T.C.
CANAKKALE ONSEKIZ MART UNIVERSITY
SCHOOL OF GRADUATE STUDIES

DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS

A NOVEL IN-VITRO DIGESTION MODEL DESIGNED BY INTEGRATION OF MICROBIOME ASSOCIATED ENZYMES

MASTER OF SCIENCES THESIS

MERVE KAPLAN

Thesis Supervisor:
Assoc. Prof. Dr. SERCAN KARAV

ÇANAKKALE – 2022

T.C.
CANAKKALE ONSEKIZ MART UNIVERSITY
SCHOOL OF GRADUATE STUDIES

DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS

A NOVEL IN-VITRO DIGESTION MODEL DESIGNED BY INTEGRATION OF MICROBIOME ASSOCIATED ENZYMES

MASTER OF SCIENCE THESIS

MERVE KAPLAN

Thesis Supervisor:
Assoc. Prof. Dr. SERCAN KARAV

This study has been supported by TÜBİTAK 1001 – Supporting Scientific and Technological Research Projects Program.
Proje No: 120z517

ÇANAKKALE – 2022

T.C.

CANAKKALE ONSEKIZ MART UNIVERSITY

SCHOOL OF GRADUATE STUDIES

The study titled "**A Novel In-Vitro Digestion Model Designed by Integration of Microbiome Associated Enzymes"**, prepared by **Merve Kaplan** under the direction of Assoc. Prof. Dr. Sercan KARAV and presented to the following jury members on **01/06/2022** was unanimously accepted as **a MASTER THESIS** at Canakkale Onsekiz Mart University, School of Graduate Studies, **Department of Molecular Biology and Genetics.**

Jury Members	**Signature**
Assoc. Prof. Dr. Sercan KARAV (Supervisor)	……
Prof. Dr. Kemal Melih TAŞKIN (Member)	……
Assoc. Prof. Dr. Mecit Halil ÖZTOP (Member)	…….

Thesis No : 10464096

Thesis Defence Date :01/06/2022

……

İSİM SOYİSMİ

Director

School of Graduate Studies

../../20..

ETİK BEYAN/ ETHICAL STATEMENT

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.

I declare that the data, information, and documents in the thesis were obtained within the framework of academic and ethical rules, I have presented all information, documents, evaluations, and results according to following scientific ethics and ethical rules, and I cited all the reference studies that I used in my thesis by making an appropriate citation. I declare that this thesis study I presented is original, otherwise I accept all the loss of rights that may happen.

(İmza)
Merve KAPLAN
(Tarih) ../../20..

ACKNOWLEDGEMENT

I have been incredibly fortunate to have had the encouragement of my supervisor, colleagues, friends, and family. I would like to thank the following people who have notoriously helped me undertake this thesis research:

Firstly, and most importantly, I would like to express my deepest appreciation to my supervisor, Assoc. Prof. Dr. Sercan Karav, for letting me be a good scientist by teaching all critical aspects of the science. My success and the completion of this thesis would not have been possible without the support and nurturing of Dr. Karav. He always provided me with patience, guidance, thoughtful comments, recommendations, and practical suggestions that cannot be underestimated. Dr. Karav's expertise in this research field significantly improved my thesis. Furthermore, I am very grateful for his friendly and personal support in my academic endeavors. I firmly maintain that the extensive knowledge that I learned from Dr. Karav will shed light on my career path in this science journey.

I opine that friendship and its support are crucial to being motivated in academic life. I am grateful to my dear friend, Ayşenur, for motivating me with her smiling face all the time wherever we are. Besides, I very much thank her for being a good friend, labmate, neighborhood, and more. I would also like to extend my sincere thanks to my other teammates Hatice, Melda, Eda, Burcu, Hasan, Taner, Ayşe, and Berfin who gave noticeable contributions to me. They considerably helped me with their collaborative effort during data collection. I cannot leave the acknowledgment part without mentioning Samet who has been instrumental during my master term. Samet always attentively listened to me and took the time to proofread the earlier drafts of my studies even though he did not understand them well. I very much appreciate him for his support, never-ending encouragement, and always being there for me.

Finally, yet importantly, I would like to express my gratitude to my family for their unfailing love, and unconditional support throughout my life. They nurtured me and my success with their unlimited love, trust, and endless patience. I very much appreciate my family's encouragement which entertained and helped me get through this difficult time most positively and efficiently.

ÖZET

MİKROBİYOM KAYNAKLI YENİ ENZİMLERİN ENTEGRE EDİLDİĞİ İN-VİTRO SİNDİRİM MODELİNİN TASARIMI

Merve KAPLAN

Çanakkale Onsekiz Mart Üniversitesi

Lisansüstü Eğitim Enstitüsü

Moleküler Biyoloji ve Genetik Anabilim Dalı Yüksek Lisans Tezi

Danışman: Doç. Dr. Sercan KARAV

01/06/2022, 118

İnsan mikrobiyotası, milyonlarca sayıda mikroorganizma içermektedir ve bu mikroorganizmalar beyin bağırsak etkileşimi ve enerji metabolizması gibi birçok önemli biyolojik fonksiyonda görev almaktadır. Özellikle, birçok komensal bakteri sahip oldukları özel enzim sistemi ile sindirim işleminde önemli roller oynamaktadır. Glikan diye isimlendirilen prebiyotik bileşenler, insan enzimleri tarafından sindirilemediklerinden bağırsağa kadar denatürasyona uğramadan ulaşmaktadırlar. Bağırsakta ise bazı bakteriler sahip oldukları enzimler (glikozidazlar vb.) sayesinde glikanları karbon kaynağı olarak kullanabilmektedir. Glikanlar gibi prebiyotiklerin sindirilme mekanizmasını daha iyi anlamak için *in-vitro* sindirim modelleri kullanılmaktadır. Fakat, kullanılan *in-vitro* sindirim modellerinde sadece insan kaynaklı enzimlerin yer alması bu modellerin mikroorganizmaların sindirimdeki etkisini inceleyen çalışmalarda kullanılmasını engellemektedir. Bu yüzden, mikrobiyal enzimlerin yer aldığı yeni sindirim modellerinin tasarımı glikan çalışmaları için kritik öneme sahiptir.

Bu tez kapsamında, öncelikle insan sindirim sisteminin farklı bölgelerinde baskın olarak bulunan mikroorganizmalar ve bu mikroorganizmalara ait glikozidaz enzimleri biyoinformatik yöntemler kullanılarak belirlenmiştir. Belirlenen ve seçilen 32 glikozidaz

rekombinant olarak uygun bir moleküler klonlama sistemi ile klonlanmış ve üretilmiştir. Üretilen rekombinant enzimler, sadece insan kökenli sindirim enzimlerini içeren standart bir *in-vitro* sindirim modeline entegre edilmiştir ve bir glikoprotein kaynağı olan whey üzerinde test edilmiştir.

Anahtar sözcükler: Mikrobiyal Enzimler, Sindirim, Glikan Metabolizması, Mikrobiyota

ABSTRACT

A NOVEL IN-VITRO DIGESTION MODEL DESIGNED BY INTEGRATION OF MICROBIOME ASSOCIATED ENZYMES

Merve KAPLAN

Canakkale Onsekiz Mart University

Master of Science Thesis in Molecular Biology and Genetics

Advisor: Assoc. Prof. Dr. Sercan KARAV

01/06/2022, 118

The human microbiota consists of much more microbial cells than human cells and they are associated with a myriad of biological functions ranging from gut-brain signaling to energy metabolism. Importantly, most gut commensals are involved in the human digestion process using their carbohydrate-active enzymes (CAZymes) like glycosidases, which are used to cleave polysaccharide chains, also called glycans, into monomers to benefit both themselves and the host. Glycans cannot be digested by human-derived enzymes due to the lack of specific enzymes. Therefore, glycans reach the colon where some bacteria can metabolize them by their unique enzymes. To better understand glycan digestion by microbial metabolism, in-vitro digestion models could be a great way to study the interaction between microbial enzymes and glycans in laboratory conditions. However, current in-vitro digestion models are not available for glycan studies due to the lacking the human enzyme specificity. Thus, the design of novel models including host and microbiome-associated enzymes is critical to paving the way for glycan research.

Within the purpose of this thesis, novel glycosidases were examined from different microorganisms, which predominate in the human digestive system, using bioinformatic tools. Then, 32 unique enzymes were recombinantly cloned with a cloning and expression system and produced. The selected enzymes were integrated into a conventional in-vitro digestion model which includes only human-associated digestion enzymes. Finally, the new

digestion model designed by recombinant microbial enzymes integration was tested on a glycoprotein source, whey.

Keywords: Microbial Enzymes, Digestion, Glycan Metabolism, Microbiota

CONTENTS

Page No

SYMBOLS AND ABBREVIATIONS

Asn	Asparagine
Ser	Serine
Thr	Threonine
dH_2O	Distilled water
et.al.	Others
g	Gram
M	Molar
Mg/mL	Milligram/milliliter
kDa	Kilodalton
%	Percent
h	Hour
s	Second
min	Minute
V	Volt
GI	Gastrointestinal
SCFA	Short chain fatty acid
HMO	Human milk oligosaccharide
SSF	Simulated saliva fluid
SGF	Simulated gastric fluid
SIF	Simulated intestinal fluid
GalNAc	N-acetylgalactosamine
HexNAc	N-acetylglucosamine
HCI	Hydrochloric acid
NaCI	Sodium chloride
KCI	Potassium chloride
KH_2PO_4	Monopotassium phosphate
$NaHCO_3$	Sodium bicarbonate
$MgCI_2(H_2O)_6$	Magnesium chloride hexahydrate
$(NH_4)_2CO_3$	Ammonium carbonate
LB	Luria Bertani
OD	Optical density
SDS PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1. Human Gastrointestinal System and Microbiota

The gastrointestinal (GI) system is a long and complex system covering the oral cavity, stomach, intestines, anus, and other connected organs. The major function of the GI system is to absorb food components through mechanical as well as chemical digestion. In addition, it also takes a significant role in many systems including the immunity. GI system and its functions have been studied for long years (Corinaldesi et al., 1987; Hawkey et al., 1992). Furthermore, the population of microorganisms symbiotically habiting in the GI system and do not cause any pathogenic disease has attracted particular interest in last couple of years (Hillman et al., 2017). With new studies and data, microbiota and its functions related to many pathways in the human body has become a hot topic amongst many fields (O'Hara & Shanahan, 2006). The microbiota is commonly defined as the collective microorganisms including bacteria, fungi, viruses, archaea, as well as eukaryotes. A massive number of microorganisms reside in the human intestinal tract, which is called gastrointestinal or gut microbiota (Turnbaugh et al., 2007). The gastrointestinal microbiota plays an important role in the regulating basic physiology in the human body due to its wide range of enzyme abilities (Figure 1). It helps the production of vitamins, mineral absorption, protection against pathogens, and immune system enhancement (Hillman et al., 2017).

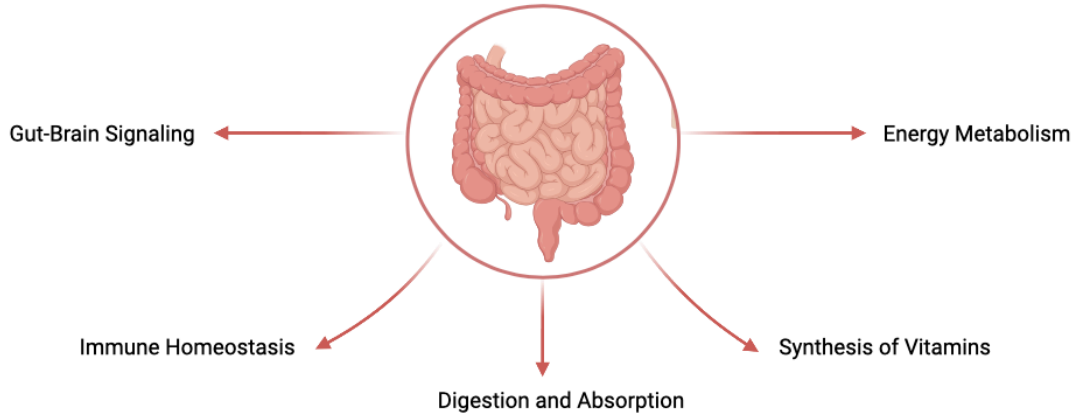


Figure 1. Functions of gut microbiota (Hillman et al., 2017).

The gastrointestinal microbiota considerably depends on a number of factors such as age, gender, health situation, as well as nutrition style (Ursell et al., 2012). Microbial variety also varies across the GI system (Figure 2) due to factors including pH, peristalsis, redox potency, adhesion, secretion of mucin, availability of nutrients, as well as bacterial antagonism (Tannock & Savage, 1974). Whilst Streptococcus is one of the most prevalentgenera in the oral cavity, some other populations such as Neisseria, Gemella, Granulicatella Veillonella, and Prevotella also exist (Aas et al., 2005). In addition to the bacterial population, some virus and fungal species are also found in the human oral phase. Bacteriophages, for instance, are the most common virus population, whereas Candida is a fungi species mostly found in this region (Dupuy et al., 2014). The esophagus is the first part where food reaches during its passage to the stomach after chewing. Despite the limited study related to esophageal microbiota, similar results to oral microbiota have been encountered in some studies. As for the stomach, it has a unique microbiota due to its acidic environment, Proteobacteria constitute the majority of the microorganisms in the stomach. In addition, the population of Streptococcus and Prevotella is also found in similar to oral and esophageal microbiota (Bik et al., 2006). The last phase of the digestive tract is the intestines which include three subparts: duodenum, jejunum, and ileum. Bacteroides, Clostridium and Streptococcus reside in these parts, which can also vary across these three subparts (Leimena et al., 2013). The large intestine contains 70% of bacteria in the entire human body and consists of mainly Bacteroides, Prevotella, and Ruminoccocus. In general, phyla, Bacteroides, Firmicutes, Proteobacteria, and Actinobacteria are four major phyla that predominantly belong to the gut. The gut, in particular, is the most predominantly populated by about 1,000 different species of known bacteria, which includes both resident and transient bacteria in a complex environment. The gut microbiota is considered to be a diverse and complex array of microbial ecosystems, which considerably affect human health in many aspects. This wide variety in the gut is caused by slow intestinal motility and lower redox potential. Gut microbiota changes significantly based on the age, diet, and lifestyle of the host (Davenport et al., 2015).

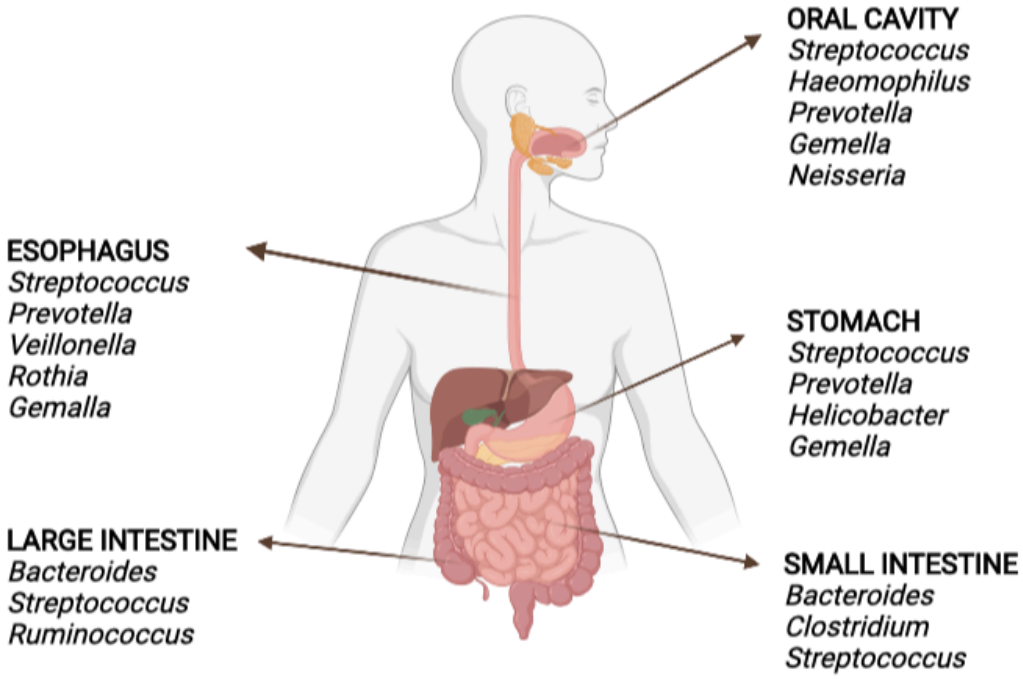


Figure 2. Distribution of gastrointestinal bacteria through the GI system (Aas et al., 2005; Dupuy et al., 2014).

Studies regarding human microbiota and its relationship with human physiology have noticeably increased in recent years with the Human Microbiome Project. This interest in human microbiota is mainly because it takes a significant part in human physiology and related diseases (various types of allergies, asthma, intestinal-related disorders, etc.) (Hillman et al., 2017). With new applied techniques, human microbiota, which was previously difficult to study and understand, has become easier and clearer to study. The relationship between humans and microbiota is comprehensively studied with developed technology including several techniques such as next-generation sequencing, total DNA characterization, and 16S gene region sequencing (Ursell et al., 2012). Within this perspective, the differences between healthy and unhealthy human microbiota have been distinguished and new treatment methods have been developed against microbiota-related diseases using the data in these hot studies (Shafquat et al., 2014).

1.2. Gastrointestinal Microbiota and Nutrition

One of the most significant roles of gastrointestinal microbiota is its role in digestion and metabolism. The density of the microbiota population generally increases from the stomach to the small intestine and from the small to the colon. This indicates a progressive increment of pH and variable digestive functions. In the colon, for instance, a very dense and diverse microbiota ferment undigested food. The gastrointestinal microbiota is efficiently involved in processing foods such as starch and dietary fiber. As a symbiotic host-microbe relationship, microbes can utilize indigestible nutrients as a carbon source to grow, whereas absorption of byproducts and the enhancement of nutrient bioavailability provides considerable benefits to the human body. Short-chain fatty acids (SCFAs), for example, are byproducts such as lactic acid, butyric acid formed after utilization of undigested poly or oligosaccharides as a carbon source by microbes. SCFAs can easily be absorbed and utilized as a source of energy by a human host, and they are responsible for 10% of human energy requirements (Gerritsen et al., 2011; Wopereis et al., 2014). In the same manner, this relationship exists between unique components of breast milk and infant gut microbiota in early life development (Bode, 2012; Karav et al., 2016).

1.3. Early Development of Microbiota and Its Interaction with Human Milk

Early colonization progress in terms of microbiota is critical to both early and lifelong human health, which influences immune development, maturation of gut, physiological functions, and metabolism (Wopereis et al., 2014). Microbial colonization in infants occurs immediately after birth from mother (vertical) and environment (horizontal) transfers (Townsend & Moore, 2019). Some pioneering bacteria enter infants' bodies and establish a new microbial ecosystem within their gut. Furthermore, initial colonization of the infant's gut mainly results from microbes in the environment covering the maternal vaginal, skin, and fecal microbiota (Wopereis et al., 2014).

The development of human microbiota in the first three months is closely dependent on some factors such as feeding type, antibiotic usage, and delivery type. For instance,

cesarean-born infants showed less diversity of the bacterial population in comparison to vaginally delivered ones (Clarke et al., 2014). Feeding type is a critical factor that affects human health in the long term since human milk includes so significant components lacking in formulas. Human milk provides optimal nutrition for infants in their early development, as it has rich nutritional content providing all the energy, bioactive components which are essential for infant growth. Breast milk composition is very dynamic and has evolved to meet optimal nutrition for infants. It includes protein, lipids, lactose, and bioactive components, which take different roles in infant health (Ballard & Morrow, 2013). In addition to this content of human milk, the third major but non-nutritional component is human milk oligosaccharides which are also called HMOs (Bode, 2012).

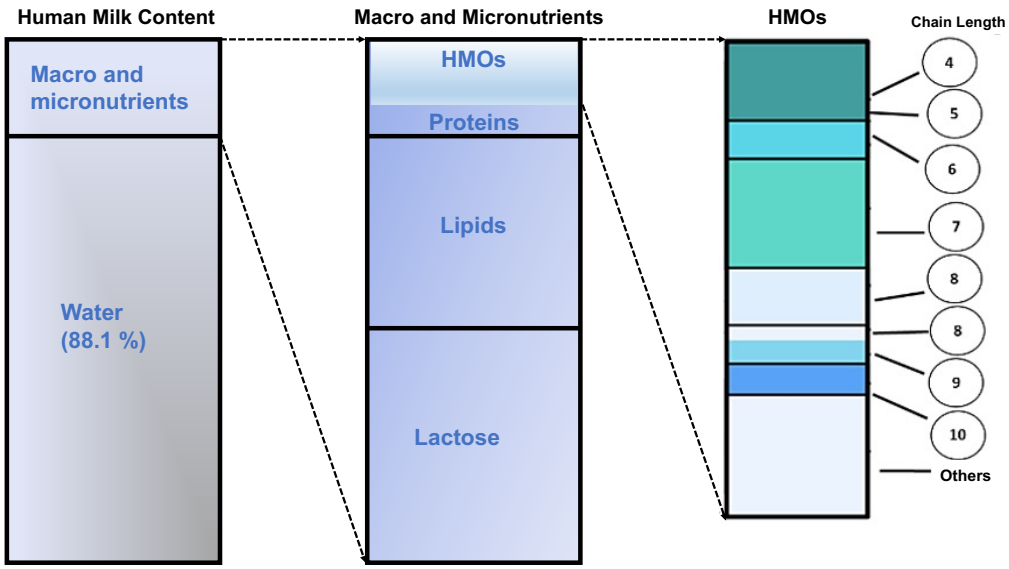


Figure 3. Human milk composition (Ballard & Morrow, 2013).

As HMOs are functional and complex carbohydrates, they play crucial roles in the infant body from innate defense to neural development, and in particular gut health (Bienenstock et al., 2013; Bode, 2012; Wiciński et al., 2020). These complex carbohydrate molecules are indigestible by human-associated enzymes, so they can reach the colon as an intact form. In the colon, they are considered to be a prebiotic and shape the gut microbiota in the infant GI tract (Walsh et al., 2020).

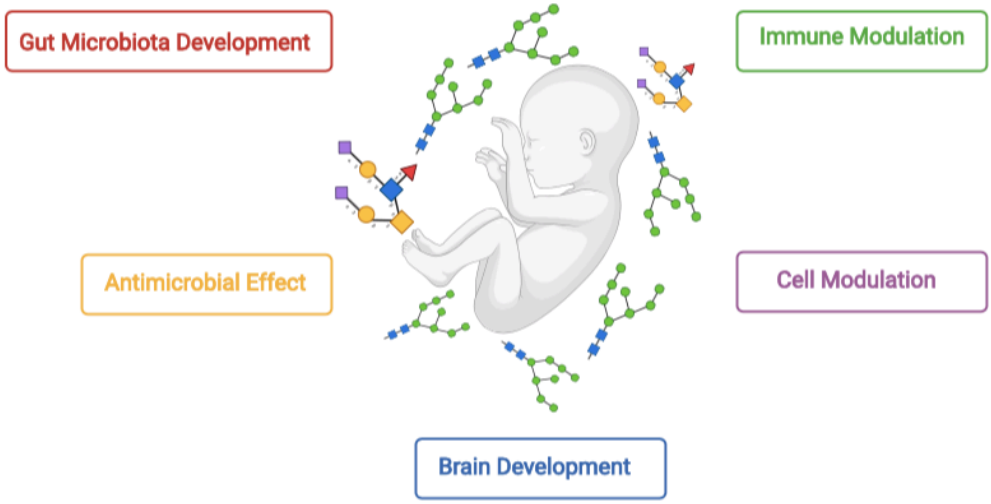


Figure 4. Functions of human milk oligosaccharides (Bienenstock et al., 2013; Bode, 2012).

1.4. Glycans and Their Interactions with Gastrointestinal Microbiota

Human milk includes not only free oligosaccharides, HMOs, but also consists of significant conjugated glycans to proteins or lipids. Most of the proteins (70%) are found as glycoprotein form in human milk such as lysozyme, lactoferrin, casein, and secretory IgA (SIgA). Protein glycosylation is a post-translational modification which takes part in crucial roles in such biological mechanisms as recognition, protein folding, and enzyme protection (Moremen et al., 2012; van Berkel et al., 1995). Protein glycosylation is found as *N*-glycosylation and *O*-glycosylation in eukaryotes; *N*-glycosylation takes place when *N*-glycans make covalent bonds with proteins at carboxamide group in asparagine (Asn) side chain residue of Asn-X-Ser / Thr via N-glycosidic bond. On the other hand, *O*-glycosylation formed when O-linked ones attached to the OH group at the side chain of serine or threonine amino acid (Varki et al., 2009). N-linked glycans include three different forms as high mannose, hybrid, and complex according to their monosaccharide sequence and branch type, whereas *O*-linked ones have eight different core structures (Parc et al., 2015). Conjugated glycans are involved in several biological mechanisms including protein folding, cell-cell or cell-host communication, antimicrobial, antiviral, and prebiotic effects (Karav et al., 2017). Furthermore, they decline leukocyte binding to endothelial cells, hinder pathogen binding to epithelial cells, inhibition rotavirus related to diarrhea in infants, and development of the cognitive ability of infants (Kunz et al., 2000; Morgan & Winick, 1980).

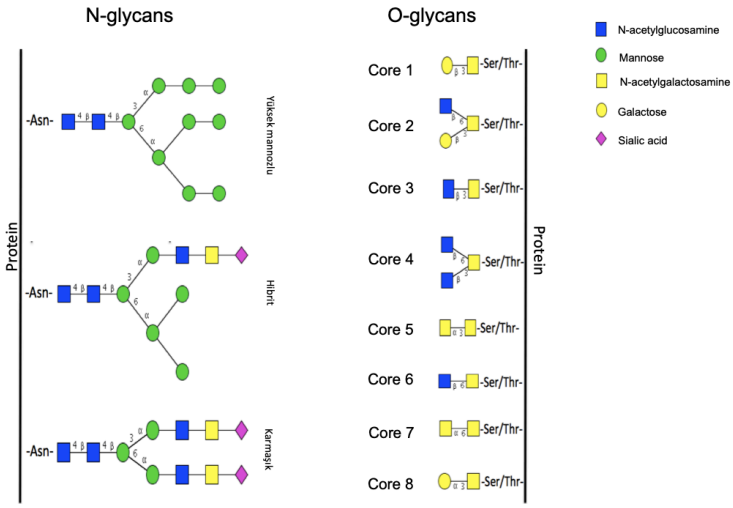


Figure 5. Glycan structures (Varki et al., 2009; Parc et al., 2015).

Conjugated glycans are similar to HMOs regarding structure including composition of monosaccharides and link type. *N*-glycans, in particular, can form complex structures and this increases the specificity of these molecules. *N*-glycans released from bovine and human milk are considered to be bifidogenic compounds, which can shape gastrointestinal microbiota like HMOs. A unique function of these conjugated glycans is that released glycans from glycoproteins are used also as a carbon source by Bifidobacteria species in the human gut due to their genomic capability (Karav et al., 2018). *Bifidobacterium infantis* (*B. infantis)* that is a probiotic extensively found in the gut of infants can release breast milk glycans from glycoproteins by Endo-*ß*-*N*-acetylglucosaminidase enzyme and then these released free glycans are used as a carbon source for *B. infantis* (Karav et al., 2016). Moreover, these molecules cause a selective growth in the microbial ecosystem, for instance, released *N*-glycans from bovine milk glycoproteins stimulate *B. infantis* adapted to the infant's gut. However, *Bifidobacterium animalis* (*B. animalis*) cannot utilize these structures. In an *in-vivo* study, pathogens cannot utilize these oligosaccharides, whereas they can degrade glycans found on the infant gastric mucosa (Karav et al., 2018). Another study showed that nineteen different *N*-glycans conjugated to lactoferrin as well as immunoglobulins enhance the *B. infantis* growth (Karav et al., 2019). *N*-glycans are also fermented to SCFAs like HMOs, this positively affects the microbial environment and lowers the pH which creates a resistance to pathogen colonization since they preferentially

grow at nearly neutral pH like 6-7 (Koropatkin et al., 2012). As the fermentation process of conjugated *N*-glycans forms end-products like acetate and lactate, this creates a disfavored environment for pathogens which degrade gastrointestinal mucin structures and significantly reduces the pathogen population (Duar et al., 2020). With these important functions, conjugated glycans shape the gut microbiota providing colonization resistance, reducing inflammation and virulence factors (Duar et al., 2020; Olin et al., 2018).

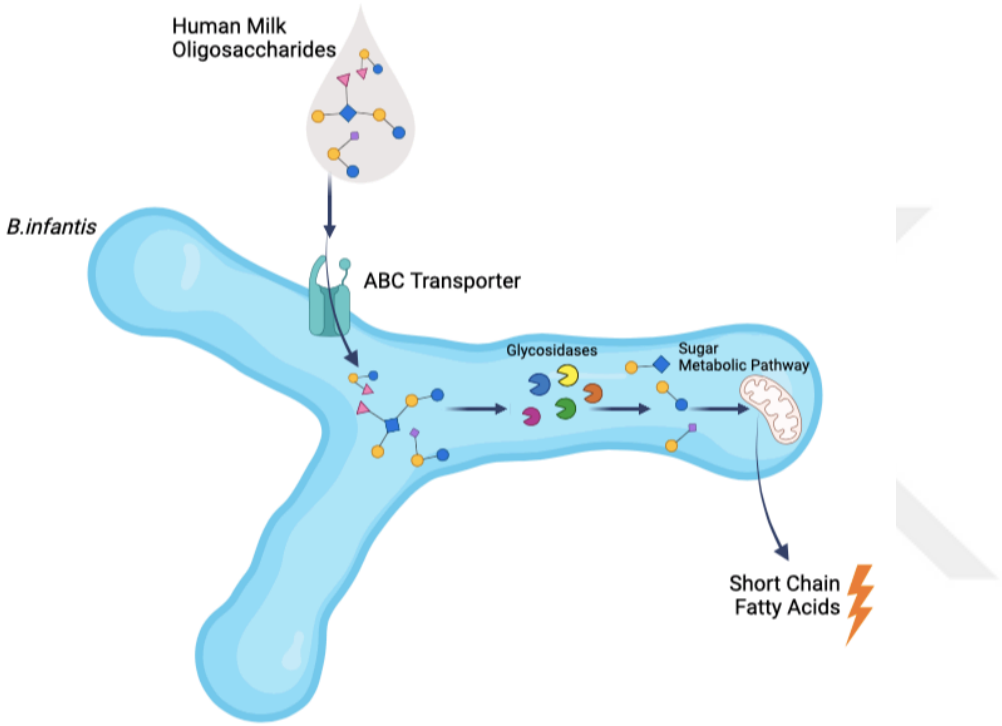


Figure 6. HMO metabolism by *B. infantis* (Chichlowski et al., 2020).

Both conjugated and free oligosaccharides are named prebiotics which are significant compounds for gut health. Prebiotics are indigestible food components which selectively enhance the activity and/or growth of certain bacteria in the GI microbiota (Gibson & Roberfroid, 1995). They exert a myriad of beneficial effects on the human body, however, interactions between prebiotics, probiotics, and pathogens make the definition of prebiotics more complex. Robert W. Hutkins indicated that many prebiotics do not show an actual prebiotic effect on the gut (Hutkins et al., 2016). This is mainly related to some metabolites formed during prebiotic fermentation in the gut, which is readily utilized by pathogens in their first colonization. Even though prebiotics show incredible beneficial effects for the

human body, they can allow the growth of pathogens related to different strategies to transfer oligosaccharides into the cell (LoCascio et al., 2009). For instance, *B. bifidum* and *B. infantis* are two bacteria that are genomically adapted to metabolize HMOs, whereas *B. breve* and *B. longum* have more strain-specific phenotypes and degrade certain HMOs (LoCascio et al., 2009; Sela et al., 2012). On the other hand, some adult-type Bifidobacteria including *B. animalis* and *B. adolescentis* cannot metabolize HMO structures. Such different utilization situations are strongly dependent on bacterial genomes and strategy. *B. infantis* has different glycosyl hydrolase enzymes and firstly takes complex oligosaccharides into its cell and then utilizes them, so there are not any metabolites formed in the microbial environment. This considerably reduces the cross-feeding potency for pathogens as the whole hydrolysis process takes place in the cell. However, *B. bifidum* firstly cleaves the linkages in oligosaccharides by using its enzymes and converting complex molecules to monosaccharides outside. Then, it takes monomers through its cell which promotes the pathogen colonization by cross-feeding process (Chaplin et al., 2015).

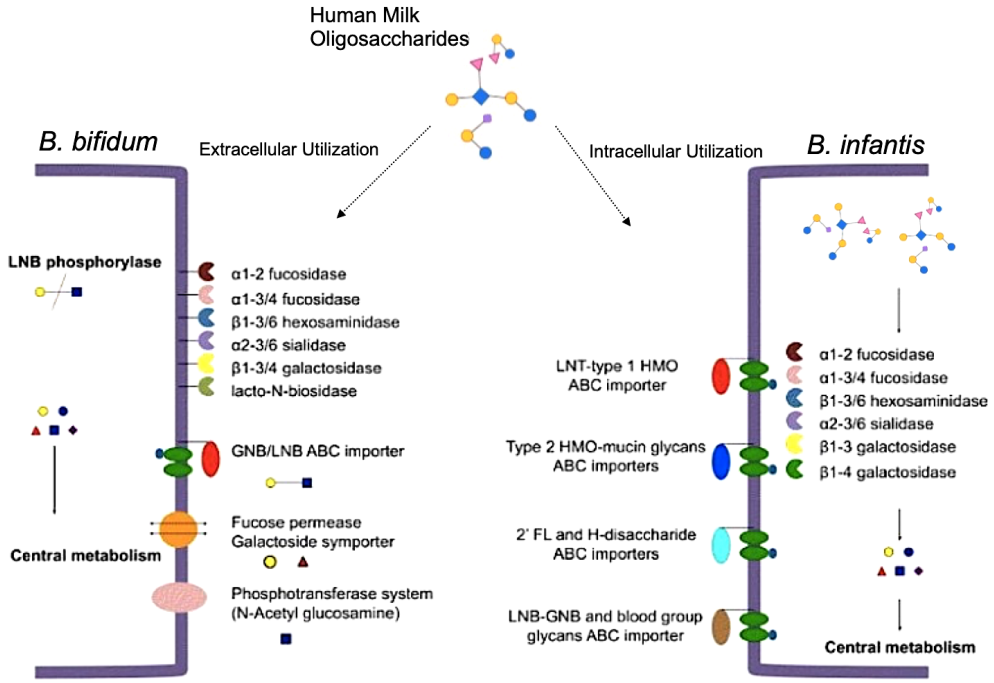


Figure 7. Two different HMO utilization mechanisms by *B. infantis* and *B. bifidum* (LoCascio et al., 2009).

Conjugated glycans can be released from glycoproteins by some chemical as well as enzymatic methods to better understand and study their functions and prebiotic activity related to gastrointestinal microbiota. Chemical methods are commonly used to release

glycans because of their advantages including easy application, lower cost, and activity on different substrates (Sojar & Bahl, 1987). Hydration and β-elimination are the most common chemical methods in alkaline conditions (Dwek, 1993). Though sodium borohydride as a reducing agent is preferred to prevent the chemical structure of oligosaccharides, peeling action may negatively affect the structure and remaining protein structure (Carlson, 1968). In addition, a loss of glycan takes place during the salt removal which is used in β-elimination (Turyan et al., 2014). Hydration is a more effective process in comparison to the β-elimination in terms of both effectiveness and released glycans variety.

The releasing of glycans by chemical methods affects both glycans and the remaining part. Moreover, the mass spectrometry analysis of released glycans is so difficult because of the high amount of salt in glycans related to the chemical method. As for enzymatic techniques, glycans are released from glycoproteins using Peptidyl-N-glycosidases (PNGases) (Altmann et al., 1995). These enzymes can release all glycans, but they cannot show activity if there is fucose linked to the N-acetylglucosamine with a 1,3 bond (Tretter et al., 1991). Furthermore, glycoproteins are denatured using detergent and high temperatures for enzyme activity. Other enzymes including Endoglycosidase F1, F2, and F3 also exert the activity regardless of any substrate denaturation. However, these enzymes are active on a so limited number of glycans (Trimble & Tarentino, 1991). Considering all these chemical and enzymatic techniques, novel enzymes are essential and of utmost importance to release glycans from glycoproteins. To further study the interaction between food components, prebiotics, and gastrointestinal microbiota, novel enzymes and models are a critical requirement in this field.

1.5. *In-Vitro* Digestion Models

GI system and its mechanism related to food digestion are highly complex since a variety of factors can affect it. However, the GI system is a major focus for many foods and health studies as nutrition is important part in human health (Bornhorst et al., 2016). Ingested foods during the human digestion process are converted to nutrients that are of utmost importance for the human body in terms of growth, energy as well as repair. The digestion of food consists of two key steps: mechanical process in which larger food components are broken down into smaller components, begins in the oral phase and continues through the gastric phase; an enzymatic process where various enzymes transform macromolecules to small ones which can be easily absorbed through bloodstream, begins in the oral phase and continues through the intestinal phase (Alminger et al., 2014; Guerra et al., 2012). In vivo approaches are used to study food digestion, which generally include feeding and the acquisition of samples of digested food from the gastric part and small intestine. However, *in-vivo* systems to study digestion have noticeable drawbacks including technical issues, ethical difficulties, high cost, and physiological differences between individuals. Therefore, *in-vitro* models are maintained to be good alternative models for GI system and digestion studies (Ménard et al., 2014). Even though the GI system is difficult to study since it has complex interactions with other physiological systems in the human body, *in-vitro* digestion models are successfully used to understand the digestion process further (Marcano et al., 2015; Minekus et al., 2014). The *in-vitro* digestion model is used firstly by DeBaun and Connors (Debaun & Connors, 1954). These models can vary regarding the GI system phase (Sek et al., 2001). Generally, common *in-vitro* digestion models try to mimic the whole digestion process of food components along with oral, gastric, and small intestine parts of the GI tract. These *in-vitro* models may vary as static, or dynamic based on the complexity.

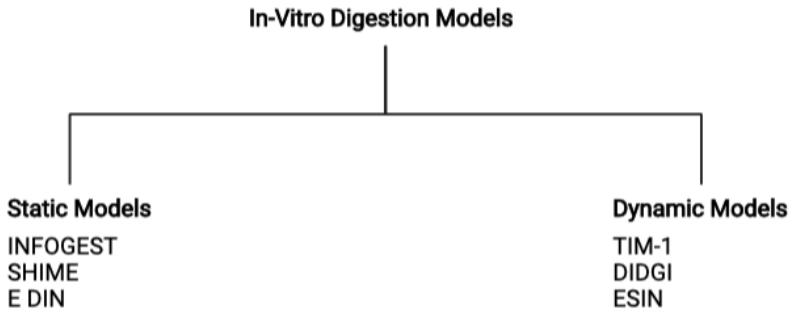


Figure 8. Categorization of *in-vitro* digestion models (Ménard et al., 2014).

1.5.1. Static *In-Vitro* Digestion Models

Static gastric models (SGM) are basic models which include a system that all components of phases are incubated at each one with appropriate enzymes and gastric juices mimicking pH, temperature, and time (Minekus et al., 2014). Food is introduced into a reaction tube which could be a beaker, test tube, or an Erlenmeyer flask). When the food is added to the test tube, digestive fluids, as well as enzymes, are also introduced to each GI phase. The temperature and pH are maintained according to the phase conditions, the pH could be in an uncontrolled situation or kept stable with a pH-stat system. To cite an instance, 1 g food sample is introduced into a test vessel and then 1 mL of simulated salivary fluid is added to the mix. The pH is 7 and the temperature is 37°C for the oral phase and it takes 2 minutes. In the same manner, 2 mL of simulated gastric fluid, as well as pepsin enzyme (2000 U/mL), are added to the test tube, and pH is adjusted to 3 using HCl (final volume is 4 mL). The gastric phase takes 120 min, after that pH is adjusted to 7 with NaOH to mimic intestinal conditions. 4 mL of simulated intestinal fluid with pancreatin and bile salts is added and incubated through 120 min. The final volume of this last phase is 8 mL, and trypsin activity is 100U/mL (Brodkorb et al., 2019).

An international network INFOGEST consists of multidisciplinary applications' professions from 32 different nations. Within the perspective of INFOGEST, an *in-vitro* digestion protocol, also named as INFOGEST methods, is well simulated to human digestion process (Brodkorb et al., 2019; Minekus et al., 2014). In addition to INFOGEST, some other static models are also used, for instance, United States Pharmacopoeia methods and Unified BARGE methods, but they do not fix for assessment of food products as they are developed for pharmaceuticals and soil or food contaminants (Brodkorb et al., 2019).

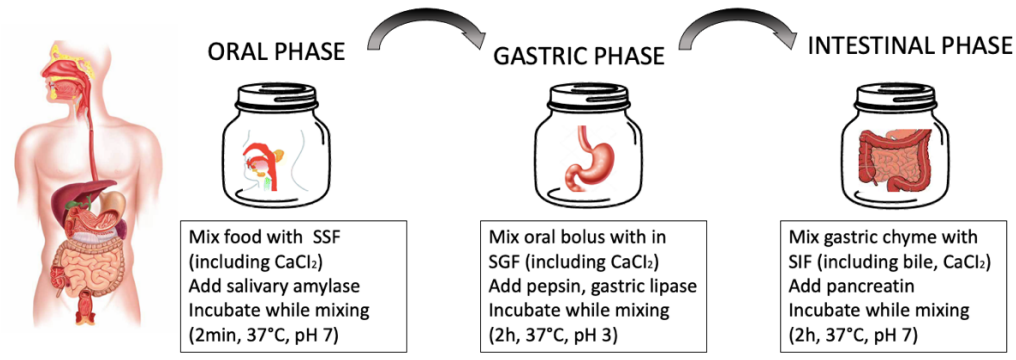


Figure 9. The basic principle of the *in-vitro* digestion model (Minekus et al., 2014).

SGMs are easy to perform and can investigate physiological processes at the molecular level (Ménard et al., 2014). Its simplicity offers that it is well-suited for *in-vitro* digestion works with the number of food samples. Static models, in particular, are commonly used to determine the food process' effect on nutrient bioaccessibility, bioavailability, and/or allergenic peptides. On the other hand, even though static *in-vitro* digestion models are easy and fast, they cannot mimic thoroughly *in-vivo* digestion process due to some reasons such as pH changes uncontrolling, lacking gradual addition of gastric fluids, and emptying (Brodkorb et al., 2019).

1.5.2. Dynamic *In-Vitro* Digestion Models

Dynamic *in-vitro* digestion models are computer-controlled models, so they have the capacity to simulate complex digestion. Dynamic digestion models are stated as mono-compartmental as well as multi-compartmental models. Many mono- compartmental ones simulate gastric digestion with its gastric contraction, mechanism of fluids, the gradual addition of enzymes, and emptying. All models generally have a main chamber with an elastic material and incorporated adding gradually gastric juice and pH controlling (Ji et al., 2021; Wang et al., 2021). They can have a variety of apparatus to simulate gastric contraction including water pressure, pistons, ropes, and rollers. Even more, some models mimic the J-shaped human stomach using 3D printing. Elastic annulus, mesh filter, and a more structural design are also used to simulate opening the pyloric valve and emptying (Ji et al., 2021).

