Characterization and Optimization Studies of EndoBI-2

According to the results of the optimization tests, temperature has a significant impact on EndoBI-2's ability to activate RNase B.
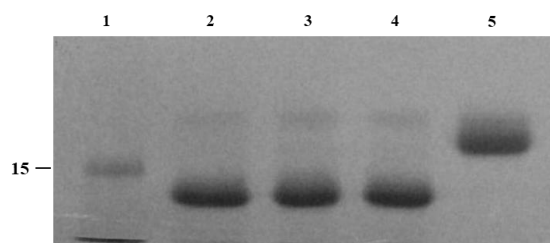
Figure 35. SDS-PAGE (4-12%) gel of recombinant EndoBI-2 enzyme optimization experiment. Lane 1: ladder (kDa); Lane 2: Denatured RNase B deglycosylated by EndoBI-2 at 37 °C pH 5; Lane 3: Denatured RNase B deglycosylated by EndoBI-2 at 30 °C pH 5; Lane 4: Denatured RNase B deglycosylated by EndoBI-2 at 24 °C pH 5 Lane 5: RNase B. 1 µL of EndoBI-2 and RNase B were used for the experiment.
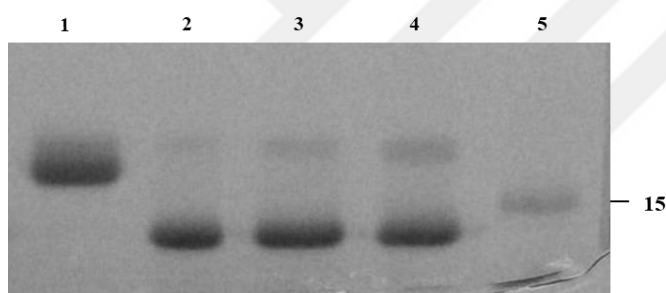
Figure 36. SDS-PAGE (4-12%) gel of recombinant EndoBI-2 enzyme optimization experiment. Lane 1: RNase B; Lane 2: Denatured RNase B deglycosylated by EndoBI-2 at 37 °C pH 5; Lane 3: Denatured RNase B deglycosylated by EndoBI-2 at 37 °C pH 6; Lane 3: Denatured RNase B deglycosylated by EndoBI-2 at 37 °C pH 7 Lane 5: Ladder (kDa). 1 µL of EndoBI-2 and RNase B were used for the experiment.

Table 5

Kinetic parameters of EndoBI-1 and EndoBI-2 on RNase B, LF, whey, soy and pea proteins.

	RNASE B	LF	WHEY	SOY	PEA
EndoBI-1	0.34+/-0.06	2.34+/-0.26	0.78+/-0.04	3.25+/-0.26	5.25+/0.24
EndoBI-2	0.33+/-0.07	1.88+/-0.32	0.26+/-0.05	5.29+/0.27	7.25+/0.35
Combined	0.34+/-0.06	2.76+/-0.28	0.88+/-0.09	5.36+/0.21	7.65+/0.26

*Kinetik Parameters (mg glycan released per min)

4.5. Enzymatic Deglycosytion of Different Glycoproteins Using EndoBI-2 Enzyme and Characterization of Released *N*-glycans by Advanced Mass Spectrometry

A

B

C
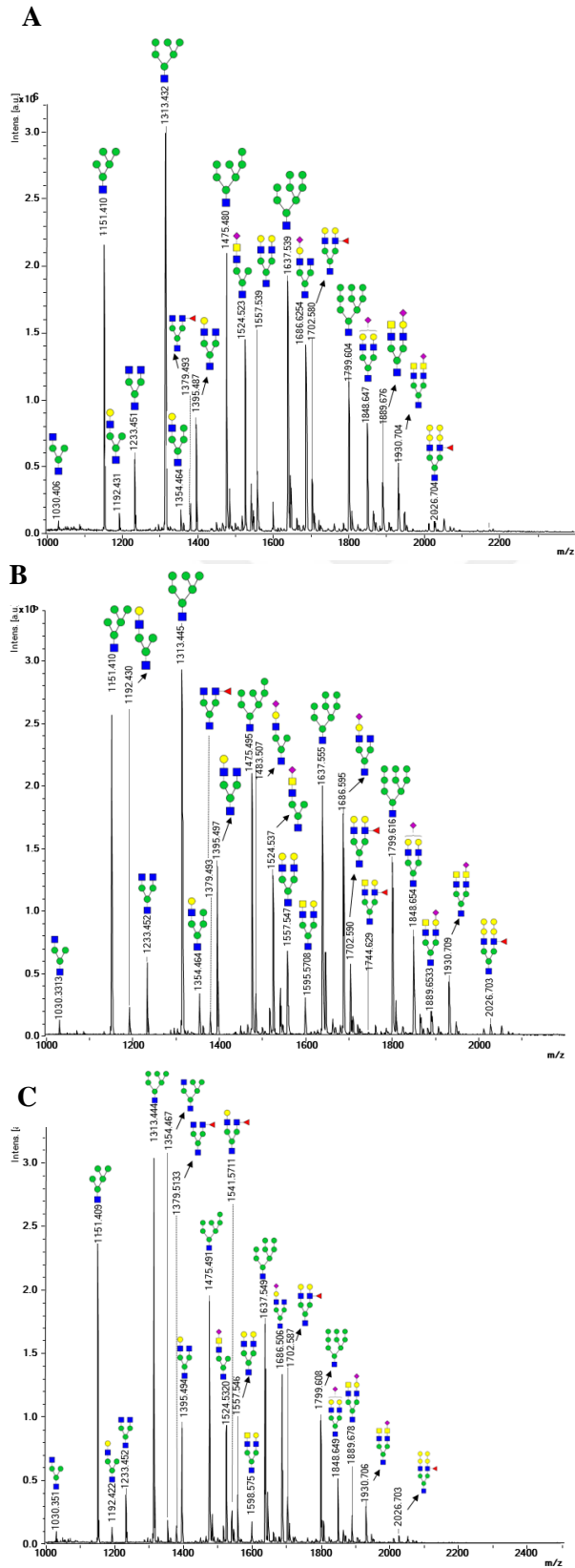
Figure 37. MALDI-MS spectrum of released *N*-glycans from enzymatic deglycosylation of bovine LF with EndoBI-1 (A), EndoBI-2 (B), and enzyme cocktail including EndoBI-1 and EndoBI-2 (C).

Table 6

N-glycan composition released from enzymatic deglycosylation of bovine LF with EndoBI-1.

Row	Composition	m/z meas.	z	m/z calc.	Mr calc.	Δ MH+ [Da]	Int.	Δ m/z [ppm]	MH+ meas.	MH+ calc.	Δ m/z [Da]	Δ MH+ [ppm]
1	Hex3HexNAc2-AA	1030.373	-1	1030.373	1031.381	-0.001	85275	-0.66	1032.388	1032.388	-0.001	-0.66
2	Hex5HexNAc1-AA	1151.412	-1	1151.400	1152.407	0.013	2212434	11.24	1153.428	1153.415	0.013	11.22
3	Hex4HexNAc2-AA	1192.446	-1	1192.426	1193.434	0.020	144579	16.36	1194.461	1194.441	0.020	16.33
4	Hex3HexNAc3-AA	1233.475	-1	1233.453	1234.460	0.023	507427	18.53	1235.491	1235.468	0.023	18.50
5	Hex6HexNAc1-AA	1313.471	-1	1313.452	1314.460	0.019	3103629	14.13	1315.486	1315.467	0.019	14.11
6	Hex5HexNAc2-AA	1354.506	-1	1354.479	1355.487	0.027	184864	19.84	1356.521	1356.494	0.027	19.81
7	Hex2HexNAc3NeuAc1-AA	1362.474	-1	1362.495		-0.021	71423					
8	Hex3HexNAc3dHex1-AA	1379.539	-1	1379.511	1380.518	0.028	219080	20.42	1381.554	1381.526	0.028	20.39
9	Hex4HexNAc3-AA	1395.534	-1	1395.505	1396.513	0.029	766914	20.80	1397.550	1397.521	0.029	20.77
10	Hex7HexNAc1-AA	1475.531	-1	1475.505	1476.513	0.026	2120330	17.45	1477.546	1477.520	0.026	17.43
11	Hex4HexNAc2NeuAc1-AA	1483.548	-1	1483.522	1484.529	0.026	315204	17.85	1485.563	1485.537	0.026	17.83
12	Hex6HexNAc2-AA	1516.561	-1	1516.532	1517.540	0.030	125528	19.58	1518.577	1518.547	0.030	19.55
13	Hex3HexNAc3NeuAc1-AA	1524.523	-1	1524.548		-0.025	1.423452+e6					
14	Hex5HexNAc3-AA	1557.589	-1	1557.558	1558.566	0.031	424867	19.98	1559.605	1559.573	0.031	19.95
15	Hex4HexNAc4-AA	1598.614	-1	1598.585	1599.593	0.029	212388	18.17	1600.629	1600.600	0.029	18.15
16	Hex8HexNAc1-AA	1637.586	-1	1637.558	1638.566	0.028	1827677	17.28	1639.601	1639.573	0.028	17.25
17	Hex3HexNAc5-AA	1639.631	-1	1639.611	1640.619	0.020	381322	11.96	1641.646	1641.626	0.020	11.95
18	Hex5HexNAc2NeuAc1-AA	1645.600	-1	1645.574	1646.582	0.026	392442	15.62	1647.615	1647.589	0.026	15.60
19	Hex7HexNAc2-AA	1678.612	-1	1678.585	1679.592	0.027	60244	16.07	1680.627	1680.600	0.027	16.05
20	Hex4HexNAc3NeuAc1-AA	1686.625	-1	1686.601	1687.609	0.024	1318435	14.52	1688.640	1688.616	0.024	14.50
21	Hex6HexNAc3-AA	1719.644	-1	1719.611	1720.619	0.033	87431	18.97	1721.659	1721.626	0.033	18.95
22	Hex5HexNAc4-AA	1760.665	-1	1760.638	1761.645	0.028	66992	15.73	1762.680	1762.653	0.028	15.72
23	Hex3HexNAc5dHex1-AA	1785.697	-1	1785.669	1786.677	0.028	68830	15.52	1787.712	1787.684	0.028	15.50
24	Hex9HexNAc1-AA	1799.640	-1	1799.611	1800.619	0.029	1075065	16.07	1801.655	1801.626	0.029	16.05
25	Hex6HexNAc2NeuAc1-AA	1807.650	-1	1807.627	1808.635	0.023	148916	12.57	1809.665	1809.642	0.023	12.56
26	Hex5HexNAc3NeuAc1-AA	1848.679	-1	1848.654	1849.662	0.025	716902	13.64	1850.694	1850.669	0.025	13.62
27	Hex7HexNAc3-AA	1881.695	-1	1881.664	1882.672	0.031	50278	16.36	1883.710	1883.679	0.031	16.34
28	Hex4HexNAc4NeuAc1-AA	1889.706	-1	1889.680	1890.688	0.025	348212	13.41	1891.721	1891.695	0.025	13.39
29	Hex5HexNAc5NeuAc1-AA	1930.704	-1	1930.707		-0.003	521923					
30	Hex6HexNAc3NeuAc1-AA	2010.727	-1	2010.707	2011.714	0.020	70803	10.04	2012.742	2012.722	0.020	10.03
31	Hex5HexNAc4NeuAc1-AA	2051.756	-1	2051.733	2052.741	0.023	98653	11.39	2053.772	2053.748	0.023	11.38
32	Hex5HexNAc4NeuGc1-AA	2067.751	-1	2067.728	2068.736	0.023	46202	11.23	2069.766	2069.743	0.023	11.22
33	Hex4HexNAc5NeuAc1-AA	2092.719	-1	2092.760		-0.041	25770					
34	Hex5HexNAc3NeuAc2-AA	2139.756	-1	2139.749		-0.007	24165					
35	Hex4HexNAc4NeuAc2-AA	2180.779	-1	2180.776	2181.783	0.003	26708	1.45	2182.794	2182.791	0.003	1.45
36	Hex3HexNAc5NeuAc2-AA	2221.799	-1	2221.802		-0.003	22101					
37	Hex12HexNAc2-AA	2488.800	-1	2488.849		-0.049	17097					
38	Hex13HexNAc2-AA	2651.806	-1	2650.902		0.905	15901					
39	Hex14HexNAc2-AA	2812.816	-1	2812.954		-0.138	13495					

*Different *N*-glycan compositions were also detected manually using FlexAnalysis software.

Table 7

N-glycan composition released from enzymatic deglycosylation of bovine LF with EndoBI-2.

Row	Composition	m/z meas.	z	m/z calc.	Mr calc.	Δ MH+ [Da]	Int.	Δ m/z [ppm]	MH+ meas.	MH+ calc.	Δ m/z [Da]	Δ MH+ [ppm]
1	Hex3HexNAc2-AA	1030.331	-1	1030.373	1031.381	-0.042	127107	-40.69	1032.346	1032.388	-0.042	-40.61
2	Hex2HexNAc3-AA	1071.423	-1	1071.400		0.023	38711					
3	Hex5HexNAc1-AA	1151.373	-1	1151.400	1152.407	-0.026	2804060	-22.81	1153.388	1153.415	-0.026	-22.77
4	Hex4HexNAc2-AA	1192.402	-1	1192.426	1193.434	-0.024	227178	-20.33	1194.417	1194.441	-0.024	-20.29
5	Hex3HexNAc3-AA	1233.432	-1	1233.453	1234.460	-0.021	577572	-16.96	1235.447	1235.468	-0.021	-16.93
6	Hex6HexNAc1-AA	1313.434	-1	1313.452	1314.460	-0.018	3260469	-13.91	1315.449	1315.467	-0.018	-13.89
7	Hex5HexNAc2-AA	1354.464	-1	1354.479	1355.487	-0.015	328423	-11.20	1356.479	1356.494	-0.015	-11.18
8	Hex2HexNAc3NeuAc1-AA	1362.484	-1	1362.495		-0.011	77500					
9	Hex3HexNAc3dHex1-AA	1379.493	-1	1379.511	1380.518	-0.018	177275	-12.69	1381.508	1381.526	-0.018	-12.68
10	Hex4HexNAc3-AA	1395.492	-1	1395.505	1396.513	-0.014	1333757	-9.99	1397.507	1397.521	-0.014	-9.97
11	Hex3HexNAc4-AA	1436.507	-1	1436.532	1437.540	-0.025	35393	-17.14	1438.522	1438.547	-0.025	-17.11
12	Hex7HexNAc1-AA	1475.490	-1	1475.505	1476.513	-0.015	2203705	-10.20	1477.505	1477.520	-0.015	-10.19
13	Hex4HexNAc2NeuAc1-AA	1483.502	-1	1483.522	1484.529	-0.019	310361	-12.88	1485.518	1485.537	-0.019	-12.86
14	Hex6HexNAc2-AA	1516.518	-1	1516.532	1517.540	-0.014	222804	-9.09	1518.533	1518.547	-0.014	-9.08
15	Hex5HexNAc3NeuAc1-AA	1524.537	-1	1524.548		-0.011	1.239688+e6					
16	Hex5HexNAc3-AA	1557.546	-1	1557.558	1558.566	-0.013	647926	-8.14	1559.561	1559.573	-0.013	-8.13
17	Hex5HexNAc4-AA	1598.571	-1	1598.585	1599.593	-0.014	280350	-8.77	1600.586	1600.600	-0.014	-8.76
18	Hex8HexNAc1-AA	1637.544	-1	1637.558	1638.566	-0.014	2030717	-8.36	1639.559	1639.573	-0.014	-8.35
19	Hex3HexNAc5-AA	1639.582	-1	1639.611	1640.619	-0.029	430576	-17.78	1641.597	1641.626	-0.029	-17.76
20	Hex5HexNAc2NeuAc1-AA	1645.554	-1	1645.574	1646.582	-0.020	603224	-12.19	1647.569	1647.589	-0.020	-12.18
21	Hex7HexNAc2-AA	1678.568	-1	1678.585	1679.592	-0.016	89797	-9.81	1680.583	1680.600	-0.016	-9.80
22	Hex4HexNAc3NeuAc1-AA	1686.581	-1	1686.601	1687.609	-0.020	1746952	-12.05	1688.596	1688.616	-0.020	-12.04
23	Hex6HexNAc3-AA	1719.595	-1	1719.611	1720.619	-0.017	82836	-9.63	1721.610	1721.626	-0.017	-9.62
24	Hex3HexNAc4dHex2-AA	1728.594	-1	1728.648	1729.656	-0.054	29028	-31.37	1730.609	1730.663	-0.054	-31.33
25	Hex4HexNAc4dHex1-AA	1744.611	-1	1744.643	1745.651	-0.032	27177	-18.07	1746.626	1746.658	-0.032	-18.05
26	Hex5HexNAc4-AA	1760.619	-1	1760.638	1761.645	-0.019	75940	-10.70	1762.634	1762.653	-0.019	-10.69
27	Hex3HexNAc5dHex1-AA	1785.647	-1	1785.669	1786.677	-0.022	57677	-12.39	1787.662	1787.684	-0.022	-12.38
28	Hex9HexNAc1-AA	1799.595	-1	1799.611	1800.619	-0.016	1290986	-8.99	1801.610	1801.626	-0.016	-8.98
29	Hex6HexNAc2NeuAc1-AA	1807.603	-1	1807.627	1808.635	-0.024	231164	-13.19	1809.618	1809.642	-0.024	-13.17
30	Hex5HexNAc4S1-AA	1840.608	-1	1840.594	1841.602	0.014	24360	7.59	1842.624	1842.610	0.014	7.58
31	Hex5HexNAc3NeuAc1-AA	1848.629	-1	1848.654	1849.662	-0.024	721933	-13.10	1850.645	1850.669	-0.024	-13.08
32	Hex7HexNAc3-AA	1881.645	-1	1881.664	1882.672	-0.019	55774	-9.96	1883.660	1883.679	-0.019	-9.95
33	Hex4HexNAc4NeuAc1-AA	1889.653	-1	1889.680	1890.688	-0.027	300618	-14.25	1891.668	1891.695	-0.027	-14.24
34	Hex6HexNAc5NeuAc-AA	1930.709	-1	1930.707		0.002	396985					
35	Hex6HexNAc3NeuAc1-AA	2010.678	-1	2010.707	2011.714	-0.028	55373	-13.97	2012.694	2012.722	-0.028	-13.95
36	Hex5HexNAc4NeuAc1-AA	2051.705	-1	2051.733	2052.741	-0.028	80897	-13.48	2053.721	2053.748	-0.028	-13.47
37	Hex5HexNAc4NeuGc1-AA	2067.696	-1	2067.728	2068.736	-0.032	34743	-15.25	2069.712	2069.743	-0.032	-15.24
38	Hex3HexNAc5NeuAc1dHex1-AA	2076.765	-1	2076.765		0.000	26459					
39	Hex4HexNAc5NeuAc1-AA	2092.779	-1	2092.760		0.019	16367					
40	Hex6HexNAc3NeuAc1dHex1-AA	2156.721	-1	2156.764		-0.043	11174					
41	Hex7HexNAc3NeuAc1-AA	2172.720	-1	2172.759		-0.039	14155					
42	Hex4HexNAc4NeuAc2-AA	2180.764	-1	2180.776		-0.012	13812					
43	Hex7HexNAc3NeuGc1-AA	2391.800	-1	2391.834		-0.034	9090					
44	Hex12HexNAc2-AA	2488.770	-1	2488.849		-0.079	11333					

*Different *N*-glycan compositions were also detected manually using FlexAnalysis software.