Dynamic models can mimic gastric mixing, gradual secretion of enzymes, emptying, as well as absorption, in contrast to static models. Dynamic *in-vitro* digestion models can mimic gastric mixing, gradual secretion of enzymes, emptying, and absorption, in contrast to static models. They are preferred to study digestion in detail covering emulsion of lipidsproperties for lipid-soluble nutrients, food and/or drug encapsulation techniques, kinetic changes, the release of proteins, and/or lipid oxidation during digestion (Corstens et al., 2018; Qazi et al., 2021). Dynamic models have not only a better accuracy rate but also provide kinetic parameters for the digestion process. Although they have critical advantages for digestion studies, they are time-consuming, so complex, even more need expensive enzymes. Dynamic models are less accessible than a static digestion model due to these reasons.

In general, both *in-vitro* digestion models; dynamic and static ones are preferred to study a wide variety of subjects such as analysis of antioxidant effects of bioactive molecules and assessing these molecules with nano properties at digestion system. Vitamins A, C, D, E, polyphenolic compounds, and carotenoids are widely studied using these models for a better understanding of their effects on human health or pharmacological activities. Furthermore, milk proteins' degradation prediction is also studied by *in-vitro* digestion models just mimicking appropriate pH and incubation conditions (Egger et al., 2019; Wada & Lönnerdal, 2015). However, an important point which is the contribution of microorganisms, and their enzymes is missed in both models. Food digestion is a complex process that considerably interacts with gastrointestinal microbiota. Especially, the digestion of complex carbohydrates including prebiotics is dependent on gastrointestinal microbiota activity as they are indigestible by human-associated digestion enzymes. Therefore, microbiota-associated enzymes are a requirement for *in-vitro* models to mimic the digestion process precisely.

1.6. Aim of the Thesis

The overall purpose of the thesis is to integrate microbiota-associated enzymes through the *in-vitro* digestion model. Within this perspective, the integration of microbial enzymes into a convention digestion model creates a novel *in-vitro* model to simulate a proper digestion process with four GI phases (oral, gastric, small, and large intestine). Current *in-vitro* digestion models are not available for significant glycan studies because of lacking microbial enzymes specificity, they only consist of human-associated enzymes. Therefore, the design of this novel model including host and microbial enzymes is extremely critical to paving the way for studies of complex carbohydrates such as glycans. With the thesis, recombinantly cloned glycosidases of target microorganisms were integrated into the conventional *in-vitro* digestion model. This model helps better understand glycan metabolism and leads to further studies to determine the impact of glycans on GI microbiota.

Within the thesis:

- Target microorganisms in the GI tract parts including oral, gastric, small, and large intestine were determined with deeper literature research.
- Glycosidases as microbial enzymes of target microorganisms were examined.
- Determined microbial enzymes were recombinantly cloned using an appropriate cloning and expression system (Expresso® Rhamnose SUMO Cloning and Expression System).
- Recombinant enzymes were purified using immobilized metal affinity chromatography and their kinetic parameters were determined.
- Recombinant enzymes were integrated through the conventional *in-vitro* model which was also performed within the thesis.
- The novel model was tested on a glycoprotein.

CHAPTER 2

PREVIOUS STUDIES

Glycosylation is a significant post-translation process that takes place in various cellular mechanisms. Most eukaryotic proteins are glycosylated form, and their glycan parts are involved in several biological mechanisms related to human health. Many studies have already indicated that glycans take a significant role in cell adhesion and activation of receptors, which explains the glycoprotein structure linked with the protection function by the host against pathogen attacks. Furthermore, they also take roles in the recognition and connection of microorganisms through cell membranes. Protein folding, conformation, immunogenicity, solubility as well as capacity to proteolysis resistance are also mechanisms in which glycans considerably take part.

Recently, glycans are also considered to be prebiotics since they selectively promote some bacteria in the human microbiota. Human milk glycans are utilized as a carbon source by Bifidobacteria which are beneficial microorganisms associated with healthy infant microbiota and selectively metabolized by probiotics (Karav et al., 2016; Karav, Bell, et al., 2015). A study by Karav et al., presented that *B. infantis* has a unique enzyme which is named Endo-B-N-acetylglucosaminidase (EndoBI-1) of the Blon_2468 gene, and this enzyme cleaves glycans from glycoproteins (Karav, Parc, et al., 2015). An *in-vivo* study showed that pathogen microorganisms cannot utilize these human milk glycans, however; can degrade glycans on the gastric mucin layer. Moreover, the study also showed that microbiota predominated by Bifidobacteria (especially *B. infantis*) utilize mainly human milk glycans, whereas in control groups' infants the focus is mucin layer glycans (Karav et al., 2018). In 2019, it was shown that *B. infantis* in healthy infant microbiota can utilize glycans conjugated to glycoproteins such as lactoferrin and glycoproteins (Karav et al., 2019). Therefore, different microorganisms use different enzymes to release glycans from glycoproteins and then utilize them as a carbon source. Within this perspective, human nutrition is so critical to shaping microbiota with undigestible carbohydrates.

Studies covering nutrients and digestion are critical to understanding the mechanism of digestion and their relationship with human microbiota. GI system, therefore, is a common

focus for many food and health studies (Bornhorst et al., 2016). *In-vivo* studies including the digestion process generally include the feeding and acquisition process, which are generally considered to be more precise to study complex human digestion. Even though *in-vivo* digestion models are preferred in some studies, they have some disadvantages such as ethical problems, high cost, and technical issues. *In-vitro* models are another commonly used to study complex digestion processes in laboratory conditions. They are generally preferred due to their simplicity, and applicability (Ménard et al., 2014). These models include three simulated phases of the GI system including oral, gastric, and small intestine. *In-vitro* digestion models are widely used in a variety of applications such as analysis of the antioxidant effect of bioactive molecules, nano properties at digestion system, pharmacological activities, and milk proteins' degradation (Marcano et al., 2015; Minekus et al., 2014; Qazi et al., 2021). Models basically use human-associated enzymes and appropriate conditions such as pH, temperature, as well as incubation duration which are mimicked to human digestion. A food component is basically integrated through the system and its *in-vitro* digestion process takes place under appropriate conditions like *in-vivo*. On the other hand, glycans and glycan-rich foods, which are indigestible by human enzymes, were not studied within studies of these models. *In-vitro* digestion models depend on only human-associated enzymes. However, in addition to those enzymes, microbial enzymes from millions of microorganisms in human gastrointestinal microbiota have a crucial role in the digestion (Karav et al., 2018, 2019).

Investigation of novel microbial enzymes and integration of them through an *in-vitro* model is so significant to further study digestion and microbiota development. The novel model within the scope of this thesis aims to contribute to several studies in this field. With this model, many glycoproteins or glycan-rich nutrients would be studied to better understand their digestion and interaction with human microbiota.

CHAPTER 3

MATERIAL & METHOD

3.1. Materials

3.1.1. Chemicals, Kits, Culture Media, and Essential Items

All kits, chemicals and other items used in this thesis are listed below (Table 1).

Table 1. List of chemicals, kits, and other items

Material Name	Code	Information
Lysogeny Broth	A507002-0250	Bacterial Media
Phusion Hot Start II High-Fidelity PCR Master Mix	F-565S	Amplification -PCR
EconoTaq PLUS 2X Master mix	30035-1	Amplification - PCR
Safe Red Loading Dye	G108-R	DNA Gel Electrophoresis
DNA Ladder	D011-500	DNA Gel Electrophoresis
DNA Loading Dye	R0611	DNA Gel Electrophoresis
Agarose	16500-500	DNA Gel Preparation
Kanamycin Monosulfate	K-120-5	Antibiotic
Expresso Rhamnose Cloning and Expression System Kit	49011-1	Molecular Cloning Kit
Qubit dsDNA BR Assay Kit	Q32850	Measurement of DNA
Qubit Protein Assay Kit	Q33211	Measurement of Protein
Glycerol	GLY001.500	Preparation of Bacterial Stock
L-Rhamnose monohydrate	83650	Protein Induction
2X Laemmli Sample Buffer	TGS10	SDS-PAGE
Sure Cast 40% Acrylamide	HC2040	SDS-PAGE
Sure Cast Ammonium Persulfate	HC2005	SDS-PAGE
Sure Cast TEMED	HC2006	SDS-PAGE
10X Running Buffer	TGS10	SDS-PAGE

Table 1 (continues)

10% SDS	S100	SDS-PAGE
Tris	TRS001.1	SDS-PAGE, DNA Gel Electrophoresis, Lysis Buffer
Acetic Acid	27225	SDS-PAGE
Methanol	947.046.25000	SDS-PAGE
Coomassie Brilliant Blue (R-250)	0472	SDS-PAGE
Imidazole	56750	Cell Lysis, Protein Purification
NaCI	31434	Cell Lysis, Protein Purification
Sodium Phosphate Monobasic Anhydrous	0471	Preparation of Sodium Phosphate Buffer
Sodium Phosphate Dibasic Dihydrate	04272	Preparation of Sodium Phosphate Buffer
HisPur Ni-NTA Resin	88222	Protein Purification
Amicon Centrifugal Tubes	UFC9010	Concentration of Protein Samples
Ethanol	920.027.2501	Protein Precipitation
Phenol	P4557	Quantification of Glycan Concentration
Sulphuric Acid	1.007.132.500	Quantification of Glycan Concentration

3.1.2. Substrates

Bacterial strains used in the thesis for the recombinant molecular cloning were provided from The Belgian Coordinated Collections of Microorganisms (BCCM/LMG) and The Global Bioresource Center (ATCC). Whey from bovine colostrum was used in the novel model as glycoprotein source. Other enzymes and chemicals (amylase, pepsin, trypsin, and chemicals) used in the conventional model digestion model were also supplied from Sigma-Aldrich.

3.1.3. Laboratory Equipment

All laboratory equipment used in this thesis is listed below (Table 2). For these, the research lab of the Molecular Biology and Genetics Department at Canakkale Onsekiz Mart University (COMU) was used.

Table 2. Laboratory equipment list and brand information

Equipment List	Brand
Autoclave	Nuve
Incubator	Indem Nuve EN 400
Fume Cupboard	Setra
-80 °C Freezer	Operon
Pure Water System	Millipore
Gel Imaging System	Vilber Lourmat
Electrophoresis and Power Supply	Thermo
Electrophoresis System	BioRad
Thermal Cycler	BioRad
NanoQuant	Tecan Infinite M200 PRO
-20 °C Freezer	Bosch
Cooling Centrifuge	Hettich
Heated Stirrer	Wisd
Minicentrifuge	CF-5 Wisd
Vibra Cell/Sonicator	Sonics
Shaker	Benchmark
Vortex	Vortex Genie 2
Block Heater	BSH1002 Benchmark
Qubit 3 Flurometer	Invitrogen
Centrifuge	Beckman Allegra X-15R
Shaker	Benchmark-Incu10LR
Microwave	Bosch
pH Meter	IsoLab
Thermometer	IsoLab
Ice Machine	Izmak

3.2. Method

The general method scheme is shown below (Figure10).

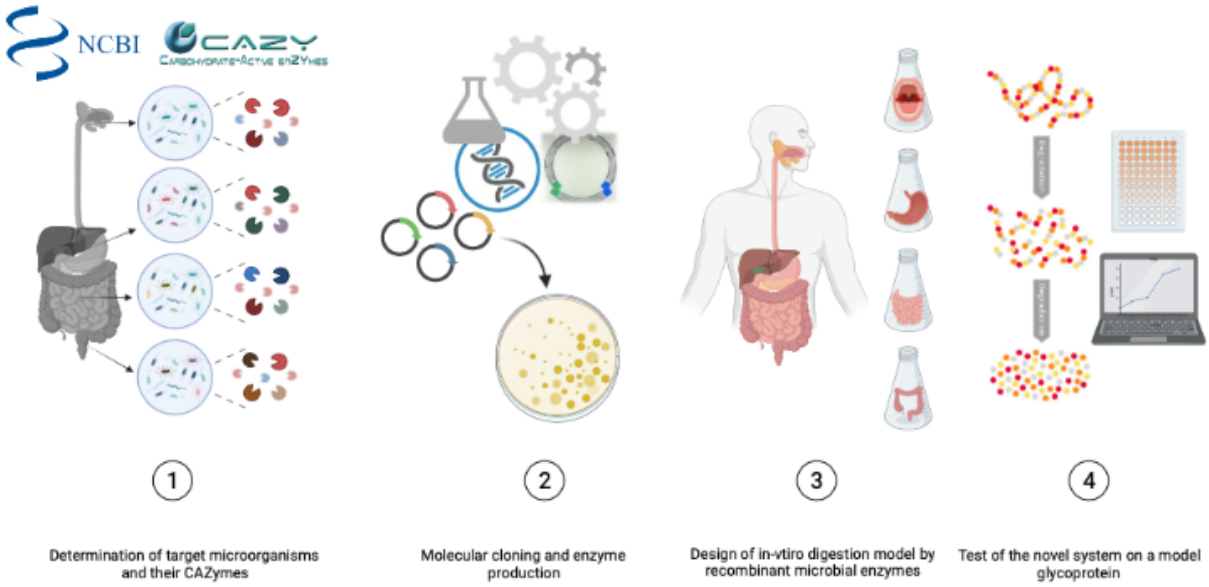


Figure 10. General method scheme.

3.2.1. Determination of Target Microorganisms and Molecular Cloning of Their Specific Enzymes for Novel *In-Vitro* Digestion Model

Determination of Target Microorganisms

All enzymes used in this thesis were searched using Carbohydrate-Active enZYmes (CAZy) and National Center for Biotechnology Information (BLAST). The information about microorganisms and target genes is shown in Table 3.

3.2.2. Primer Design and In-Silico Analysis of Target Genes Prior to Molecular Cloning

In-Silico Signal Peptide/Transmembrane Domain Analyses

Signal peptides and transmembrane domains of target genes were analyzed before the molecular cloning experiment to increase protein expression. The determined signal peptides and transmembrane domains in target genes were excluded from sequences of

genes. The amino acid sequences of target genes determined from the Integrated Microbial Genomes and NCBI database were used in Signal 5.0, and TMHMM 2.0 program for the signal peptide and transmembrane domain analysis, respectively.

Primer Design

The design of primers was performed by excluding some amino acid sequences according to the results of signal peptide and transmembrane domain analysis from online tools. All primers used in this method were designed based on sticky ends for each fusion tag and continue with the specificity to interest genes. Primer sequences' specificity is shown below. Primers were designed to produce different genes used in this thesis. Primer concentration was prepared as 100 μM using sterile water, whereas new stocks were prepared as 10 μM for the PCR amplification step.

Fusion to an N-terminal 6xHis tag (pRham™ N-His Kan Vector):

Forward primer (defined vector sequence includes 6 His codons):

5'-CAT CAT CAC CAC CAT CAC XXX_2 XXX_3 XXX_4 XXX_5 XXX_6 XXX_7 XXX_8

(XXX_2-XXX_8 represents codons 2 through 8 of the target coding region).

Reverse primer (defined vector sequence includes Stop anticodon):

5'-GTG GCG GCC GCT CTA TTA XXX_n XXX_{n-1} XXX_{n-2} XXX_{n-3} XXX_{n-4} XXX_{n-5} XXX_{n-6}

(XXX_n - XXX_{n-6} represents the sequence **complementary** to the last 7 codons of the target coding region).

Fusion to a C-terminal 6xHis tag (pRham C-His Kan Vector):

Forward primer (defined vector sequence includes Start codon):

5'-GAA GGA GAT ATA CAT ATG XXX_2 XXX_3 XXX_4 XXX_5 XXX_6 XXX_7 XXX_8

(XXX_2-XXX_8 represents codons 2 through 8 of the target coding region).

Reverse primer (defined vector sequence includes 6 His anticodons):

5'-GTG ATG GTG GTG ATG ATG XXX_n XXX_{n-1} XXX_{n-2} XXX_{n-3} XXX_{n-4} XXX_{n-5} XXX_{n-6}

(XXX_n - XXX_{n-6} represents the sequence **complementary** to the last 7 codons of the target coding region).

Fusion to a N-terminal SUMO 6xHis tag (pRham N-his SUMO Kan Vector):

Forward primer (defined sequence includes the last 6 codons of SUMO):

5'- CGC GAA CAG ATT GGA GGT XXX_2 XXX_3 XXX_4 XXX_5 XXX_6 XXX_7 XXX_8

(XXX_2-XXX_8 represents codons 2 through 8 of the target coding region).

Reverse primer (vector sequence includes Stop anticodon):

5'-GTG GCG GCC GCT CTA TTA XXX_n XXX_{n-1} XXX_{n-2} XXX_{n-3} XXX_{n-4} XXX_{n-5} XXX_{n-6}

XXX_n - XXX_{n-6} represents the **reverse complement** of the sequence of the last 7 codons of the target coding region. XXX_n is the reverse complement of the final codon of the protein. The stop codon of the target gene need not be included, as the vector encodes stop codons.

Figure 11. Primer design according to molecular cloning kit A) Fusion to an N-terminal 6xHis tag, B) Fusion to a C-terminal 6xHis tag (Lucigen).

3.2.3. Molecular Cloning

The molecular cloning experiment was performed using an advanced kit which is Expresso Rhamnose Cloning and Expression System (Lucigen). This molecular cloning kit helps to achieve faster and more reliable results in comparison to other molecular cloning methods. Expresso Rhamnose Cloning and Expression System is also named an *in-vivo* cloning system since all process takes place in cells. There is not any enzymatic ligation process, amplified gene and vector in the kit can be easily mixed with competent cells. The vector in the kit is 18 nucleotides long and has sticky ends on both sides. Primers used in this method were designed based on sticky ends, which provides a strong binding of primers to the template. When molecular cloning is performed, protein production can be increased with a promoter; rhamnose. In addition, the protein purification method was conducted by immobilized metal affinity chromatography (IMAC) using NTA-Ni charged columns with 6xHistidine in three distinct vectors including pRhamTM N-His SUMO, pRhamTM N-His, and pRhamTM C-His).

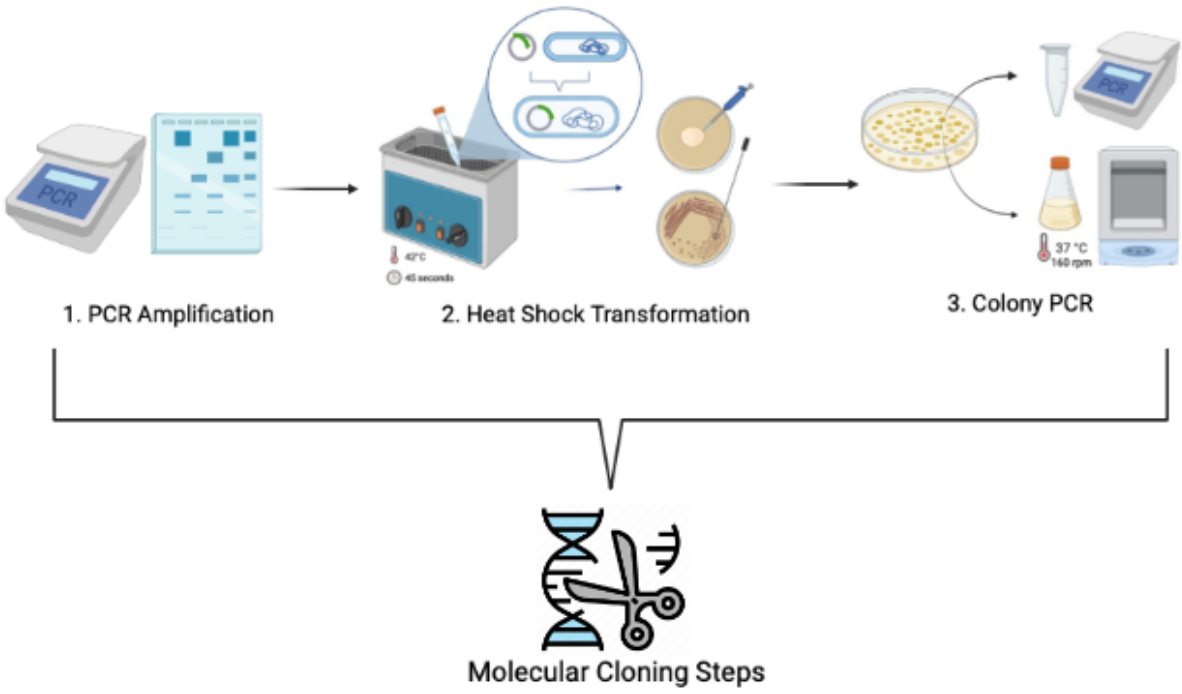


Figure 12. Molecular cloning steps.

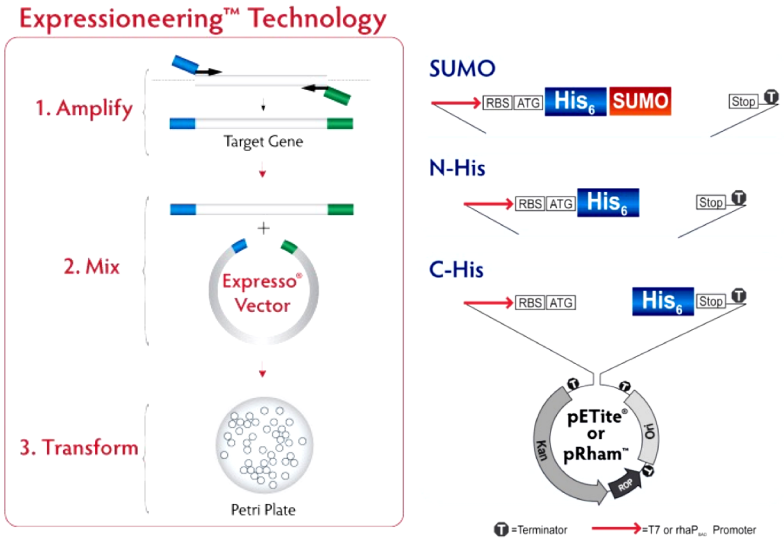


Figure 13. Molecular cloning kit (Expresso Rhamnose Cloning and Expression System) technology and its vectors (Lucigen).

PCR Amplification

The interest genes were amplified using PCR. The amplification step was performed in 50 µL containing 2 µL template DNA, 1 µL forward primer, 1 µL reverse primer, 0.2 µM, 25 µL Master Mix, 21 µL DNase/RNase free distilled water. Once the mixture was ready, PCR tubes were placed into a thermal cycler. PCR steps was as shown below:

Stage	Temperature	Time - Cycle
First Denaturation	95 °C	5 minutes
Release of Genomic DNA	95 °C	20 seconds – 40 cycles
Annealing	60 °C	30 seconds
Elongation	72 °C	1 minute
Final Extension	72 °C	10 minutes

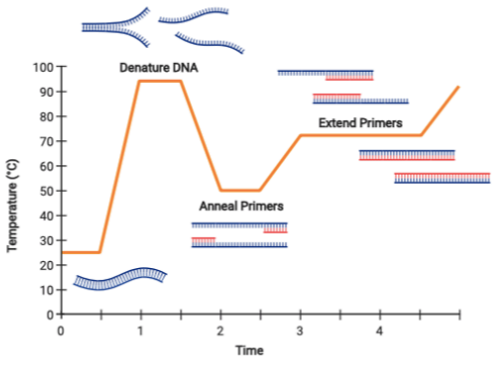


Figure 14. PCR stages.

Agarose Gel Electrophoresis

Agarose gel electrophoresis was used to control PCR products. Firstly, Safe Red loading dye was mixed with PCR products as well as DNA ladder (the ratio was 1:5). The experiment was run on 1% agarose gel at 100 V for 60 min using 1X TBE buffer. After gel electrophoresis, PCR products were visualized using a gel documentation system ST4 1100 (Vilber Lourmat, France). All PCR products' concentrations were measured with Qubit 3 Fluorometer using its dsDNA assay kit.

Preparation of Lysogeny Broth (LB) Media

To prepare LB agar media for the molecular cloning step, 6 g agarose and 12.5 g LB were mixed with 500 mL dH_2O and autoclaved at 121 °C during 20 min. After autoclaving, 15 mg kanamycin (30 µg/mL) was dissolved in 1 mL dH_2O and transferred into the bottle containing 500 mL LB agar media. Then, the media was poured into plates. To prepare LB medium for colony PCR step, 10 g LB was mixed with 400 mL distilled water and autoclaved at 121 °C for 20 min. 12 mg kanamycin (30 µg/mL) was dissolved in 1 mL distilled water and transferred into the bottle containing 400 mL LB medium. The media was stored at 4 °C until the colony PCR step.

Heat Shock Transformation

First, the recovery medium in the kit which is used to heal cells rapidly after transformation was taken from -80°C and placed into 37°C before the cloning of E. cloni 10G cells at -80°C thawed at the ice. His-tagged PCR products' concentration was mixed with 2 µL of pRhamTM Vector DNA and mixed with 40 µL of E. cloni 10G cell. The prepared mixture was introduced to falcon tubes (15 mL) which were put on ice for 30 min. The thermal shock process was in a 42 °C water bath for 45 seconds for vector insertion and PCR product into the cell. Samples were taken to the ice for 2 min to close competent cell pores. After 2 min, 960 µL recovery medium was introduced into each tube, and tubes were incubated at 37 °C during 1 hour at 250 rpm. Then, all samples were firstly spread as 100 µL to LB agar plates and remained parts centrifugated at 4000 rpm for 10 min. The pellet part of the samples was dissolved with 100 µL recovery medium by pipetting then spread

into LB agar plates. E. cloni cells were used as a negative control on a different LB agar plate and all were incubated at 37°C overnight.

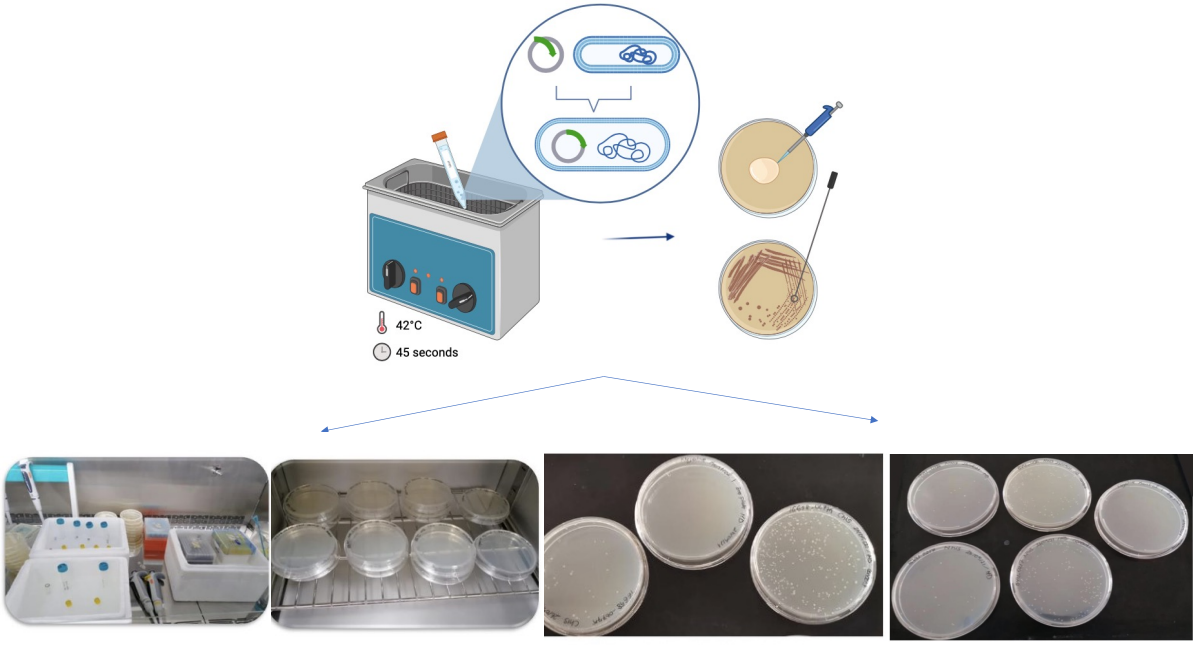


Figure 15. Heat shock steps.

Colony PCR

Colony PCR was performed to confirm the transformants whether take the recombinant genes or not. After overnight incubation of cells on LB agar plates, colonies were selected randomly. Target genes were amplified by PCR with the sequencing primers in the Expresso molecular cloning kit. The half part of each colony was transferred into a PCR tube as well as used as a template for PCR amplification, while the other part was inoculated to LB kanamycin liquid media (5 mL) and incubated at 250 rpm 37 °C overnight. PCR process was performed as previously mentioned in step 3.2.5, the only difference was primers were sequencing primers provided with the kit.

Agarose gel electrophoresis was performed to control PCR products after colony PCR using Safe Red loading dye, PCR products, as well as DNA ladder. The gel was run at 1% agarose at 100 V during 60 min using 1 X TBE buffer. ST4 1100 (Vilber Lourmat, France) was used to visualize samples. According to the results, successful transformants including recombinant genes were determined and their cultures in 5 mL LB were used to

prepare glycerol stocks. 500 µL 60% glycerol and 1500 µL culture were mixed in cryotubes to prepare 15% glycerol stocks. All stocks were stored at -80°C.

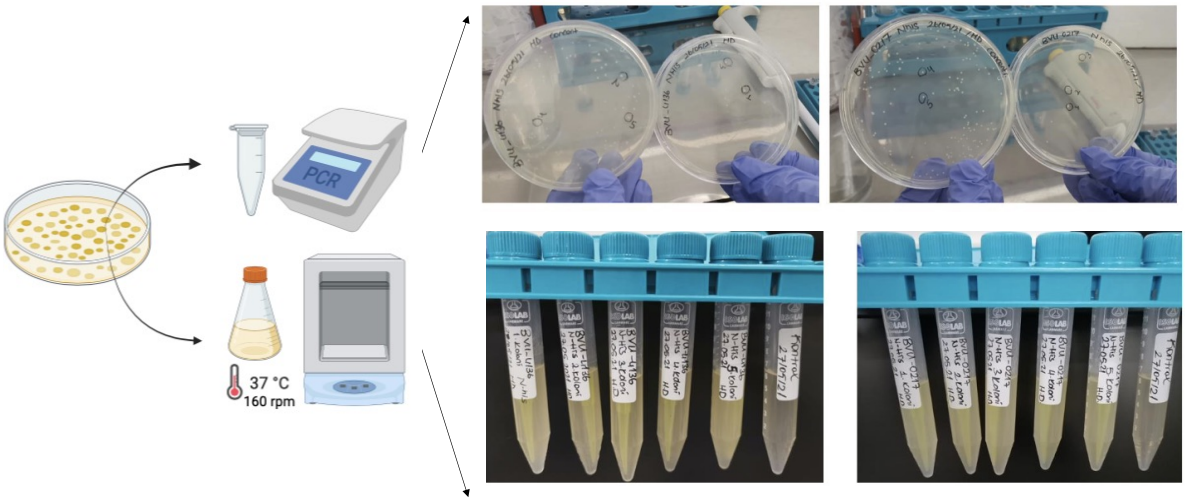


Figure 16. Colony PCR.

3.2.4. Protein Production and Purification

L-Rhamnose Induction

Firstly, a preculture was prepared by inoculating bacterial stocks into an 8 mL LB medium prepared. To prepare 20 % L-rhamnose, 0.5 g L-rhamnose was dissolved in 2.5 mL dH_2O and stored at -20°C before use. Prepared cultures were incubated at 37 °C overnight at 160 rpm. The following day, 2.5 mL culture was transferred to 250 mL fresh LB medium (1:100). Then, cultures in the fresh medium were incubated at 37 °C, 160 rpm about 4-5 hours (until OD reaches 0.5-0.6). The optical density of cells was measured at 600 nm with a spectrophotometer. When it was reached to 0.5-0.6, 2.5 mL 20% rhamnose as final concentration 0.2% were added into 250 mL LB. Cultures were incubated overnight under the conditions of 160 rpm, 24°C. Following the incubation, they were centrifugated at 4000 rpm, 4 °C for 20 min. Pellet parts were stored at -20°C.

Cell Lysis

The pellets were placed to -80°C to freeze and taken to room temperature to thaw. Then, they were washed with five mL dH_2O and centrifugated at 4000 rpm, 4 °C for 15 min. Pellet parts after centrifugation were dissolved with 6300 µL lysis buffer and 63 µL protease inhibitor (1:100) for each 50 mL culture pellet. The prepared samples were incubated on ice during 30 min by vortexing in every 10 min. After 30 min, samples were sonicated with a sonicator; the pulse mode of a sonicator was cycled on 10 s and cycled off 59 s, and the amplitude was 37%. The sonication was performed as six pulses for 10 s with a one min cooling step. After that, samples were centrifugated at 4000 rpm, 4 °C for 30 min. Supernatant parts after centrifugation were used in protein purification and taken 100 µL for SDS-PAGE. Cell lysates' concentration was measured with a Qubit 3 Fluorometer.

Protein Purification

1 mL Ni-NTA resin was used in the protein purification steps, it was centrifugated at 700 g for 2 min and its buffer was removed. 2 mL equilibration buffer was mixed with the resin and centrifugated at 700 g during 2 min and again süpernatant part were discarded. Cell lysates were also mixed with equilibration buffer (1:1) and added to the falcons containing resin. Prepared samples were incubated at room temperature, 150 rpm for 30 min. After that, samples were centrifugated at 700 g for 2 min, and süpernatant parts were removed. The remained resin parts were washed with 5 mL wash buffer and centrifugated at 700 g for 2 min, this step was repeated until the samples' concentrations decrease to the baseline so there are no potential contaminants. When it comes to elution steps, 1 mL elution buffer was added to tubes and centrifugated at 700 g for 2 min. The elution was repeated three times and each one was taken into sterile tubes. 100 µL of each was taken for the SDS-PAGE experiment and their concentrations were measured with a Qubit 3 Fluorometer. Eluted samples were collected in an Amicon Centrifugal Tube based on their molecular weight, and they were concentrated. Purified samples were stored at -80°C.

Figure 17. Protein purification steps.

SDS-PAGE Analysis

SDS-PAGE was performed according to the Laemmli protocol and Bio-Rad protein gel system. First of all, 4% (stacking) and 12% (separating) gels were prepared. Stacking gel (4%) was prepared using 40% Acrylamide/Bisacrylamide, 1 M Tris pH 6.8, 10% SDS, 10% Ammonium Persulfate, TEMED, distilled water) and separating gel (12%) was prepared using 40% Acrylamide/Bisacrylamide, 1.5 M Tris pH 8.8, 10% SDS, 10% Ammonium Persulfate. SeeBlue™ Pre-stained Protein Standard (Invitrogen) was used as the standard ladder. 10 µL of each sample was mixed with 10 µL of Laemmli Sample Buffer (2X) was added. After the prepared samples were incubated at 95°C for 5 minutes, they were loaded into the prepared gel. Tris-Gly SDS Running Buffer was the running buffer as well as samples were run at 80-120 V conditions. Proteins separated based on their molecular weights were incubated with Coomassie Brilliant Blue at 50 rpm for 30 minutes. Then, the gel was destained with the Destaining solution (50% distilled water, 40% Methanol, 10% Glacial Acetic Acid) and gel images were taken.

3.2.5. Integration of Obtained Microbiome-Based Enzymes to a Conventional *In-Vitro* Digestion Model

For a conventional model, i*n-vitro* digestion protocol was prepared according to the INFOGEST model (Minekus et al., 2014). 2 µL of each recombinant enzyme was integrated into appropriate phases during *in-vitro* digestion. The last colon part where the microbial mechanism of digestion takes place included only microbial enzymes and performed at 37 °C, pH 8. For the experiment, three groups were planned; first was experimental group including whey, human digestion fluids, enzymes, and microbial enzymes, second group was control 1 consisting of human digestion fluids and enzymes, no microbial enzymes, and third group was control 2 including only digestion fluids, no enzyme.

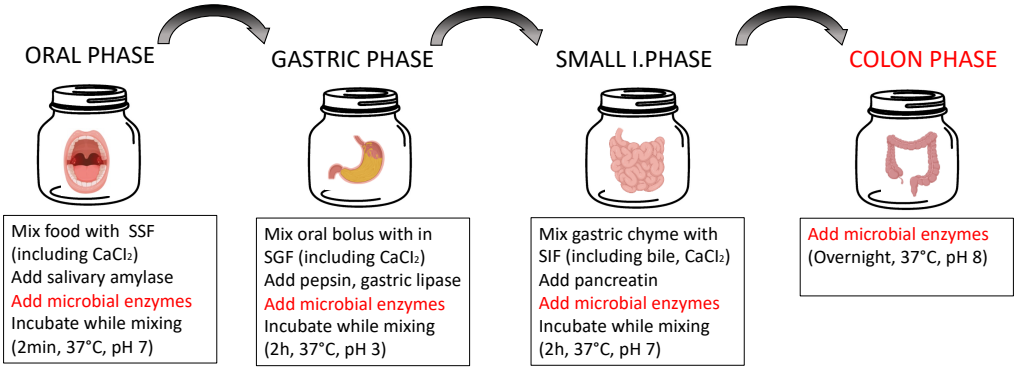


Figure 18. Design of the novel model by integrating microbial enzymes through a conventional *in-vitro* digestion system.

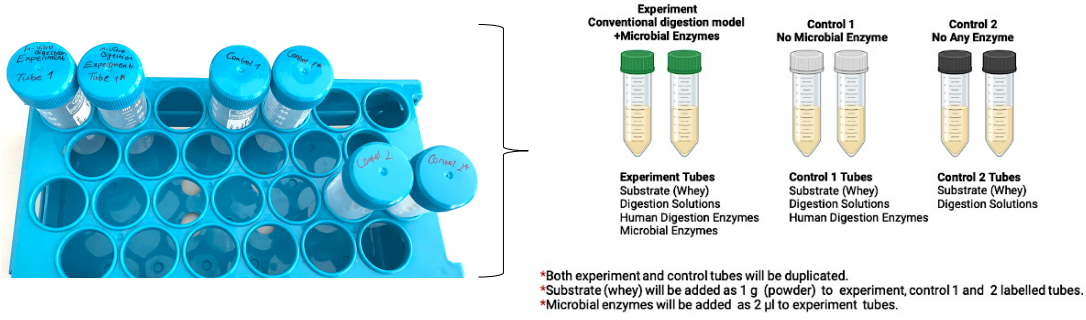


Figure 19. Experiment and control groups that are used in the test of the novel system on a glycoprotein source.

3.2.6. Digestion of a Glycoprotein Source by Using Novel *In-Vitro* Digestion Model

Whey from bovine colostrum (1 g) which was filtered using 10 kDa Amicon tube to remove contaminants including free oligosaccharides and lactose was used as a glycoprotein source that includes a high concentration of glycans. All digestion solutions including simulated salivary fluid (SSF), simulated gastric fluid (SGF), and simulated intestinal fluid (SIF) were prepared based on the content given Appendix 3). All digestion fluids and enzymes were preincubated at 37 °C before the experiment. 2mL of sample were removed after each phase and kept at -20°C until the next usage. In general, in the oral phase, a 1 g whey from bovine colostrum which does not contain lactose and free oligosaccharides, was added to a 3.5 mL SSF stock solution and mixed. Then, human salivary a-amylase (EC 3.2.1.1, 15000 U mL^{-1}) and 25 µL, 0.3 M $CaCI_2$ were added to the mixture. Finally, 975 µL distilled water was added to the mixture and mixed well. The incubation time for the oral phase took 2 minutes at 37°C shaking by hand. In the gastric phase, the final ratio of food to SGF solution was at 50:50 (v/v) after adding other components. A 10 mL liquid sample was added to a 7.5 mL SGF solution and then 1.6 mL pepsin (from porcine gastric mucosa 3200-4500 U mg^{-1}) was mixed with the mixture. 5 µL, 0.3 M $CaCI_2$, and 1 M HCI for keeping pH at 3.0 and 0.695 µL distilled water were added to the final mixture. The incubation time for the gastric phase was 2 hours at 37°C, 100 rpm. In the intestinal phase, the final ratio of gastric chyme to SIF stock solution was at 50:50 (v/v) after adding other chemicals and distilled water. 1 M NaOH was required to adjust pH at 7. 20 mL of gastric chyme from the previous phase was mixed with 11 mL of SIF solution. 5 mL pancreatin solution (from porcine pancreas, 800 U mL^{-1}), 2.5 mL, 160 mM fresh bile, 40 µL, and 0.3 M $CaCI_2$ were added to the mixture. Finally, 0.15 mL of 1 M NaOH was added to adjust pH at 7.0 and 1.31 mL of distilled water was mixed with the final solution. The incubation time for intestinal digestion took 2 hours at 37°C, 100 rpm. During experiment, 2 µL of each recombinant enzyme was integrated into appropriate phases. The last colon part where the microbial mechanism of digestion takes place includes only microbial enzymes and it was incubated overnight under the conditions of 37°C, pH 8, and 100 rpm.

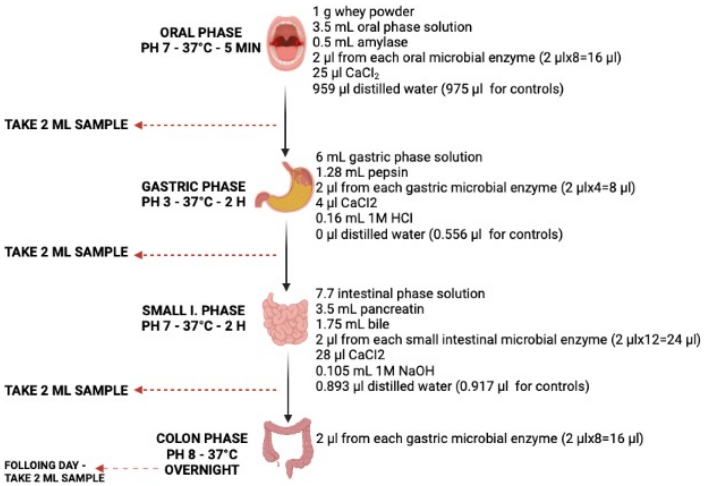


Figure 20. Flow of novel *in-vitro* digestion model.

The samples taken from each phase (1 mL) was mixed with cold ethanol (1:4; v/v) and incubated at -20°C for 1 h to precipitate proteins. After the incubation, samples were centrifugated during 30 min under the conditions of 4°C and 4000 rpm. The supernatant parts were removed and dried using a vacuum evaporator machine. The dry samples were dissolved with 600 µL dH_2O and used in phenol sulphuric acid assay to be quantified. As for the phenol-sulphuric acid assay, each 25 µL sample was mixed firstly with 25 µL phenol (1:1; v/v) and then 125 µL sulphuric acid in a plate. After the 20 min incubation at room conditions, concentrations were measured at OD_{490nm}. Data was analyzed statistically according to the one-way ANOVA variance analysis along with Tukey's multiple comparisons statistical test to assess the statistical significance of the data at $p<0.05$ using NCSS 12 statistical software.

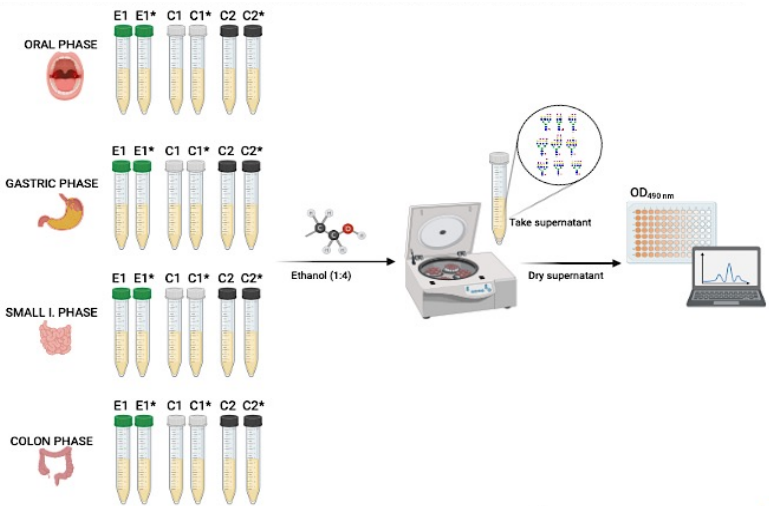


Figure 21. Phenol-sulphuric acid assay for quantification of released glycans.