Table 8

N-glycan composition released from enzymatic deglycosylation of bovine LF with enzyme cocktail including EndoBI-1 and EndoBI-2.

Row	Composition	m/z meas.	z	m/z calc.	Mr calc.	Δ MH+ [Da]	Int.	Δ m/z [ppm]	MH+ meas.	MH+ calc.	Δ m/z [Da]	Δ MH+ [ppm]
1	Hex3HexNAc2-AA	1030.351	-1	1030.3733	1031.3811	-0.022	93467	-21.78	1032.366	1032.388	-0.022	-21.74
2	Hex2HexNAc3-AA	1071.031	-1	1071.3998		-0.369	42352					
3	Hex5HexNAc1-AA	1151.393	-1	1151.3995	1152.4074	-0.007	2417570	-6.06	1153.408	1153.415	-0.007	-6.05
4	Hex4HexNAc2-AA	1192.422	-1	1192.4261	1193.4339	-0.004	132181	-3.16	1194.437	1194.441	-0.004	-3.16
5	Hex3HexNAc3-AA	1233.453	-1	1233.4526	1234.4605	0.001	372300	0.51	1235.468	1235.468	0.001	0.50
6	Hex5HexNAc2-AA	1354.482	-1	1354.4789	1355.4867	0.003	189591	2.50	1356.497	1356.494	0.003	2.49
7	Hex2HexNAc3NeuAc1-AA	1362.480	-1	1362.4952		-0.015	50110					
8	Hex3HexNAc3dHex1-AA	1379.513	-1	1379.5105	1380.5184	0.003	134487	1.98	1381.528	1381.526	0.003	1.98
9	Hex4HexNAc3-AA	1395.512	-1	1395.5055	1396.5133	0.007	857227	4.78	1397.527	1397.521	0.007	4.77
10	Hex7HexNAc1-AA	1475.512	-1	1475.5052	1476.5130	0.006	1862927	4.37	1477.527	1477.520	0.006	4.36
11	Hex4HexNAc2NeuAc1-AA	1483.505	-1	1483.5215		-0.017	215167					
12	Hex6HexNAc2-AA	1516.537	-1	1516.5317	1517.5396	0.005	137959	3.54	1518.552	1518.547	0.005	3.53
13	Hex3HexNAc3NeuAc1-AA	1524.532	-1	1524.5481		-0.016	892563					
14	Hex4HexNAc3dHex1-AA	1541.571	-1	1541.5634	1542.5712	0.008	91530	4.99	1543.586	1543.578	0.008	4.99
15	Hex5HexNAc3-AA	1557.567	-1	1557.5583	1558.5661	0.008	360597	5.32	1559.582	1559.573	0.008	5.31
16	Hex4HexNAc4-AA	1598.595	-1	1598.5848	1599.5927	0.010	164066	6.40	1600.610	1600.600	0.010	6.39
17	Hex8HexNAc1-AA	1637.567	-1	1637.5580	1638.5658	0.009	1643489	5.58	1639.582	1639.573	0.009	5.57
18	Hex5HexNAc2NeuAc1-AA	1645.562	-1	1645.5743		-0.012	36232					
19	Hex6HexNAc2dHex1-AA	1661.559	-1	1662.5896		-0.031	86155					
20	Hex7HexNAc2-AA	1678.591	-1	1678.5846	1679.5924	0.007	59105	4.07	1680.606	1680.600	0.007	4.06
21	Hex4HexNAc3NeuAc1-AA	1686.605	-1	1686.6009	1687.6087	0.005	1250028	2.73	1688.621	1688.616	0.005	2.73
22	Hex6HexNAc3-AA	1719.621	-1	1719.6111	1720.6189	0.010	52605	5.55	1721.636	1721.626	0.010	5.54
23	Hex5HexNAc4-AA	1760.644	-1	1760.6377	1761.6455	0.007	47520	3.81	1762.659	1762.653	0.007	3.80
24	Hex3HexNAc5dHex1-AA	1785.671	-1	1785.6693	1786.6771	0.002	38850	0.90	1787.686	1787.684	0.002	0.90
25	Hex9HexNAc1-AA	1799.619	-1	1799.6108	1800.6187	0.008	850930	4.60	1801.634	1801.626	0.008	4.59
26	Hex6HexNAc2NeuAc1-AA	1807.622	-1	1807.6272		-0.005	157626					
27	Hex5HexNAc3NeuAc1-AA	1848.649	-1	1848.6537		-0.005	482185					
28	Hex7HexNAc3-AA	1881.664	-1	1881.6639		0.000	34409					
29	Hex4HexNAc4NeuAc1-AA	1889.678	-1	1889.6803		-0.002	200057					
30	Hex3HexNAc5NeuAc1-AA	1930.706	-1	1930.7068		-0.001	268994					
31	Hex6HexNAc3NeuAc1-AA	2010.709	-1	2010.7065		0.003	44257					
32	Hex5HexNAc4NeuAc1-AA	2051.736	-1	2051.7331	2052.7409	0.003	62030	1.47	2053.751	2053.748	0.003	1.47
33	Hex5HexNAc4NeuGc1-AA	2067.734	-1	2067.7280	2068.7358	0.006	29674	2.72	2069.749	2069.743	0.006	2.72
34	Hex3HexNAc5NeuAc1dHex1-AA	2076.752	-1	2076.7647		-0.013	26471					
35	Hex4HexNAc5NeuAc1-AA	2092.770	-1	2092.7596		0.010	17259					
36	Hex4HexNAc4NeuAc2-AA	2180.759	-1	2180.7757		-0.017	18322					
37	Hex12HexNAc2-AA	2488.821	-1	2488.8487		-0.028	13231					

*Different *N*-glycan compositions were also detected manually using FlexAnalysis software.
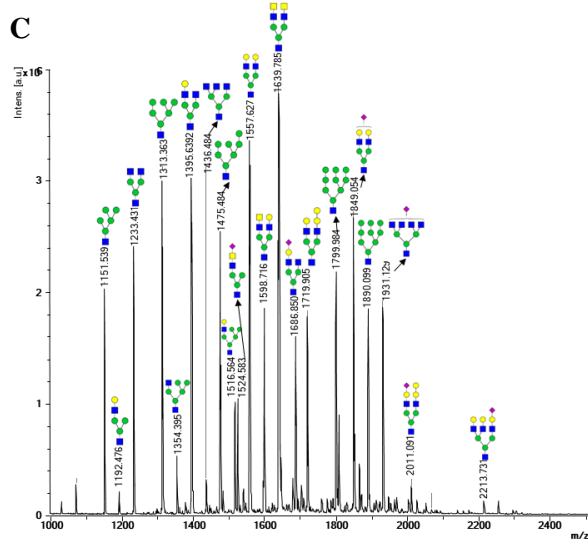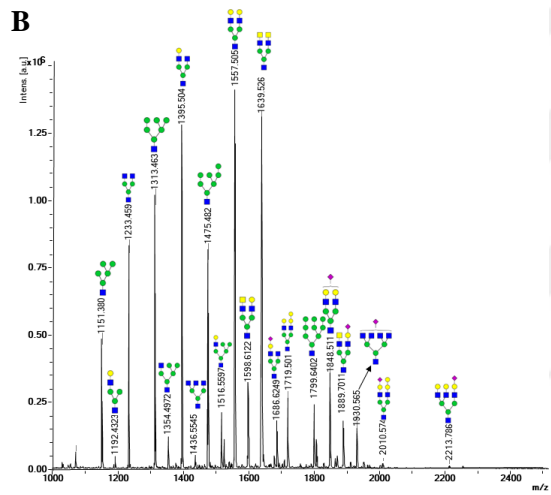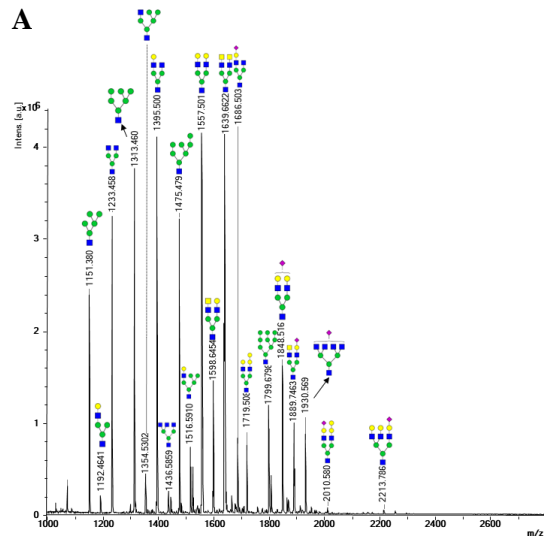
Figure 38. MALDI-MS spectrum of released *N*-glycans from enzymatic deglycosylation of bovine LPO with EndoBI-1 (A), EndoBI-2 (B), and enzyme cocktail including EndoBI-1 and EndoBI-2 (C).

Table 9

N-glycan composition released from enzymatic deglycosylation of bovine LPO with EndoBI-1.

Row	Composition	m/z meas.	z	m/z calc.	Mr calc.	Δ MH+ [Da]	Int.	Δ m/z [ppm]	MH+ meas.	MH+ calc.	Δ m/z [Da]	Δ MH+ [ppm]
1	Hex3HexNAc2-AA	1030.391	-1	1030.373	1031.381	0.018	122759	17.50	1032.406	1032.388	0.018	17.46
2	Hex2HexNac3-AA	1071.332	-1	1071.399		-0.067	298580					
3	Hex5HexNAc1-AA	1151.433	-1	1151.400	1152.407	0.033	2301461	29.00	1153.448	1153.415	0.033	28.94
4	Hex4HexNAc2-AA	1192.464	-1	1192.426	1193.434	0.038	205120	31.91	1194.479	1194.441	0.038	31.86
5	Hex3HexNAc3-AA	1233.495	-1	1233.453	1234.460	0.042	3273287	34.09	1235.510	1235.468	0.042	34.04
6	Hex6HexNAc1-AA	1313.500	-1	1313.452	1314.460	0.047	3847497	36.02	1315.515	1315.467	0.047	35.97
7	Hex5HexNAc2-AA	1354.530	-1	1354.479	1355.487	0.051	465864	37.85	1356.545	1356.494	0.051	37.79
8	Hex2HexNAc3NeuAc1-AA	1362.481	-1	1362.495		-0.014	12150					
9	Hex4HexNAc3-AA	1395.556	-1	1395.505	1396.513	0.050	4238535	36.01	1397.571	1397.521	0.050	35.96
10	Hex7HexNAc4-AA	1436.586	-1	1436.532	1437.540	0.054	279661	37.49	1438.601	1438.547	0.054	37.44
11	Hex7HexNAc1-AA	1475.560	-1	1475.505	1476.513	0.054	3216585	36.93	1477.575	1477.520	0.054	36.88
12	Hex4HexNAc2NeuAc1-AA	1483.573	-1	1483.522	1484.529	0.052	118241	34.96	1485.588	1485.537	0.052	34.91
13	Hex6HexNAc2-AA	1516.591	-1	1516.532	1517.540	0.059	664012	39.05	1518.606	1518.547	0.059	39.00
14	Hex3HexNAc3NeuAc1-AA	1524.505	-1	1524.548		-0.043	483236					
15	Hex5HexNAc3-AA	1557.611	-1	1557.558	1558.566	0.053	4190384	34.15	1559.627	1559.573	0.053	34.10
16	Hex4HexNAc4-AA	1598.645	-1	1598.585	1599.593	0.061	1395348	37.88	1600.660	1600.600	0.061	37.83
17	Hex7HexNAc1dHex1-AA	1621.523	-1	1621.563		-0.040	15915					
18	Hex8HexNAc1-AA	1637.622	-1	1637.558	1638.566	0.064	1917442	39.25	1639.637	1639.573	0.064	39.20
19	Hex3HexNAc5-AA	1639.662	-1	1639.611	1640.619	0.051	3671640	30.98	1641.677	1641.626	0.051	30.94
20	Hex5HexNAc2NeuAc1-AA	1645.622	-1	1645.574	1646.582	0.048	209946	29.04	1647.637	1647.589	0.048	29.01
21	Hex1HexNAc6dHex1-AA	1664.784	-1	1664.643		0.141	197483					
22	Hex7HexNAc2-AA	1678.650	-1	1678.585	1679.592	0.065	123776	38.91	1680.665	1680.600	0.065	38.86
23	Hex4HexNAc3NeuAc1-AA	1686.662	-1	1686.601	1687.609	0.061	758855	36.45	1688.677	1688.616	0.061	36.40
24	Hex6HexNAc3-AA	1719.678	-1	1719.611	1720.619	0.067	684515	39.06	1721.693	1721.626	0.067	39.02
25	Hex5HexNAc5dHex1-AA	1785.728	-1	1785.669	1786.677	0.058	51971	32.64	1787.743	1787.684	0.058	32.60
26	Hex9HexNAc1-AA	1799.680	-1	1799.611	1800.619	0.069	1046879	38.35	1801.695	1801.626	0.069	38.30
27	Hex6HexNAc2NeuAc1-AA	1807.691	-1	1807.627	1808.635	0.064	370995	35.42	1809.706	1809.642	0.064	35.38
28	Hex5HexNAc3NeuAc1-AA	1848.717	-1	1848.654	1849.662	0.063	1441173	34.02	1850.732	1850.669	0.063	33.98
29	Hex4HexNAc4NeuAc1-AA	1889.746	-1	1889.680	1890.688	0.066	867701	34.92	1891.761	1891.695	0.066	34.89
30	Hex3HexNAc5NeuAc1-AA	1930.569	-1	1930.707		-0.138	373334					
31	Hex5HexNAc5-AA	1963.782	-1	1963.717	1964.725	0.065	43672	33.07	1965.797	1965.732	0.065	33.03
32	Hex6HexNAc3Neu1-AA	1968.744	-1	1968.696	1969.704	0.048	44648	24.41	1970.759	1970.711	0.048	24.38
33	Hex9HexNAc2-AA	2002.512	-1	2002.690		-0.178	31295					
34	Hex6HexNAc3NeuAc1-AA	2010.775	-1	2010.707	2011.714	0.068	78271	33.92	2012.790	2012.722	0.068	33.88
35	Hex5HexNAc4NeuAc1-AA	2051.640	-1	2051.733		-0.093	30274					
36	Hex5HexNAc3NeuAc2-AA	2139.670	-1	2139.749		-0.079	33006					
37	Hex6HexNAc3NeuAc1dHex1-AA	2156.655	-1	2156.764		-0.109	35993					
38	Hex6HexNAc4NeuAc1-AA	2213.870	-1	2213.786	2214.794	0.084	52737	38.01	2215.885	2215.801	0.084	37.97
39	Hex5HexNAc5NeuAc1-AA	2254.896	-1	2254.812	2255.820	0.084	47056	37.14	2256.911	2256.828	0.084	37.11

*Different *N*-glycan compositions were also detected manually using FlexAnalysis software.

Table 10

N-glycan composition released from enzymatic deglycosylation of bovine LPO with EndoBI-2.

Row	Composition	m/z meas.	z	m/z calc.	Mr calc.	Δ MH+ [Da]	Int.	Δ m/z [ppm]	MH+ meas.	MH+ calc.	Δ m/z [Da]	Δ MH+ [ppm]
1	Hex3HexNAc2-AA	1030.358	-1	1030.373	1031.381	-0.015	26643	-14.49	1032.373	1032.388	-0.015	-14.46
2	Hex2HexNac3-AA	1071.329	-1	1071.399		-0.070	61975					
3	Hex5HexNAc1-AA	1151.403	-1	1151.400	1152.407	0.003	455774	2.57	1153.418	1153.415	0.003	2.57
4	Hex4HexNAc2-AA	1192.432	-1	1192.426	1193.434	0.006	48232	5.22	1194.447	1194.441	0.006	5.22
5	Hex3HexNAc3-AA	1233.465	-1	1233.453	1234.460	0.012	797349	9.67	1235.480	1235.468	0.012	9.66
6	Hex6HexNAc1-AA	1313.470	-1	1313.452	1314.460	0.018	973943	13.45	1315.485	1315.467	0.018	13.43
7	Hex5HexNAc2-AA	1354.497	-1	1354.479	1355.487	0.018	118799	13.53	1356.512	1356.494	0.018	13.51
8	Hex4HexNAc3-AA	1395.528	-1	1395.505	1396.513	0.023	1303594	16.17	1397.543	1397.521	0.023	16.14
9	Hex3HexNAc4-AA	1436.554	-1	1436.532	1437.540	0.022	57695	15.64	1438.570	1438.547	0.022	15.61
10	Hex7HexNAc1-AA	1475.531	-1	1475.505	1476.513	0.025	771492	17.17	1477.545	1477.520	0.025	17.14
11	Hex4HexNAc2NeuAc1-AA	1483.540	-1	1483.522	1484.529	0.018	26574	12.26	1485.555	1485.537	0.018	12.24
12	Hex6HexNAc2-AA	1516.560	-1	1516.532	1517.540	0.028	183192	18.47	1518.575	1518.547	0.028	18.44
13	Hex3HexNAc3NeuAc1-AA	1524.507	-1	1524.548		-0.041	106386					
14	Hex5HexNAc3-AA	1557.585	-1	1557.558	1558.566	0.027	1445574	17.02	1559.600	1559.573	0.027	16.99
15	Hex4HexNAc4-AA	1598.612	-1	1598.585	1599.593	0.027	302936	17.12	1600.627	1600.600	0.027	17.10
16	Hex8HexNAc1-AA	1637.589	-1	1637.558	1638.566	0.031	408669	19.13	1639.604	1639.573	0.031	19.11
17	Hex3HexNAc5-AA	1639.637	-1	1639.611	1640.619	0.026	1216366	15.65	1641.652	1641.626	0.026	15.63
18	Hex5HexNAc2NeuAc1-AA	1645.592	-1	1645.574	1646.582	0.018	49551	10.90	1647.607	1647.589	0.018	10.89
19	Hex1HexNAc6dHex1-AA	1664.784	-1	1664.643		0.141	18996					
20	Hex7HexNAc2-AA	1678.614	-1	1678.585	1679.592	0.030	44268	17.60	1680.629	1680.600	0.030	17.58
21	Hex4HexNAc3NeuAc1-AA	1686.625	-1	1686.601	1687.609	0.024	164017	14.26	1688.640	1688.616	0.024	14.24
22	Hex6HexNAc3-AA	1719.639	-1	1719.611	1720.619	0.028	246351	16.36	1721.654	1721.626	0.028	16.34
23	Hex4HexNAc2NeuAc1dHex1-AA	1791.484	-1	1791.632		-0.148	22130					
24	Hex9HexNAc1-AA	1799.640	-1	1799.611	1800.619	0.029	217655	16.32	1801.655	1801.626	0.029	16.31
25	Hex6HexNAc2NeuAc1-AA	1807.650	-1	1807.627	1808.635	0.022	95154	12.39	1809.665	1809.642	0.022	12.37
26	Hex5HexNAc3NeuAc1-AA	1848.675	-1	1848.654	1849.662	0.022	322101	11.63	1850.690	1850.669	0.022	11.62
27	Hex4HexNAc4NeuAc1-AA	1889.701	-1	1889.680	1890.688	0.021	164372	11.03	1891.716	1891.695	0.021	11.01
28	Hex9HexNAc4-AA	1922.504	-1	1922.690		-0.186	16046					
29	Hex3HexNAc5NeuAc1-AA	1930.565	-1	1930.707		-0.142	146615					
30	Hex5HexNAc5-AA	1963.737	-1	1963.717	1964.725	0.020	15449	10.25	1965.752	1965.732	0.020	10.24
31	Hex7HexNAc2NeuAc1-AA	1969.543	-1	1969.680		-0.164	19587					
32	Hex9HexNAc2-AA	2002.689	-1	2002.690	2003.698	-0.002	15990	-0.79	2004.704	2004.705	-0.002	-0.79
33	Hex6HexNAc3NeuAc1-AA	2010.723	-1	2010.707	2011.714	0.017	24253	8.35	2012.738	2012.722	0.017	8.34
34	Hex6HexNAc4NeuAc1-AA	2213.800	-1	2213.786	2214.794	0.014	15321	6.36	2215.815	2215.801	0.014	6.36
35	Hex5HexNAc5NeuAc1-AA	2254.832	-1	2254.812	2255.820	0.020	13638	8.66	2256.847	2256.828	0.020	8.65