CHAPTER 4

RESEARCH FINDINGS

4.1. Bioinformatic Analysis for the Determination of Target Genes

Table 3. Microorganisms and their genes are recombinantly cloned and produced

	GeneBank ID / Accession Number **Locus tag** **Target Phase**	**Microorganism**	**bp**	**kDa**
1	ATP38112.1 CR531_08240 Oral – Mouth	*Lactobacillus salivarus subsp. salivarius (Ligilactobacillus salivarius)* ATCC 11741	2238 bp	97.85 kDa
2	ATP36889.1 CR531_01355 Oral – Mouth	*Lactobacillus salivarus subsp. salivarius (Ligilactobacillus salivarius)* ATCC 11741	600 bp	31.35 kDa
3	ATP37244.1 CR531_03290 Oral – Mouth	*Lactobacillus salivarus subsp. salivarius (Ligilactobacillus salivarius)* ATCC 11741	1659 bp	76.51 kDa
4	SQH52440.1 NCTC11324_01490 Oral – Mouth	*Streptococcus intermedius* ATCC 27335	1398 bp	66.16 kDa
5	ATP38122.1 CR531_08290 Oral – Esophagus	*Lactobacillus salivarus subsp. salivarius (Ligilactobacillus salivarius)* ATCC 11741	978 bp	49.21 kDa
6	ATP37586.1 CR531_05275 Oral – Esophagus	*Lactobacillus salivarus subsp. salivarius (Ligilactobacillus salivarius)* ATCC 11741	2196 bp	95.71 kDa
7	SQH51076.1 NCTC11324_00070 Oral – Esophagus	*Streptococcus intermedius* ATCC 27335	1806 bp	82.16 kDa
8	SQH51655.1 NCTC11324_00672 Oral – Esophagus	*Streptococcus intermedius* ATCC 27335	1884 bp	84.49 kDa

9	CAR86329.1 LGG_00434 Gastric	*Lactobacillus rhamnosus GG*	1815 bp	67.124 kDa
10	AAO77566.1 BT_2459 Gastric	*Bacteroides thetaiotaomicron* ATCC 29148	1863 bp	74.62 kDa
11	ABR41745.1 BVU_4143 Gastric	*Bacteroides vulgatus* ATCC 8482	1641 bp	62.812 kDa
12	AAO75562.1 BT_0455 Gastric	*Bacteroides thetaiotaomicron* ATCC 29148	1635 bp	63.299 kDa
13	ACD04858.1 Amuc_1032 Small Intestine – Duodenum	*Akkermansia muciniphila* ATCC BAA-835	1506 bp	69.25 kDa
14	BAQ98211.1 BBBF_1004 Small Intestine – Duodenum	*Bifidobacterium bifidum* ATCC 29521	2529 bp	102.16 kDa
15	BAQ97897.1 BBBF_0690 Small Intestine – Duodenum	*Bifidobacterium bifidum* ATCC 29521	1197 bp	53.11 kDa
16	ERK41518.1 HMPREF0495_02198 Small Intestine – Duodenum	*Levilactobacillus brevis (Lactobacillus brevis*) ATCC14869	1554 bp	69.92 kDa
17	ACJ51836.1 Blon_0732 Small Intestine – Jejunum	*Bifidobacterium longum subsp. infantis* ATCC 15697	1956 bp	75.847 kDa
18	ACJ53413.1 Blon_2355 Small Intestine – Jejunum	*Bifidobacterium longum subsp. infantis* ATCC 15697	1956 bp	75.284 kDa
19	QQA29671.1 I6G58_17045 Small Intestine – Jejunum	*Bacteroides uniformis FDAARGOS_901* ATCC 8492	1668 bp	65.26 kDa
20	BAQ30021.1 BBKW_1886	*Bifidobacterium catenulatum subsp. kashiwanohense* JCM 15439	2013 bp	77.293 kDa

	Small Intestine – Jejunum			
21	ABR38247.1 BVU_0537 Small Intestine – Ileum	*Bacteroides vulgatus* ATCC 8482	2256 bp	87.48 kDa
22	SQF24907.1 NCTC12958_01101 Small Intestine – Ileum	*Streptococcus thermophilus* ATCC 19258	1086 bp	51.45 kDa
23	SQF25661.1 NCTC12958_01892 Small Intestine – Ileum	*Streptococcus thermophilus* ATCC 19258	1329 bp	63.13 kDa
24	SQF24918.1 NCTC12958_01112 Small Intestine – Ileum	*Streptococcus thermophilus* ATCC 19258	1284 bp	56.76 kDa
25	ACD04701.1 Amuc_0868 Large Intestine – Proximal Colon	*Akkermansia muciniphila* ATCC BAA-835	1590 bp	71.63 kDa
26	ACD04208.1 Amuc_0369 Large Intestine – Proximal Colon	*Akkermansia muciniphila* ATCC BAA-835	1941 bp	84.02 kDa
27	BAQ97280.1 BBBF_0073 Large Intestine – Proximal Colon	*Bifidobacterium bifidum* ATCC 29521	1323 bp	61.21 kDa
28	CAH09389.1 BF9343_3608 Large Intestine – Proximal Colon	*Bacteroides fragilis* ATCC 25285	1590 bp	61.335 kDa
29	ABR38963.1 BVU_1273 Large Intestine – Distal Colon	*Bacteroides vulgatus* ATCC 8482	1038 bp	41.143 kDa
30	AAO76145.1 BT_1038 Large Intestine – Distal Colon	*Bacteroides thetaiotaomicron* ATCC 29148	1023 bp	40.573 kDa
31	ACJ53522.1 Blon_2468 Large Intestine – Distal Colon	*Bifidobacterium longum subsp. infantis* ATCC 15697	1450 bp	56.12 kDa

32	ACJ51376.1 Blon_0248 Large Intestine – Distal Colon	*Bifidobacterium longum subsp. infantis* ATCC 15697	1350 bp	52.641 kDa

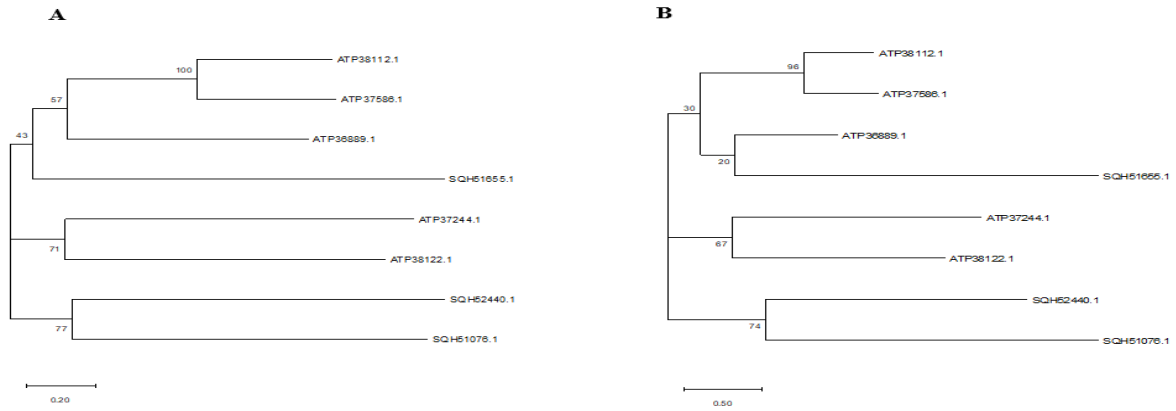


Figure 22. Neighbor-Joining (A) and Maximum Likelihood (B) phylogenetic trees of target enzymes from the oral phase.

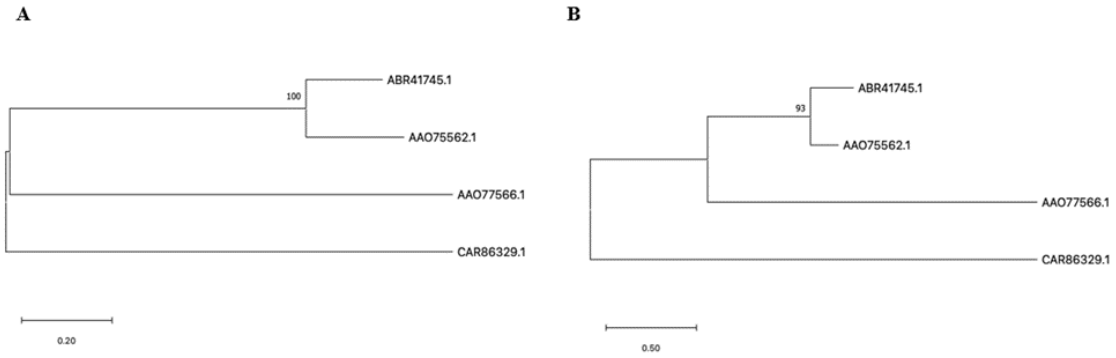


Figure 23. Neighbor-Joining (A) and Maximum Likelihood (B) phylogenetic trees of target enzymes from the gastric phase.

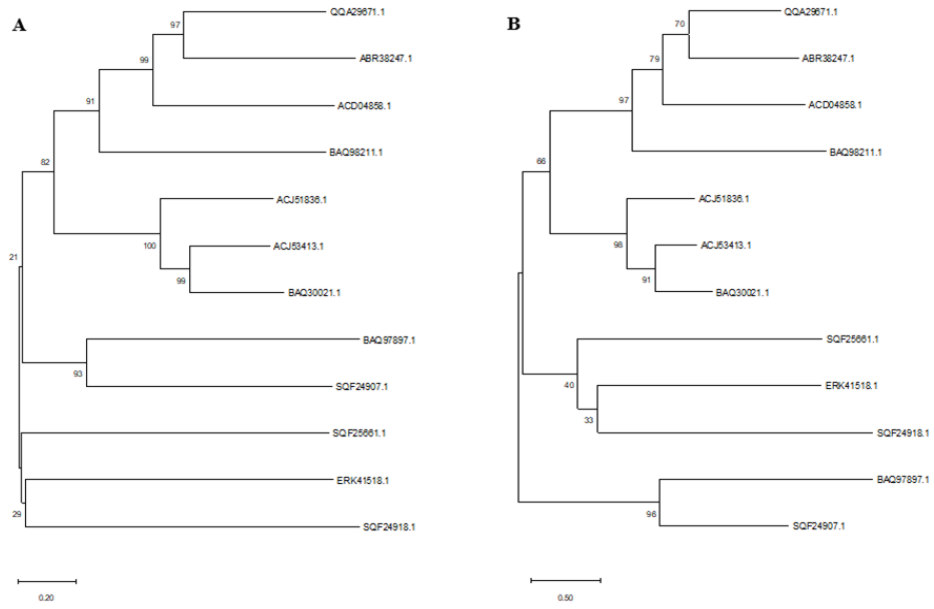


Figure 24. Neighbor-Joining (A) and Maximum Likelihood (B) phylogenetic trees of target enzymes from the small intestine phase.

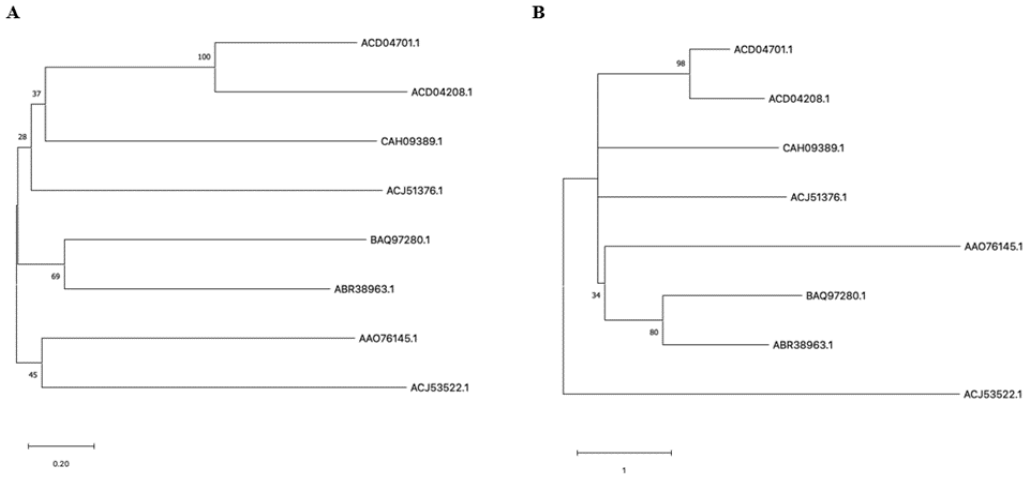


Figure 25. Neighbor-Joining (A) and Maximum Likelihood (B) phylogenetic trees of target enzymes from the colon phase.

4.2. PCR Amplification of Target Genes

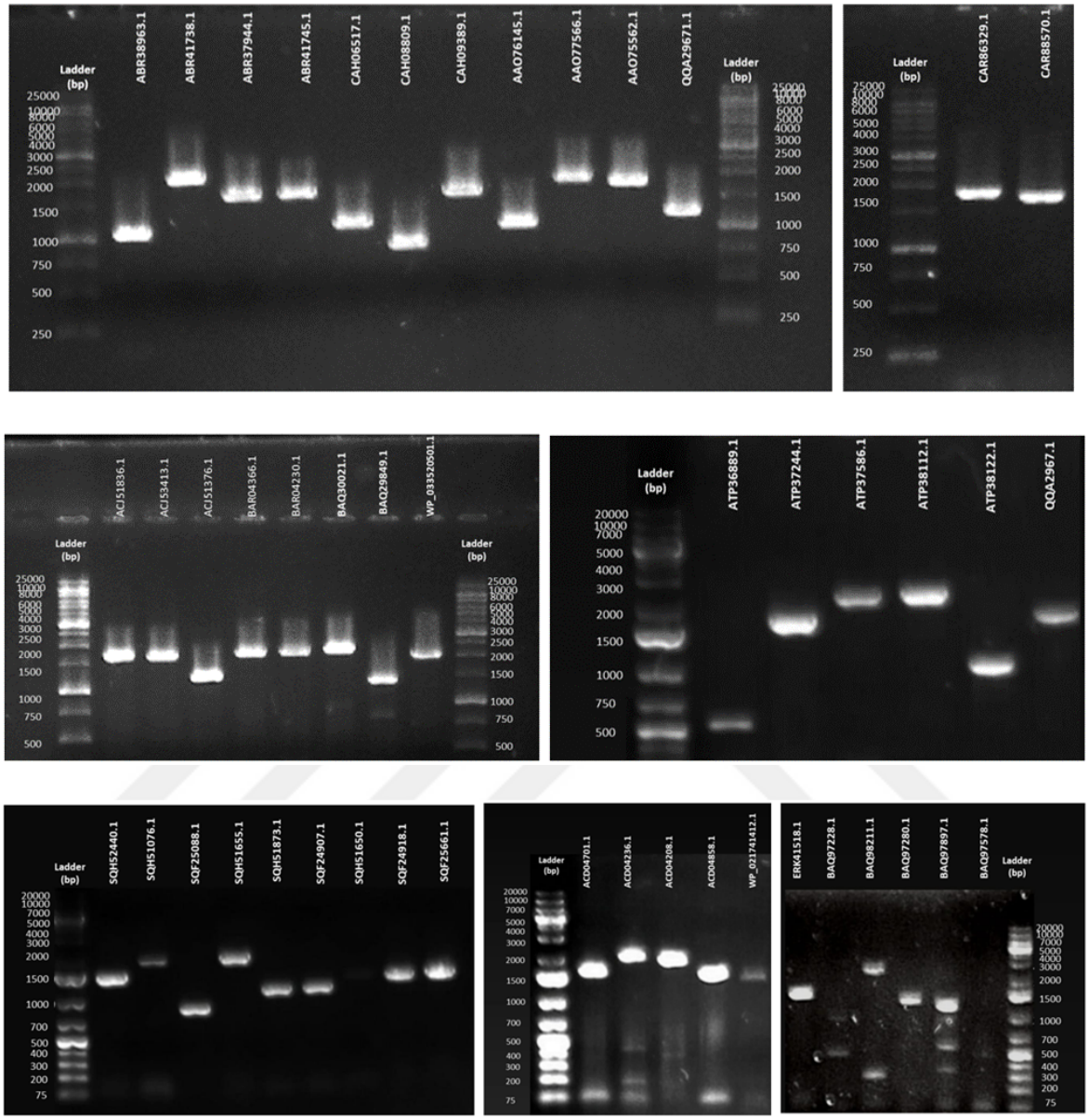


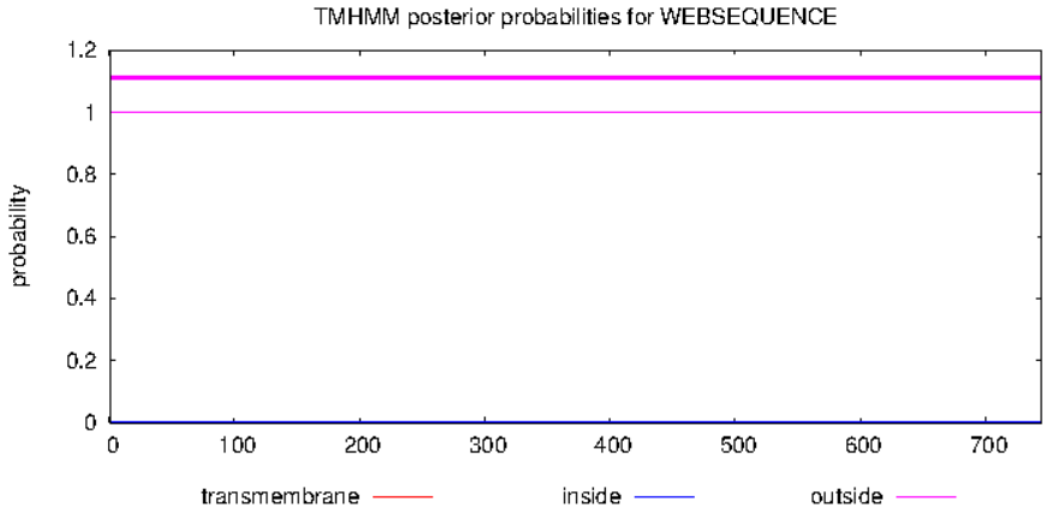
Figure 26. Agarose gel electrophoresis results after PCR amplification of target genes.

4.3. Signal Peptide/Transmembrane Domain Analysis and Primer Information of Each Target Enzyme

1. **ATP38112.1 – CR531_08240**
   Lactobacillus salivarus subsp. Salivarius (Ligilactobacillus salivarius) **ATCC 11741**

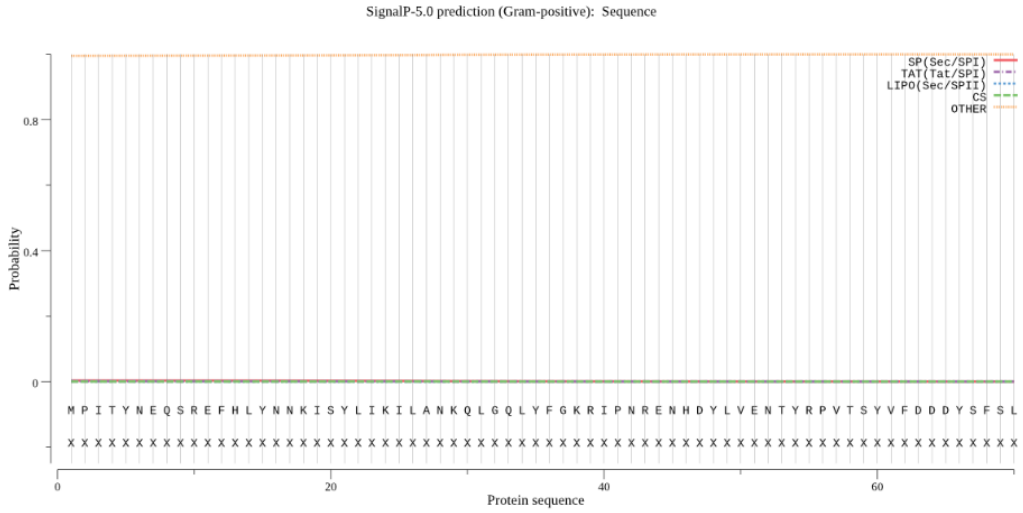
TMHMM result

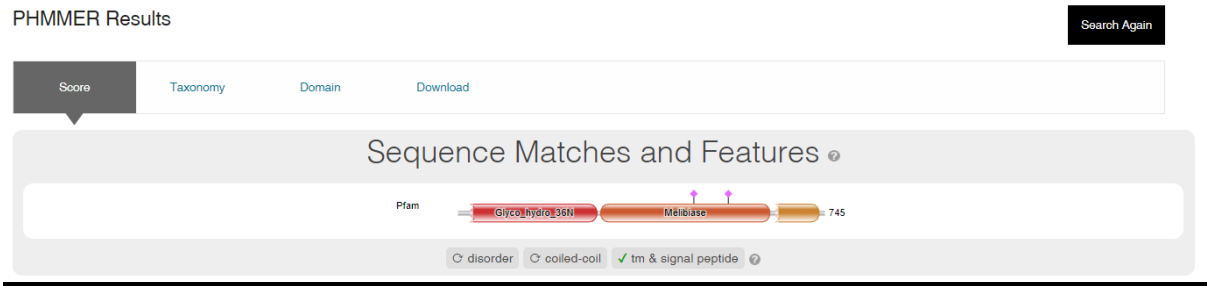
```
# WEBSEQUENCE Length: 745
# WEBSEQUENCE Number of predicted TMHs:  0
# WEBSEQUENCE Exp number of AAs in TMHs: 0.01569
# WEBSEQUENCE Exp number, first 60 AAs:  0.01367
# WEBSEQUENCE Total prob of N-in:        0.00103
WEBSEQUENCE     TMHMM2.0        outside      1   745
```



Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.0029	0.0004	0.0006	0.996

Download: PNG / EPS / Tabular





-Primer

#6. CR531_08240_FS

5'-CGC GAA CAG ATT GGA GGT CCA ATT ACA TAT AAC GAA CAA AGC-3' - 42 bp

Oligo No	220222-2-6	GC	%43	Tm(Basic)	66.46°C	Total nmol	35.66nmol
Skala	50 nmol	MW	12957.52	Conc	1650.43ng/µl	Total ng	462121.34ng
Saflastirma	DSLT	A260	54.8	OD	15.3	100 uM stok - ul TE	356.6

#2. CR531_08240_RS

5'-GTG GCG GCC GCT CTA TTA GTC ACT TAA TTC CCC ACA-3' - 36 bp

Oligo No	220223-2-21	GC	%53	Tm(Basic)	67.86°C	Total nmol	42.30nmol
Skala	50 nmol	MW	10947.14	Conc	1653.65ng/µl	Total ng	463021.03ng
Saflastirma	DSLT	A260	49.9	OD	14.0	100 µM stok - µl TE	423.0

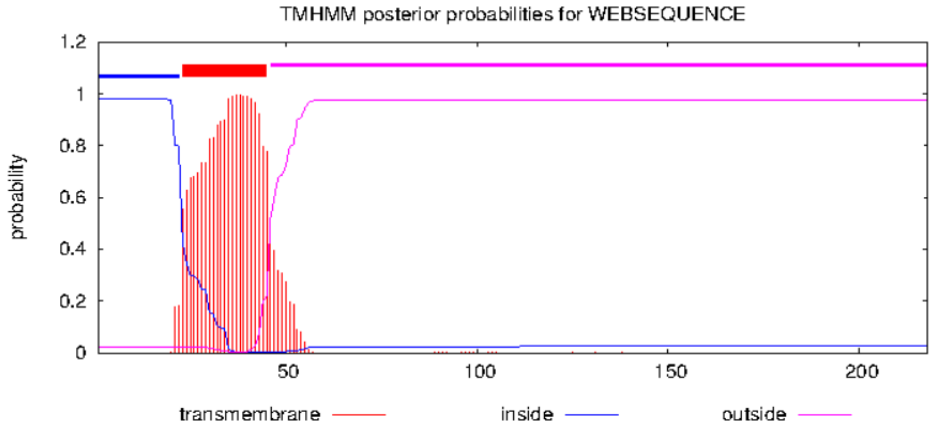
Figure 27. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of ATP38112.1.

2. ATP36889.1 – CR531_01355

***Lactobacillus salivarus subsp. Salivarius (Ligilactobacillus salivarius*) ATCC 11741**

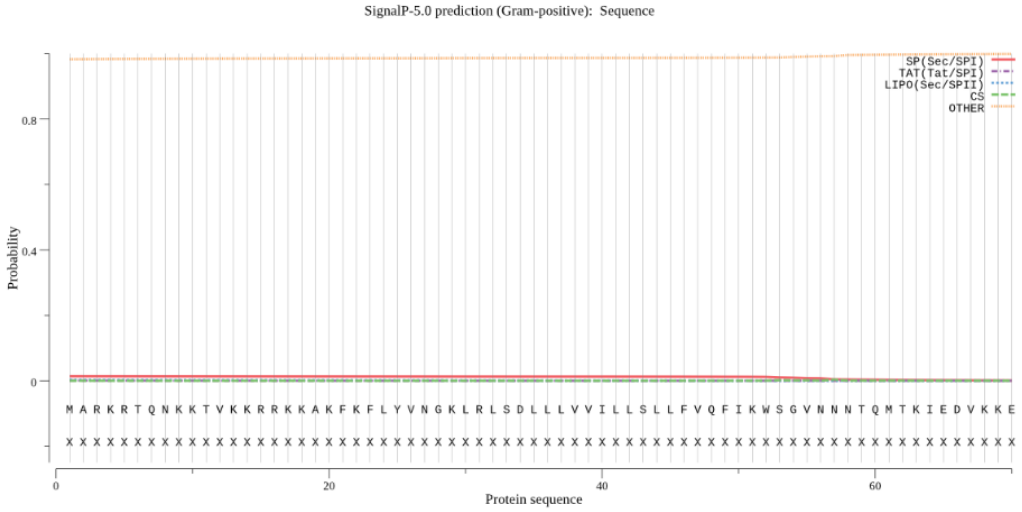
TMHMM result

```
# WEBSEQUENCE Length: 218
# WEBSEQUENCE Number of predicted TMHs:  1
# WEBSEQUENCE Exp number of AAs in TMHs: 22.23448
# WEBSEQUENCE Exp number, first 60 AAs:  22.20511
# WEBSEQUENCE Total prob of N-in:        0.97919
# WEBSEQUENCE POSSIBLE N-term signal sequence
WEBSEQUENCE     TMHMM2.0        inside       1    22
WEBSEQUENCE     TMHMM2.0        TMhelix     23    45
WEBSEQUENCE     TMHMM2.0        outside     46   218
```



Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.0132	0.0002	0.0019	0.9846

Download: PNG / EPS / Tabular



-Primer

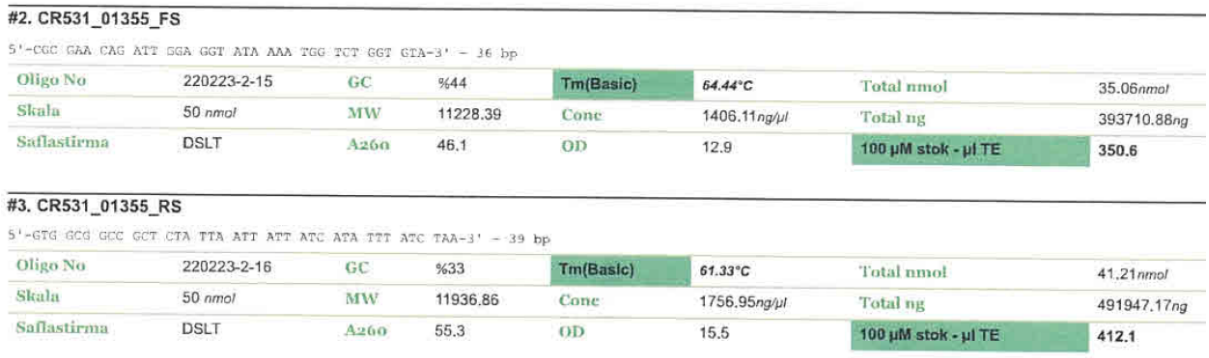


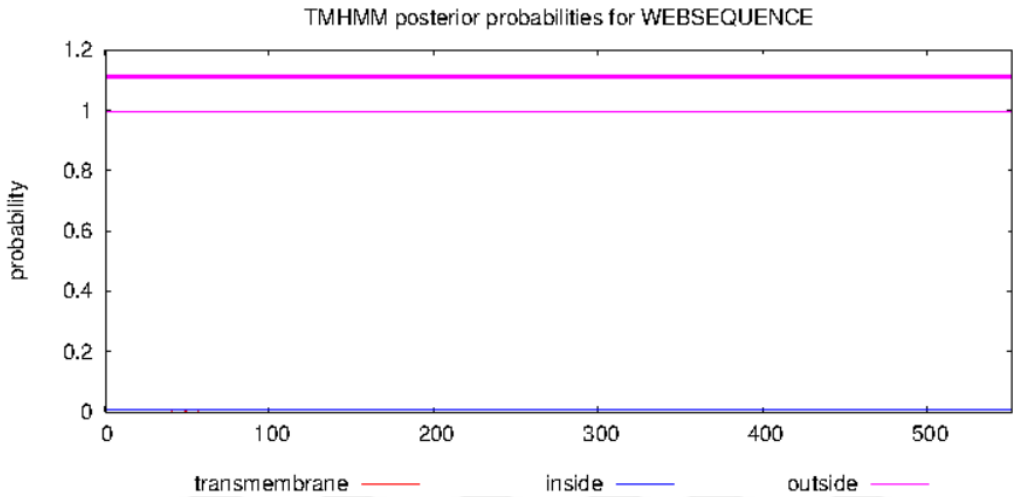
Figure 28. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of ATP36889.1.

3. ATP37244.1 – CR531_03290

Lactobacillus salivarus subsp. Salivarius (Ligilactobacillus salivarius) **ATCC 11741**

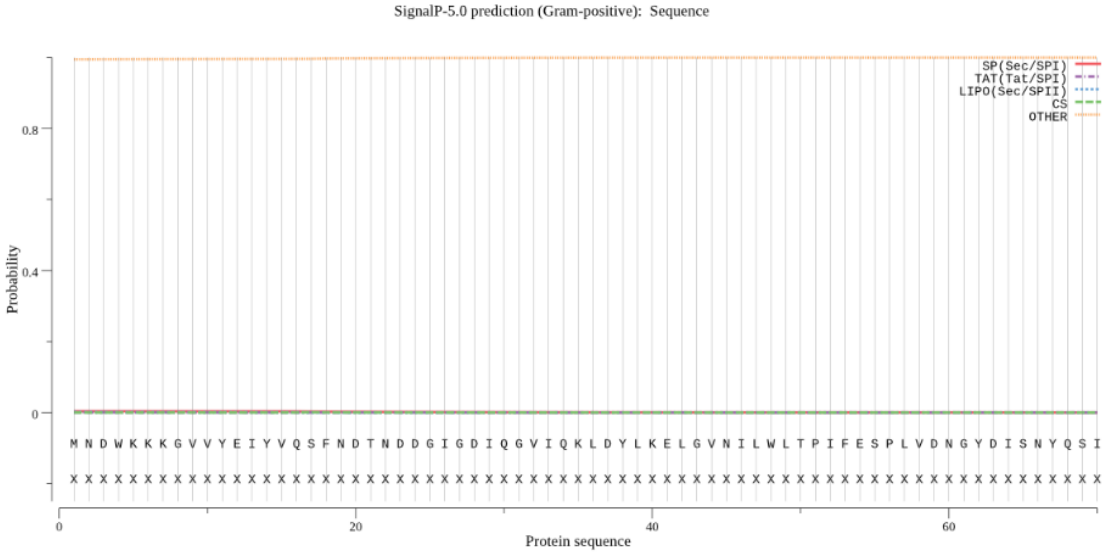
TMHMM result

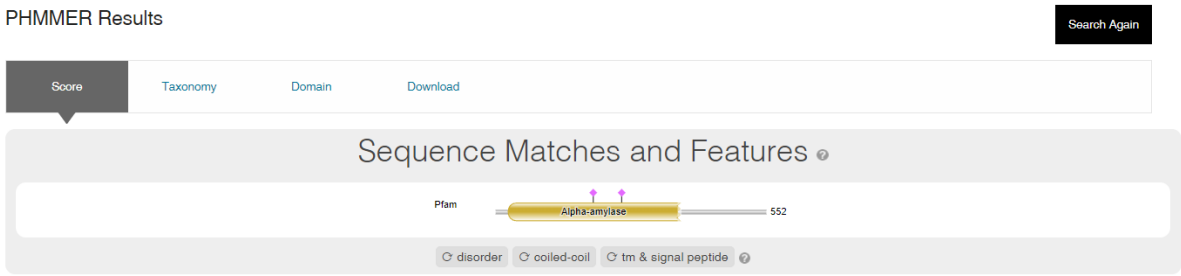
```
# WEBSEQUENCE Length: 552
# WEBSEQUENCE Number of predicted TMHs:  0
# WEBSEQUENCE Exp number of AAs in TMHs: 0.00944999999999998
# WEBSEQUENCE Exp number, first 60 AAs:  0.0077
# WEBSEQUENCE Total prob of N-in:        0.00583
WEBSEQUENCE     TMHMM2.0        outside      1   552
```



Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.0032	0.0001	0.0008	0.996

Download: PNG / EPS / Tabular





-Primer

#4. CR531_03290_FS

5'-CGC GAA CAG ATT GGA GGT AAT GAT TGG AAA AAG AAA GGT-3' - 39 bp

Oligo No	220223-2-17	GC	%41	Tm(Basic)	64.48°C	Total nmol	42.20nmol
Skala	50 nmol	MW	12226.07	Conc	1842.74ng/µl	Total ng	515968.03ng
Saflastirma	DSLT	A260	62.0	OD	17.4	100 µM stok - µl TE	422.0

#5. CR531_03290_RS

5'-GTG GCG GCC GCT CTA TTA ACC AAT TAT ATA AGC TCT CCC-3' - 39 bp

Oligo No	220223-2-18	GC	%49	Tm(Basic)	67.63°C	Total nmol	42.14nmol
Skala	50 nmol	MW	11877.76	Conc	1787.48ng/µl	Total ng	500494.98ng
Saflastirma	DSLT	A260	54.5	OD	15.3	100 µM stok - µl TE	421.4

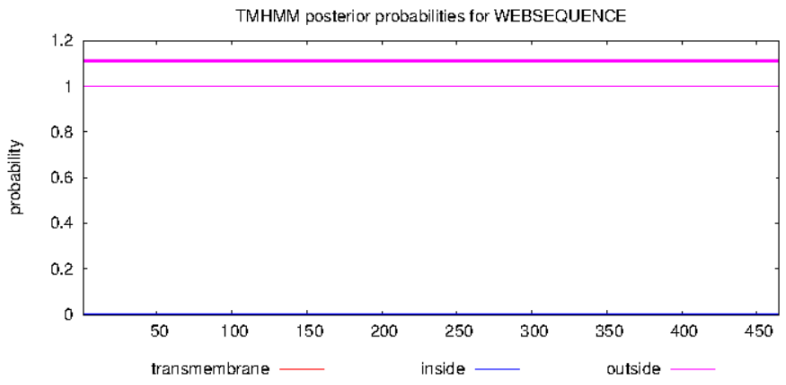
Figure 29. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of ATP37244.1.

4. SQH52440.1 – NCTC11324_01490

Streptococcus intermedius **ATCC 27335**

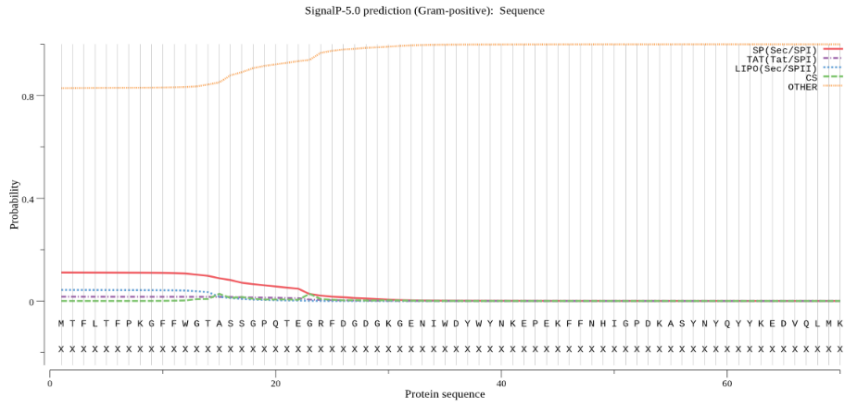
TMHMM result

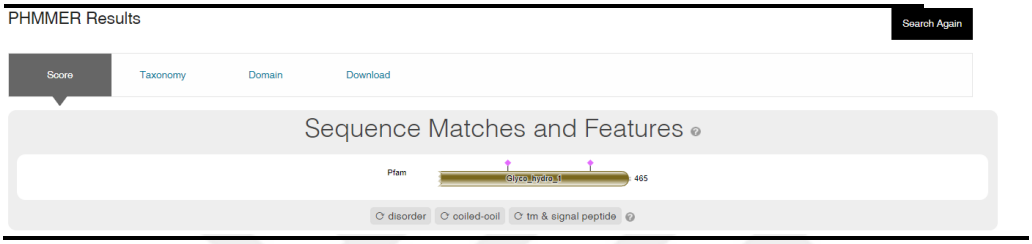
```
# WEBSEQUENCE Length: 465
# WEBSEQUENCE Number of predicted TMHs:  0
# WEBSEQUENCE Exp number of AAs in TMHs: 0.00115
# WEBSEQUENCE Exp number, first 60 AAs:  0
# WEBSEQUENCE Total prob of N-in:        0.00050
WEBSEQUENCE     TMHMM2.0        outside     1   465
```



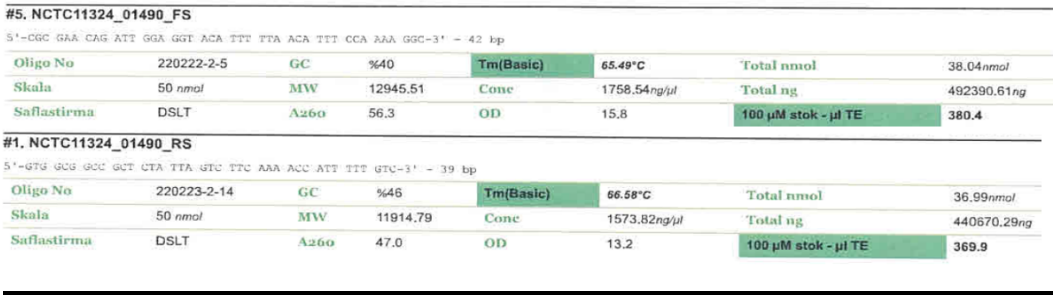
Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.1112	0.0164	0.0428	0.8296

Download: PNG / EPS / Tabular





-Primer



#5. NCTC11324_01490_FS

5'-CGC GAA CAG ATT GGA GGT ACA TTT TTA ACA TTT CCA AAA GGC-3' - 42 bp

Oligo No	220222-2-5	GC	%40	Tm(Basic)	65.49°C	Total nmol	38.04nmol
Skala	50 nmol	MW	12945.51	Conc	1758.54ng/µl	Total ng	492390.61ng
Saflastirma	DSLT	A260	56.3	OD	15.8	100 µM stok - µl TE	380.4

#1. NCTC11324_01490_RS

5'-GTG GCG GCC GCT CTA TTA GTC TTC AAA ACC ATT TTT GTC-3' - 39 bp

Oligo No	220223-2-14	GC	%46	Tm(Basic)	66.58°C	Total nmol	36.99nmol
Skala	50 nmol	MW	11914.79	Conc	1573.82ng/µl	Total ng	440670.29ng
Saflastirma	DSLT	A260	47.0	OD	13.2	100 µM stok - µl TE	369.9

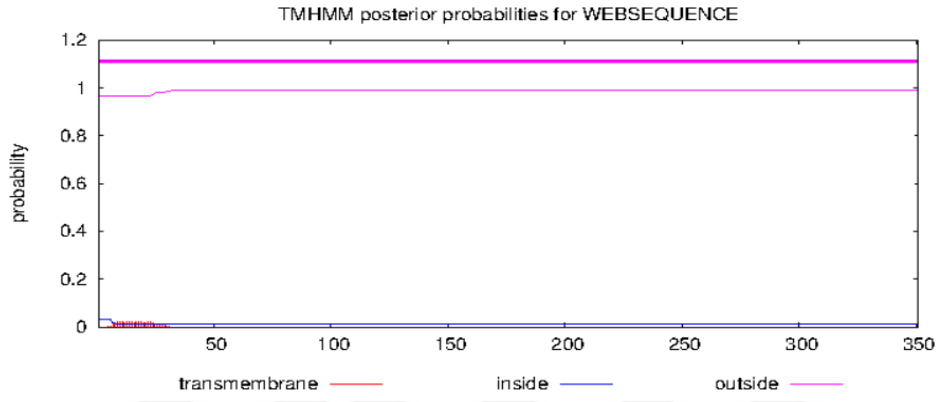
Figure 30. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of SQH52440.1.