*Different *N*-glycan compositions were also detected manually using FlexAnalysis software.

Table 11

N-glycan composition released from enzymatic deglycosylation of bovine LPO with enzyme cocktail including EndoBI-1 and EndoBI-2.

Row	Composition	m/z meas.	z	m/z calc.	Mr calc.	Δ MH+ [Da]	Int.	Δ m/z [ppm]	MH+ meas.	MH+ calc.	Δ m/z [Da]	Δ MH+ [ppm]
1	Hex3HexNAc2-AA	1030.447	-1	1030.373	1031.381	0.073	121874	71.26	1032.462	1032.388	0.073	71.12
2	Hex2HexNac3-AA	1071.807	-1	1071.399		0.408	268392					
3	Hex5HexNAc1-AA	1151.495	-1	1151.400	1152.407	0.095	2057701	82.76	1153.510	1153.415	0.095	82.61
4	Hex4HexNAc2-AA	1192.532	-1	1192.426	1193.434	0.106	210072	89.22	1194.548	1194.441	0.106	89.07
5	Hex3HexNAc3-AA	1233.562	-1	1233.453	1234.460	0.109	2492270	88.35	1235.577	1235.468	0.109	88.20
6	Hex6HexNAc1-AA	1313.573	-1	1313.452	1314.460	0.120	3084501	91.57	1315.588	1315.467	0.120	91.43
7	Hex5HexNAc2-AA	1354.395	-1	1354.479		-0.084	515333					
8	Hex2HexNAc3NeuAc1-AA	1362.405	-1	1362.495		-0.090	71478					
9	Hex3HexNAc3dHex1-AA	1379.429	-1	1379.516		-0.087	99311					
10	Hex4HexNAc3-AA	1395.639	-1	1395.505	1396.513	0.134	3052355	95.82	1397.654	1397.521	0.134	95.69
11	Hex3HexNAc4-AA	1436.484	-1	1436.546		-0.062	317992					
12	Hex7HexNAc1-AA	1475.484	-1	1475.505		-0.021	2.666547+e6					
13	Hex4HexNAc2NeuAc1-AA	1483.487	-1	1483.522		-0.035	198245					
14	Hex6HexNAc2-AA	1516.568	-1	1516.532		0.036	953101					
15	Hex3HexNAc3NeuAc1-AA	1524.583	-1	1524.548		0.035	983182					
16	Hex5HexNAc3-AA	1557.627	-1	1557.558		0.069	3.410327+e6					
17	Hex4HexNAc4-AA	1598.716	-1	1598.585		0.131	1.726066+e6					
18	Hex7HexNAc1dHex1-AA	1621.523	-1	1621.563		-0.040	86625					
19	Hex8HexNAc1-AA	1637.749	-1	1637.558		0.191	2.133517+e6					
20	Hex3HexNAc5-AA	1639.785	-1	1639.625		0.160	3.09775+e6					
21	Hex5HexNAc2NeuAc1-AA	1645.757	-1	1645.574		0.183	358858					
22	Hex7HexNAc2-AA	1678.834	-1	1678.585		0.249	294148					
23	Hex4HexNAc3NeuAc1-AA	1686.850	-1	1686.611		0.239	1.437837+e6					
24	Hex5HexNAc3dHex1-AA	1702.875	-1	1703.634		-0.759	265588					
25	Hex6HexNAc3-AA	1719.905	-1	1719.640		0.265	1.610896+e6					
26	Hex3HexNAc5dHex1-AA	1786.031	-1	1785.669		0.362	121002					
27	Hex9HexNAc1-AA	1799.984	-1	1799.611		0.373	1.971367+e6					
28	Hex6HexNAc2NeuAc1-AA	1808.006	-1	1807.627		0.379	751263					
29	Hex5HexNAc3Neu1-AA	1824.021	-1	1824.654		-0.633	70890					
30	Hex5HexNAc3NeuAc1-AA	1849.054	-1	1848.654		0.400	2.455671+e6					
31	Hex6HexNAc3dHex1-AA	1865.062	-1	1865.669		-0.607	405147					
32	Hex4HexNAc4NeuAc1-AA	1890.099	-1	1889.680		0.419	1.667919+e6					
33	Hex5HexNAc4dHex1-AA	1906.099	-1	1906.696		-0.597	101194					
34	Hex6HexNAc4	1922.088	-1	1922.690		-0.602	95874					
35	Hex3HexNAc5NeuAc1-AA	1931.129	-1	1930.707		0.422	1.718287+e6					
36	Hex4HexNAc5dHex1-AA	1947.123	-1	1947.722		-0.599	153645					
37	Hex5HexNAc5-AA	1964.136	-1	1963.717		0.419	115125					
38	Hex7HexNAc2NeuAc1-AA	1970.099	-1	1969.680		0.419	147117					
39	Hex9HexNAc2-AA	2003.067	-1	2002.690		0.377	116460					
40	Hex6HexNAc3Neu1-AA	2011.091	-1	2010.707		0.385	246651					
41	Hex7HexNAc3dHex1-AA	2027.070	-1	2027.722		-0.652	117207					
42	Hex5HexNAc4NeuAc1-AA	2052.069	-1	2051.733		0.336	98705					
43	Hex6HexNAc4dHex1-AA	2068.039	-1	2068.748		-0.709	49805					
44	Hex4HexNAc5NeuAc1-AA	2093.027	-1	2092.760		0.267	43213					
45	Hex7HexNAc3NeuAc1-AA	2171.792	-1	2172.759		-0.967	42986					
46	Hex5HexNAc4Neuac1dHex1-AA	2197.766	-1	2197.791		-0.025	24709					
47	Hex6HexNAc4Neu1-AA	2213.727	-1	2213.786		-0.059	134474					
48	Hex5HexNAc5NeuAc1-AA	2254.594	-1	2254.812		-0.218	124432					

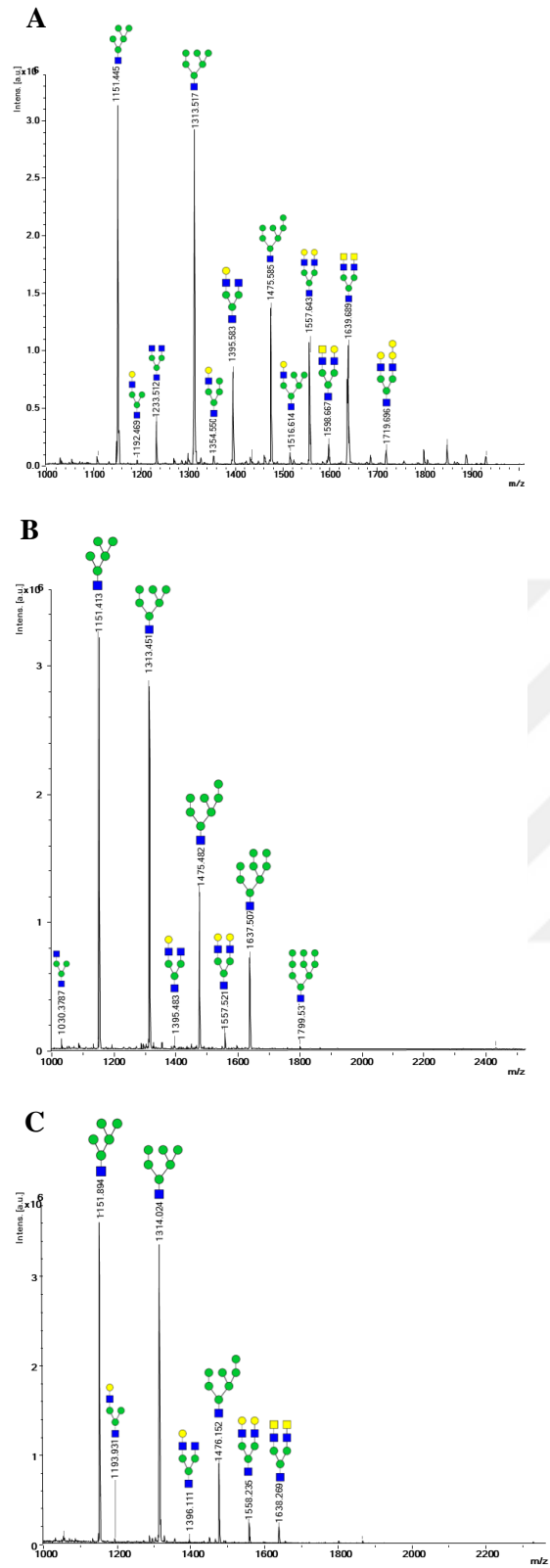*Different *N*-glycan compositions were also detected manually using FlexAnalysis software.

Figure 39. MALDI-MS spectrum of released *N*-glycans obtained from enzymatic deglycosylation of RNase B with EndoBI-1 (A), EndoBI-2 (B), and enzyme cocktail including EndoBI-1 and EndoBI-2 (C).

Table 12

N-glycan composition released from enzymatic deglycosylation of RNase B with EndoBI-1.

Row	Composition	m/z meas.	z	m/z calc.	Mr calc.	Δ MH+ [Da]	Int.	Δ m/z [ppm]	MH+ meas.	MH+ calc.	Δ m/z [Da]	Δ MH+ [ppm]
1	Hex3HexNAc2-AA	1030.381	-1	1030.373	1031.381	0.008	58537	7.48	1032.396	1032.388	0.008	7.46
2	Hex5HexNAc1-AA	1151.425	-1	1151.400	1152.407	0.025	2495784	21.86	1153.440	1153.415	0.025	21.82
3	Hex4HexNAc2-AA	1192.447	-1	1192.426	1193.434	0.020	37338	17.19	1194.462	1194.441	0.020	17.16
4	Hex3HexNAc3-AA	1233.487	-1	1233.453	1234.460	0.034	287762	27.69	1235.502	1235.468	0.034	27.64
5	Hex6HexNAc1-AA	1313.491	-1	1313.452	1314.460	0.039	2257060	29.74	1315.507	1315.467	0.039	29.69
6	Hex5HexNAc2-AA	1354.520	-1	1354.479	1355.487	0.041	68872	30.10	1356.535	1356.494	0.041	30.06
7	Hex4HexNAc3-AA	1395.550	-1	1395.505	1396.513	0.044	543405	31.83	1397.565	1397.521	0.044	31.79
8	Hex7HexNAc1-AA	1475.552	-1	1475.505	1476.513	0.047	916802	31.76	1477.567	1477.520	0.047	31.72
9	Hex6HexNAc2-AA	1516.579	-1	1516.532	1517.540	0.048	78658	31.46	1518.595	1518.547	0.048	31.42
10	Hex5HexNAc3-AA	1557.608	-1	1557.558	1558.566	0.049	703181	31.61	1559.623	1559.573	0.049	31.57
11	Hex4HexNAc4-AA	1598.630	-1	1598.585	1599.593	0.045	114229	28.23	1600.645	1600.600	0.045	28.19
12	Hex8HexNAc1-AA	1637.608	-1	1637.558	1638.566	0.050	466659	30.39	1639.623	1639.573	0.050	30.36
13	Hex3HexNAc5-AA	1639.656	-1	1639.611	1640.619	0.044	535556	26.93	1641.671	1641.626	0.044	26.90
15	Hex7HexNAc2-AA	1678.632	-1	1678.585	1679.592	0.047	23868	28.12	1680.647	1680.600	0.047	28.09
17	Hex6HexNAc3-AA	1719.661	-1	1719.611	1720.619	0.050	92244	28.98	1721.676	1721.626	0.050	28.95
18	Hex9HexNAc1-AA	1799.659	-1	1799.611	1800.619	0.049	93449	26.96	1801.674	1801.626	0.049	26.93
19	Hex9HexNAc2-AA	2002.741	-1	2002.690		0.034	17019					
20	Hex11HexNAc2-AA	2325.866	-1	2326.796		-0.930	11014					
21	Hex12HexNAc2-AA	2488.974	-1	2488.849		0.125	11897					
22	Hex13HexNAc2-AA	2651.114	-1	2650.902		0.213	9311					

*Different *N*-glycan compositions were also detected manually using FlexAnalysis software.

Table 13

N-glycan composition released from enzymatic deglycosylation of RNase B with EndoBI-2.

Row	Composition	m/z meas.	z	m/z calc.	Mr calc.	Δ MH+ [Da]	Int.	Δ m/z [ppm]	MH+ meas.	MH+ calc.	Δ m/z [Da]	Δ MH+ [ppm]
1	Hex3HexNAc2-AA	1030.379	-1	1030.373	1031.381	0.005	87881	5.28	1032.394	1032.388	0.005	5.27
2	Hex2HexNac3-AA	1070.955	-1	1071.400		-0.445	31616					
3	Hex5HexNAc1-AA	1151.422	-1	1151.400	1152.407	0.023	3438079	19.72	1153.437	1153.415	0.023	19.69
4	Hex4HexNAc2-AA	1192.447	-1	1192.426	1193.434	0.021	43410	17.82	1194.462	1194.441	0.021	17.79
5	Hex6HexNAc1-AA	1313.488	-1	1313.452	1314.460	0.036	2990258	27.36	1315.503	1315.467	0.036	27.32
6	Hex5HexNAc2-AA	1354.512	-1	1354.479	1355.487	0.033	70201	24.67	1356.527	1356.494	0.033	24.64
7	Hex4HexNAc3-AA	1395.534	-1	1395.505	1396.513	0.029	31595	20.55	1397.549	1397.521	0.029	20.52
8	Hex7HexNAc1-AA	1475.547	-1	1475.505	1476.513	0.042	1183955	28.59	1477.562	1477.520	0.042	28.55
9	Hex6HexNAc2-AA	1516.567	-1	1516.532	1517.540	0.036	29964	23.56	1518.583	1518.547	0.036	23.53
10	Hex5HexNAc3-AA	1557.602	-1	1557.558	1558.566	0.044	127743	28.03	1559.617	1559.573	0.044	27.99
11	Hex8HexNAc1-AA	1637.603	-1	1637.558	1638.566	0.045	662580	27.35	1639.618	1639.573	0.045	27.31
12	Hex9HexNAc1-AA	1799.657	-1	1799.611	1800.619	0.046	32349	25.62	1801.672	1801.626	0.046	25.59
13	Hex11HexNAc2-AA	2326.595	-1	2326.796		-0.201	15368					
14	Hex12HexNAc2-AA	2488.611	-1	2488.849		-0.238	15989					
15	Hex13HexNAc2-AA	2649.951	-1	2650.902		-0.951	11858					

*Different *N*-glycan compositions were also detected manually using FlexAnalysis software.

Table 14

N-glycan composition released from enzymatic deglycosylation of RNase B with enzyme cocktail including EndoBI-1 and EndoBI-2.

Row	Composition	m/z meas.	z	m/z calc.	Mr calc.	Δ MH+ [Da]	Int.	Δ m/z [ppm]	MH+ meas.	MH+ calc.	Δ m/z [Da]	Δ MH+ [ppm]
1	Hex2HexNAc3-AA	1071.08	-1	1071.400		-0.3248	52134					
2	Hex5HexNAc1-AA	1151.45	-1	1151.400	1152.407	0.055	3672355	47.75	1153.470	1153.415	0.055	47.66
3	Hex6HexNAc1-AA	1313.52	-1	1313.452	1314.460	0.070	3442115	53.55	1315.538	1315.467	0.070	53.47
4	Hex4HexNAc3-AA	1395.58	-1	1395.505	1396.513	0.072	51640	51.69	1397.593	1397.521	0.072	51.61
5	Hex7HexNAc1-AA	1475.59	-1	1475.505	1476.513	0.083	842452	56.33	1477.603	1477.520	0.083	56.25
6	Hex5HexNAc3-AA	1557.64	-1	1557.558	1558.566	0.081	224070	52.20	1559.655	1559.573	0.081	52.13
7	Hex8HexNAc1-AA	1637.64	-1	1637.558	1638.566	0.085	177202	52.03	1639.658	1639.573	0.085	51.96
8	Hex6HexNAc3-AA	1719.70	-1	1719.611		0.085	23512					
9	Hex9HexNAc1-AA	1799.69	-1	1799.611	1800.619	0.080	36785	44.30	1801.706	1801.626	0.080	44.25

*Different *N*-glycan compositions were also detected manually using FlexAnalysis software.

CHAPTER 5
RESULTS AND RECOMMENDATIONS

5.1. Structural Characterization Studies

To date, the 3D structures of numerous ENGases from different species have been investigated in the literature. However, unique and very effective enzymes EndoBI-1 and EndoBI-2 are not structurally described in databases. Comparative modeling and bioinformatic analysis could yield priceless structural data on these two enzymes. To understand the unique features and structure-function connection of the EndoBI-2, as well as its homology EndoBI-1 enzyme with high similarity, were analyzed structurally. With this direction, the EndoBI-1 and EndoBI-2 enzymes' full and partial comparative models were created, compared, and assessed. Confirmed models belonging to the EndoBI-1 and EndoBI-2 enzymes were used for Fuc-GlcNAc docking analysis for the elucidation of the substrate specificity of the enzymes.