5. ATP38122.1 – CR531_08290

Lactobacillus salivarus subsp. Salivarius (Ligilactobacillus salivarius) **ATCC 11741**

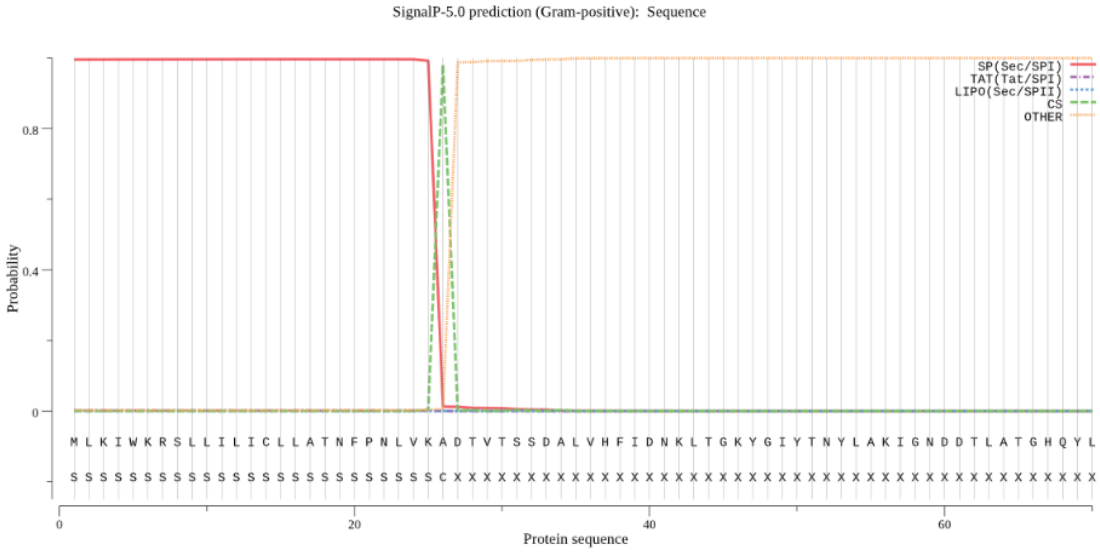
TMHMM result

```
# WEBSEQUENCE Length: 351
# WEBSEQUENCE Number of predicted TMHs:  0
# WEBSEQUENCE Exp number of AAs in TMHs: 0.42169
# WEBSEQUENCE Exp number, first 60 AAs:  0.4189
# WEBSEQUENCE Total prob of N-in:        0.03137
WEBSEQUENCE     TMHMM2.0        outside     1   351
```



Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.9959	0.001	0.0009	0.0022

Download: PNG / EPS / Tabular



-Primer

#3. CR531_08290_FS

5'-CGC GAA CAG ATT GGA GGT GAT ACA GTT ACT TCT AGT-3' - 36 bp

Oligo No	220223-2-22	GC	%44	Tm(Basic)	64.44°C	Total nmol	34.13nmol
Skala	50 nmol	MW	11139.32	Conc	1357.64ng/µl	Total ng	380139.03ng
Saflastirma	DSLT	A260	43.6	OD	12.2	100 µM stok - µl TE	341.3

#4. CR531_08290_RS

5'-GTG GCG GCC GCT CTA TTA TTG ATT TAA TGC TAC TAA AGC-3' - 39 bp

Oligo No	220223-2-23	GC	%44	Tm(Basic)	65.53°C	Total nmol	45.22nmol
Skala	50 nmol	MW	11987.86	Conc	1936.07ng/µl	Total ng	542099.27ng
Saflastirma	DSLT	A260	59.7	OD	16.7	100 µM stok - µl TE	452.2

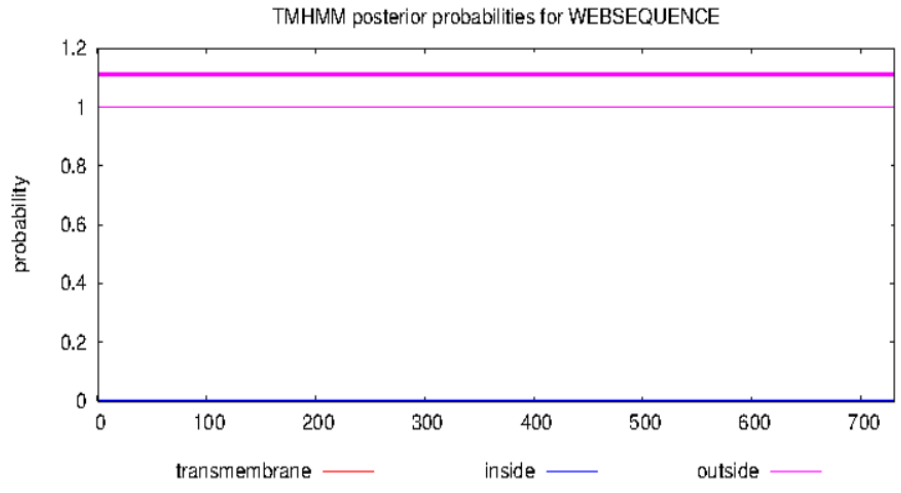
Figure 31. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of ATP38122.1.

6. ATP37586.1 – CR531_05275

Lactobacillus salivarus subsp. Salivarius (Ligilactobacillus salivarius) **ATCC 11741**

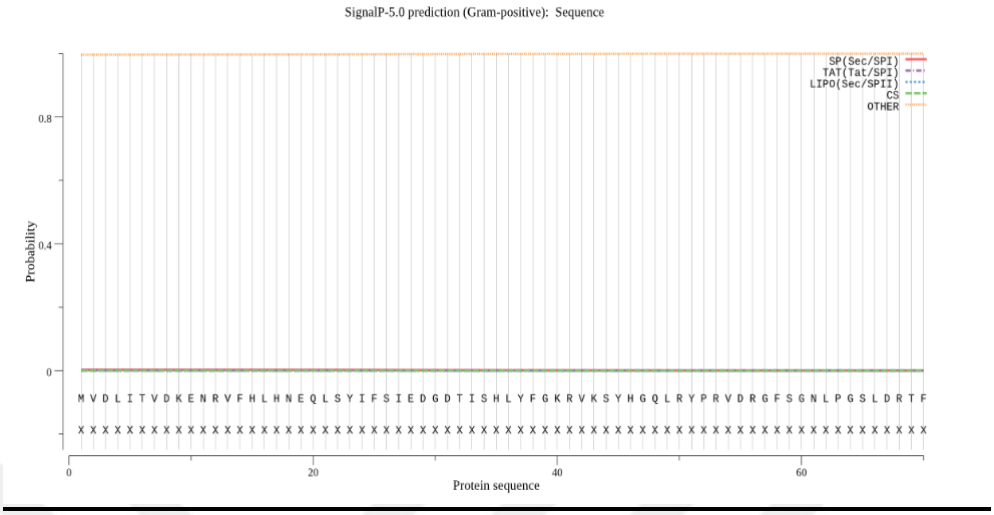
TMHMM result

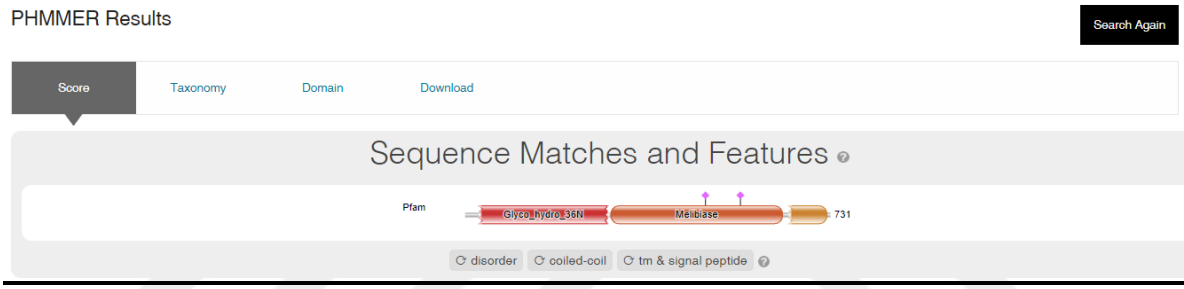
```
# WEBSEQUENCE Length: 731
# WEBSEQUENCE Number of predicted TMHs:  0
# WEBSEQUENCE Exp number of AAs in TMHs: 0.00590999999999998
# WEBSEQUENCE Exp number, first 60 AAs:  0.00303
# WEBSEQUENCE Total prob of N-in:        0.00035
WEBSEQUENCE     TMHMM2.0        outside     1   731
```



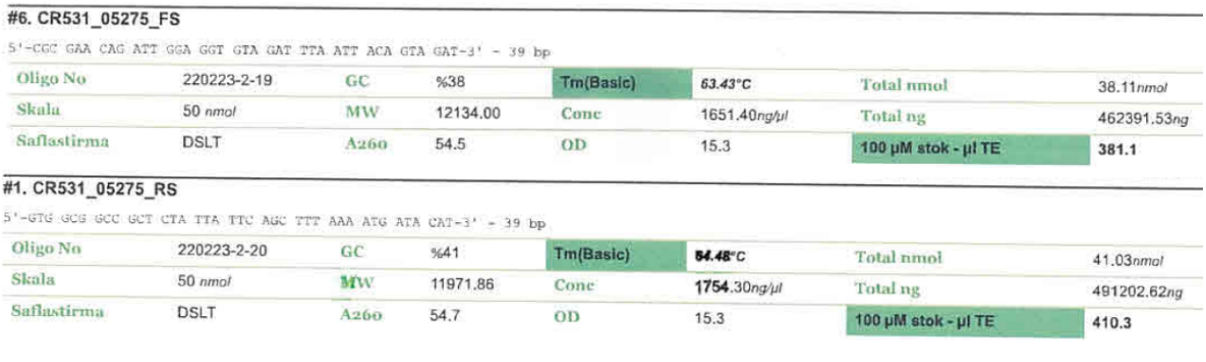
Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.0023	0.0003	0.0004	0.997

Download: PNG / EPS / Tabular





-Primer



#6. CR531_05275_FS

5'-CGC GAA CAG ATT GGA GGT GTA GAT TTA ATT ACA GTA GAT-3' - 39 bp

Oligo No	220223-2-19	GC	%38	Tm(Basic)	63.43°C	Total nmol	38.11nmol
Skala	50 nmol	MW	12134.00	Conc	1651.40ng/µl	Total ng	462391.53ng
Saflastirma	DSLT	A260	54.5	OD	15.3	100 µM stok - µl TE	381.1

#1. CR531_05275_RS

5'-GTG GCG GCC GCT CTA TTA TTC AGC TTT AAA ATG ATA CAT-3' - 39 bp

Oligo No	220223-2-20	GC	%41	Tm(Basic)	64.48°C	Total nmol	41.03nmol
Skala	50 nmol	MW	11971.86	Conc	1754.30ng/µl	Total ng	491202.62ng
Saflastirma	DSLT	A260	54.7	OD	15.3	100 µM stok - µl TE	410.3

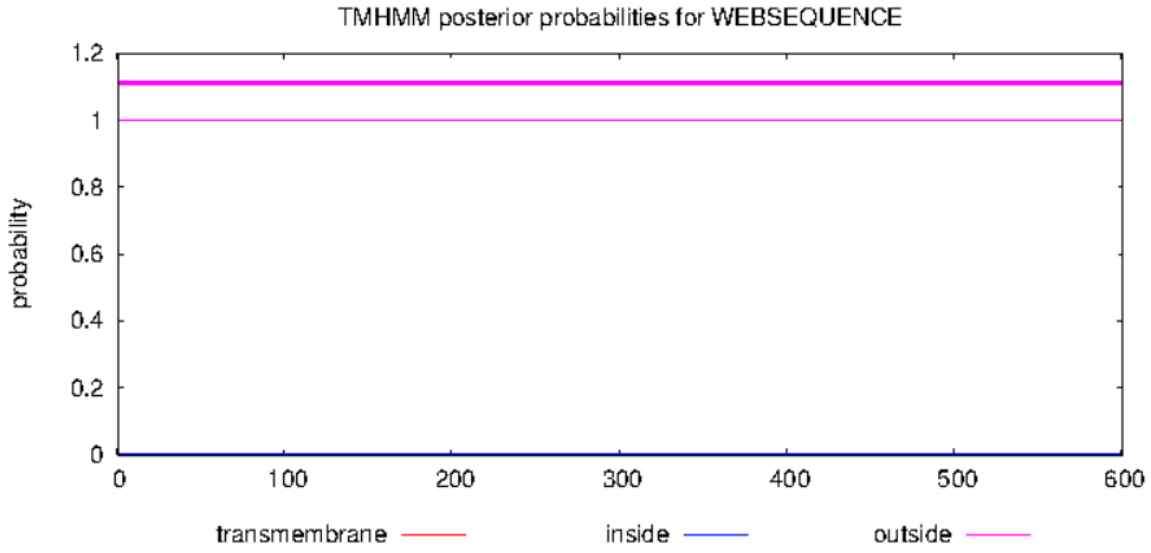
Figure 32. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of ATP37586.1.

7. SQH51076.1 – NCTC11324_00070

Streptococcus intermedius **ATCC 27335**

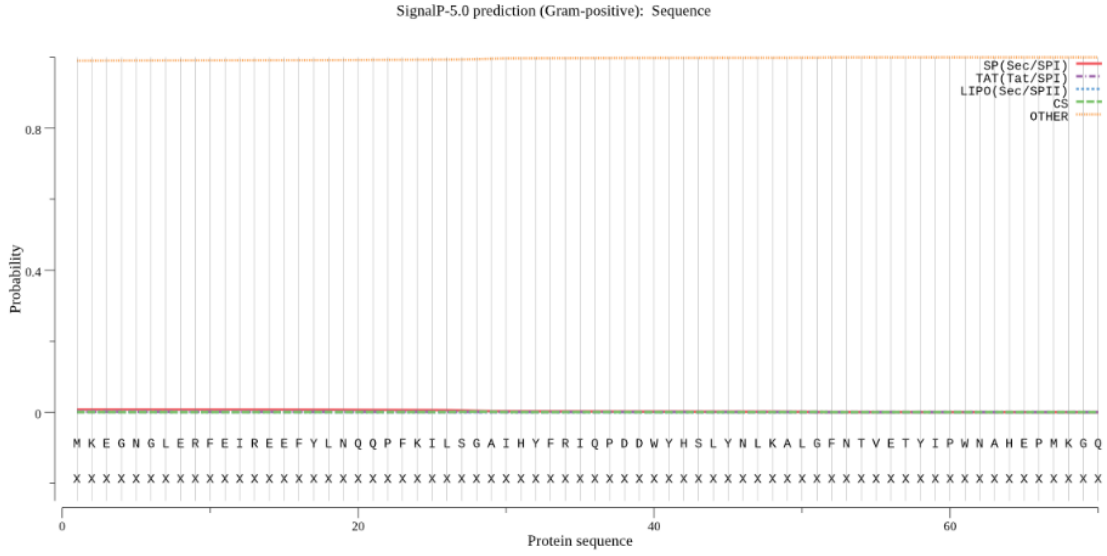
TMHMM result

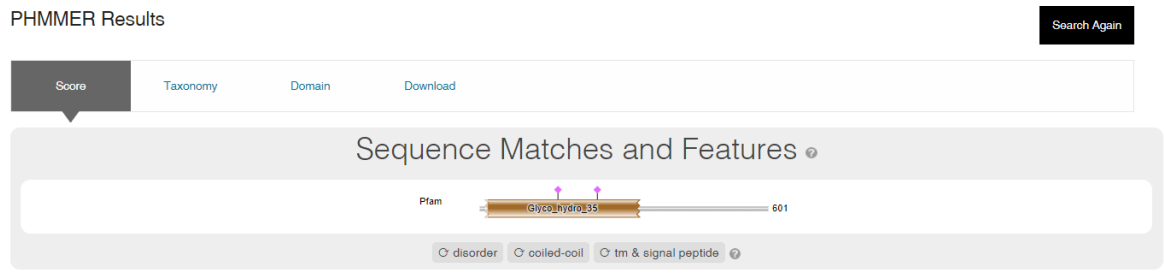
```
# WEBSEQUENCE Length: 601
# WEBSEQUENCE Number of predicted TMHs:  0
# WEBSEQUENCE Exp number of AAs in TMHs: 0.03563
# WEBSEQUENCE Exp number, first 60 AAs:  0.00442
# WEBSEQUENCE Total prob of N-in:        0.00075
WEBSEQUENCE     TMHMM2.0        outside      1   601
```



Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.0073	0.0006	0.0006	0.9915

Download: PNG / EPS / Tabular





-Primer

#1. NCTC11324_00070_FS

5'-CGC GAA CAG ATT GGA GGT AAA GAG GGA AAT GGT TTG-3' - 36 bp

Oligo No	220223-2-8	GC	%47	Tm(Basic)	65.58°C	Total nmol	36.15nmol
Skala	50 nmol	MW	11302.44	Conc	1459.40ng/µl	Total ng	408631.33ng
Saflastirma	DSLT	A260	48.1	OD	13.5	100 µM stok - µl TE	361.5

#2. NCTC11324_00070_RS

5'-GTG GCG GCC GCT CTA TTA TAA TTT GTC CCC CTT TAT TGG-3' - 39 bp

Oligo No	220223-2-9	GC	%49	Tm(Basic)	67.63°C	Total nmol	36.30nmol
Skala	50 nmol	MW	11921.78	Conc	1545.65ng/µl	Total ng	432781.58ng
Saflastirma	DSLT	A260	45.6	OD	12.8	100 µM stok - µl TE	363.0

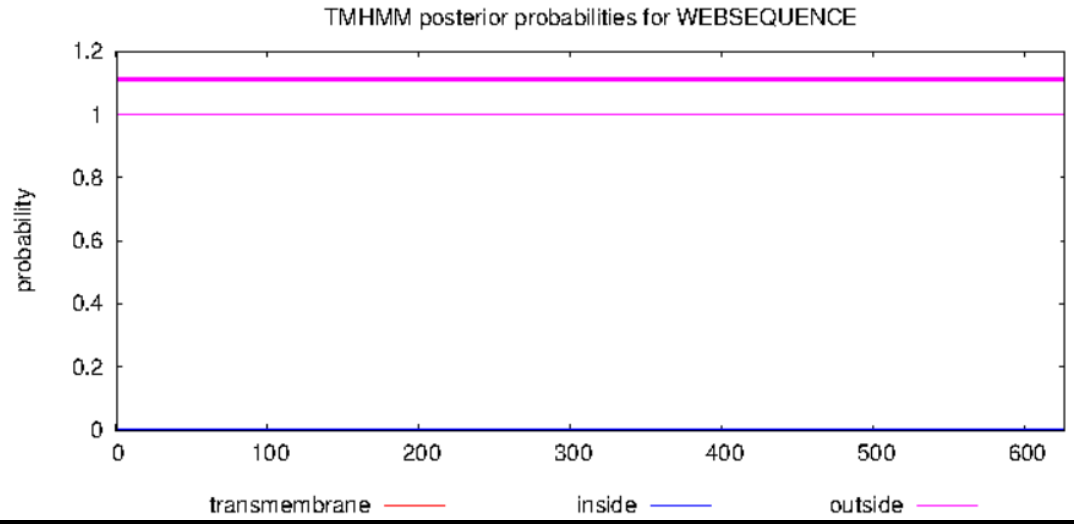
Figure 33. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of SQH51076.1.

8. SQH51655.1 – NCTC11324_00672

Streptococcus intermedius **ATCC 27335**

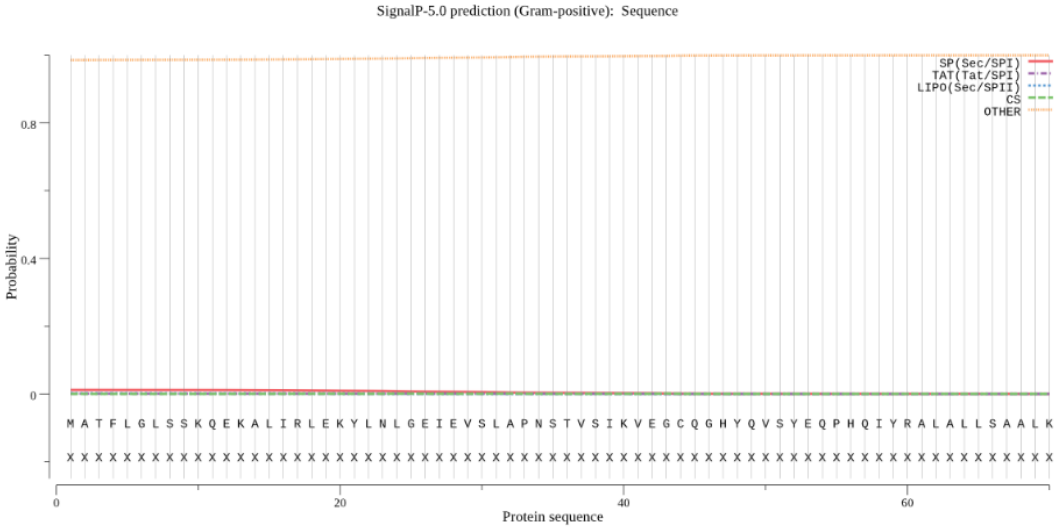
TMHMM result

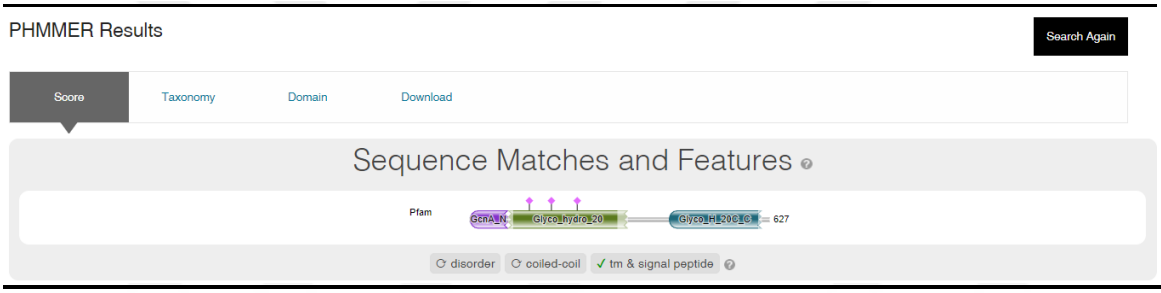
```
# WEBSEQUENCE Length: 627
# WEBSEQUENCE Number of predicted TMHs:  0
# WEBSEQUENCE Exp number of AAs in TMHs: 0.00692999999999999
# WEBSEQUENCE Exp number, first 60 AAs:  0
# WEBSEQUENCE Total prob of N-in:        0.00038
WEBSEQUENCE     TMHMM2.0        outside      1   627
```



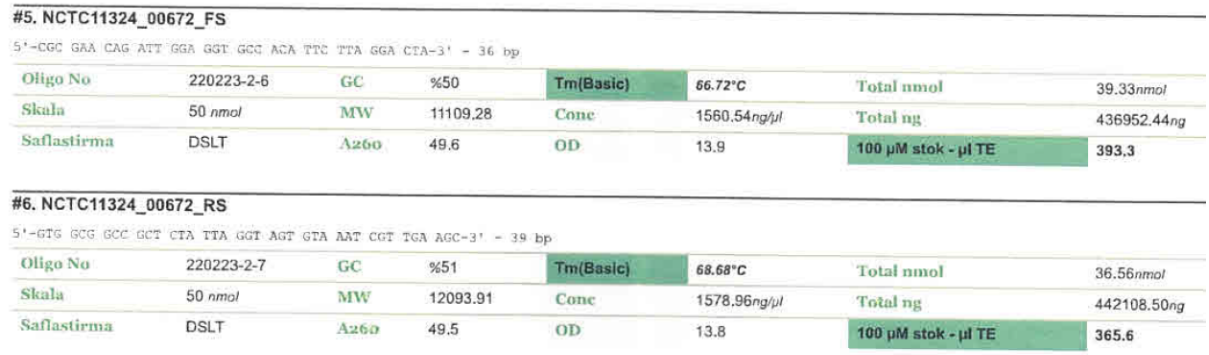
Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.0115	0.0006	0.0016	0.9863

Download: PNG / EPS / Tabular





-Primer



#5. NCTC11324_00672_FS

5'-CGC GAA CAG ATT GGA GGT GCC ACA TTC TTA GGA CTA-3' - 36 bp

Oligo No	220223-2-6	GC	%50	Tm(Basic)	66.72°C	Total nmol	39.33nmol
Skala	50 nmol	MW	11109.28	Conc	1560.54ng/µl	Total ng	436952.44ng
Saflastirma	DSLT	A260	49.6	OD	13.9	100 µM stok - µl TE	393.3

#6. NCTC11324_00672_RS

5'-GTG GCG GCC GCT CTA TTA GGT AGT GTA AAT CGT TGA AGC-3' - 39 bp

Oligo No	220223-2-7	GC	%51	Tm(Basic)	68.68°C	Total nmol	36.56nmol
Skala	50 nmol	MW	12093.91	Conc	1578.96ng/µl	Total ng	442108.50ng
Saflastirma	DSLT	A260	49.5	OD	13.8	100 µM stok - µl TE	365.6

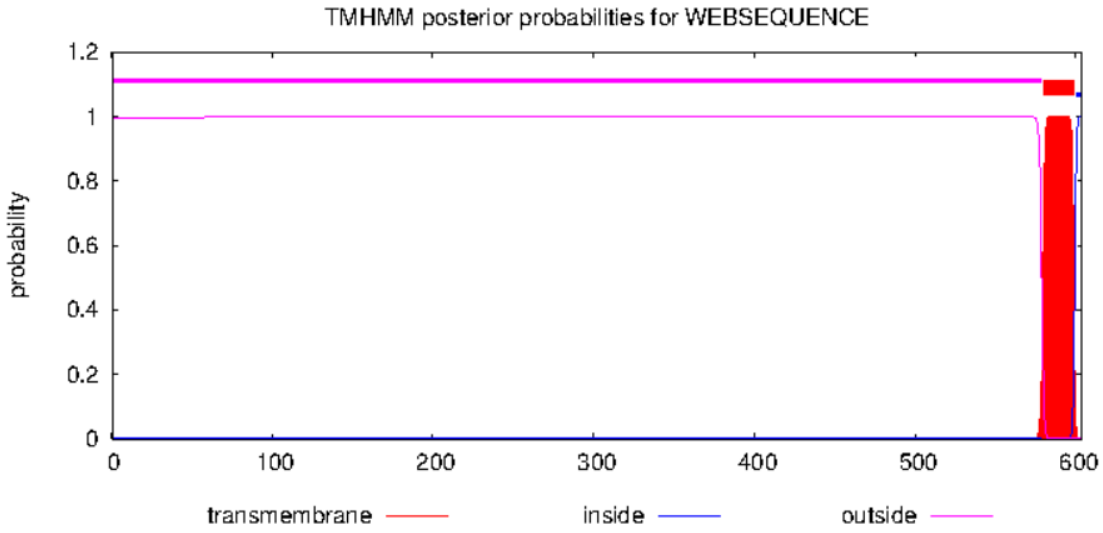
Figure 34. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of SQH51655.1.

9. CAR86329.1 — LGG 00434

Lactobacillus rhamnosus GG

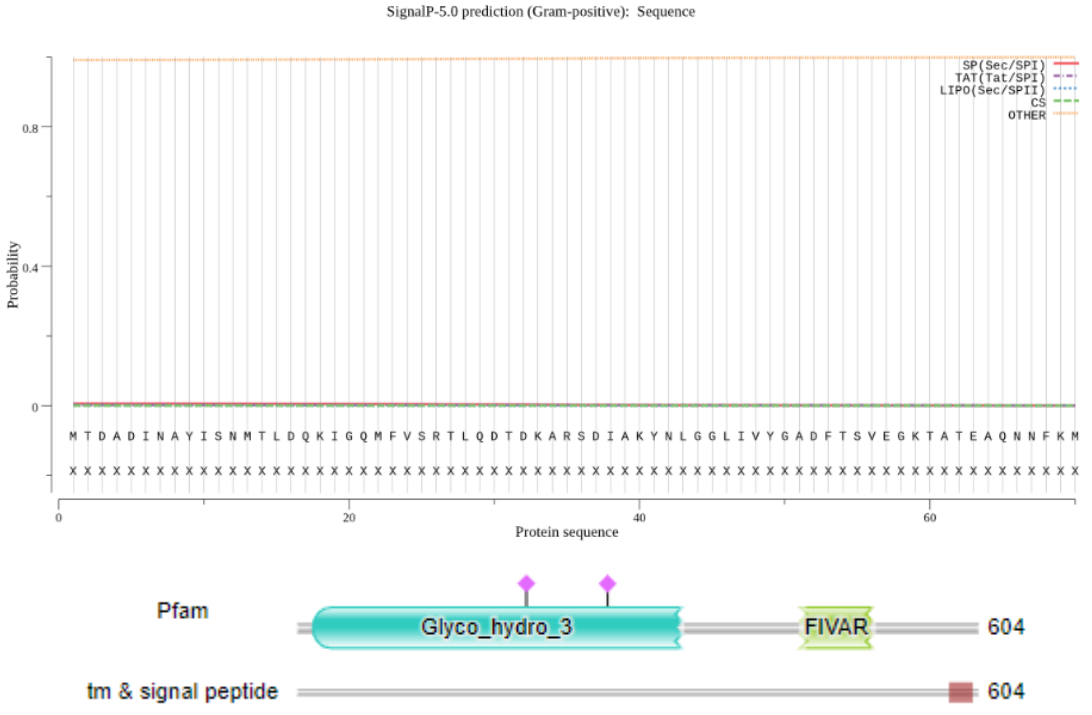
TMHMM result

```
# WEBSEQUENCE Length: 604
# WEBSEQUENCE Number of predicted TMHs:  1
# WEBSEQUENCE Exp number of AAs in TMHs: 20.6397
# WEBSEQUENCE Exp number, first 60 AAs:  0.02516
# WEBSEQUENCE Total prob of N-in:        0.00295
WEBSEQUENCE     TMHMM2.0        outside      1   579
WEBSEQUENCE     TMHMM2.0        TMhelix    580   599
WEBSEQUENCE     TMHMM2.0        inside     600   604
```



Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.0056	0.0015	0.0006	0.9923

Download: PNG / EPS / Tabular



-Primer

#1. LGG_00434FN

5'-CAT CAT CAC CAC CAT CAC ACC GAT GCC GAT ATT AAT GCC-3' - 39 bp

Oligo No	210504-1-75	GC	%49	Tm(Basic)	67.63°C	Total nmol	48.05nmol
Skala	50 nmol	MW	11784.70	Conc	2022.15ng/µl	Total ng	566203.01ng
Saflastirma	DSLT	A260	63.1	OD	17.7	100 µM stok - µl TE	480.5

#2. LGG_00434RN

5'-GTG GCG GCC GCT CTA TTA TGC CCT TTT TTT TCG AAT ATA-3' - 39 bp

Oligo No	210504-1-76	GC	%44	Tm(Basic)	65.53°C	Total nmol	41.02nmol
Skala	50 nmol	MW	11920.80	Conc	1746.51ng/µl	Total ng	489024.11ng
Saflastirma	DSLT	A260	52.1	OD	14.6	100 µM stok - µl TE	410.2

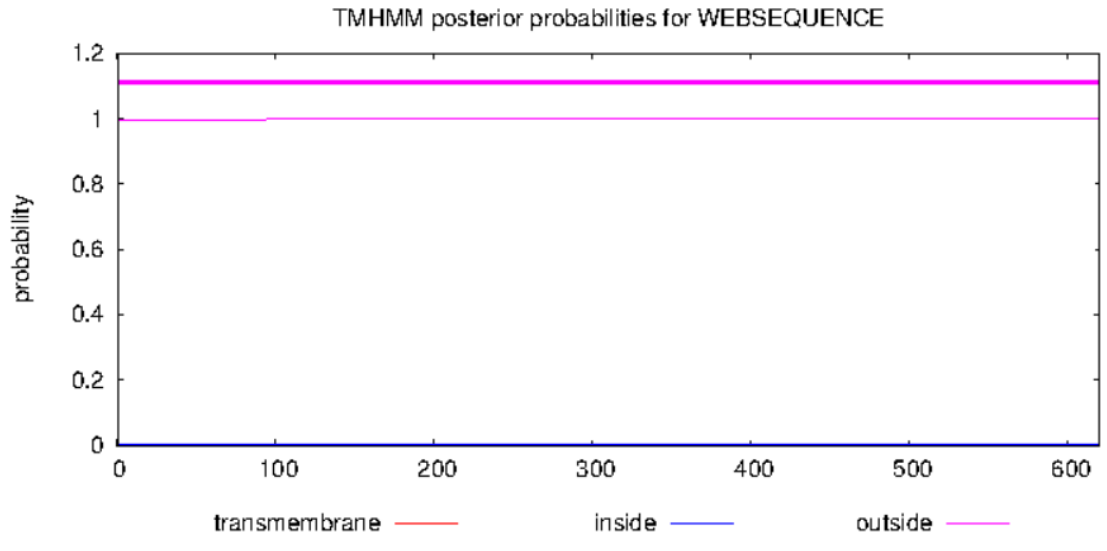
Figure 35. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of CAR86329.1.

10. AAO77566.1 – BT_2459

Bacteroides thetaiotaomicron **ATCC 29148**

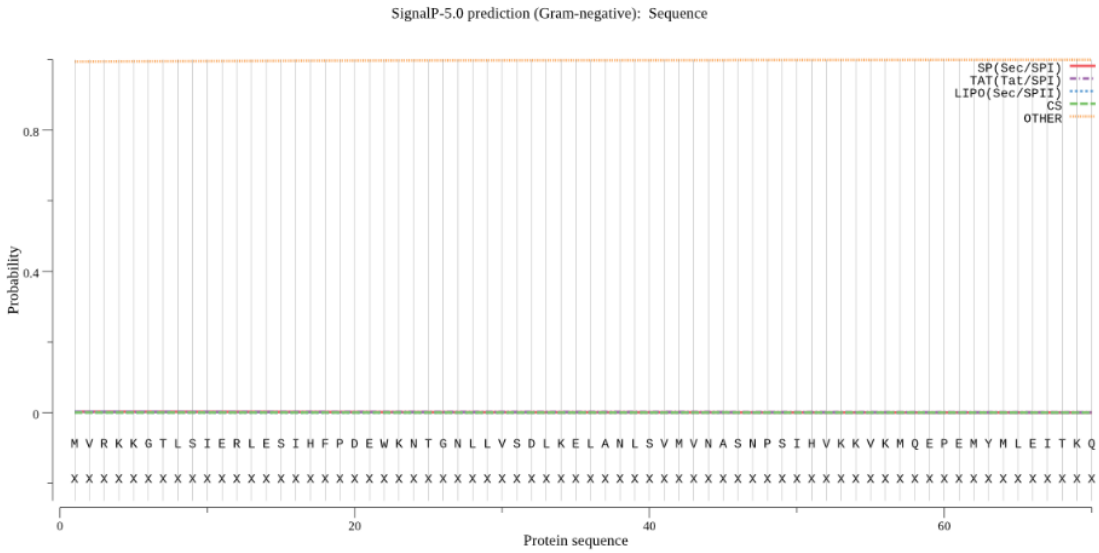
TMHMM result

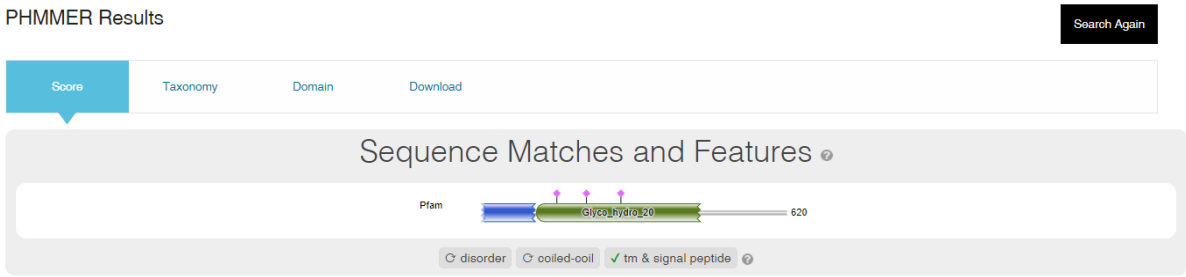
```
# WEBSEQUENCE Length: 620
# WEBSEQUENCE Number of predicted TMHs:  0
# WEBSEQUENCE Exp number of AAs in TMHs: 0.0754999999999998
# WEBSEQUENCE Exp number, first 60 AAs:  0.00018
# WEBSEQUENCE Total prob of N-in:        0.00432
WEBSEQUENCE     TMHMM2.0        outside      1   620
```



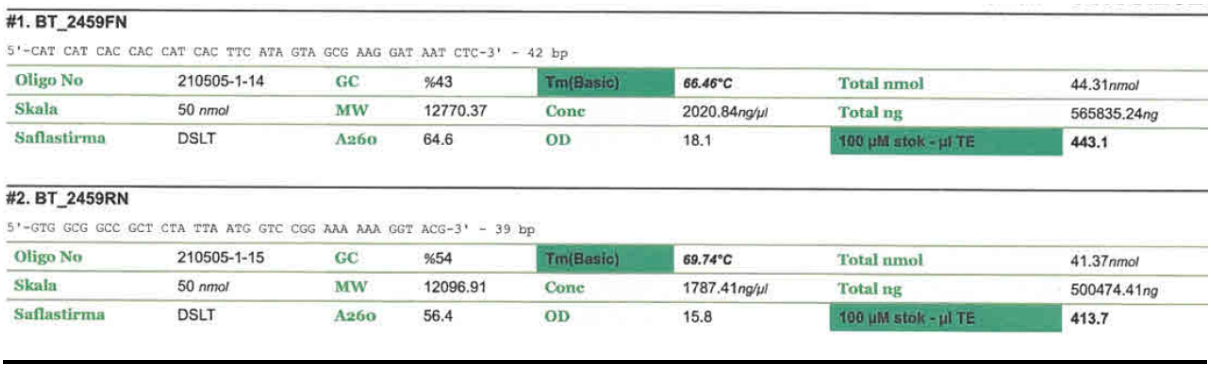
Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.0022	0.0015	0.0003	0.9961

Download: PNG / EPS / Tabular





-Primer



#1. BT_2459FN

5'-CAT CAT CAC CAC CAT CAC TTC ATA GTA GCG AAG GAT AAT CTC-3' - 42 bp

Oligo No	210505-1-14	GC	%43	Tm(Basic)	66.46°C	Total nmol	44.31nmol
Skala	50 nmol	MW	12770.37	Conc	2020.84ng/µl	Total ng	565835.24ng
Saflastirma	DSLT	A260	64.6	OD	18.1	100 µM stok - µl TE	443.1

#2. BT_2459RN

5'-GTG GCG GCC GCT CTA TTA ATG GTC CGG AAA AAA GGT ACG-3' - 39 bp

Oligo No	210505-1-15	GC	%54	Tm(Basic)	69.74°C	Total nmol	41.37nmol
Skala	50 nmol	MW	12096.91	Conc	1787.41ng/µl	Total ng	500474.41ng
Saflastirma	DSLT	A260	56.4	OD	15.8	100 µM stok - µl TE	413.7

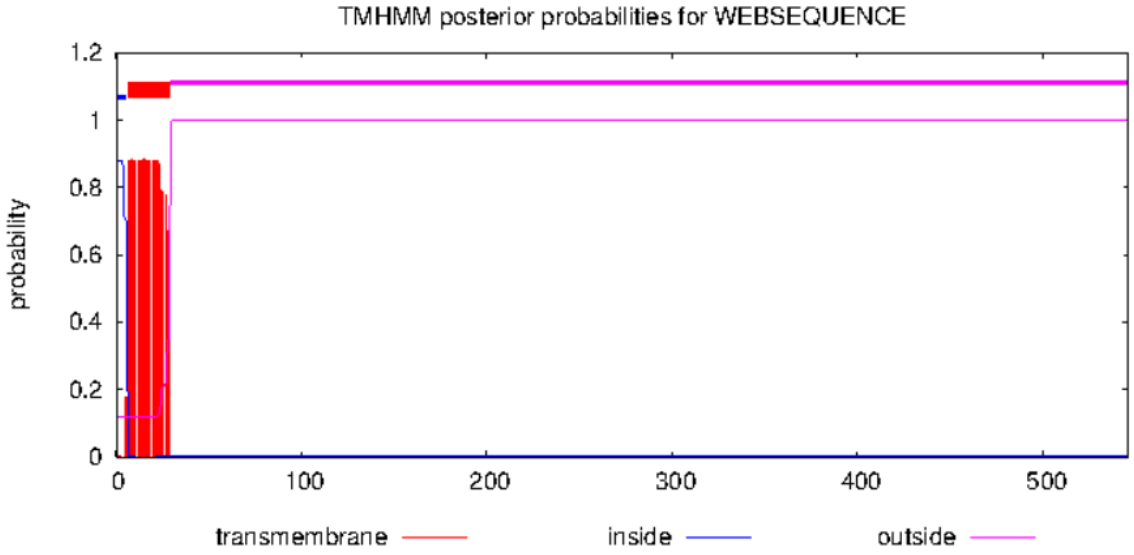
Figure 36. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of AAO77566.1.

11. ABR41745.1 -- BVU_4143

***Bacteroides vulgatus* ATCC 8482**

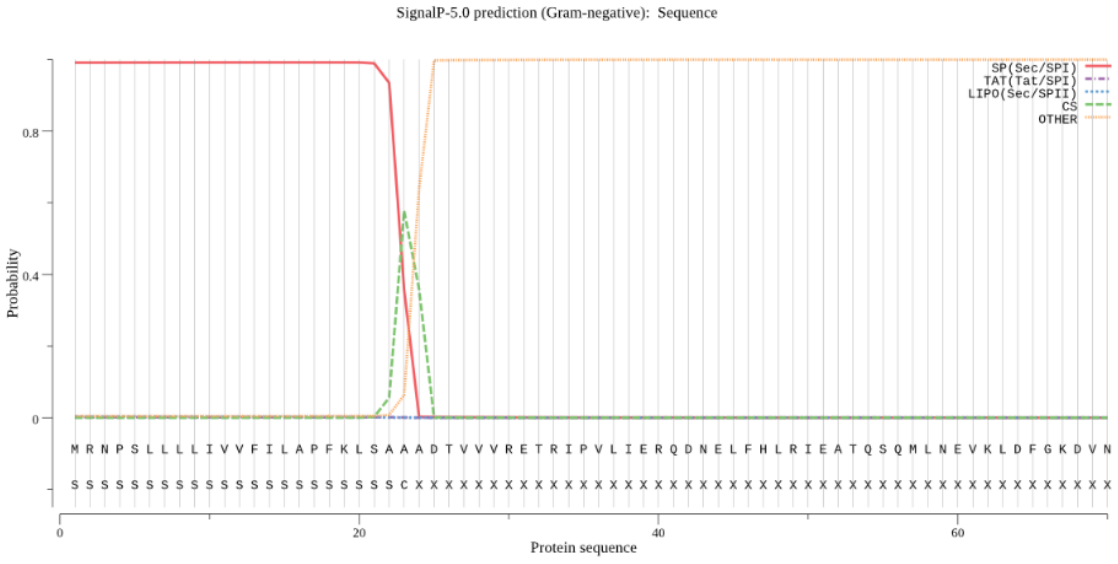
TMHMM result

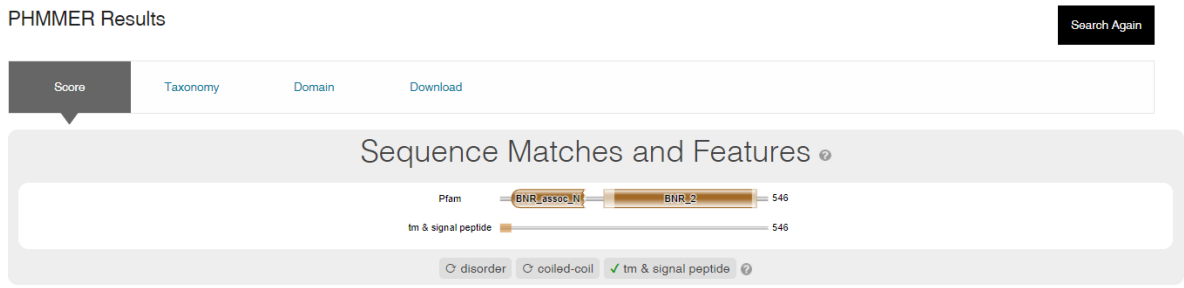
```
# WEBSEQUENCE Length: 546
# WEBSEQUENCE Number of predicted TMHs:  1
# WEBSEQUENCE Exp number of AAs in TMHs: 19.91472
# WEBSEQUENCE Exp number, first 60 AAs:  19.89672
# WEBSEQUENCE Total prob of N-in:        0.88092
# WEBSEQUENCE POSSIBLE N-term signal sequence
WEBSEQUENCE     TMHMM2.0        inside       1     6
WEBSEQUENCE     TMHMM2.0        TMhelix      7    29
WEBSEQUENCE     TMHMM2.0        outside     30   546
```



Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.9905	0.0021	0.003	0.0044

Download: PNG / EPS / Tabular





-Primer

#3. BVU_4143FN

5'-CAT CAT CAC CAC CAT CAC TTT GGT CTT AAT AAT ATC TTT CAG-3' - 42 bp

Oligo No	210505-1-4	GC	%36	Tm(Basic)	63.53°C	Total nmol	34.12nmol
Skala	50 nmol	MW	12717.35	Conc	1549.55ng/µl	Total ng	433874.72ng
Saflastirma	DSLT	A260	48.2	OD	13.5	100 µM stok - µl TE	341.2

#4. BVU_4143RN

5'-GTG GCG GCC GCT CTA TTA ATG AGA AAC CCT AGC TTA TTA CTG-3' - 42 bp

Oligo No	210505-1-5	GC	%48	Tm(Basic)	68.41°C	Total nmol	32.58nmol
Skala	50 nmol	MW	12904.44	Conc	1501.59ng/µl	Total ng	420445.41ng
Saflastirma	DSLT	A260	46.6	OD	13.0	100 µM stok - µl TE	325.8

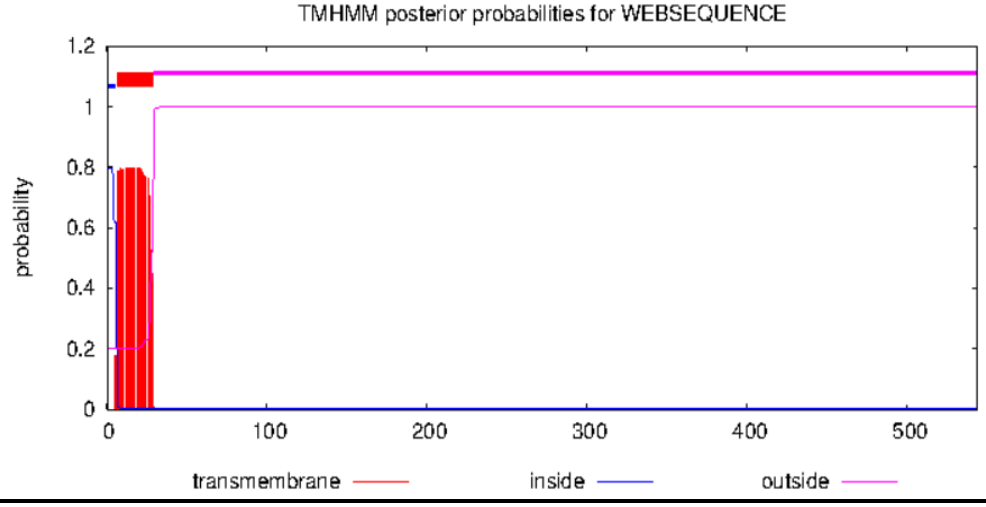
Figure 37. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of ABR41745.1.