A unique and thermostable ENGase with a wide range of selectivity is the EndoBI-2 enzyme. These characteristics make it an excellent tool for glycobiology research. Furthermore, this particular enzyme is created by a bacterium that is safe for food, opening up a variety of opportunities for the food sector, particularly the market for infant products. Despite receiving a lot of attention over the last few years, these enzymes have not yet been the subject of structural analysis. This study provided valuable information on the structure/function relationship between these enzymes.

In future research, other methods may be used to generate EndoBI-1 and EndoBI-2 enzymes with various substrate-binding affinities. These enzymes' plus subsite CBMs could be changed to boost substrate binding affinity. Furthermore, EndoBI enzyme active sites can be modified to yield catalytically inactive proteins that can be employed as selective glycan binding tools for purification. Despite the fact that this research sheds light on the structure/function link and plus subsite substrate identification of EndoBI enzymes, high-quality crystal structures are still required, particularly for minus subsite detection and thermostability research.

5.2. Molecular Cloning, Protein Expression and Purification Studies

In this thesis, the EndoBI-2 enzyme was cloned from the gut microbiome associated with *B. infantis* using an *in-vivo* cloning system without any limitations such as restriction enzyme requirement. Desired gene region (BLIF_1310) amplified with specific primers successfully cloned with the N-His Sumo tag. Based on the conserved domain analysis obtained as a result of the structural analysis of EndoBI-2, this enzyme, which was cloned without the transmembrane and signal peptide region, was thus produced by expressing it with high efficiency in protein expression and production study.

The enzyme activity studies of the EndoBI-2 produced in high yield and purity were verified on the RNase B model glycoprotein, which contains solely high mannose *N*-glycans.

5.3. Enzyme Kinetic Characterization and Optimization Studies of EndoBI-2

According to the results of the optimization tests, the temperature has a significant impact on EndoBI-2's ability to activate RNase B. The enzyme's activity was demonstrated to be slightly affected by pH but to positively correlate with rising temperature. The results suggest that the enzyme can maintain its high activity on a pH 5 on the RNase B model glycoprotein.

EndoBI-1 and EndoBI-2 were tested on various plant and milk-based substrates to understand the specificities of both enzymes. Their synergy was also investigated. Based on the results, it was shown that EndoBI-1 had more activities on milk proteins, whereas EndoBI-2 had more glycans from plant proteins. Their combined application yielded more *N*-glycans than the individual enzyme's activity.

5.4. Enzymatic Deglycosytion of Different Glycoproteins Using EndoBI-2 Enzyme and Characterization of Released *N*-glycans by Advanced Mass Spectrometry

According to an MS analysis, all EndoBI-2, as opposed to EndoBI-1, and an enzyme cocktail that includes both endoglycosidases can release the same types of glycans but to varying intensity. Based on the detected *N*-glycan compositions by MALDI-MS analysis, 39 different *N*-glycan compositions were detected in enzymatic deglycosylation of bovine LF protein using EndoBI-1, whereas 44 *N*-glycan compositions were confirmed deglycosylation of bovine LF using EndoBI-2. However, 37 *N*-glycan compositions were detected as a result of deglycosylation with an enzyme cocktail that contains EndoBI-1 and EndoBI-2 enzymes. The relative abundance between sialylated and neutral *N*-glycans of both enzymes on bovine LF was similar. Given that sialic acid is able to protect newborns against infections and is a crucial nutrient for brain development and cognition, an optimal release of sialylated *N*-glycans may be desired for gut health. Similar to bovine LF results; 39, 35, and 48 different *N*-glycan compositions were detected based on the results of enzymatic deglycosylation of bovine milk LPO with EndoBI-1, EndoBI-2, and enzyme cocktail, respectively. Additionally, *N*-glycan compositions of RNase B model glycoprotein using the enzymatic deglycosylation of each enzyme and the cocktail. Hence, 22, 15, and 9 different *N*-glycan compositions were detected based on the results of enzymatic deglycosylation of EndoBI-1, EndoBI-2, and enzyme cocktail, respectively.

Furthermore, when the *N*-glycan profiles obtained from the deglycosylation of each glycoprotein using each enzyme were analyzed, it was seen that the glycan structures of the RNaseB protein were composed of structures belonging to the HM glycan type. When the *N*-glycans released from the LPO protein, an important milk protein, were examined, it was concluded that this glycoprotein had a glycan profile in a different spectrum. In this profile structure, hybrid and complex glycan structures, especially HM glycan type, were also found in the glycan profile of this model protein. Similarly, when the glycan spectra of the LF model protein were examined, it was concluded that a profile consisting of a wide variety of compositions including all glycan types was obtained.

In conclusion that deglycosylation is an important way to obtain bioactive ingredients. Due to the structural similarity to the HMOs, the release of *N*-glycans is a promising tool for potential use for products especially infant formula to promote the colonization of infant gut microbiome with beneficial organisms like *B. infantis*. and also,

the use of novel enzymes EndoBI-2 in the production of *N*-glycans that can be used in for example infant formula has great potential to increase nutritional values. So, with this thesis, a novel and very effective enzyme EndoBI-2 was recombinantly produced and characterized extensively. The structural analysis of the released glycans and the kinetic characteristics indicate that EndoBI-2 can enable the wide-scale release of complex, and potential bioactive glycans from a range of glycoprotein. As a unique enzyme EndoBI-2 can be choosed for these objectives.

REFERENCE

Ballard, O., & Morrow, A. L. (2013). "Human milk composition: nutrients and bioactive factors". *Pediatric Clinics of North America*, *60*(1), 49–74. https://doi.org/ 10.1016/j.pcl.2012.10.002

Bienert, S., Waterhouse, A., Beer, T. A. P. De, Tauriello, G., Studer, G., Bordoli, L., & Schwede, T. (2017). *The SWISS-MODEL Repository — new features and functionality*. *45*(November 2016), 313–319. https://doi.org/10.1093/nar/gkw1132

Carlson, D. M. (1968). "Structures and immunochemical properties of oligosaccharides isolated from pig submaxillary mucins". *Journal of Biological Chemistry*, *243*(3), 616–626. https://doi.org/ 10.1016/S0021-9258(18)93649-5

Cheng, L., Akkerman, R., Kong, C., Walvoort, M. T. C., & de Vos, P. (2021). More than sugar in the milk: human milk oligosaccharides as essential bioactive molecules in breast milk and current insight in beneficial effects. In Critical Reviews in Food Science and Nutrition. https://doi.org/10.1080/10408398.2020.1754756

Duman, H., Kaplan, M., Arslan, A., Sahutoglu, A. S., Kayili, H. M., Frese, S. A., & Karav, S. (2021). "Potential Applications of Endo-β-N-Acetylglucosaminidases From Bifidobacterium longum Subspecies infantis in Designing Value-Added, Next-Generation Infant Formulas". *Frontiers in Nutrition*, *8*(April), 1–12. https://doi.org/10.3389/fnut.2021.646275

Dwek, R. A., Edge, C. J., Harvey, D. J., Wormald, M. R., & Parekh, R. B. (1993). "Analysis of Glycoprotein-Associated Oligosaccharides". *Annual Review of Biochemistry*, 62(1), 65-100. https://doi.org/10.1146/annurev.bi.62.070193.000433

Edge, A. S. B., Faltynek, C. R., Hof, L., Reichert, L. E., & Weber, P. (1981). "Deglycosylation of glycoproteins by trifluoromethanesulfonic acid". *Analytical Biochemistry*, *118*(1), 131–137. https://doi.org/ 10.1016/0003-2697(81)90168-8

Frese, S. A., Hutton, A. A., Contreras, L. N., Shaw, C. A., Palumbo, M. C., Casaburi, G., Xu, G., Davis, J. C. C., Lebrilla, C. B., Henrick, B. M., Freeman, S. L., Barile, D., German, J. B., Mills, D. A., Smilowitz, J. T., & Underwood, M. A. (2017). "Persistence of Supplemented Bifidobacterium longum subsp. infantis EVC001 in Breastfed Infants". *MSphere*. https://doi.org/10.1128/msphere.00501-17

Garrido, D, Nwosu, C., Ruiz-Moyano, S., Aldredge, D., German, J. B., Lebrilla, C. B., & Mills, D. A. (2012). "Endo-beta-N-acetylglucosaminidases from infant gut-associated

bifidobacteria release complex N-glycans from human milk glycoproteins". *Molecular & Cellular Proteomics : MCP*, *11*(9), 775–785. https://doi.org/10.1074/mcp.M112.018119

Hamosh, M. (2001). "Bioactive factors in human milk". *Pediatric Clinics of North America*, *48*(1), 69–86. https://doi.org/ 10.1016/S0031-3955(05)70286-8

Humphrey, W., Dalke, A., & Schulten, K. (1996). "VMD: Visual molecular dynamics". *Journal of Molecular Graphics*, *14*(1), 33–38. https://doi.org/10.1016/0263-7855(96)00018-5

Karav, S, Parc, A. L., Moura Bell, J. M., Rouquie, C., Mills, D. A., Barile, D., & Block, D. E. (2015). "Kinetic characterization of a novel endo-beta-N-acetylglucosaminidase on concentrated bovine colostrum whey to release bioactive glycans". *Enzyme and Microbial Technology*, *77*, 46–53. https://doi.org/10.1016/j.enzmictec.2015.05.007

Karav, Sercan. (2019). "Application of a Novel Endo-β-N-Acetylglucosaminidase to Isolate an Entirely New Class of Bioactive Compounds: N-Glycans". In *Enzymes in Food Biotechnology* (pp. 389–404). Elsevier. https://doi.org/10.1016/B978-0-12-813280-7.00022-0