12. AAO75562.1 - BT_0455

Bacteroides thetaiotaomicron **ATCC 29148**

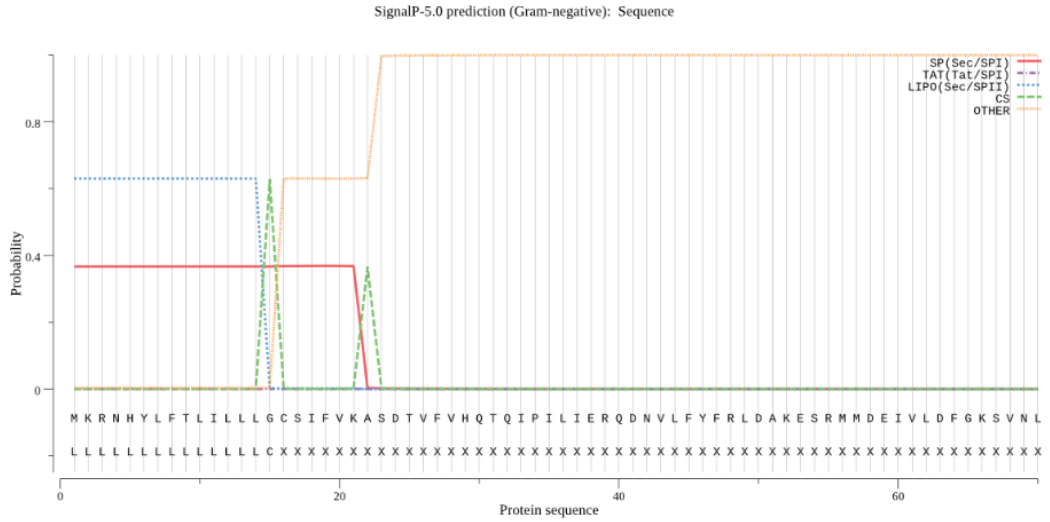
TMHMM result

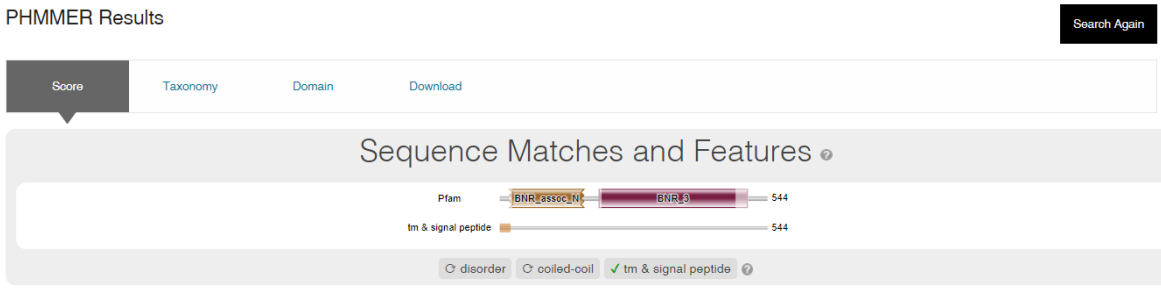
```
# WEBSEQUENCE Length: 544
# WEBSEQUENCE Number of predicted TMHs:  1
# WEBSEQUENCE Exp number of AAs in TMHs: 17.88302
# WEBSEQUENCE Exp number, first 60 AAs:  17.87916
# WEBSEQUENCE Total prob of N-in:        0.79720
# WEBSEQUENCE POSSIBLE N-term signal sequence
WEBSEQUENCE     TMHMM2.0        inside       1     6
WEBSEQUENCE     TMHMM2.0        TMhelix      7    29
WEBSEQUENCE     TMHMM2.0        outside     30   544
```



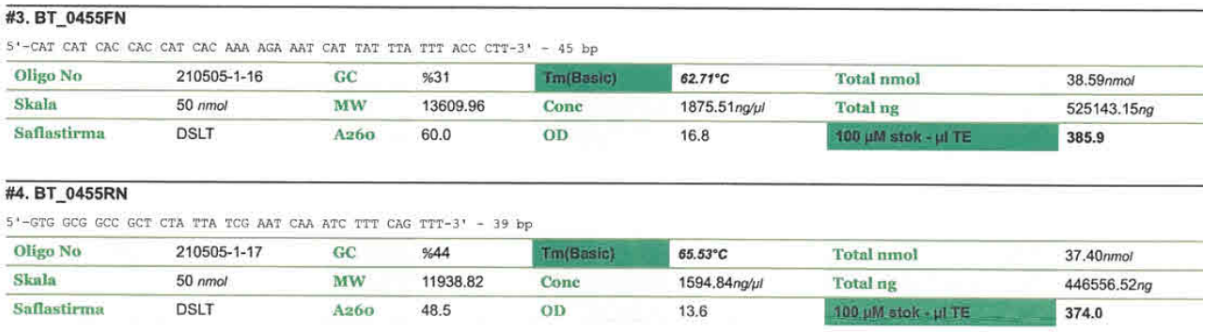
Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.3685	0.0003	0.6276	0.0037

Download: PNG / EPS / Tabular





-Primer



#3. BT_0455FN

5'-CAT CAT CAC CAC CAT CAC AAA AGA AAT CAT TAT TTA TTT ACC CTT-3' - 45 bp

Oligo No	210505-1-16	GC	%31	Tm(Basic)	62.71°C	Total nmol	38.59nmol
Skala	50 nmol	MW	13609.96	Conc	1875.51ng/µl	Total ng	525143.15ng
Saflastirma	DSLT	A260	60.0	OD	16.8	100 µM stok - µl TE	385.9

#4. BT_0455RN

5'-GTG GCG GCC GCT CTA TTA TCG AAT CAA ATC TTT CAG TTT-3' - 39 bp

Oligo No	210505-1-17	GC	%44	Tm(Basic)	65.53°C	Total nmol	37.40nmol
Skala	50 nmol	MW	11938.82	Conc	1594.84ng/µl	Total ng	446556.52ng
Saflastirma	DSLT	A260	48.5	OD	13.6	100 µM stok - µl TE	374.0

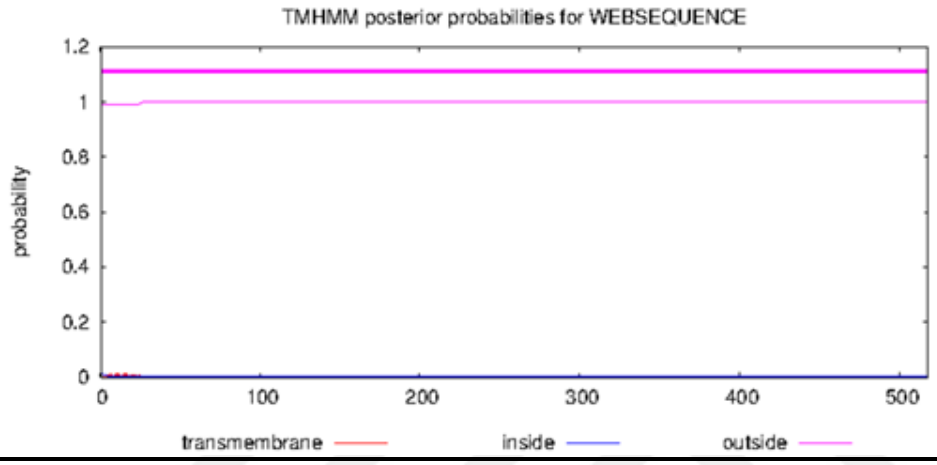
Figure 38. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of AAO75562.1.

13. ACD04858.1 – Amuc_1032

Akkermansia muciniphila **ATCC BAA-835**

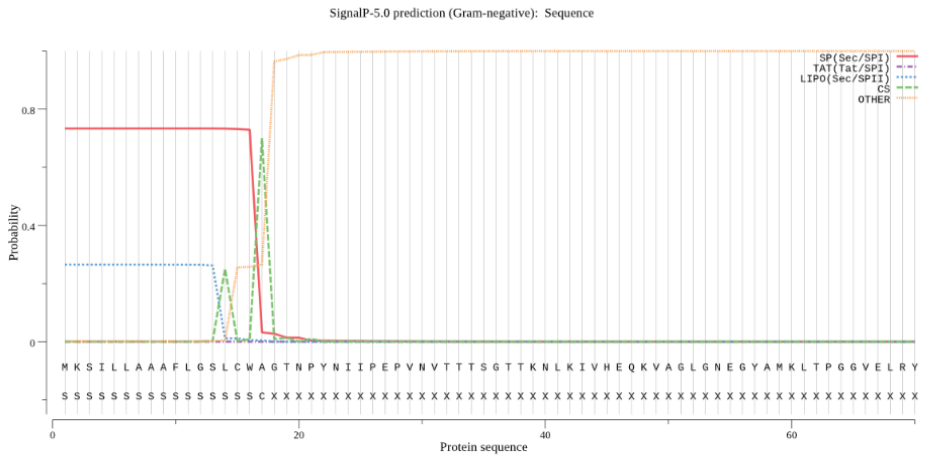
TMHMM result

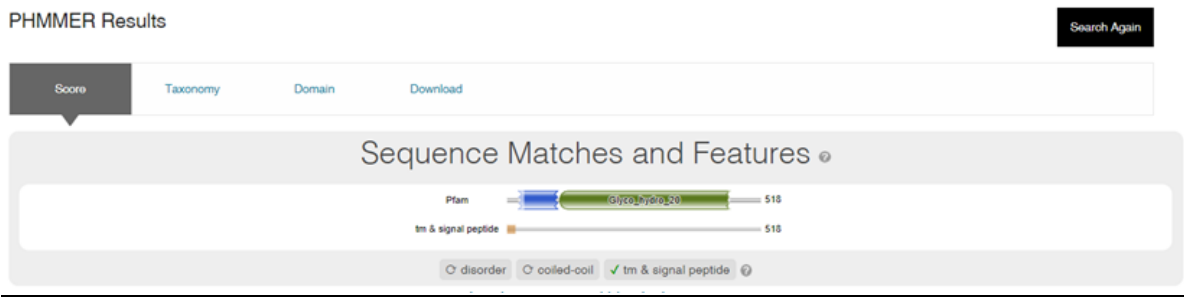
```
# WEBSEQUENCE Length: 518
# WEBSEQUENCE Number of predicted TMHs:  0
# WEBSEQUENCE Exp number of AAs in TMHs: 0.22983
# WEBSEQUENCE Exp number, first 60 AAs:  0.22574
# WEBSEQUENCE Total prob of N-in:        0.01117
WEBSEQUENCE     TMHMM2.0        outside     1   518
```



Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.7328	0.0001	0.2656	0.0015

Download: PNG / EPS / Tabular





-Primer

#1. Amuc_1032_FS

5'-CGC GAA CAG ATT GGA GGT GGA ACC AAT CCC TAC AAC-3' - 36 bp

Oligo No	220330-1-51	GC	%53	Tm(Basic)	67.86°C	Total nmol	36.04nmol
Skala	50 nmol	MW	11072.25	Conc	1425.00ng/µl	Total ng	399000.00ng
Saflastirma	DSLT	A260	45.7	OD	12.8	100 µM stok - µl TE	360.4

#2. Amuc_1032_RS

5'-GTG GCG GCC GCT CTA TTA TTT GTC CGC GTT TTG CTG-3' - 36 bp

Oligo No	220330-1-52	GC	%56	Tm(Basic)	69.00°C	Total nmol	37.24nmol
Skala	50 nmol	MW	11047.19	Conc	1469.11ng/µl	Total ng	411350.09ng
Saflastirma	DSLT	A260	42.1	OD	11.8	100 µM stok - µl TE	372.4

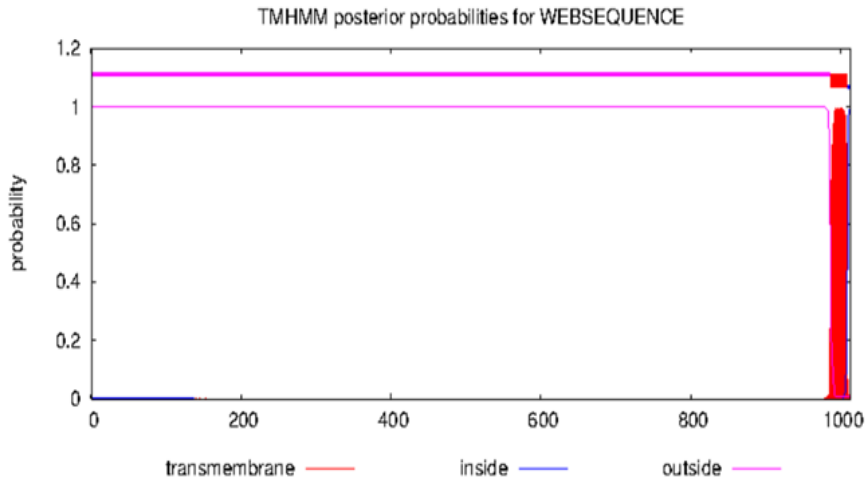
Figure 39. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of ACD04858.1.

14. BAQ98211.1 – BBBF_1004

***Bifidobacterium bifidum* ATCC 29521**

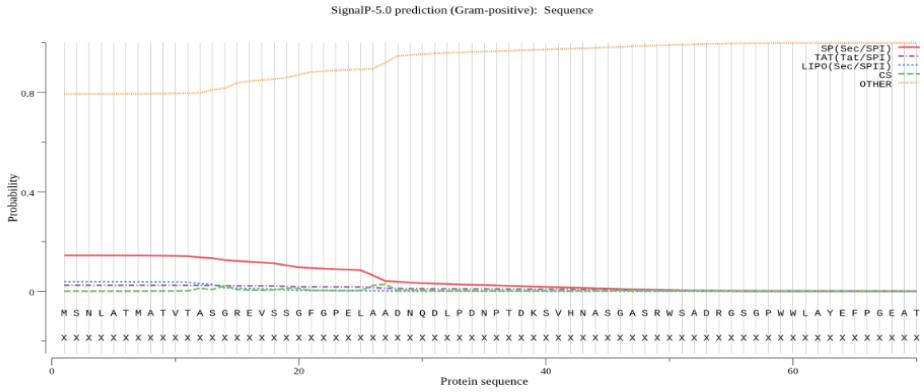
TMHMM result

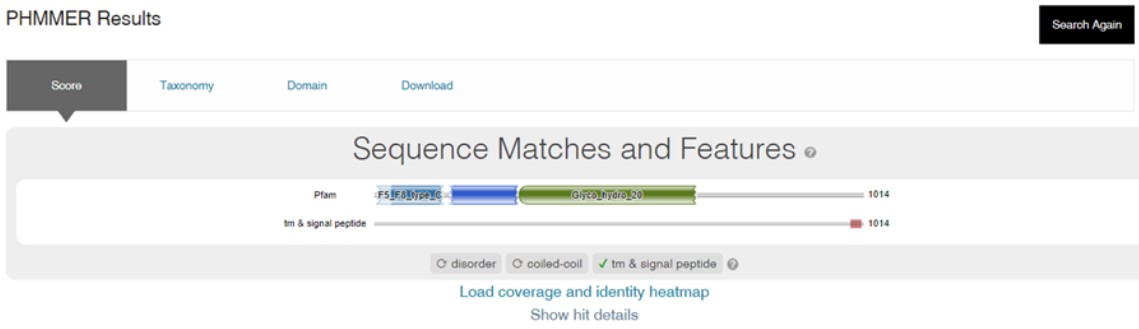
```
# WEBSEQUENCE Length: 1014
# WEBSEQUENCE Number of predicted TMHs:  1
# WEBSEQUENCE Exp number of AAs in TMHs: 22.157
# WEBSEQUENCE Exp number, first 60 AAs:  0.00032
# WEBSEQUENCE Total prob of N-in:        0.00034
WEBSEQUENCE     TMHMM2.0        outside     1   986
WEBSEQUENCE     TMHMM2.0        TMhelix   987  1009
WEBSEQUENCE     TMHMM2.0        inside   1010  1014
```



Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.1437	0.0235	0.0382	0.7945

Download: PNG / EPS / Tabular





-Primer

#3. BBBF_1004_FS

5'-CGC GAA CAG ATT GGA GGT CGC ACG ATG GGG ACC ATC-3' - 36 bp

Oligo No	220330-1-35	GC	%61	Tm(Basic)	71.28°C	Total nmol	37.15nmol
Skala	50 nmol	MW	11160.28	Conc	1480.83ng/µl	Total ng	414633.20ng
Saflastirma	DSLT	A260	46.7	OD	13.1	100 µM stok - µl TE	371.5

#1. BBBF_1004_RS

5'-GTG GCG GCC GCT CTA TTA GGC ACC GGT CTC GGC GAC GAC ATC-3' - 42 bp

Oligo No	220329-3-10	GC	%67	Tm(Basic)	76.22°C	Total nmol	38.91nmol
Skala	50 nmol	MW	12908.36	Conc	1793.82ng/µl	Total ng	502270.82ng
Saflastirma	DSLT	A260	53.6	OD	15.0	100 µM stok - µl TE	389.1

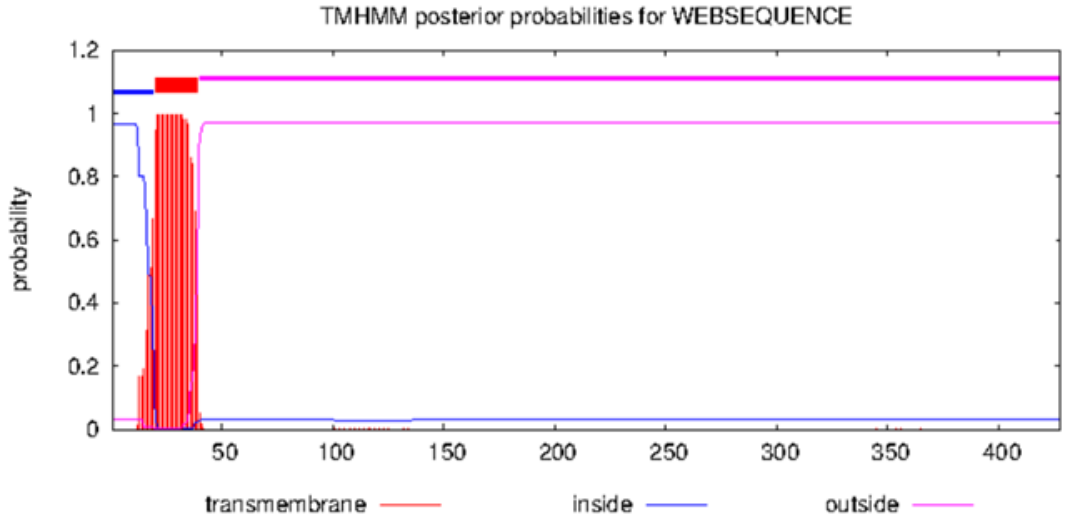
Figure 40. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of BAQ98211.1.

15. BAQ97897.1 – BBBF_0690

Bifidobacterium bifidum **ATCC 29521**

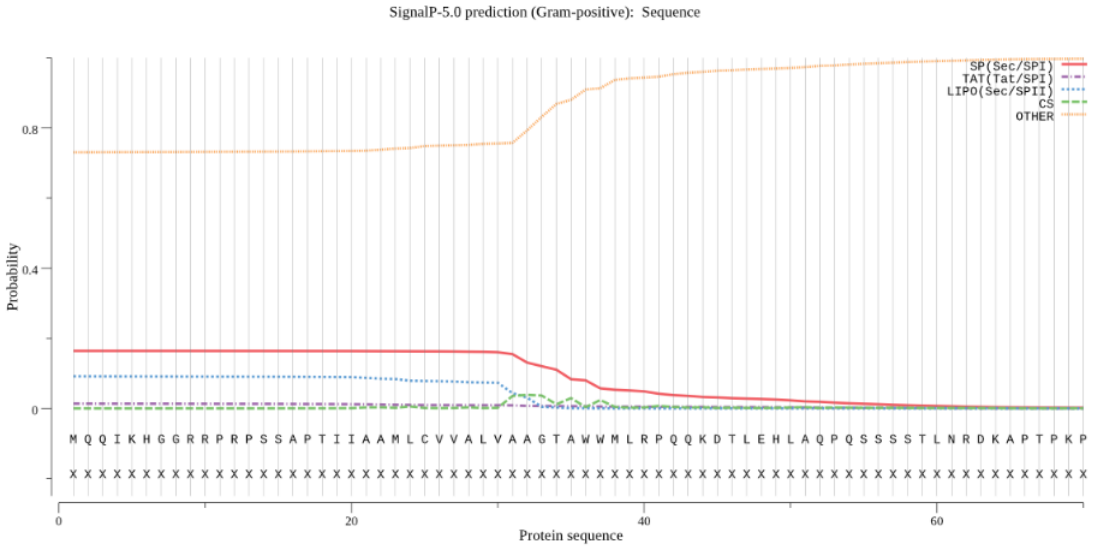
TMHMM result

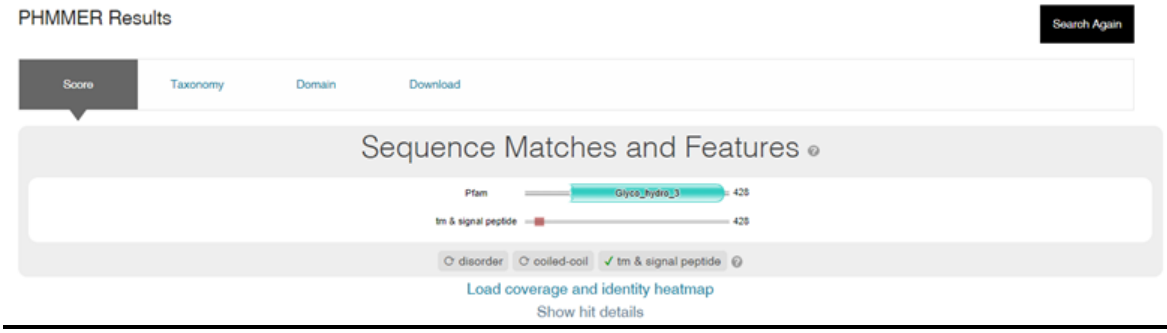
```
# WEBSEQUENCE Length: 428
# WEBSEQUENCE Number of predicted TMHs:  1
# WEBSEQUENCE Exp number of AAs in TMHs: 21.3314
# WEBSEQUENCE Exp number, first 60 AAs:  21.29086
# WEBSEQUENCE Total prob of N-in:        0.96830
# WEBSEQUENCE POSSIBLE N-term signal sequence
WEBSEQUENCE     TMHMM2.0        inside       1    19
WEBSEQUENCE     TMHMM2.0        TMhelix     20    39
WEBSEQUENCE     TMHMM2.0        outside     40   428
```



Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.1637	0.0139	0.0927	0.7297

Download: PNG / EPS / Tabular





-Primer

#6. BBBF_0690_FS

5'-CGC GAA CAG ATT GGA GGT GCG GCA GGC ACG GCA TGG-3' - 36 bp

Oligo No	220330-1-38	GC	%67	Tm(Basic)	73.56°C	Total nmol	39.78nmol
Skala	50 nmol	MW	11241.32	Conc	1597.23ng/µl	Total ng	447225.01ng
Saflastirma	DSLT	A260	50.0	OD	14.0	100 µM stok - µl TE	397.8

#1. BBBF_0690_RS

5'-GTG GCG GCC GCT CTA TTA CTT CGC CAG TCC CAT CTG-3' - 36 bp

Oligo No	220330-1-39	GC	%61	Tm(Basic)	71.28°C	Total nmol	42.96nmol
Skala	50 nmol	MW	10955.11	Conc	1680.83ng/µl	Total ng	470633.45ng
Saflastirma	DSLT	A260	48.8	OD	13.7	100 µM stok - µl TE	429.6

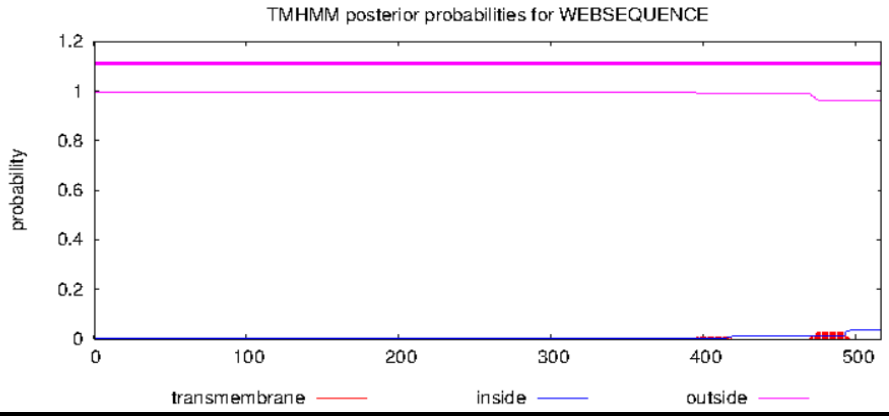
Figure 41. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of BAQ97897.1.

16. ERK41518.1 – HMPREF0495_02198

Levilactobacillus brevis (Lactobacillus brevis) **ATCC 148**

TMHMM result

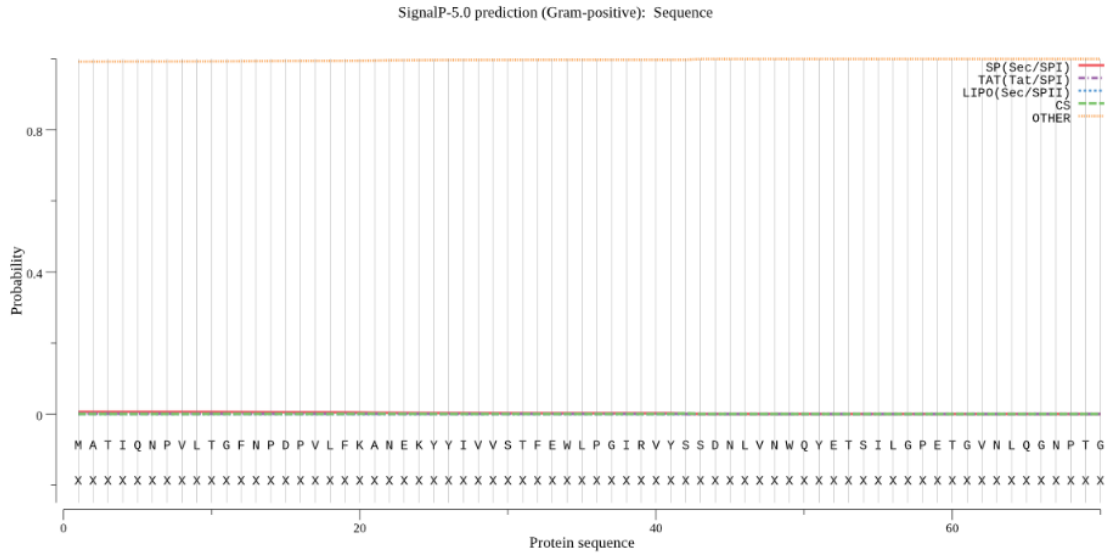
```
# WEBSEQUENCE Length: 517
# WEBSEQUENCE Number of predicted TMHs:  0
# WEBSEQUENCE Exp number of AAs in TMHs: 0.80367
# WEBSEQUENCE Exp number, first 60 AAs:  0.00316
# WEBSEQUENCE Total prob of N-in:        0.00400
WEBSEQUENCE     TMHMM2.0        outside     1   517
```

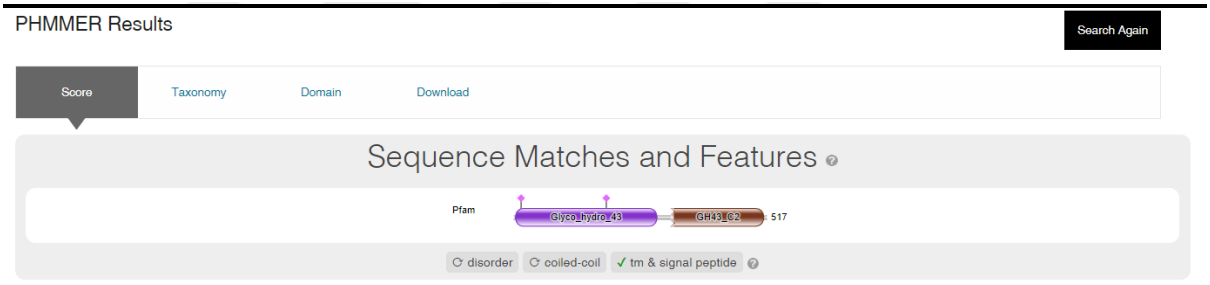


69

Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.0064	0.0004	0.0004	0.9928

Download: PNG / EPS / Tabular





-Primer

#1. HMPREF0495_02198_FS

5'-CGC GAA CAG ATT GGA GGT GCA ACG ATT CAA AAT CCA-3' - 36 bp

Oligo No	220330-1-57	GC	%47	Tm(Basic)	*65.58°C*	Total nmol	38.66*nmol*
Skala	50 *nmol*	MW	11111.30	Conc	1534.29*ng/µl*	Total ng	429600.11*ng*
Saflastirma	DSLT	A260	50.1	OD	14.0	100 µM stok - µl TE	386.6

#2. HMPREF0495_02198_RS

5'-GTG GCG GCC GCT CTA TTA TCG CGT TAC TTT ATA GGA-3' - 36 bp

Oligo No	220330-1-58	GC	%50	Tm(Basic)	*66.72°C*	Total nmol	39.86*nmol*
Skala	50 *nmol*	MW	11073.24	Conc	1576.29*ng/µl*	Total ng	441360.32*ng*
Saflastirma	DSLT	A260	48.2	OD	13.5	100 µM stok - µl TE	398.6

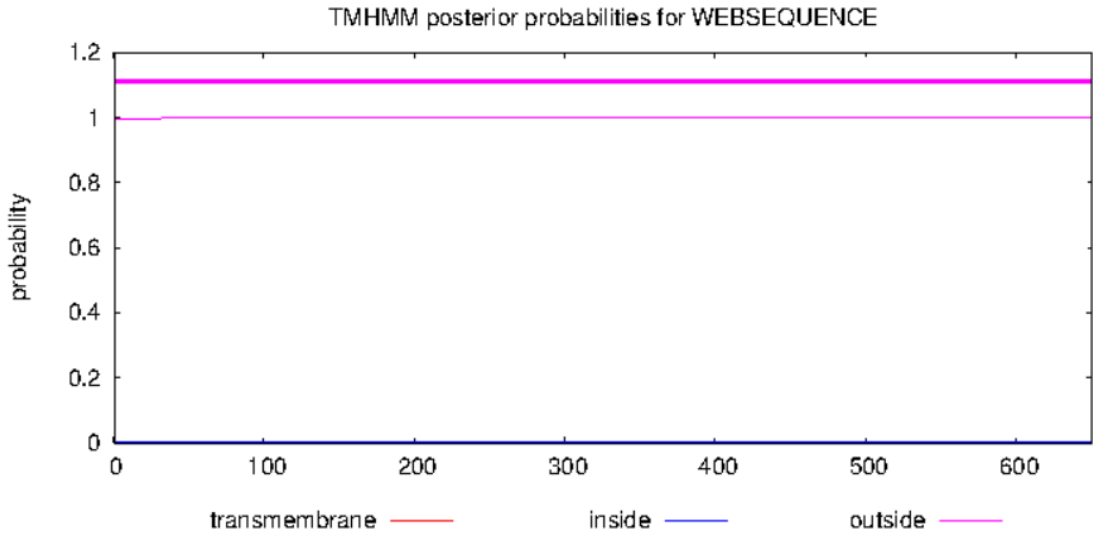
Figure 42. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of ERK41518.1.

17. ACJ51836.1 – Blon_0732

Bifidobacterium longum subsp. infantis ATCC 15697

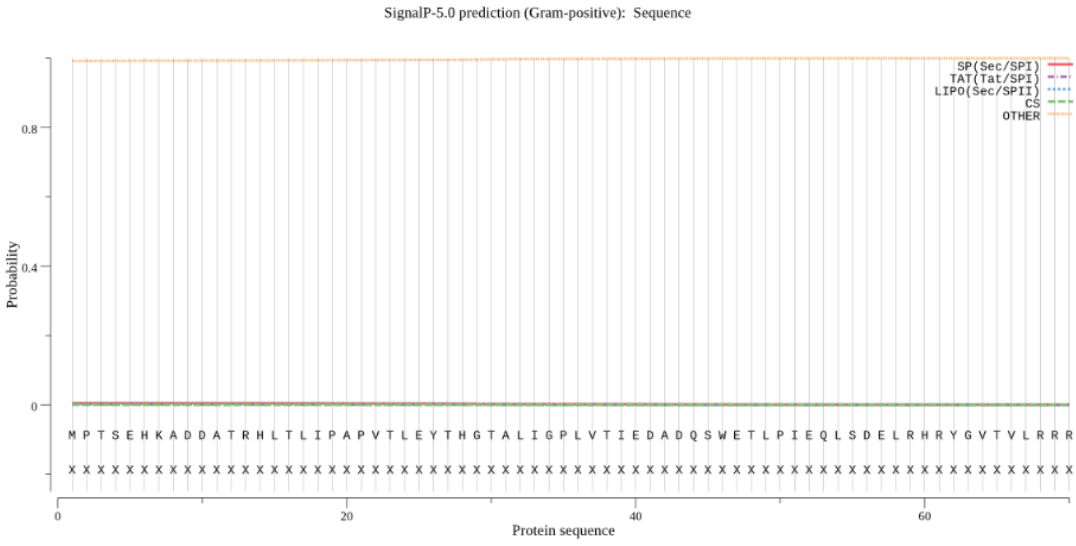
TMHMM result

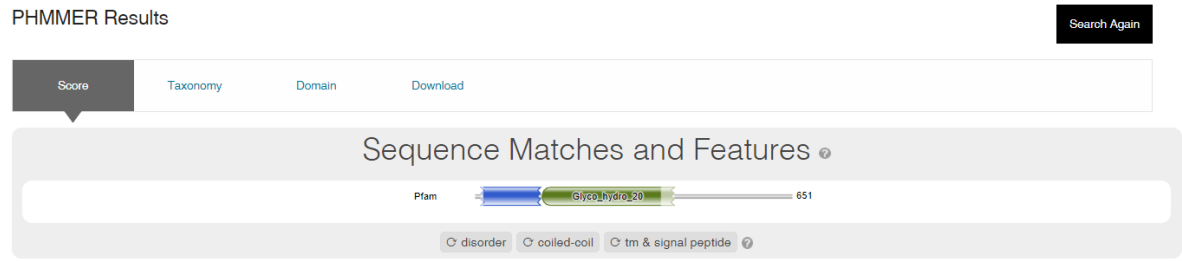
```
# WEBSEQUENCE Length: 651
# WEBSEQUENCE Number of predicted TMHs:  0
# WEBSEQUENCE Exp number of AAs in TMHs: 0.0378200000000001
# WEBSEQUENCE Exp number, first 60 AAs:  0.03282
# WEBSEQUENCE Total prob of N-in:        0.00222
WEBSEQUENCE     TMHMM2.0        outside      1   651
```



Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.0044	0.0007	0.0014	0.9934

Download: PNG / EPS / Tabular





-Primer

#1. Blon_0732FC

5'-GAA GGA GAT ATA CAT ATG CCC ACT TCC GAA CAT AAG GCC-3' - 39 bp

Oligo No	210505-1-20	GC	%46	Tm(Basic)	66.58°C	Total nmol	35.79nmol
Skala	50 nmol	MW	11977.86	Conc	1530.91ng/µl	Total ng	428655.52ng
Saflastirma	DSLT	A260	49.9	OD	14.0	100 µM stok - µl TE	357.9

#2. Blon_0732RC

5'-GTG ATG GTG GTG ATG ATG CAG AGC GCC CCT ACG CAG AAT-3' - 39 bp

Oligo No	210505-1-21	GC	%56	Tm(Basic)	70.79°C	Total nmol	46.46nmol
Skala	50 nmol	MW	12112.91	Conc	2009.73ng/µl	Total ng	562725.74ng
Saflastirma	DSLT	A260	63.3	OD	17.7	100 µM stok - µl TE	464.6

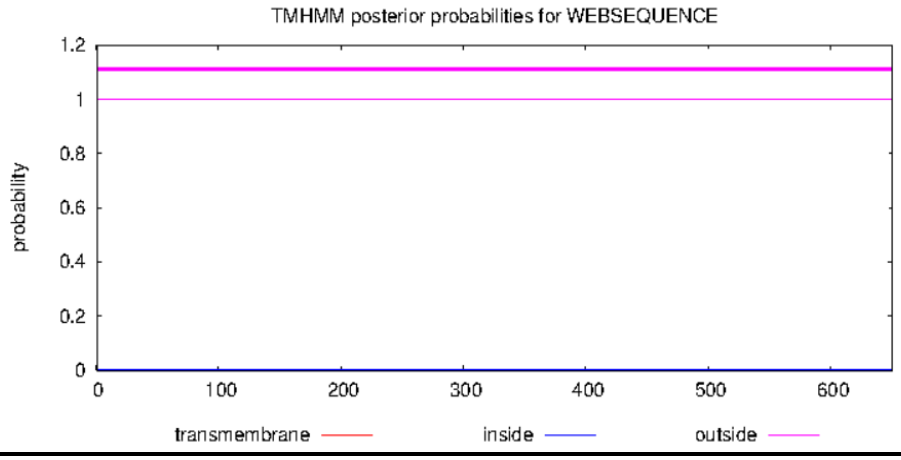
Figure 43. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of ACJ51836.1.

18. ACJ53413.1 – Blon_2355

Bifidobacterium longum subsp. infantis ATCC 15697

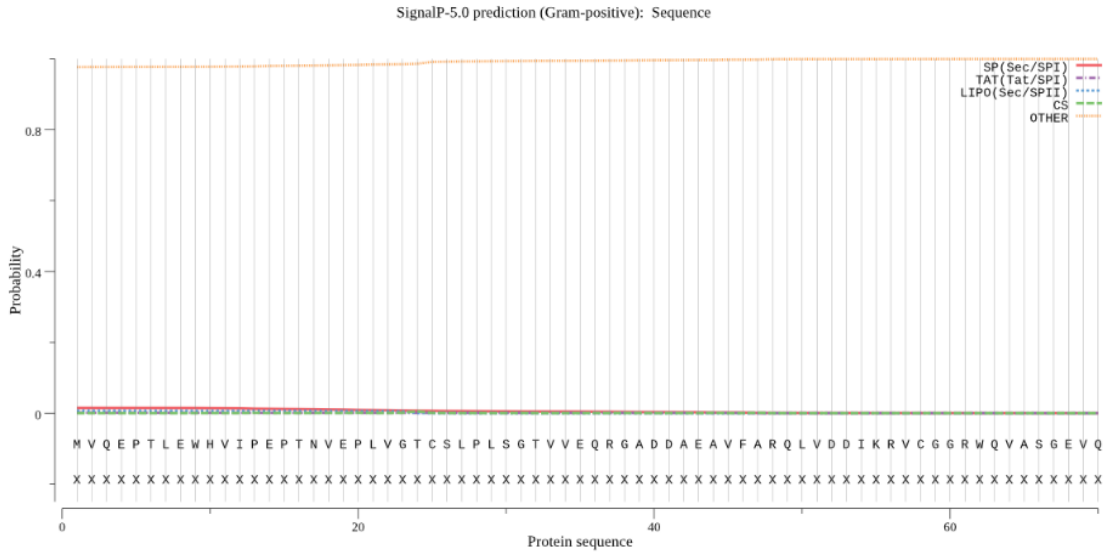
TMHMM result

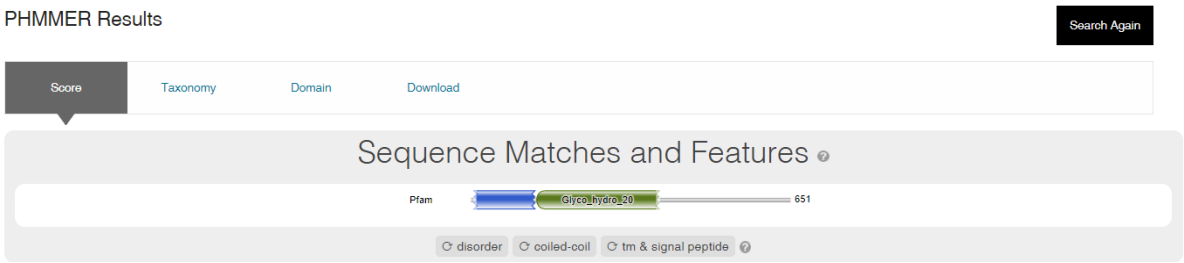
```
# WEBSEQUENCE Length: 651
# WEBSEQUENCE Number of predicted TMHs:  0
# WEBSEQUENCE Exp number of AAs in TMHs: 0.00156
# WEBSEQUENCE Exp number, first 60 AAs:  0
# WEBSEQUENCE Total prob of N-in:        0.00022
WEBSEQUENCE     TMHMM2.0        outside     1   651
```



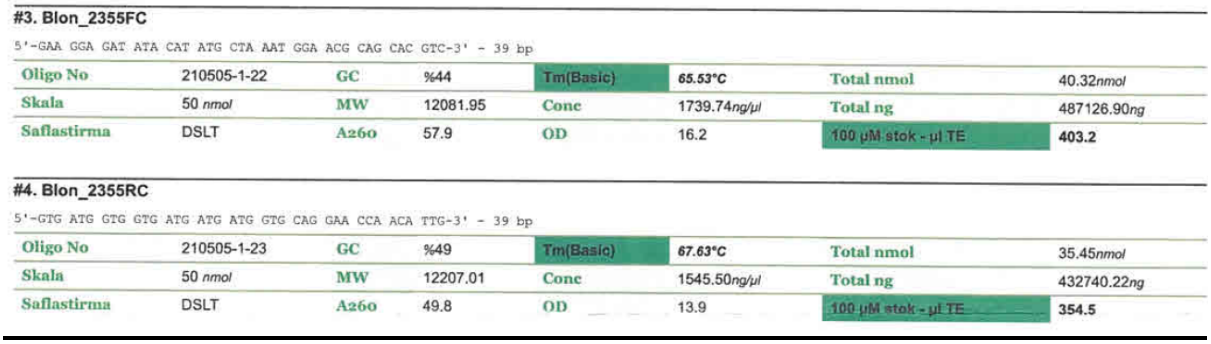
Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.0144	0.001	0.0068	0.9778

Download: PNG / EPS / Tabular





-Primer



#3. Blon_2355FC

5'-GAA GGA GAT ATA CAT ATG CTA AAT GGA ACG CAG CAC GTC-3' - 39 bp

Oligo No	210505-1-22	GC	%44	Tm(Basic)	65.53°C	Total nmol	40.32nmol
Skala	50 nmol	MW	12081.95	Conc	1739.74ng/µl	Total ng	487126.90ng
Saflastirma	DSLT	A260	57.9	OD	16.2	100 µM stok - µl TE	403.2

#4. Blon_2355RC

5'-GTG ATG GTG GTG ATG ATG ATG GTG CAG GAA CCA ACA TTG-3' - 39 bp

Oligo No	210505-1-23	GC	%49	Tm(Basic)	67.63°C	Total nmol	35.45nmol
Skala	50 nmol	MW	12207.01	Conc	1545.50ng/µl	Total ng	432740.22ng
Saflastirma	DSLT	A260	49.8	OD	13.9	100 µM stok - µl TE	354.5

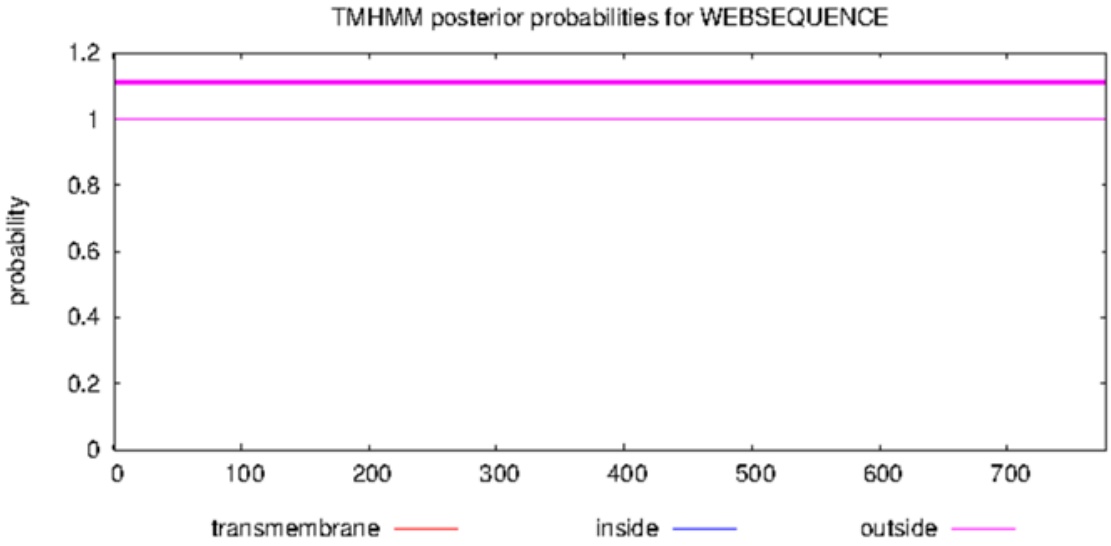
Figure 44. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of ACJ53413.1.

19. QQA29671.1 – I6G58_17045

Bacteroides uniformis FDAARGOS_901 **ATCC8492**

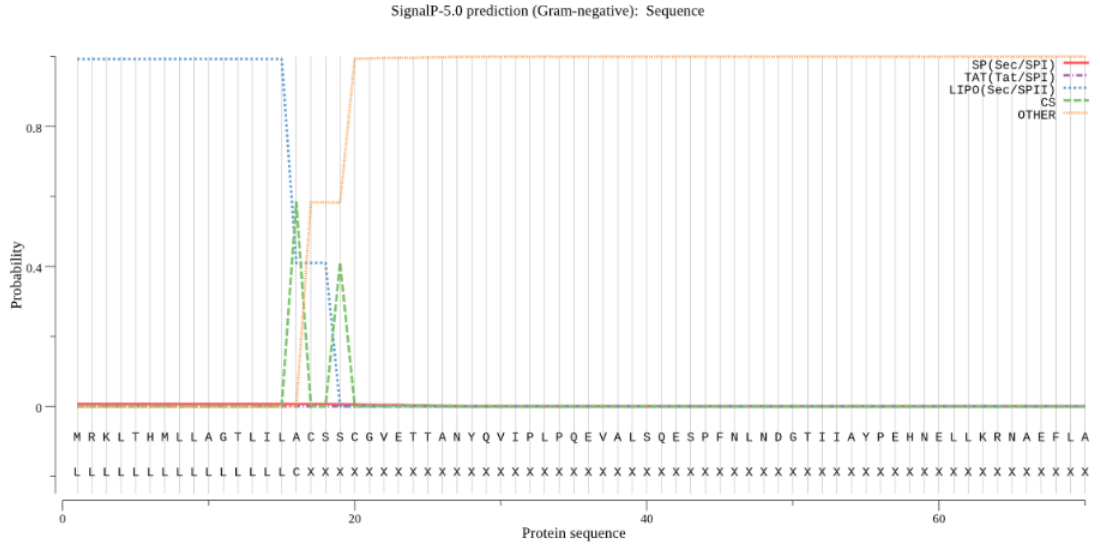
TMHMM result

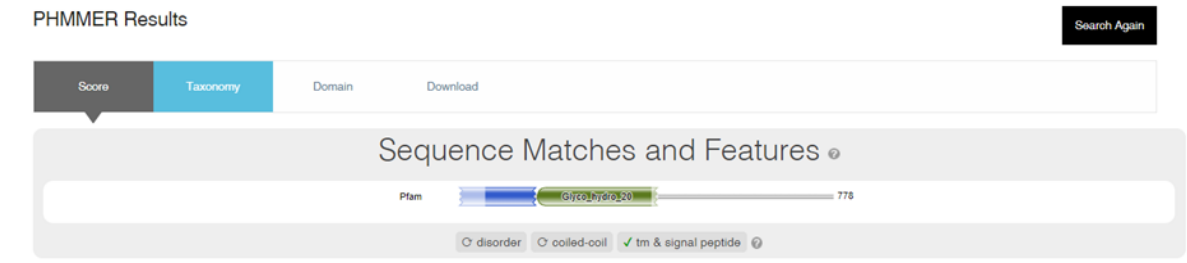
```
# WEBSEQUENCE Length: 778
# WEBSEQUENCE Number of predicted TMHs:  0
# WEBSEQUENCE Exp number of AAs in TMHs: 0.00176
# WEBSEQUENCE Exp number, first 60 AAs:  0
# WEBSEQUENCE Total prob of N-in:        0.00003
WEBSEQUENCE     TMHMM2.0        outside      1   778
```



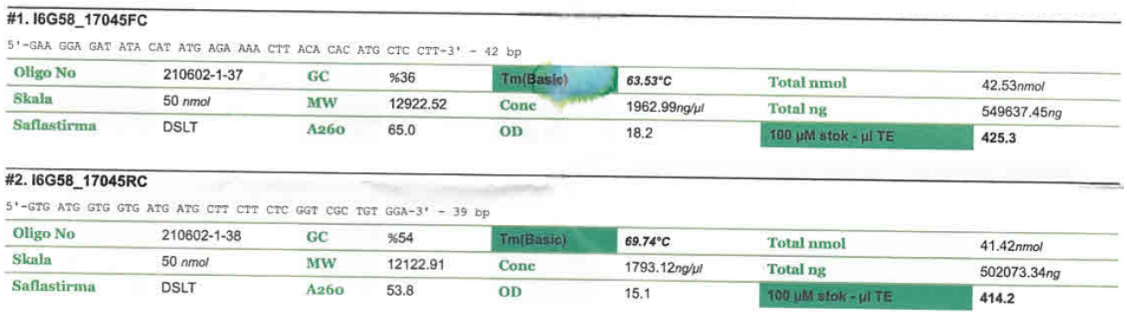
Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.007	0.0001	0.9924	0.0004

Download: PNG / EPS / Tabular





-Primer



#1. I6G58_17045FC

5'-GAA GGA GAT ATA CAT ATG AGA AAA CTT ACA CAC ATG CTC CTT-3' - 42 bp

Oligo No	210602-1-37	GC	%36	Tm(Basic)	63.53°C	Total nmol	42.53nmol
Skala	50 nmol	MW	12922.52	Conc	1962.99ng/µl	Total ng	549637.45ng
Saflastirma	DSLT	A260	65.0	OD	18.2	100 µM stok - µl TE	425.3

#2. I6G58_17045RC

5'-GTG ATG GTG GTG ATG ATG CTT CTT CTC GGT CGC TGT GGA-3' - 39 bp

Oligo No	210602-1-38	GC	%54	Tm(Basic)	69.74°C	Total nmol	41.42nmol
Skala	50 nmol	MW	12122.91	Conc	1793.12ng/µl	Total ng	502073.34ng
Saflastirma	DSLT	A260	53.8	OD	15.1	100 µM stok - µl TE	414.2

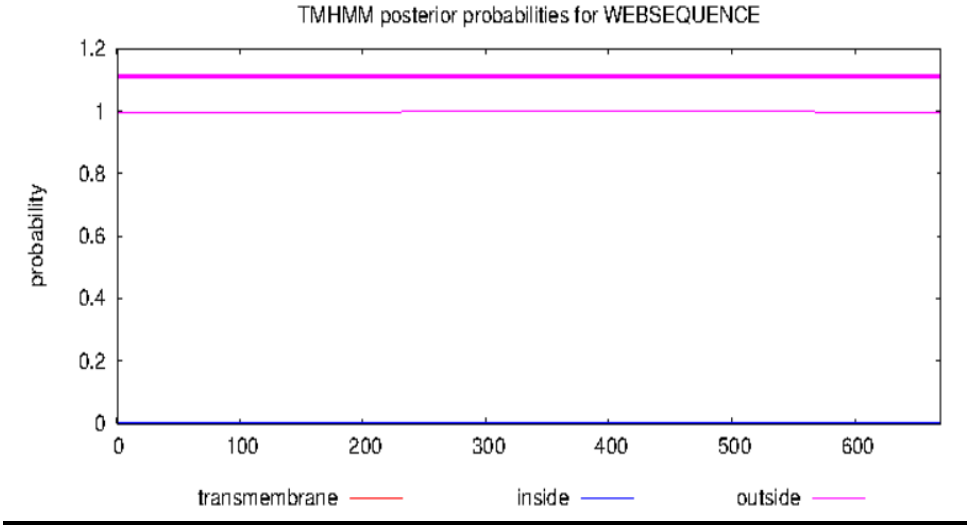
Figure 45. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of QQA29671.1.

20. BAQ30021.1 - BBKW_1886

Bifidobacterium catenulatum subsp. kashiwanohense JCM 15439

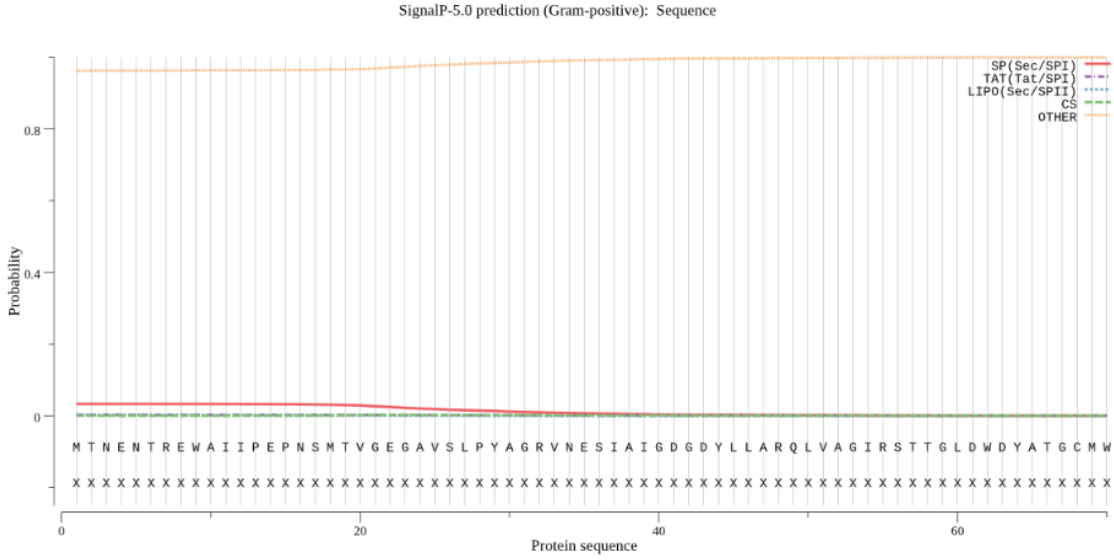
TMHMM result

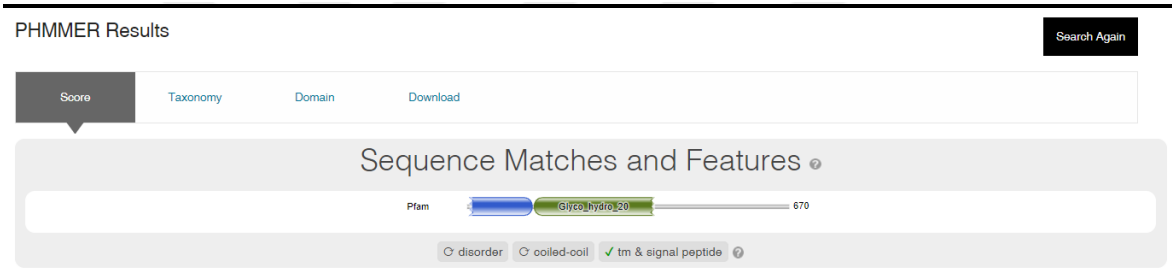
```
# WEBSEQUENCE Length: 670
# WEBSEQUENCE Number of predicted TMHs:  0
# WEBSEQUENCE Exp number of AAs in TMHs: 0.03036
# WEBSEQUENCE Exp number, first 60 AAs:  0.00331
# WEBSEQUENCE Total prob of N-in:        0.00255
WEBSEQUENCE     TMHMM2.0        outside     1   670
```



Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.0333	0.002	0.0014	0.9633

Download: PNG / EPS / Tabular





-Primer

#5. BBKW_1886FC

5'-GAA GGA GAT ATA CAT ATG TTT GCT GAT TCG CCG CAA CTC-3' - 39 bp

Oligo No	210505-1-30	GC	%44	Tm(Basic)	65.53°C	Total nmol	37.54nmol
Skala	50 nmol	MW	12005.88	Conc	1609.76ng/µl	Total ng	450732.70ng
Saflastirma	DSLT	A260	51.2	OD	14.3	100 µM stok - µl TE	375.4

#6. BBKW_1886RC

5'-GTG ATG GTG GTG ATG ATG ATG ACA AAC GAG AAT ACC CGT-3' - 39 bp

Oligo No	210505-1-31	GC	%46	Tm(Basic)	66.58°C	Total nmol	39.72nmol
Skala	50 nmol	MW	12159.99	Conc	1724.91ng/µl	Total ng	482975.97ng
Saflastirma	DSLT	A260	56.3	OD	15.8	100 µM stok - µl TE	397.2

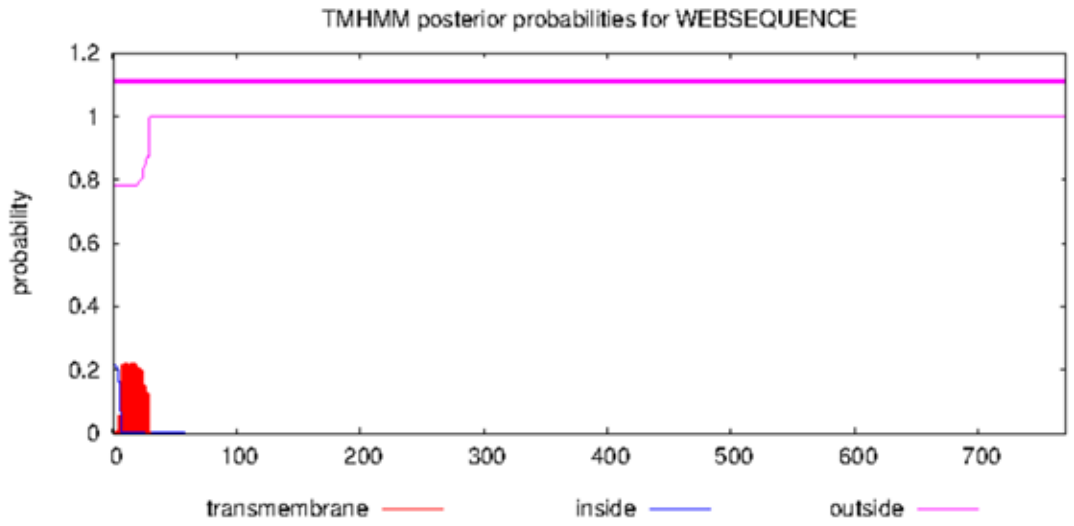
Figure 46. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of BAQ30021.1.

21. ABR38247.1 – BVU_0537

***Bacteroides vulgatus* ATCC 8482**

TMHMM result

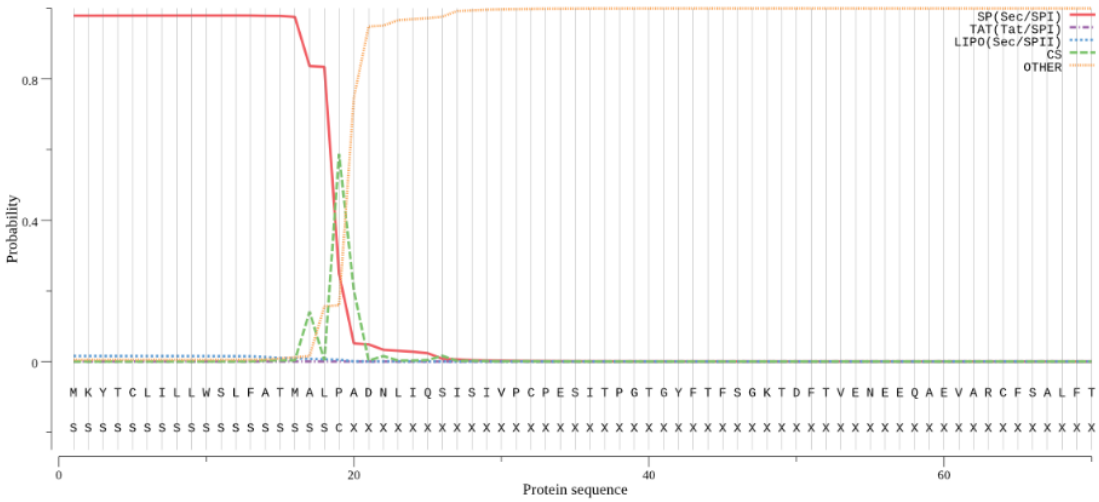
```
# WEBSEQUENCE Length: 771
# WEBSEQUENCE Number of predicted TMHs:  0
# WEBSEQUENCE Exp number of AAs in TMHs: 4.74050999999997
# WEBSEQUENCE Exp number, first 60 AAs:  4.73833
# WEBSEQUENCE Total prob of N-in:        0.21964
WEBSEQUENCE     TMHMM2.0        outside      1   771
```

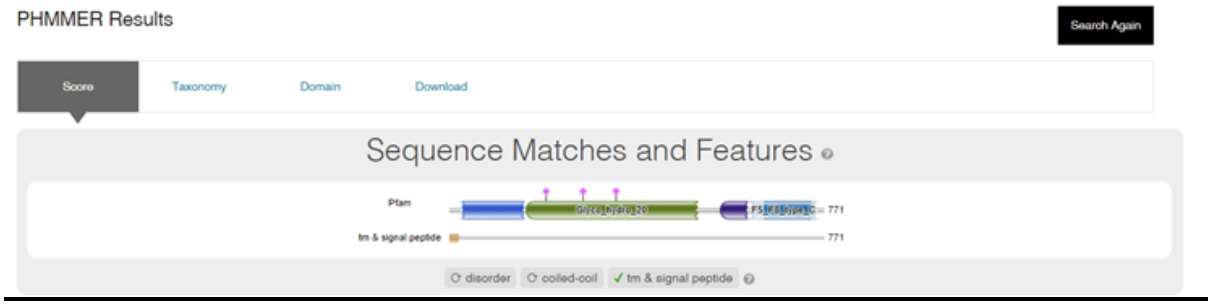


Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.9788	0.0003	0.0156	0.0053

Download: PNG / EPS / Tabular

SignalP-5.0 prediction (Gram-negative): Sequence





-Primer

#3. BVU_0537_FN

5'-CAT CAT CAC CAC CAT CAC GCT GAC AAC CTA ATC CAA-3' - 36 bp

Oligo No	220330-1-71	GC	%47	Tm(Basic)	65.58°C	Total nmol	39.16nmol
Skala	50 nmol	MW	10831.09	Conc	1514.74ng/µl	Total ng	424128.01ng
Saflastirma	DSLT	A260	47.8	OD	13.4	100 µM stok - µl TE	391.6

#4. BVU_0537_RN

5'-GTG GCG GCC GCT CTA TTA TTC TAT CAT TAC CTC ATC-3' - 36 bp

Oligo No	220330-1-72	GC	%47	Tm(Basic)	65.58°C	Total nmol	43.48nmol
Skala	50 nmol	MW	10928.14	Conc	1696.92ng/µl	Total ng	475136.52ng
Saflastirma	DSLT	A260	50.7	OD	14.2	100 µM stok - µl TE	434.8

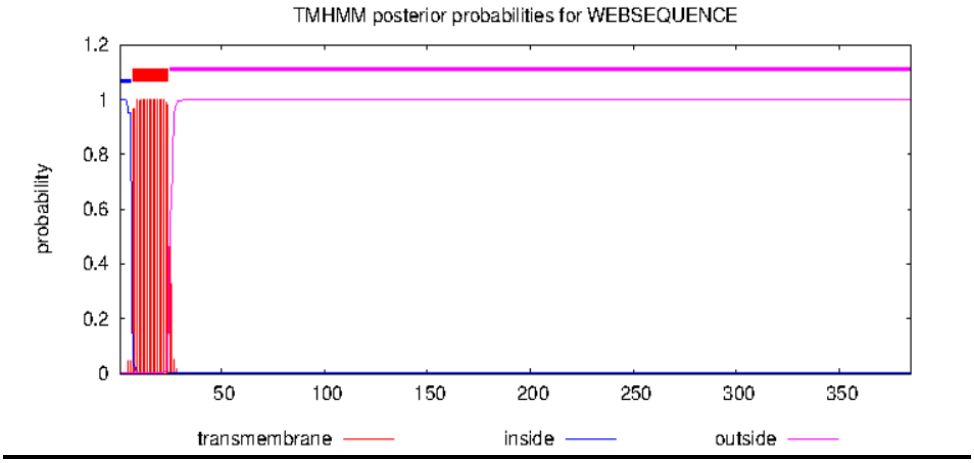
Figure 47. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of ABR38247.1.

22. SQF24907.1 – NCTC12958_01101

Streptococcus thermophilus **ATCC 19258**

TMHMM result

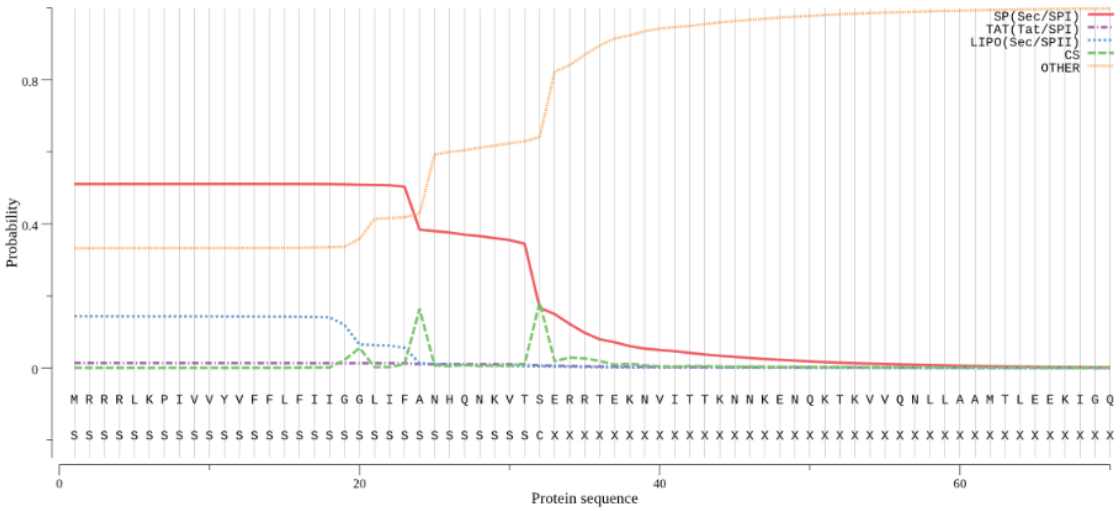
```
# WEBSEQUENCE Length: 385
# WEBSEQUENCE Number of predicted TMHs:  1
# WEBSEQUENCE Exp number of AAs in TMHs: 18.8839699999999
# WEBSEQUENCE Exp number, first 60 AAs:  18.87829
# WEBSEQUENCE Total prob of N-in:        0.99806
# WEBSEQUENCE POSSIBLE N-term signal sequence
WEBSEQUENCE     TMHMM2.0        inside       1     6
WEBSEQUENCE     TMHMM2.0        TMhelix      7    24
WEBSEQUENCE     TMHMM2.0        outside     25   385
```



Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.5097	0.0139	0.1437	0.3328

Download: PNG / EPS / Tabular

SignalP-5.0 prediction (Gram-positive): Sequence



PHMMER Results

Search Again

Score | Taxonomy | Domain | Download

Sequence Matches and Features

disorder | coiled-coil | ✓ tm & signal peptide

-Primer

#1. NCTC12958_01101_FS

5'-CGC GAA CAG ATT GGA GGT AAC CAT CAG AAT AAA GTA ACA TCA-3' - 42 bp

Oligo No	220222-2-1	GC	%40	Tm(Basic)	65.49°C	Total nmol	35.40nmol
Skala	50 nmol	MW	12981.55	Conc	1641.44ng/µl	Total ng	459602.62ng
Saflastirma	DSLT	A260	55.4	OD	15.5	100 µM stok - µl TE	354.0

#2. NCTC12958_01101_RS

5'-GTG GCG GCC GCT CTA TTA TTT CTC TAG AGT TGA TAA CAA TCC-3' - 42 bp

Oligo No	220222-2-2	GC	%45	Tm(Basic)	67.44°C	Total nmol	33.13nmol
Skala	50 nmol	MW	12870.42	Conc	1522.89ng/µl	Total ng	426409.97ng
Saflastirma	DSLT	A260	46.7	OD	13.1	100 µM stok - µl TE	331.3

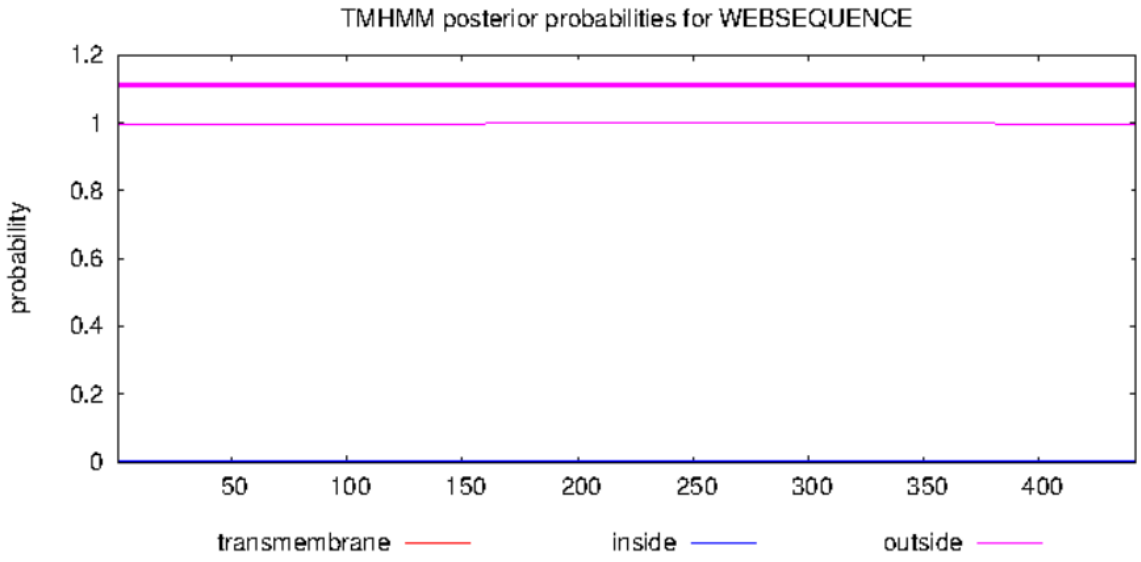
Figure 48. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of SQF24907.1.

23. SQF25661.1 – NCTC12958_01892

Streptococcus thermophilus **ATCC 19258**

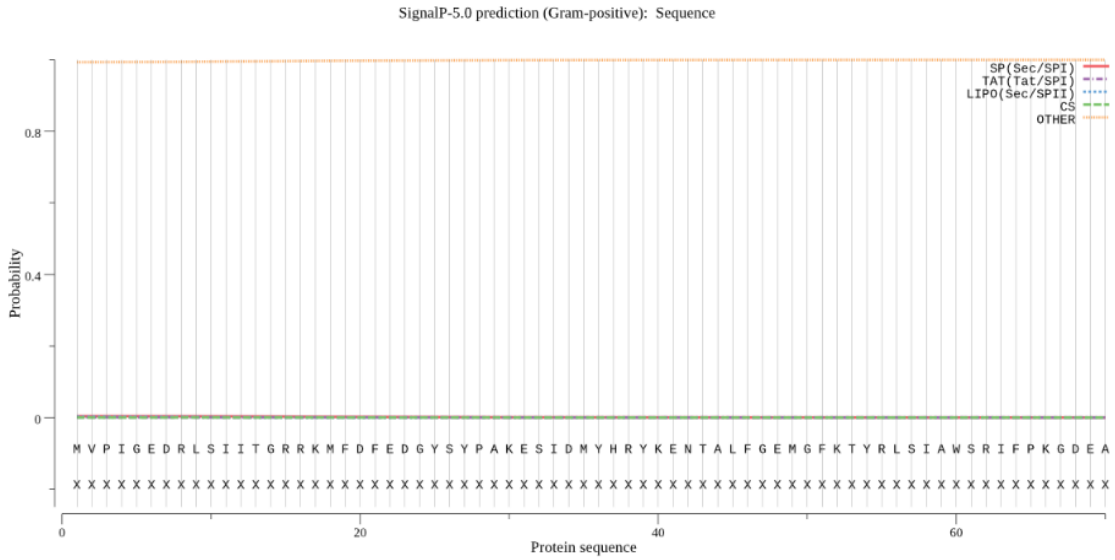
TMHMM result

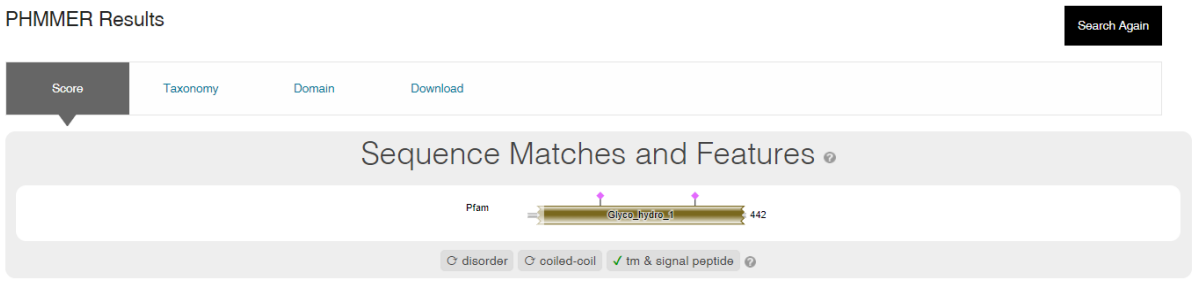
```
# WEBSEQUENCE Length: 442
# WEBSEQUENCE Number of predicted TMHs:  0
# WEBSEQUENCE Exp number of AAs in TMHs: 0.05177
# WEBSEQUENCE Exp number, first 60 AAs:  0.00054
# WEBSEQUENCE Total prob of N-in:        0.00339
WEBSEQUENCE     TMHMM2.0        outside      1   442
```



Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.0036	0.0004	0.0017	0.9943

Download: PNG / EPS / Tabular





-Primer

#3. NCTC12958_01892_FS

5'-CGC GAA CAG ATT GGA GGT GTA CCG ATT GGT GAG GAT-3' - 36 bp

Oligo No	220223-2-4	GC	%53	Tm(Basic)	67.86°C	Total nmol	38.22nmol
Skala	50 nmol	MW	11245.37	Conc	1534.85ng/µl	Total ng	429759.36ng
Saflastirma	DSLT	A260	49.4	OD	13.8	100 µM stok - µl TE	382.2

#4. NCTC12958_01892_RS

5'-GTG GCG GCC GCT CTA TTA TTC TAC AGA GGC ACC GTT-3' - 36 bp

Oligo No	220223-2-5	GC	%56	Tm(Basic)	69.00°C	Total nmol	43.79nmol
Skala	50 nmol	MW	11043.20	Conc	1727.13ng/µl	Total ng	483595.44ng
Saflastirma	DSLT	A260	52.1	OD	14.6	100 µM stok - µl TE	437.9

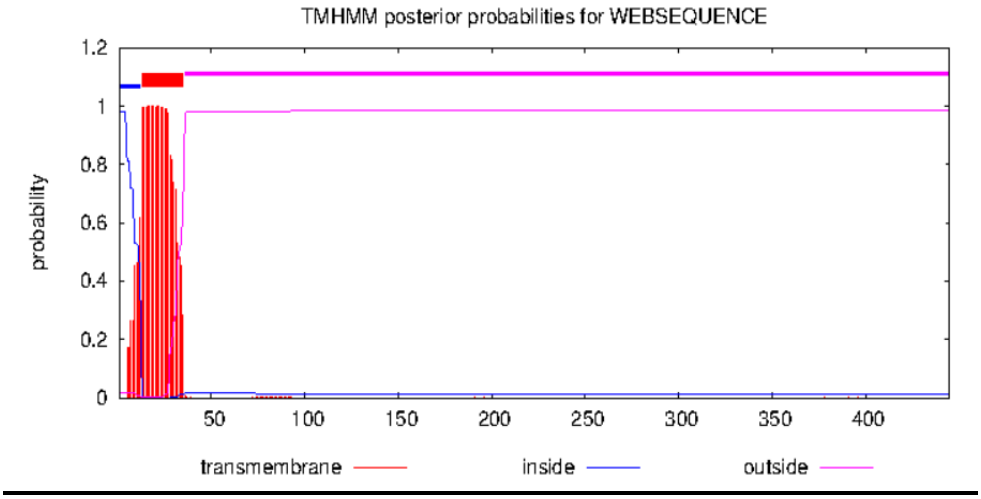
Figure 49. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of SQF25661.1.

24. SQF24918.1 – NCTC12958_01112

Streptococcus thermophilus **ATCC 19258**

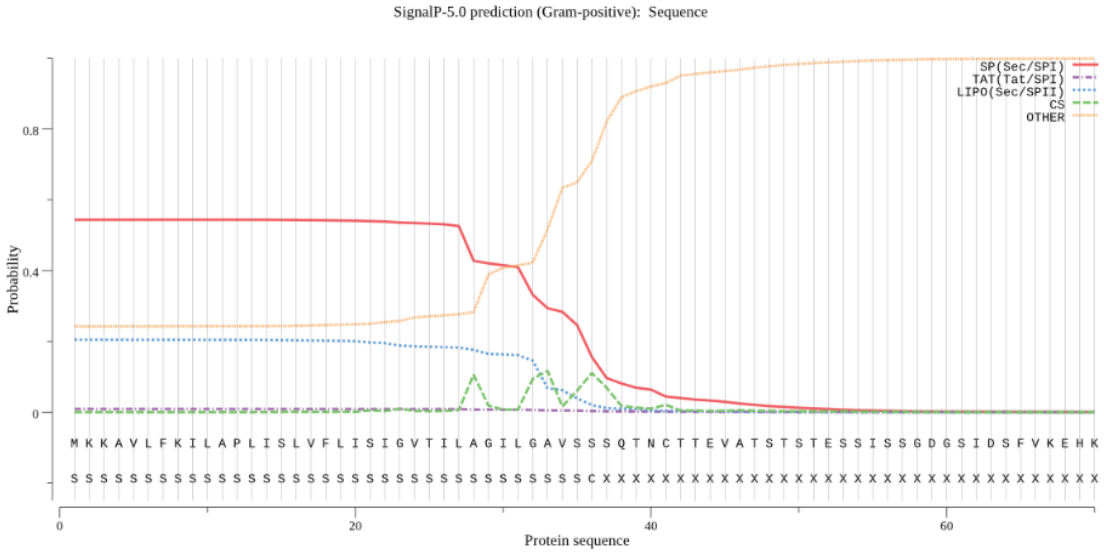
TMHMM result

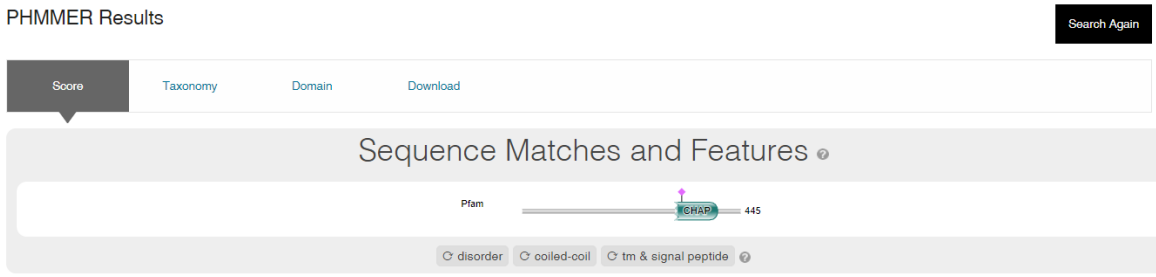
```
# WEBSEQUENCE Length: 445
# WEBSEQUENCE Number of predicted TMHs:  1
# WEBSEQUENCE Exp number of AAs in TMHs: 22.77875
# WEBSEQUENCE Exp number, first 60 AAs:  22.69423
# WEBSEQUENCE Total prob of N-in:        0.98154
# WEBSEQUENCE POSSIBLE N-term signal sequence
WEBSEQUENCE     TMHMM2.0        inside       1    12
WEBSEQUENCE     TMHMM2.0        TMhelix     13    35
WEBSEQUENCE     TMHMM2.0        outside     36   445
```



Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.5439	0.0096	0.2043	0.2422

Download: PNG / EPS / Tabular





-Primer

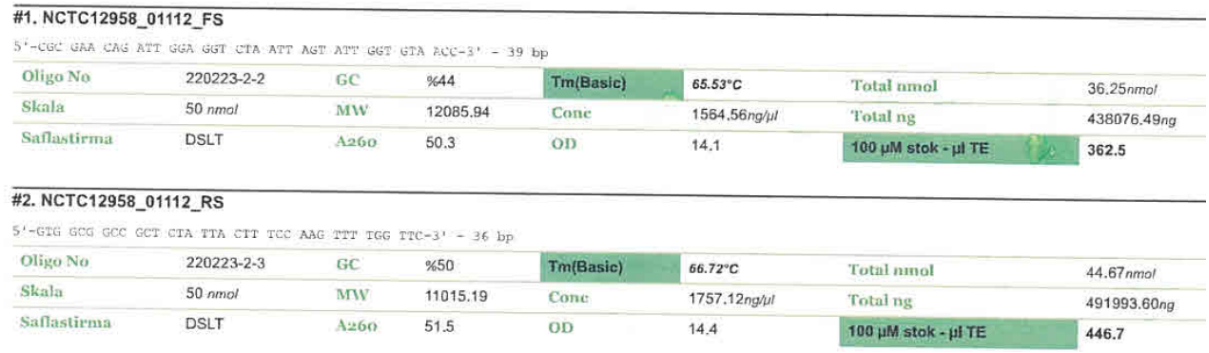


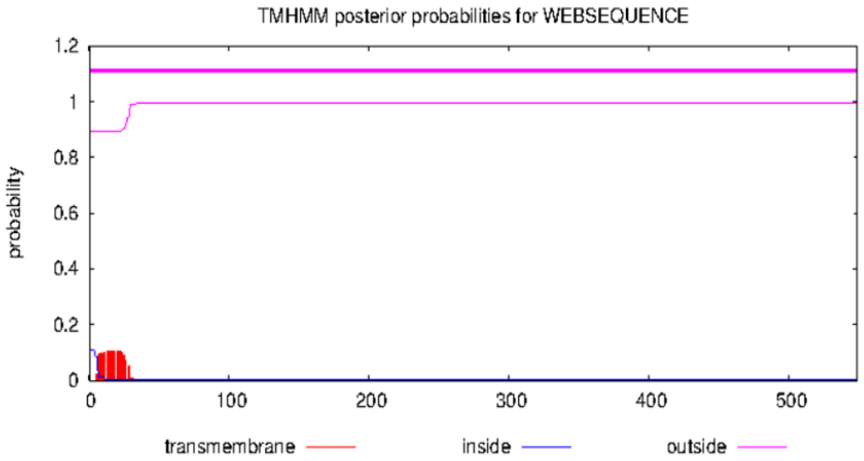
Figure 50. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of SQF24918.1.