Karav, Sercan, Casaburi, G., Arslan, A., Kaplan, M., Sucu, B., & Frese, S. (2019). "N-glycans from human milk glycoproteins are selectively released by an infant gut symbiont in vivo". *Journal of Functional Foods*, *61*, 103485. https://doi.org/10.1016/j.jff.2019.103485

Karav, Sercan, Le Parc, A., Leite Nobrega de Moura Bell, J. M., Frese, S. A., Kirmiz, N., Block, D. E., Barile, D., & Mills, D. A. (2016). "Oligosaccharides released from milk glycoproteins are selective growth substrates for infant-associated bifidobacteria". *Applied and Environmental Microbiology*, 82(12), 3622-3630. https://doi.org/10.1128/AEM.00547-16

Lafite, P., & Daniellou, R. (2012). "Rare and unusual glycosylation of peptides and proteins". In *Natural Product Reports*, 29(7), 729. https://doi.org/10.1039/c2np20030a

LoCascio, R. G., Niñonuevo, M. R., Kronewitter, S. R., Freeman, S. L., German, J. B., Lebrilla, C. B., & Mills, D. A. (2009). "A versatile and scalable strategy for glycoprofiling bifidobacterial consumption of human milk oligosaccharides". *Microbial Biotechnology*, 2(3), 333-342. https://doi.org/10.1111/j.1751-7915.2008.00072.x

Mussatto, S. I., & Mancilha, I. M. (2007). "Non-digestible oligosaccharides: A review". In

Carbohydrate Polymers, 68(3), 587-597. https://doi.org/10.1016/j.carbpol.2006.12.011

Pekdemir, B., Duman, H., Arslan, A., Kaplan, M., Karyelioğlu, M., Özer, T., Kayılı, H. M., Salih, B., Henrick, B. M., Duar, R. M., & Karav, S. (2022). "Immobilization of a Bifidobacterial Endo-ß-N-Acetylglucosaminidase to Generate Bioactive Compounds for Food Industry". *Frontiers in Bioengineering and Biotechnology*, *10*(July), 1–10. https://doi.org/10.3389/fbioe.2022.922423

Pokusaeva, K., Fitzgerald, G. F., & Van Sinderen, D. (2011). "Carbohydrate metabolism in Bifidobacteria". In *Genes and Nutrition*, 6(3), 285-306. https://doi.org/10.1007/s12263-010-0206-6

Sela, David A, Garrido, D., Lerno, L., Wu, S., Tan, K., Eom, H.-J., Joachimiak, A., Lebrilla, C. B., & Mills, D. A. (2012). "Bifidobacterium longum subsp. infantis ATCC 15697 α-fucosidases are active on fucosylated human milk oligosaccharides". *Applied and Environmental Microbiology*, *78*(3), 795–803. https://doi.org/ 10.1128/AEM.06762-11

Sliwoski, G., Kothiwale, S., Meiler, J., & Lowe, E. W. (2014). "Computational methods in drug discovery". *Pharmacological Reviews*, *66*(1), 334–395. https://doi.org/10.1124/pr.112.007336

Sojar, H. T., & Bahl, O. P. (1987). "A chemical method for the deglycosylation of proteins". *Archives of Biochemistry and Biophysics*, *259*(1), 52–57. https://doi.org/ 10.1016/0003-9861(87)90469-3

Stanley, P. (2016). "What Have We Learned from Glycosyltransferase Knockouts in Mice?" In *Journal of Molecular Biology*, 428(16), 3166-3182. https://doi.org/10.1016/j.jmb.2016.03.025

Steinmetz, E. (2011). "Expresso® Cloning and Expression Systems: Expressioneering™ Technology streamlines recombinant protein expression". *Nature Methods*, 8(6), iii-iv. https://doi.org/10.1038/nmeth.f.344

Szabo, Z., Guttman, A., & Karger, B. L. (2010). "Rapid release of N-linked glycans from glycoproteins by pressure-cycling technology". *Analytical Chemistry*, *82*(6), 2588–2593. https://doi.org/10.1021/ac100098e

Trimble, R. B., & Tarentino, A. L. (1991). "Identification of distinct endoglycosidase (endo) activities in Flavobacterium meningosepticum: endo F1, endo F2, and endo F3. Endo F1 and endo H hydrolyze only high mannose and hybrid glycans". *The Journal of Biological Chemistry*, *266*(3), 1646. https://doi.org/ 10.1016/S0021-9258(18)52343-7

Turyan, I., Hronowski, X., Sosic, Z., & Lyubarskaya, Y. (2014). "Comparison of two

approaches for quantitative O-linked glycan analysis used in characterization of recombinant proteins". *Analytical Biochemistry*, *446*, 28–36. https://doi.org/ 10.1016/j.ab.2013.10.019

Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Darvill, A. G., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., & Seeberger, P. H. (2017). Essentials of glycobiology, third edition. In *Cold Spring Harbor Laboratory Press*.
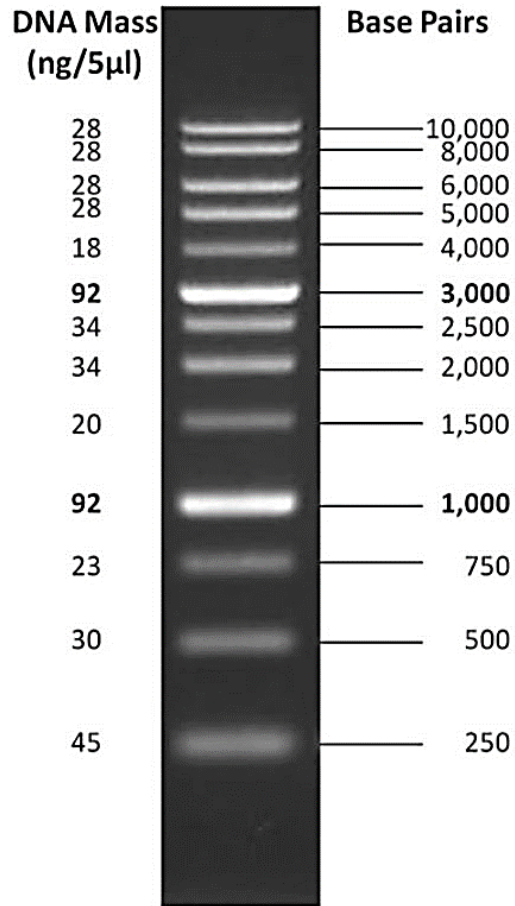
Wang, W. L., Du, Y. M., Wang, W., Conway, L. P., Cai, Z. P., Voglmeir, J., & Liu, L. (2017). "Comparison of the bifidogenic activity of human and bovine milk N-glycome". *Journal of Functional Foods*, *33*, 40–51.https://doi.org/ 10.1016/j.jff.2017.03.017
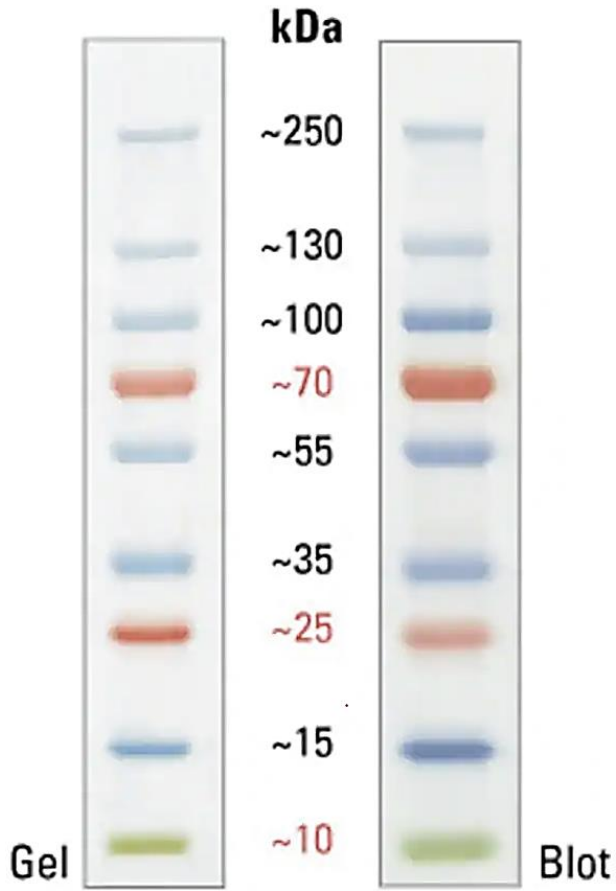
APPENDICES

APPENDIX 1

DNA LADDER USED IN RECOMBINANT DNA ANALYSIS

DNA Mass (ng/5µl)		Base Pairs
28		10,000
28		8,000
28		6,000
28		5,000
18		4,000
92		**3,000**
34		2,500
34		2,000
20		1,500
92		**1,000**
23		750
30		500
45		250

1% TAE Agarose Gel

I

APPENDIX 2

PROTEIN LADDER USED IN PROTEIN ANALYSIS

kDa

~250

~130

~100

~70

~55

~35

~25

~15

~10

Gel

Blot

APPENDIX 3

BUFFERS USED IN RECOMBINANT PROTEIN PRODUCTION STUDIES

	Lysis Buffer (pH 8.0)	Equilibration Buffer (pH 7.4)	Wash Buffer (pH 7.4)	Elution Buffer (pH 7.4)
Ingredients				
Imidazole	1 mM	10 mM	25 mM	250 mM
Tris-HCl	50 mM	-	-	-
NaCl	200 mM	300 mM	300 mM	300 mM
NaH_2PO_4	-	20 mM	20 mM	20 mM
SDS	1%	-	-	-

APPENDIX 4
ORAL PRESENTATION WITHIN THE SCOPE OF THIS THESIS

BIOGRAPHY