25. ACD04701.1 – Amuc_0868

Akkermansia muciniphila **ATCC BAA-835**

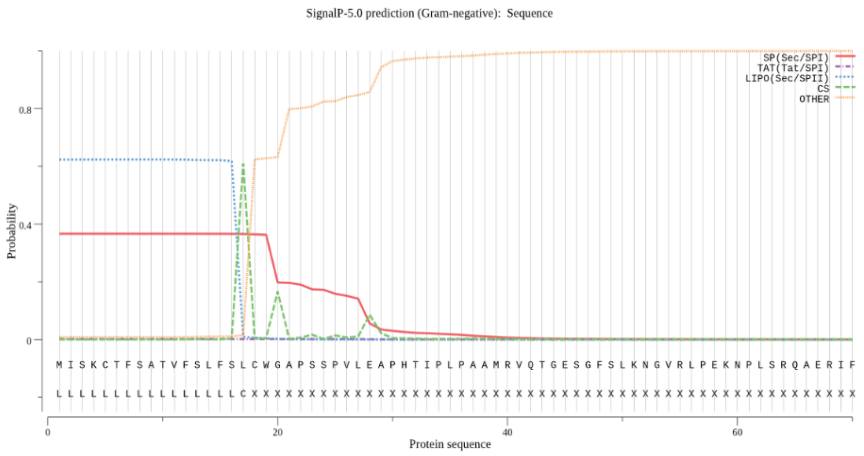
TMHMM result

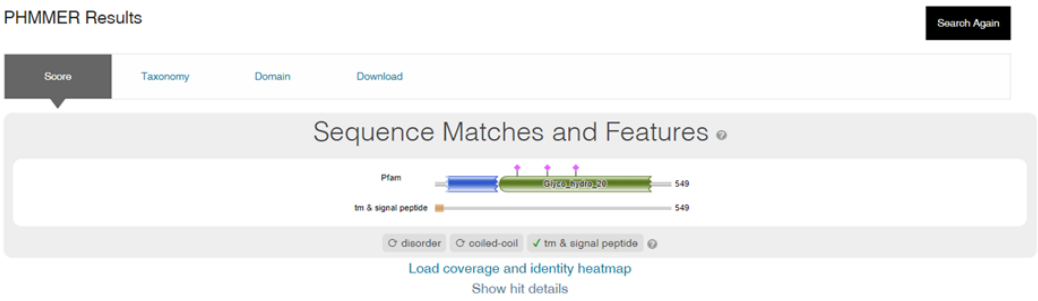
```
# WEBSEQUENCE Length: 549
# WEBSEQUENCE Number of predicted TMHs:  0
# WEBSEQUENCE Exp number of AAs in TMHs: 2.28024
# WEBSEQUENCE Exp number, first 60 AAs:  2.2697
# WEBSEQUENCE Total prob of N-in:        0.10861
WEBSEQUENCE     TMHMM2.0        outside     1   549
```



Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.3661	0.0016	0.6242	0.0081

Download: PNG / EPS / Tabular





Primer

#2. Amuc_0868_FS

5'-CGC GAA CAG ATT GGA GGT GCC CCA TCC TCT CCG GTT-3' - 36 bp

Oligo No	220330-1-46	GC	%61	Tm(Basic)	71.28°C	Total nmol	41.20*nmol*
Skala	50 *nmol*	MW	11013.16	Conc	1620.49*ng/µl*	Total ng	453738.19*ng*
Saflastirma	DSLT	A260	48.6	OD	13.6	100 µM stok - µl TE	412.0

#3. Amuc_0868_RS

5'-GTG GCG GCC GCT CTA TTA ACG GCG TTC CCG CGT AAT-3' - 36 bp

Oligo No	220330-1-47	GC	%61	Tm(Basic)	71.28°C	Total nmol	41.06*nmol*
Skala	50 *nmol*	MW	11044.18	Conc	1619.49*ng/µl*	Total ng	453456.30*ng*
Saflastirma	DSLT	A260	48.1	OD	13.5	100 µM stok - µl TE	410.6

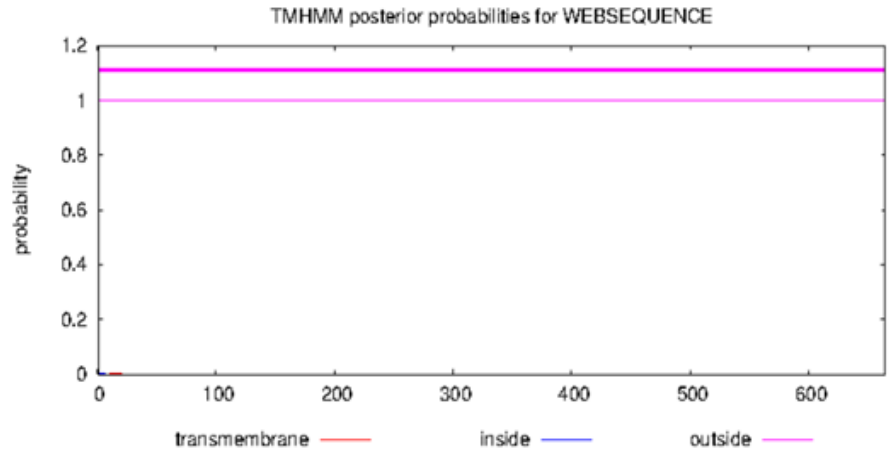
Figure 51. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of ACD04701.1.

26. ACD04208.1 –Amuc_0369

Akkermansia muciniphila **ATCC BAA-835**

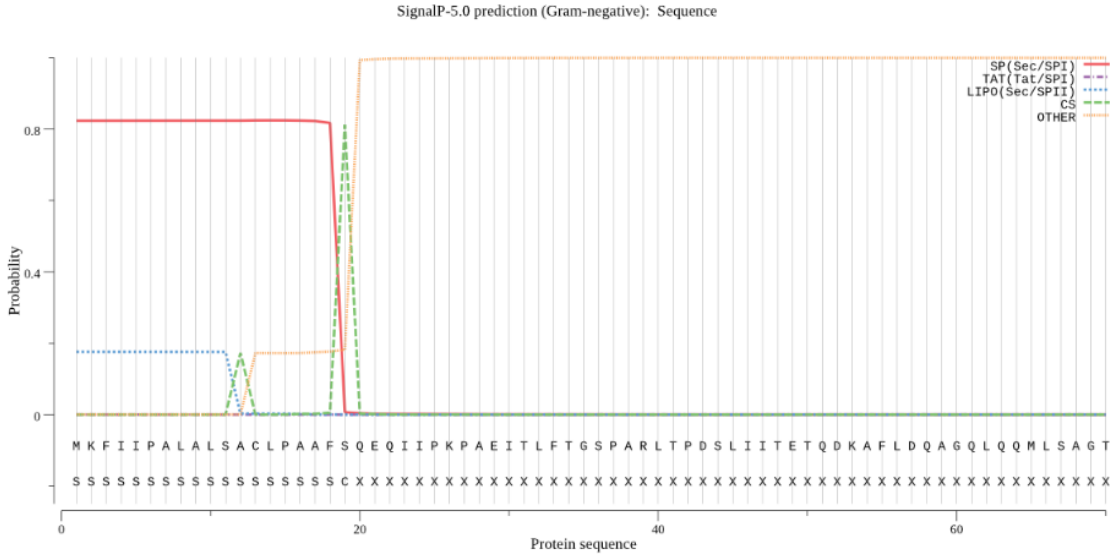
TMHMM result

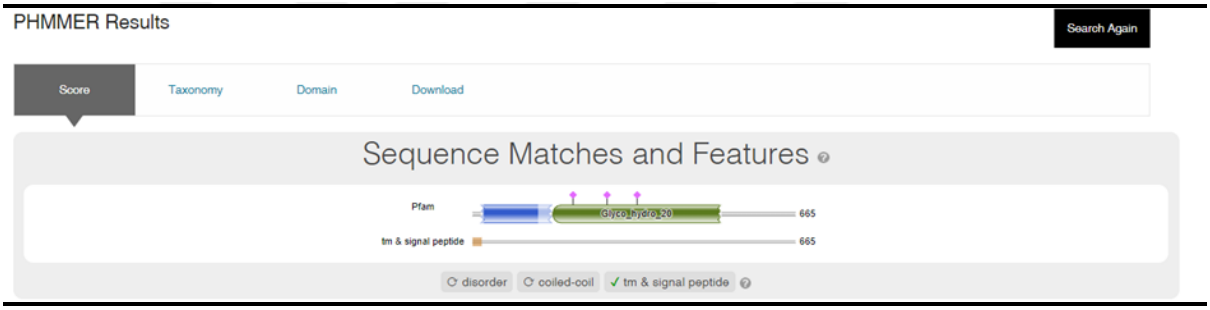
```
# WEBSEQUENCE Length: 665
# WEBSEQUENCE Number of predicted TMHs:  0
# WEBSEQUENCE Exp number of AAs in TMHs: 0.0344500000000001
# WEBSEQUENCE Exp number, first 60 AAs:  0.03378
# WEBSEQUENCE Total prob of N-in:        0.00180
WEBSEQUENCE     TMHMM2.0        outside     1   665
```



Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.8239	0.0002	0.1751	0.0008

Download: PNG / EPS / Tabular





-Primer

#5. Amuc_0369_FS

5'-CGC GAA CAG ATT GGA GGT CAA GAA CAA ATC ATC CCG-3' - 36 bp

Oligo No	220330-1-49	GC	%50	Tm(Basic)	66.72°C	Total nmol	33.99nmol
Skala	50 nmol	MW	11096.28	Conc	1346.90ng/µl	Total ng	377132.48ng
Saflastirma	DSLT	A260	44.0	OD	12.3	100 µM stok - µl TE	339.9

#6. Amuc_0369_RS

5'-GTG GCG GCC GCT CTA TTA TGG CGT TTC AAA ATA CAC-3' - 36 bp

Oligo No	220330-1-50	GC	%50	Tm(Basic)	66.72°C	Total nmol	36.56nmol
Skala	50 nmol	MW	11051.23	Conc	1442.84ng/µl	Total ng	403995.44ng
Saflastirma	DSLT	A260	44.3	OD	12.4	100 µM stok - µl TE	365.6

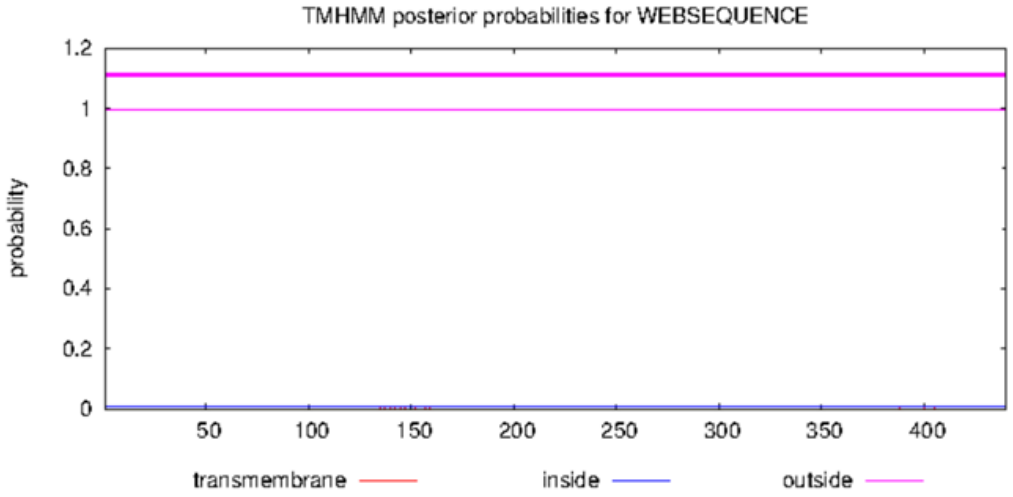
Figure 52. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of ACD04208.1.

27. BAQ97280.1 – BBBF_0073

Bifidobacterium bifidum **ATCC 29521**

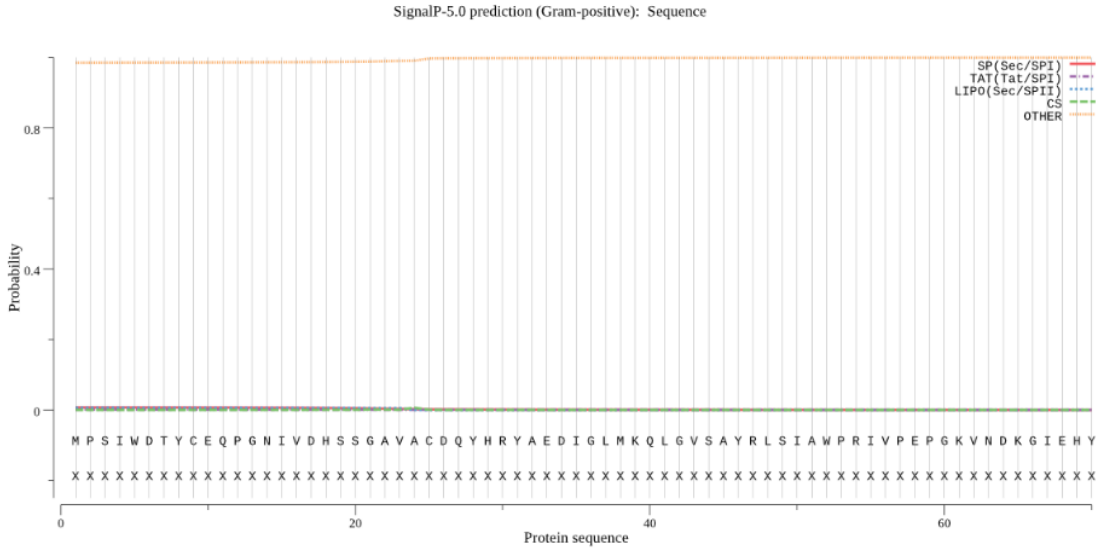
TMHMM result

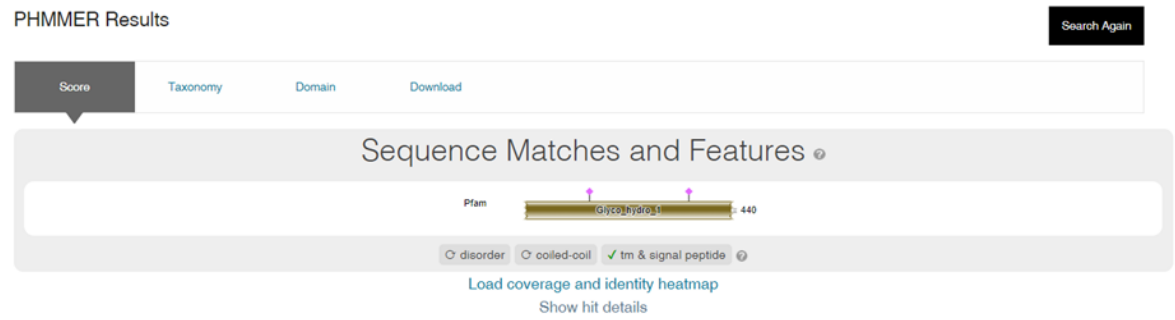
```
# WEBSEQUENCE Length: 440
# WEBSEQUENCE Number of predicted TMHs:  0
# WEBSEQUENCE Exp number of AAs in TMHs: 0.02262
# WEBSEQUENCE Exp number, first 60 AAs:  0.00061
# WEBSEQUENCE Total prob of N-in:        0.00627
WEBSEQUENCE     TMHMM2.0        outside      1   440
```



Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.0069	0.0008	0.0062	0.9861

Download: PNG / EPS / Tabular





-Primer

#4. BBBF_0073_FS

5'-CGC GAA CAG ATT GGA GGT CCC TCA ATA TGG GAC ACC-3' - 36 bp

Oligo No	220330-1-36	GC	%56	Tm(Basic)	69.00°C	Total nmol	46.80nmol
Skala	50 nmol	MW	11079.24	Conc	1851.96ng/µl	Total ng	518548.90ng
Saflastirma	DSLT	A260	58.5	OD	16.4	100 µM stok - µl TE	468.0

#5. BBBF_0073_RS

5'-GTG GCG GCC GCT CTA TTA CTG CAG ATA GCC ATC GGG-3' - 36 bp

Oligo No	220330-1-37	GC	%61	Tm(Basic)	71.28°C	Total nmol	40.97nmol
Skala	50 nmol	MW	11093.22	Conc	1623.33ng/µl	Total ng	454532.17ng
Saflastirma	DSLT	A260	49.1	OD	13.8	100 µM stok - µl TE	409.7

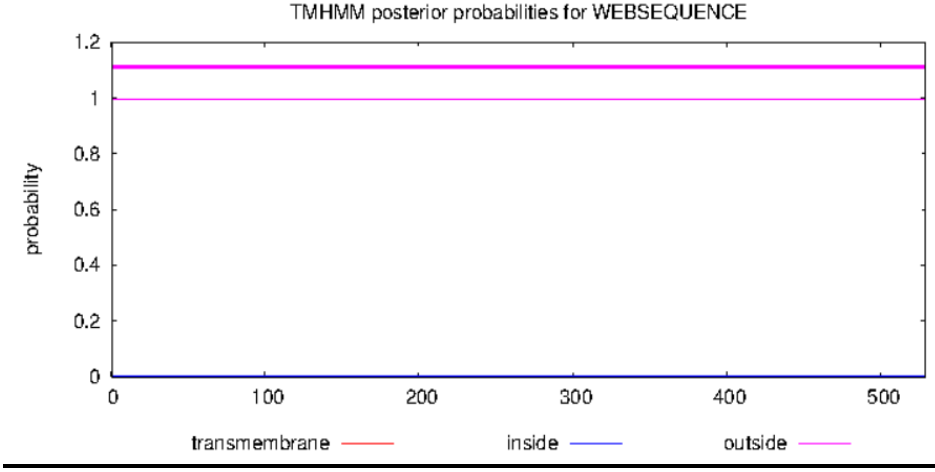
Figure 53. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of BAQ97280.1.

28. CAH09389.1 – BF9343_3608

Bacteroides fragilis **ATCC 25285**

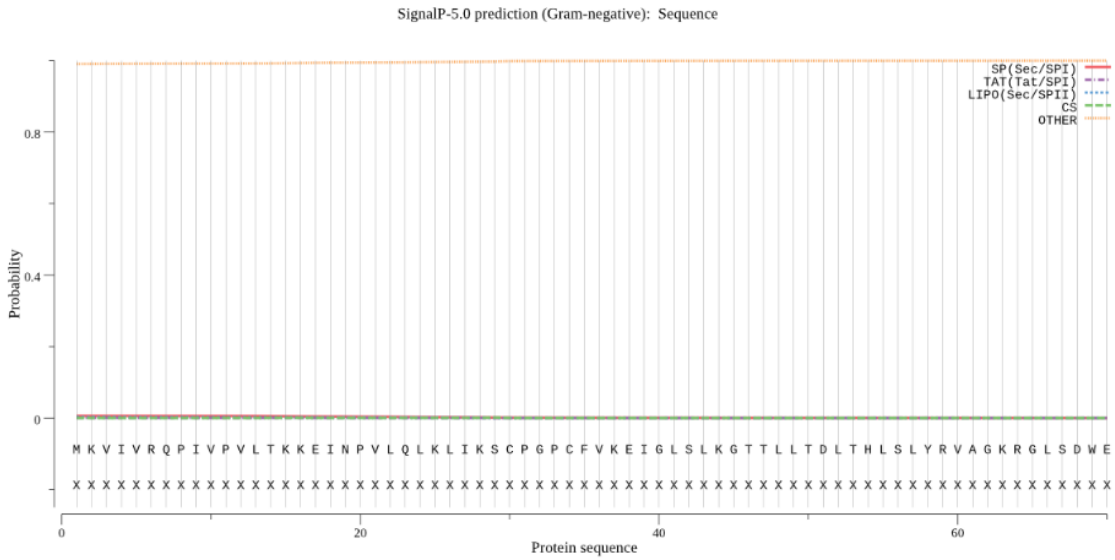
TMHMM result

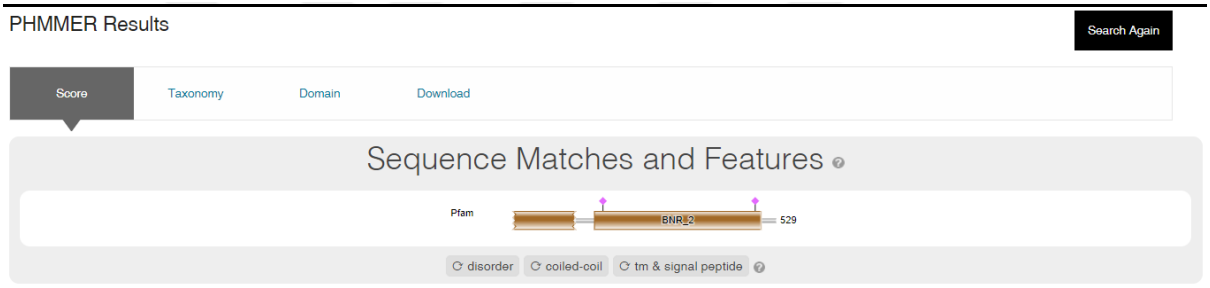
```
# WEBSEQUENCE Length: 529
# WEBSEQUENCE Number of predicted TMHs:  0
# WEBSEQUENCE Exp number of AAs in TMHs: 0.02158
# WEBSEQUENCE Exp number, first 60 AAs:  0.00175
# WEBSEQUENCE Total prob of N-in:        0.00377
WEBSEQUENCE     TMHMM2.0        outside      1   529
```



Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.0057	0.0004	0.0018	0.9921

Download: PNG / EPS / Tabular





-Primer

#3. BF9343_3608FN

5'-CAT CAT CAC CAC CAT CAC CTA TCT CTC TTT ATA TGA TTT ACC-3' - 42 bp

Oligo No	210505-1-10	GC	%38	Tm(Basic)	64.51°C	Total nmol	40.22nmol
Skala	50 nmol	MW	12613.26	Conc	1811.71ng/µl	Total ng	507279.53ng
Saflastirma	DSLT	A260	55.2	OD	15.5	100 µM stok - µl TE	402.2

#4. BF9343_3608RN

5'-GTG GCG GCC GCT CTA TTA GTG AAA GTT ATT GTT AGA CAG-3' - 39 bp

Oligo No	210505-1-11	GC	%46	Tm(Basic)	66.58°C	Total nmol	36.93nmol
Skala	50 nmol	MW	12092.93	Conc	1594.99ng/µl	Total ng	446595.90ng
Saflastirma	DSLT	A260	50.5	OD	14.1	100 µM stok - µl TE	369.3

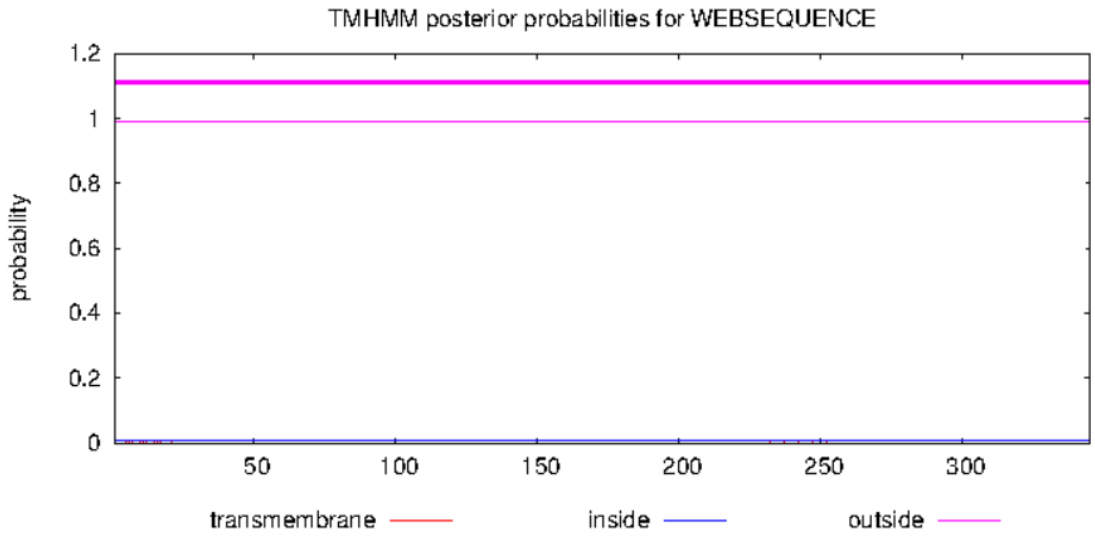
Figure 54. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of CAH09389.1.

29. ABR38963.1 - BVU_1273

Bacteroides vulgatus 8482

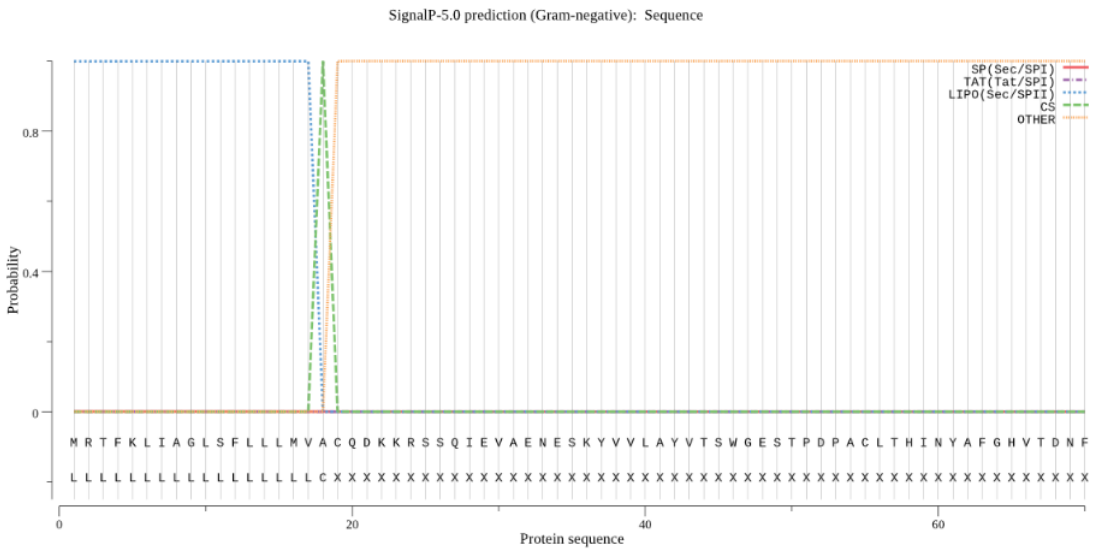
TMHMM result

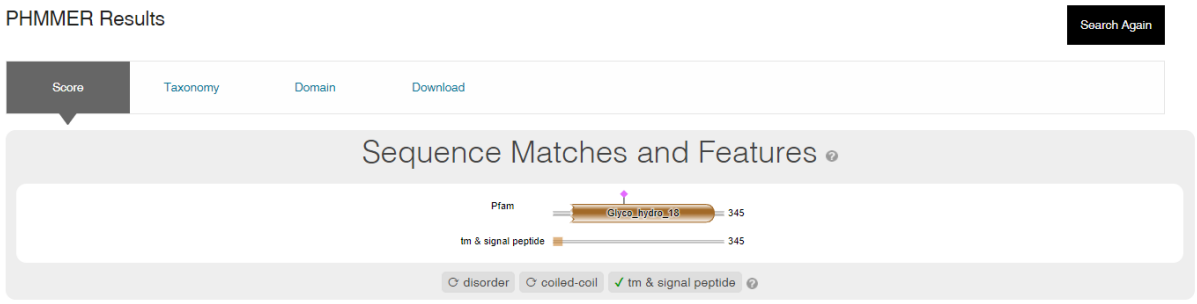
```
# WEBSEQUENCE Length: 345
# WEBSEQUENCE Number of predicted TMHs:  0
# WEBSEQUENCE Exp number of AAs in TMHs: 0.03031
# WEBSEQUENCE Exp number, first 60 AAs:  0.01694
# WEBSEQUENCE Total prob of N-in:        0.00896
WEBSEQUENCE     TMHMM2.0        outside      1   345
```



Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.0009	0.0001	0.9988	0.0002

Download: PNG / EPS / Tabular





-Primer

#5. BVU_1273FN

5'-CAT CAT CAC CAC CAT CAC AGA ACC TTT AAA TTA ATA GCA GGT-3' - 42 bp

Oligo No	210504-1-85	GC	%38	Tm(Basic)	64.51°C	Total nmol	41.35nmol
Skala	50 nmol	MW	12778.40	Conc	1887.29ng/µl	Total ng	528441.69ng
Saflastirma	DSLT	A260	60.9	OD	17.1	100 µM stok - µl TE	413.5

#6. BVU_1273RN

5'-GTG GCG GCC GCT CTA TTA AGG ATT CAT TAC ACC TTG ATA GAC-3' - 42 bp

Oligo No	210504-1-86	GC	%48	Tm(Basic)	68.41°C	Total nmol	41.37nmol
Skala	50 nmol	MW	12904.44	Conc	1906.43ng/µl	Total ng	533799.41ng
Saflastirma	DSLT	A260	59.4	OD	16.6	100 µM stok - µl TE	413.7

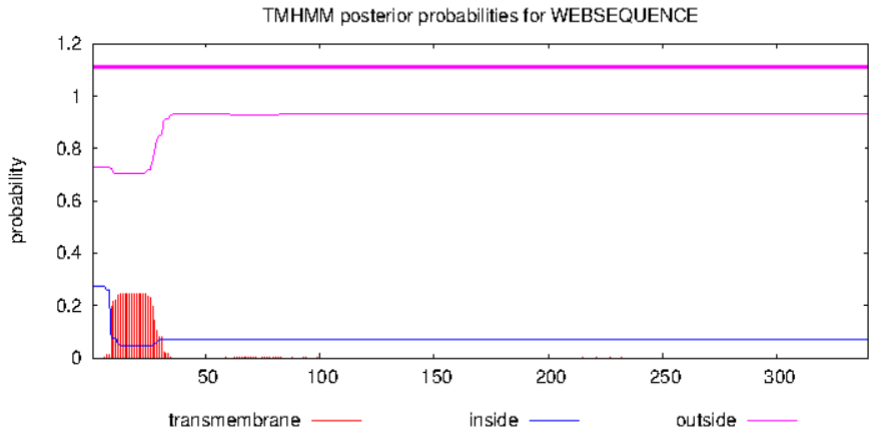
Figure 55. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of ABR38963.1.

30.AAO76145.1 – BT_1038

Bacteroides thetaiotaomicron **ATCC 29148**

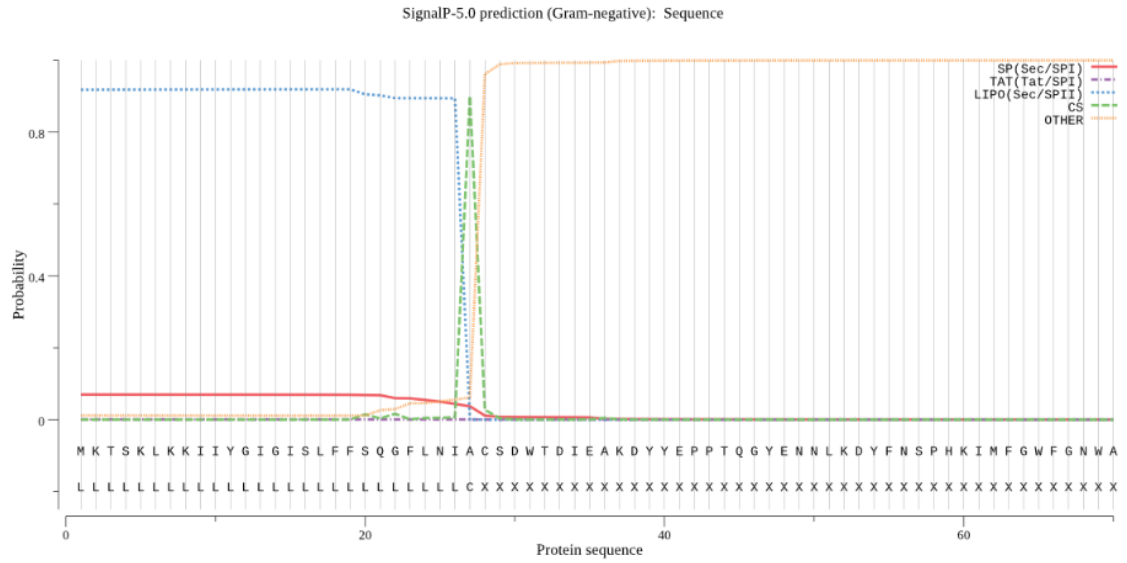
TMHMM result

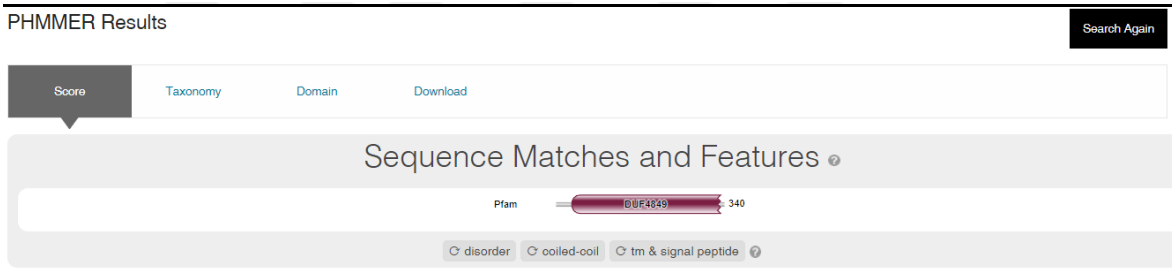
```
# WEBSEQUENCE Length: 340
# WEBSEQUENCE Number of predicted TMHs:  0
# WEBSEQUENCE Exp number of AAs in TMHs: 5.06742
# WEBSEQUENCE Exp number, first 60 AAs:  5.01057
# WEBSEQUENCE Total prob of N-in:        0.27293
WEBSEQUENCE     TMHMM2.0        outside     1   340
```



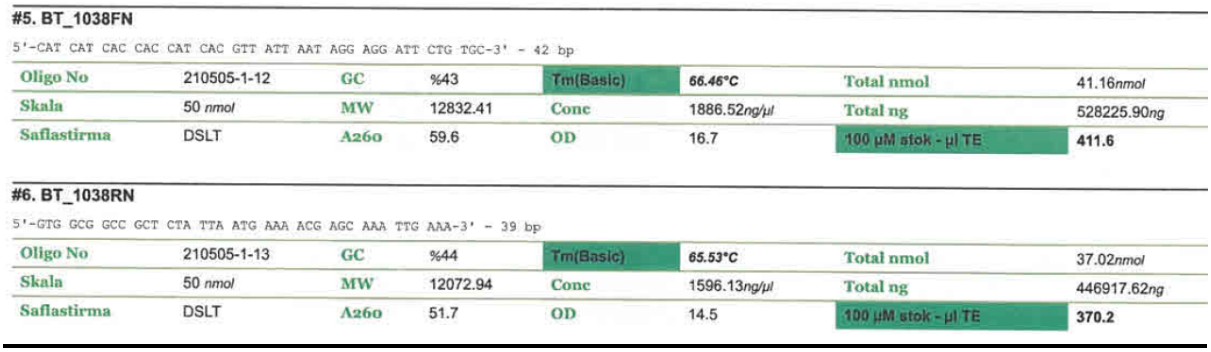
Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.0702	0.0008	0.9178	0.0112

Download: PNG / EPS / Tabular





-Primer



#5. BT_1038FN

5'-CAT CAT CAC CAC CAT CAC GTT ATT AAT AGG AGG ATT CTG TGC-3' - 42 bp

Oligo No	210505-1-12	GC	%43	Tm(Basic)	66.46°C	Total nmol	41.16nmol
Skala	50 nmol	MW	12832.41	Conc	1886.52ng/µl	Total ng	528225.90ng
Saflastirma	DSLT	A260	59.6	OD	16.7	100 µM stok - µl TE	411.6

#6. BT_1038RN

5'-GTG GCG GCC GCT CTA TTA ATG AAA ACG AGC AAA TTG AAA-3' - 39 bp

Oligo No	210505-1-13	GC	%44	Tm(Basic)	65.53°C	Total nmol	37.02nmol
Skala	50 nmol	MW	12072.94	Conc	1596.13ng/µl	Total ng	446917.62ng
Saflastirma	DSLT	A260	51.7	OD	14.5	100 µM stok - µl TE	370.2

Figure 56. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of AAO76145.1.

31. ACJ53522.1 -Blon_2468

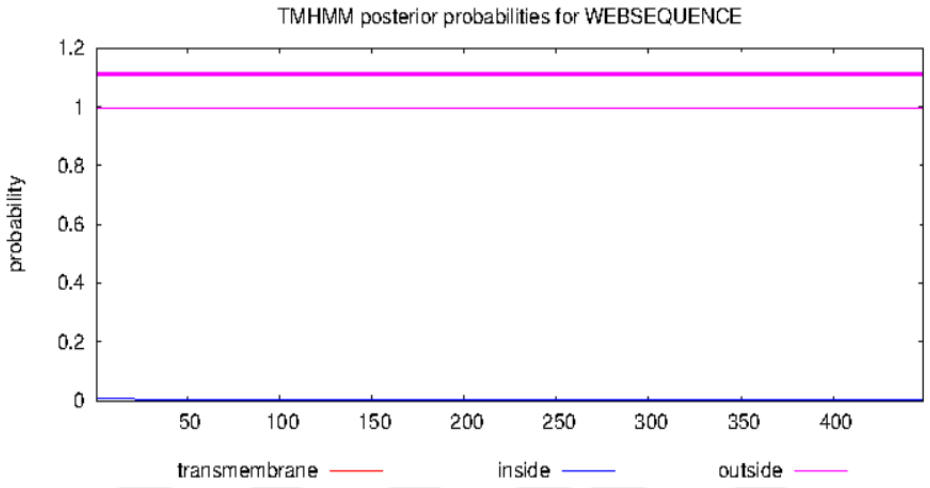
Bifidobacterium longum subsp. infantis ATCC 15697

32.ACJ51376.1 –Blon_0248

Bifidobacterium longum subsp. infantis ATCC 15697

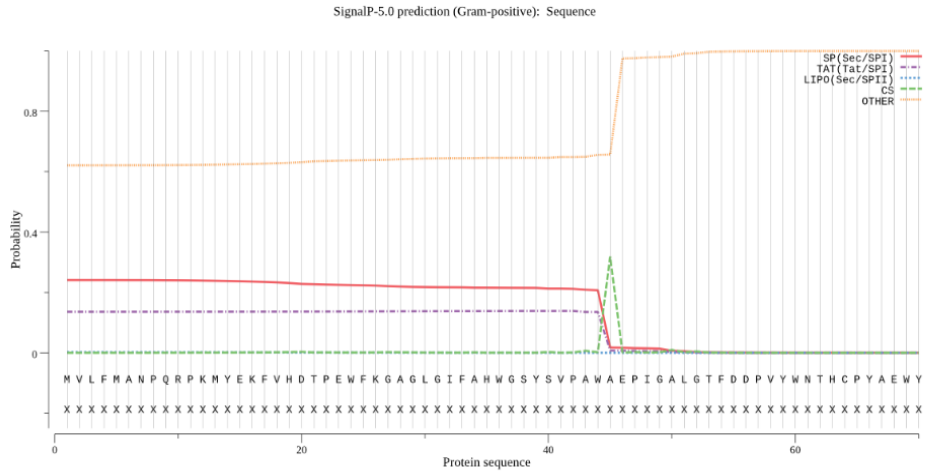
TMHMM result

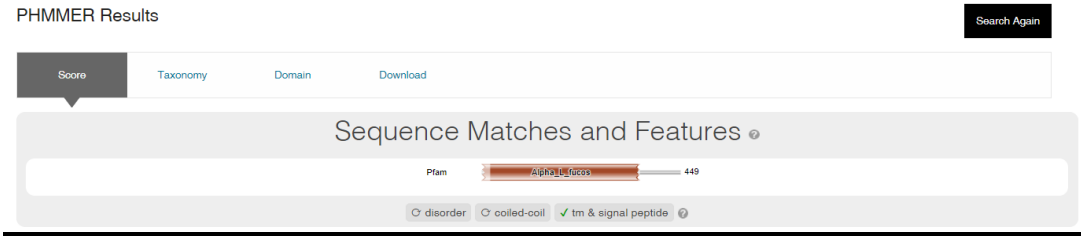
```
# WEBSEQUENCE Length: 449
# WEBSEQUENCE Number of predicted TMHs:  0
# WEBSEQUENCE Exp number of AAs in TMHs: 0.0402500000000001
# WEBSEQUENCE Exp number, first 60 AAs:  0.04002
# WEBSEQUENCE Total prob of N-in:        0.00529
WEBSEQUENCE     TMHMM2.0        outside      1   449
```



Protein type	Signal peptide (Sec/SPI)	TAT signal peptide (Tat/SPI)	Lipoprotein signal peptide (Sec/SPII)	Other
Likelihood	0.2433	0.1387	0.002	0.616

Download: PNG / EPS / Tabular





-Primer

#5. Blon_0248FC

5'-GAA GGA GAT ATA CAT ATG GTG TTG TTC ATG GCC AAT CCA-3' - 39 bp

Oligo No	210505-1-24	GC	%41	Tm(Basic)	64.48°C	Total nmol	39.34nmol
Skala	50 nmol	MW	12069.94	Conc	1695.83ng/µl	Total ng	474832.66ng
Saflastirma	DSLT	A260	55.3	OD	15.5	100 µM stok - µl TE	393.4

#6. Blon_0248RC

5'-GTG ATG GTG GTG ATG ATG GCG ACG GAC GAA GTG CAC GAC-3' - 39 bp

Oligo No	210505-1-25	GC	%59	Tm(Basic)	71.84°C	Total nmol	41.87nmol
Skala	50 nmol	MW	12217.98	Conc	1827.03ng/µl	Total ng	511567.11ng
Saflastirma	DSLT	A260	58.4	OD	16.3	100 µM stok - µl TE	418.7

Figure 57. Transmembrane/Signal Peptide/Domain Analysis Results and Primer Information of ACJ51376.1.

4.4.Transformation, Colony PCR, Induction and Purification of Each Target Enzyme

1. ATP38112.1 – CR531_08240

Lactobacillus salivarus subsp. Salivarius (Ligilactobacillus salivarius) ATCC 11741 / DSM 20555 / JCM 1231 / LMG9477

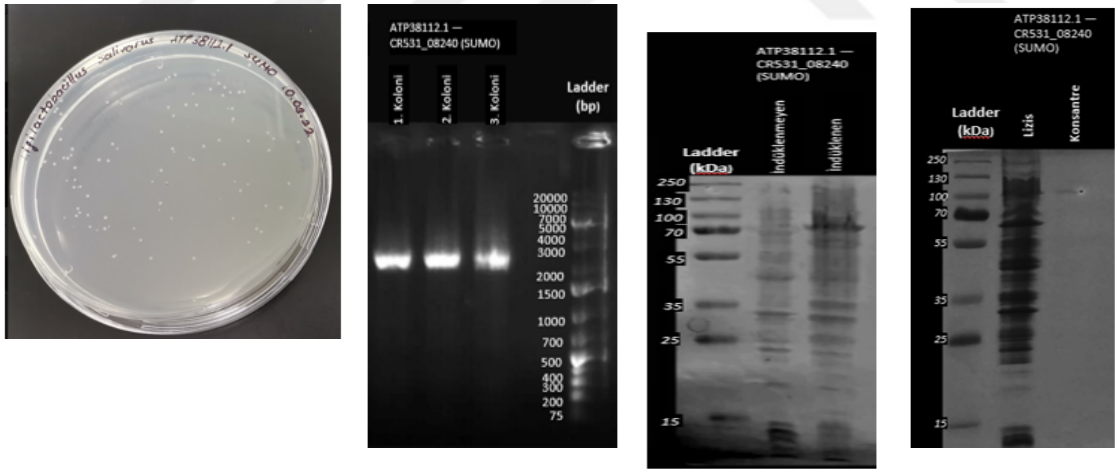


Figure 58. Transformation, colony PCR, induction, and purification results of ATP38112.1 respectively from left to right.

2. ATP36889.1 – CR531_01355

Lactobacillus salivarus subsp. Salivarius (Ligilactobacillus salivarius) ATCC 11741 / DSM 20555 / JCM 1231 / LMG9477

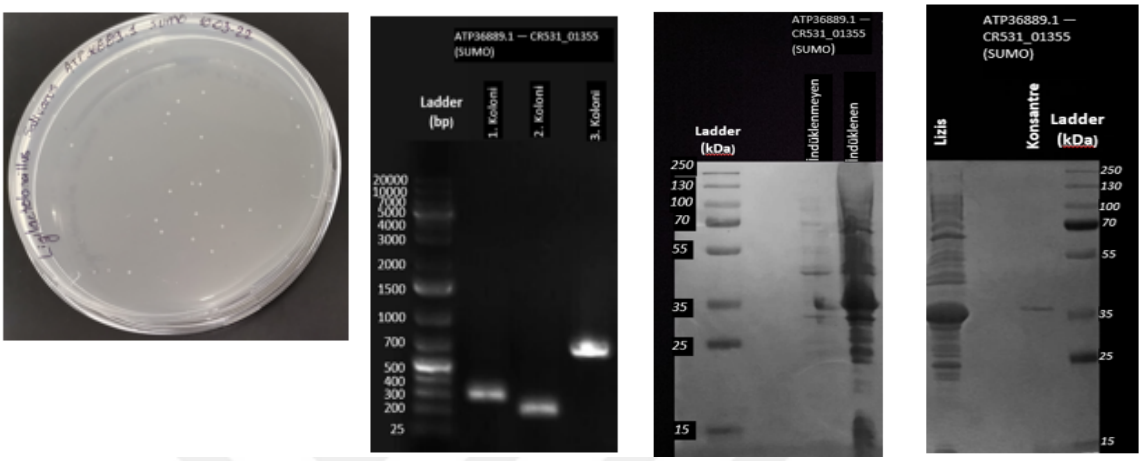


Figure 59. Transformation, colony PCR, induction and purification results of ATP36889.1 respectively from left to right.

3. ATP37244.1 – CR531_03290

Lactobacillus salivarus subsp. Salivarius (Ligilactobacillus salivarius) ATCC 11741 / DSM 20555 / JCM 1231 / LMG9477

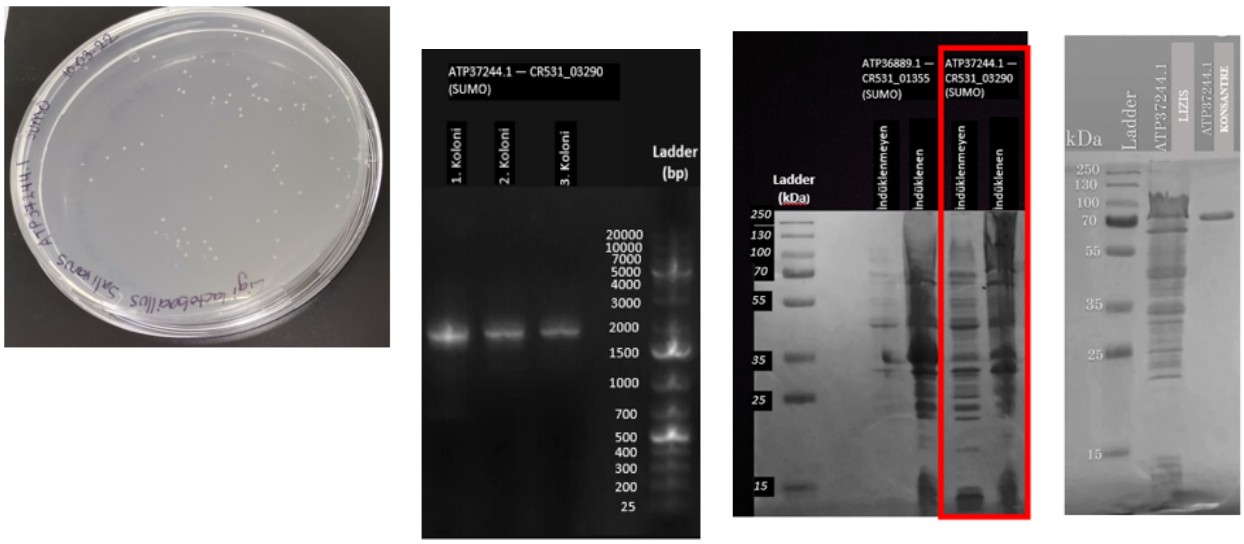


Figure 60. Transformation, colony PCR, induction and purification results of ATP37244.1 respectively from left to right.

4. SQH52440.1 – NCTC11324_01490

Streptococcus intermedius ATCC 27335/ DSM 20573/NCTC 11324

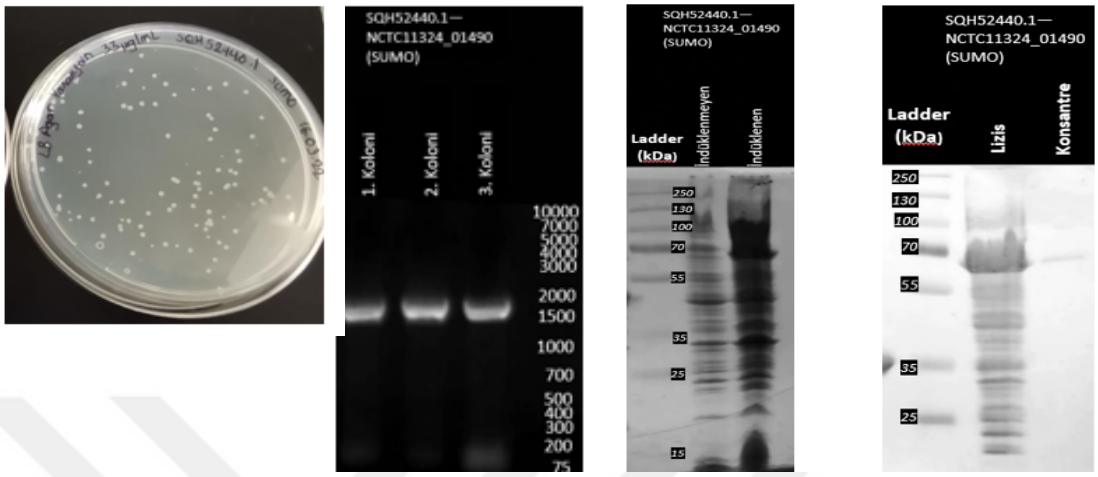


Figure 61. Transformation, colony PCR, induction and purification results of SQH52440.1 respectively from left to right.

5. ATP38122.1 – CR531_08290

Lactobacillus salivarus subsp. Salivarius (Ligilactobacillus salivarius) ATCC 11741 / DSM 20555 / JCM 1231 / LMG9477

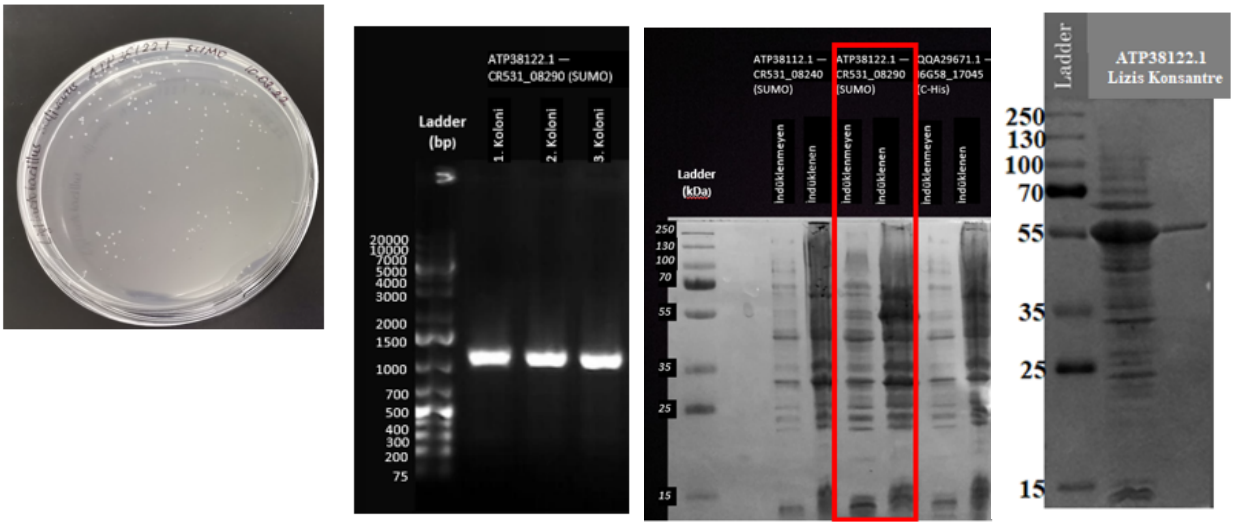


Figure 62. Transformation, colony PCR, induction and purification results of ATP38122.1 respectively from left to right.

6. ATP37586.1 – CR531_05275

Lactobacillus salivarus subsp. Salivarius (Ligilactobacillus salivarius) ATCC 11741 / DSM 20555 / JCM 1231 / LMG9477

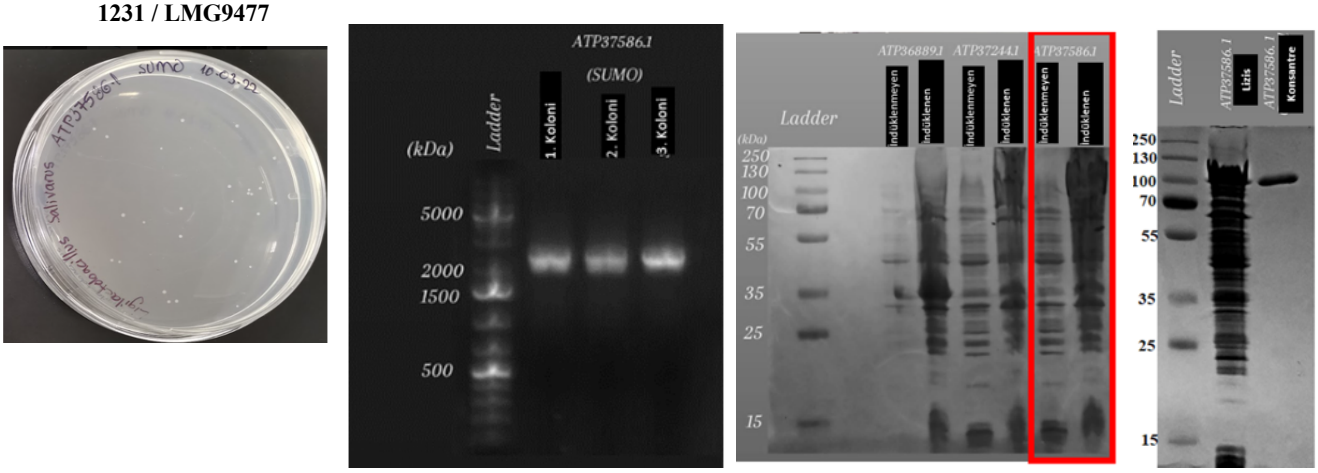


Figure 63. Transformation, colony PCR, induction and purification results of ATP37586.1 respectively from left to right.

7. SQH51076.1 – NCTC11324_00070

Streptococcus intermedius ATCC 27335/ DSM 20573/NCTC 11324

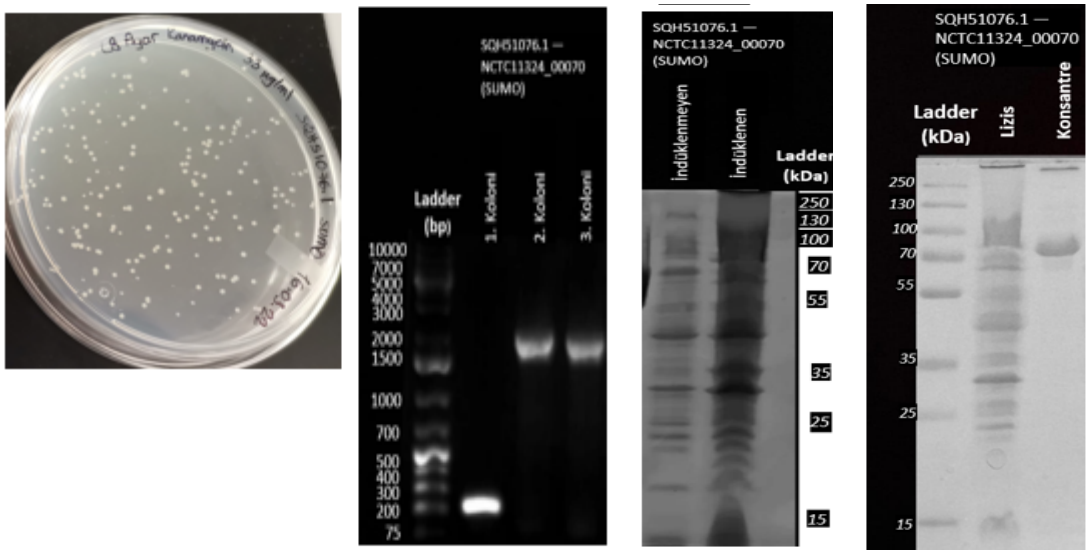


Figure 64. Transformation, colony PCR, induction and purification results of SQH51076.1 respectively from left to right.

8. SQH51655.1 – NCTC11324_00672

Streptococcus intermedius ATCC 27335/ DSM 20573/NCTC 11324

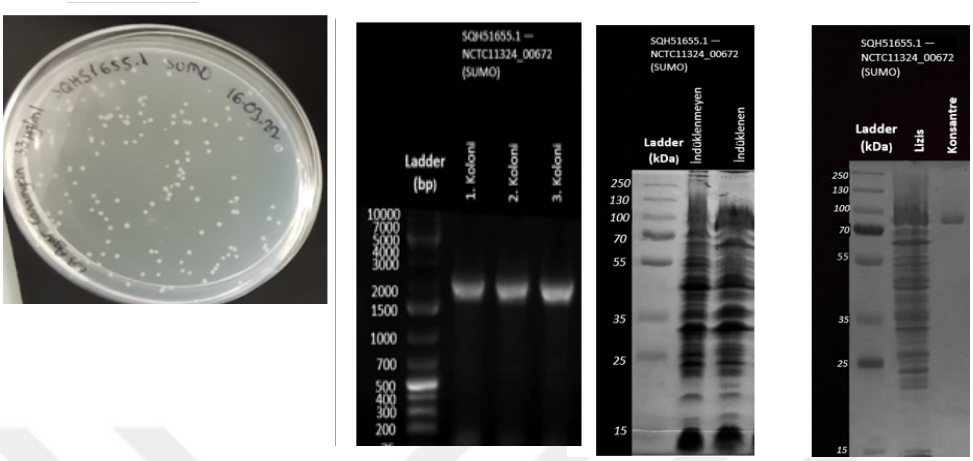


Figure 65. Transformation, colony PCR, induction and purification results of SQH51655.1 respectively from left to right.

9. CAR86329.1 — LGG 00434

Lactobacillus rhamnosus GG

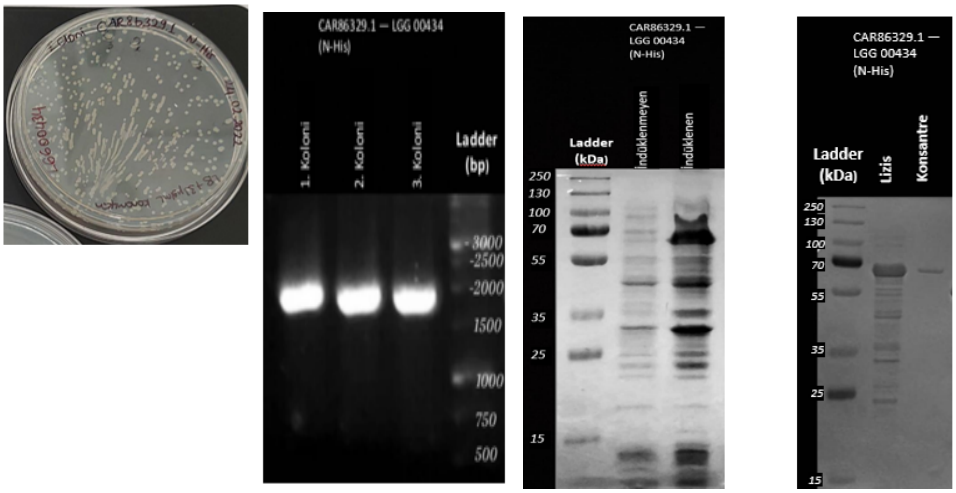


Figure 66. Transformation, colony PCR, induction and purification results of CAR86329.1 respectively from left to right.

10. AAO77566.1 – BT_2459

Bacteroides thetaiotaomicron ATCC 29148

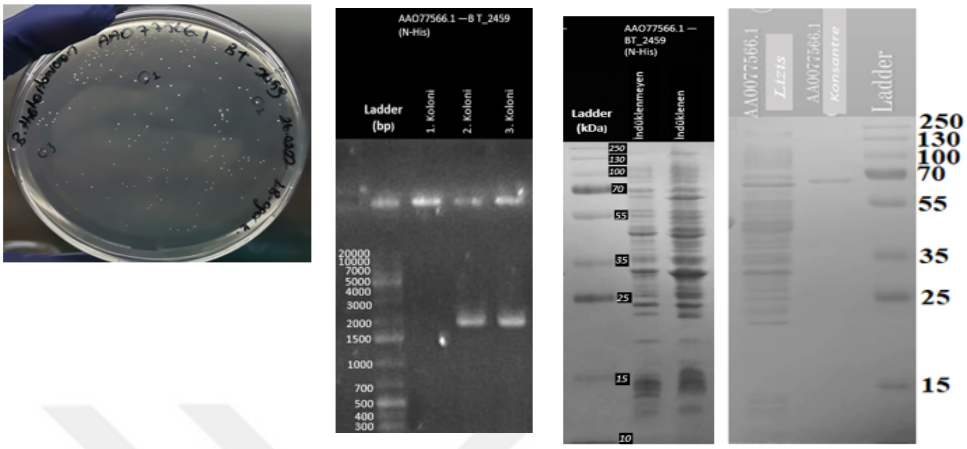


Figure 67. Transformation, colony PCR, induction and purification results of AAO77566.1 respectively from left to right.

11. ABR41745.1 -- BVU_4143

Bacteroides vulgatus 8482

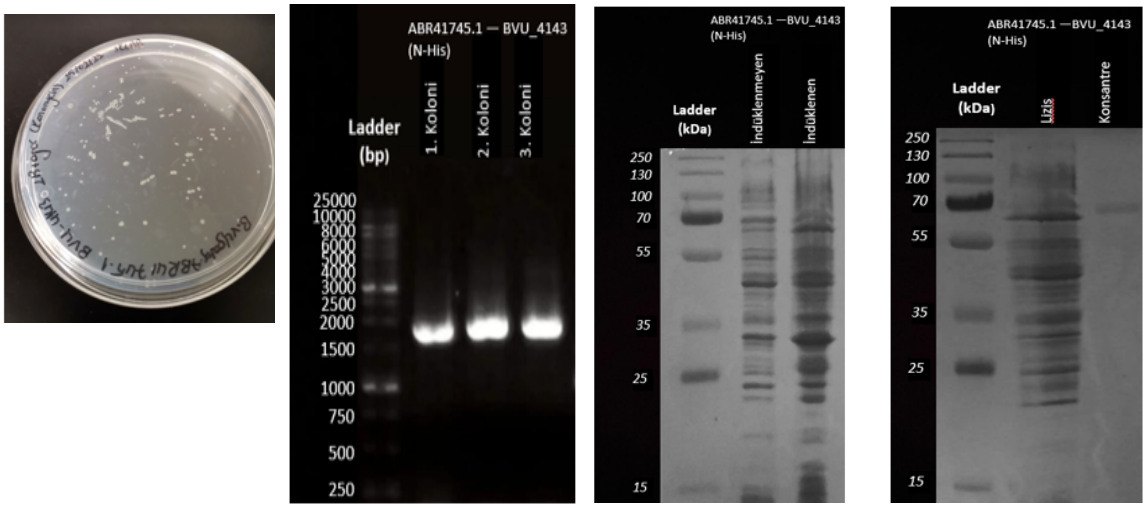


Figure 68. Transformation, colony PCR, induction and purification results of ABR41745.1 respectively from left to right.

12. AAO75562.1 - BT_0455

Bacteroides thetaiotaomicron ATCC 29148

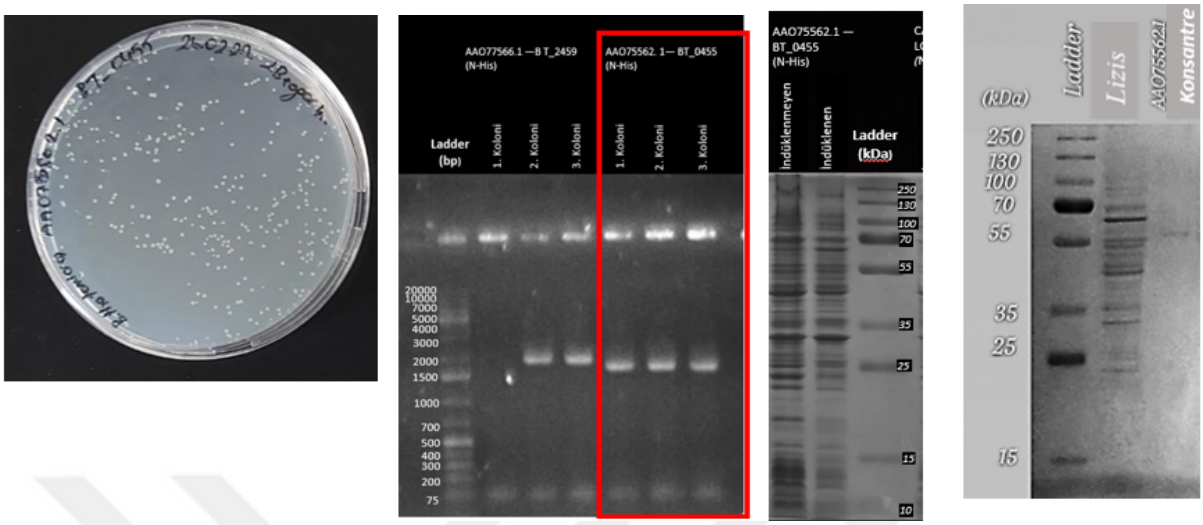


Figure 69. Transformation, colony PCR, induction and purification results of AA075562.1 respectively from left to right.

13. ACD04858.1 – Amuc_1032

Akkermansia muciniphila ATCC BAA-835 / DSM 22959 / LMG 27907

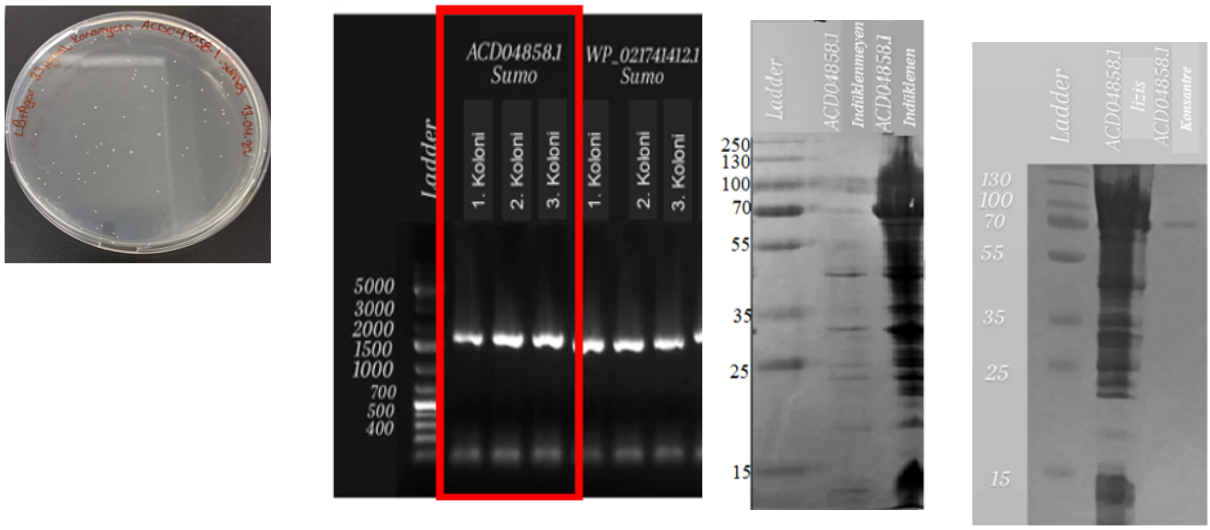


Figure 70. Transformation, colony PCR, induction and purification results of ACD04858.1 respectively from left to right.

14. BAQ98211.1 – BBBF_1004

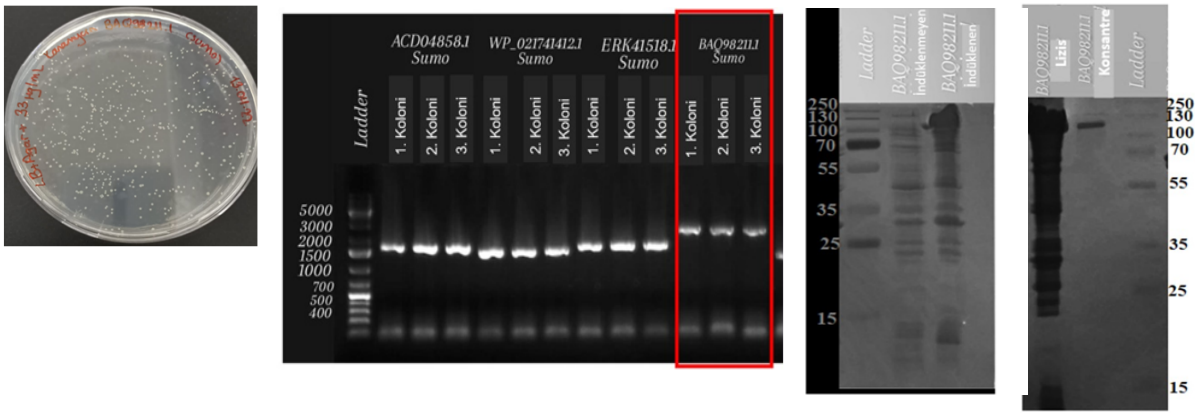


Figure 71. Transformation, colony PCR, induction and purification results of BAQ98211.1 respectively from left to right.

15. BAQ97897.1 – BBBF_0690

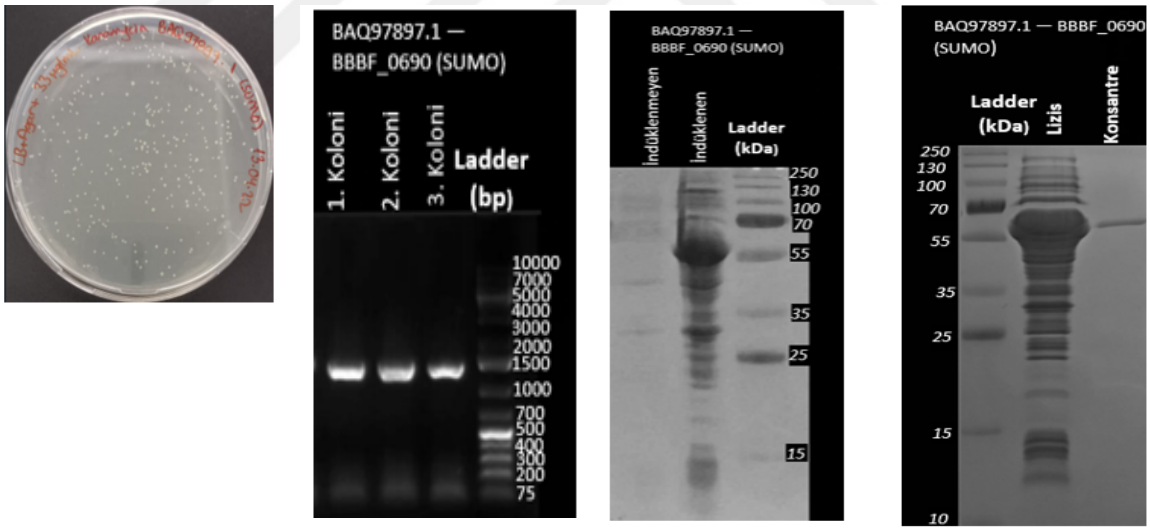


Figure 72. Transformation, colony PCR, induction and purification results of BAQ97897.1 respectively from left to right.

16. ERK41518.1 – HMPREF0495_02198

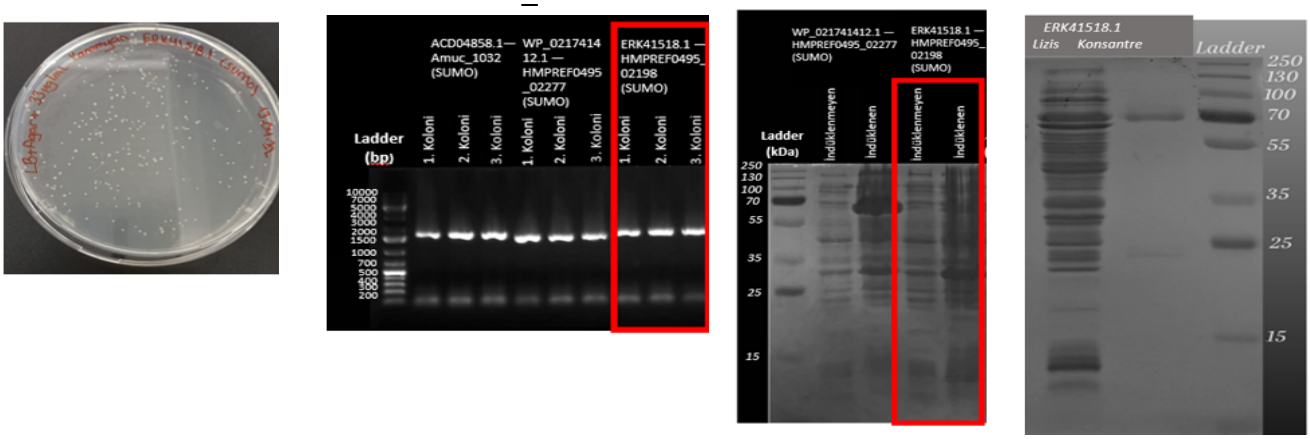


Figure 73. Transformation, colony PCR, induction and purification results of ERK41518.1 respectively from left to right.

17. ACJ51836.1 – Blon_0732

Bifidobacterium longum subsp. infantis ATCC 15697

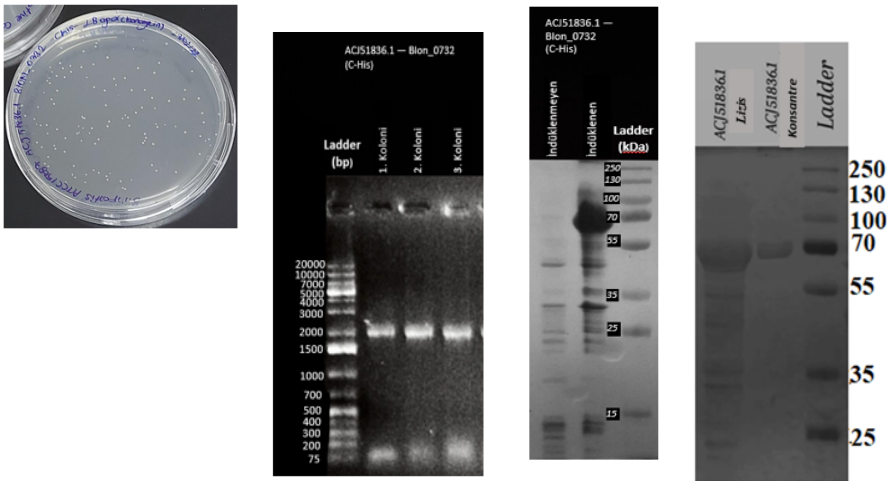


Figure 74. Transformation, colony PCR, induction and purification results of ACJ51836.1 respectively from left to right.

18. ACJ53413.1 – Blon_2355

Bifidobacterium longum subsp. infantis ATCC 15697

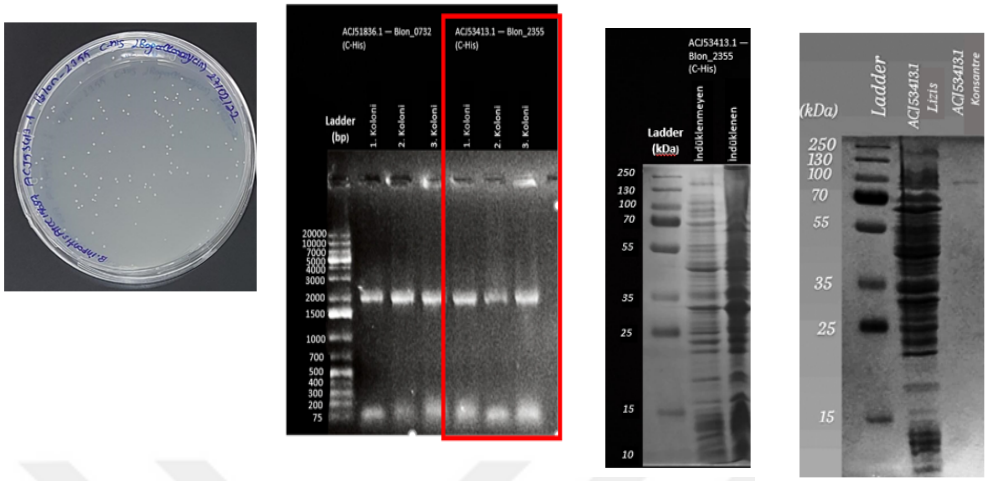


Figure 75. Transformation, colony PCR, induction and purification results of ACJ53413.1 respectively from left to right.

19. QQA29671.1 – I6G58_17045

Bacteroides uniformis FDAARGOS_901 ATCC8492

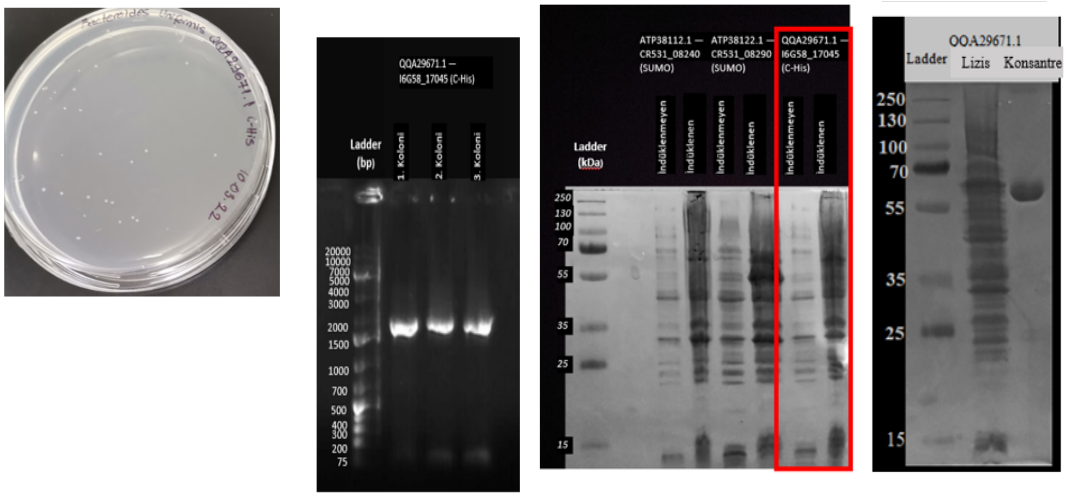


Figure 76. Transformation, colony PCR, induction and purification results of QQA29671.1 respectively from left to right.

20. BAQ30021.1 - BBKW_1886

Bifidobacterium catenulatum subsp. kashiwanohense JCM 15439

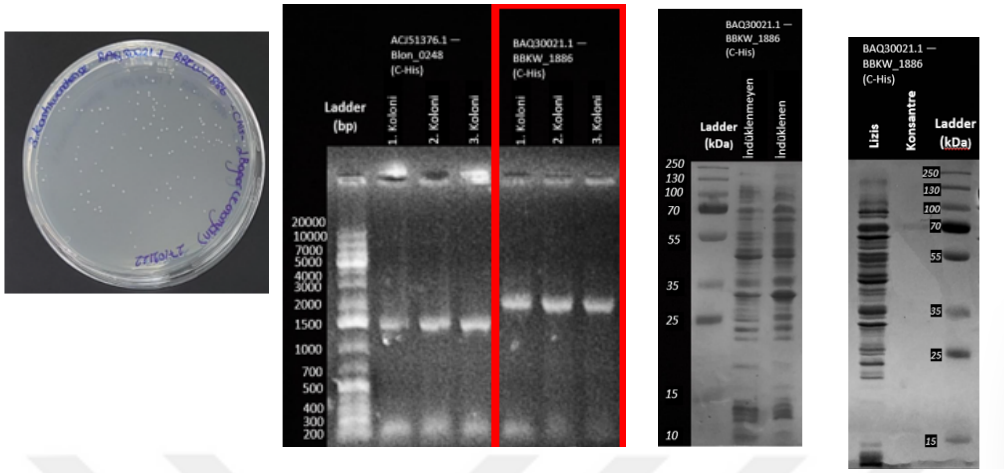


Figure 77. Transformation, colony PCR, induction and purification results of BAQ30021.1 respectively from left to right.

21. ABR38247.1 – BVU_0537

Bacteroides vulgatus 8482

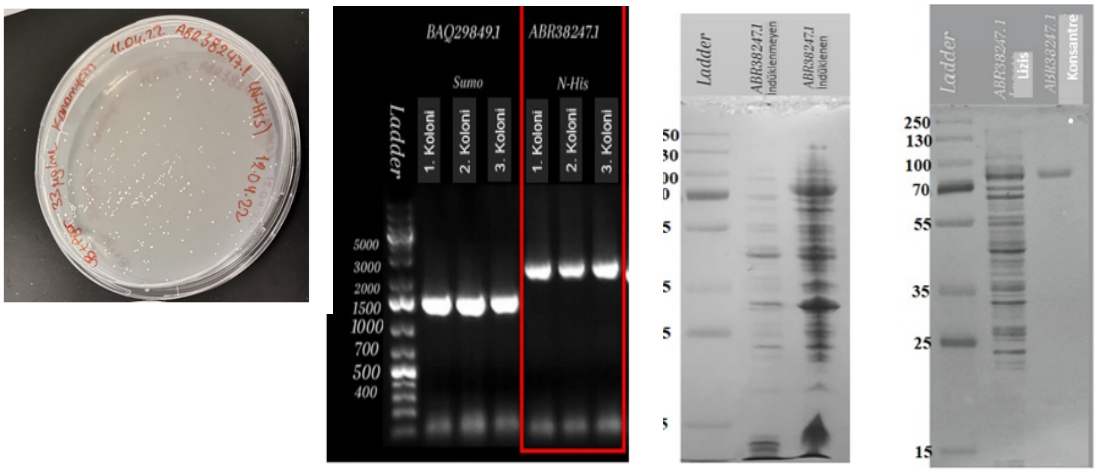


Figure 78. Transformation, colony PCR, induction and purification results of ABR38247.1 respectively from left to right.

22. SQF24907.1 – NCTC12958_01101

Streptococcus thermophilus ATCC 19258/ DSM 20617

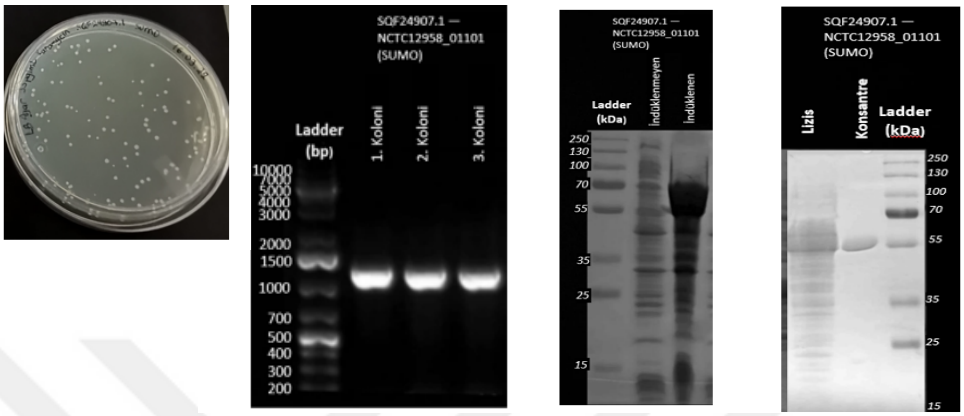


Figure 79. Transformation, colony PCR, induction and purification results of SQF24907.1 respectively from left to right.

23. SQF25661.1 – NCTC12958_01892

Streptococcus thermophilus ATCC 19258/ DSM 20617

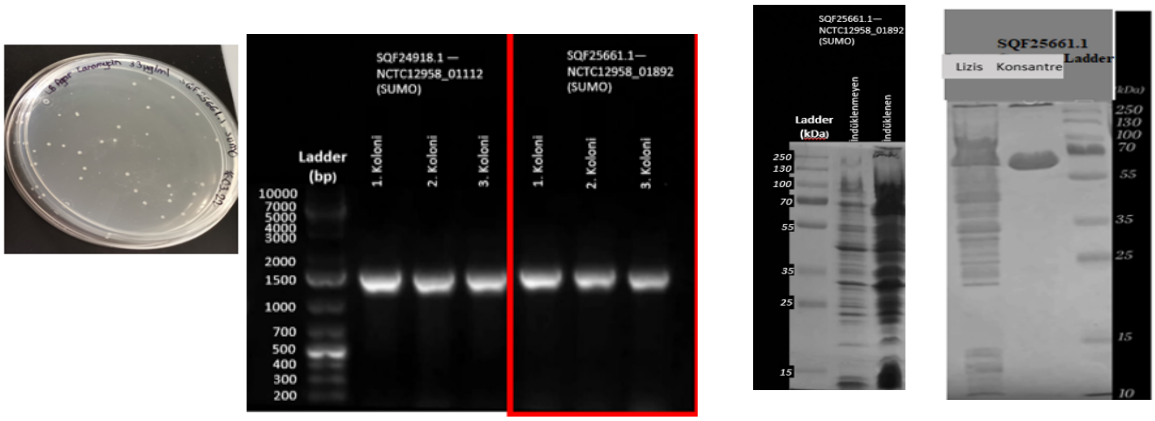


Figure 80. Transformation, colony PCR, induction and purification results of SQF25661.1 respectively from left to right.

24. SQF24918.1 – NCTC12958_01112

Streptococcus thermophilus ATCC 19258/ DSM 20617

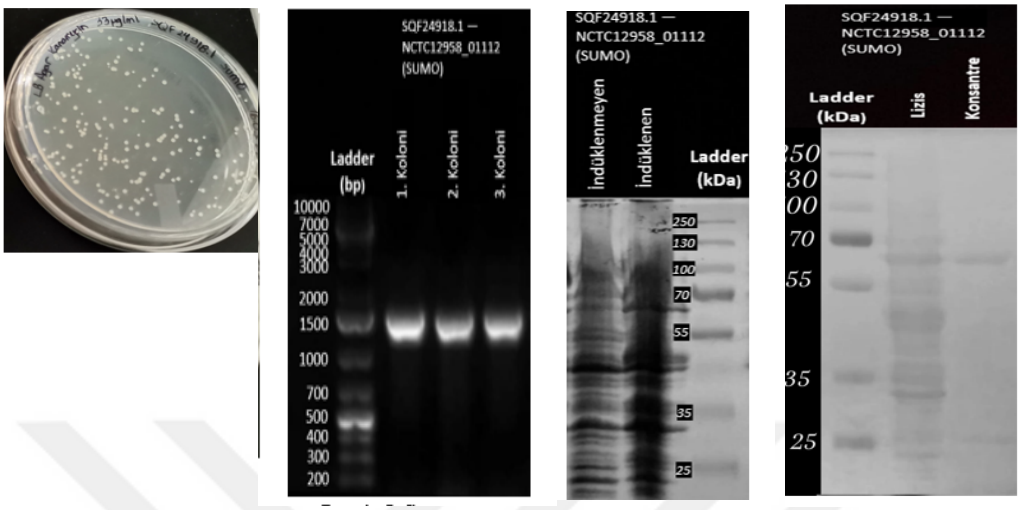


Figure 81. Transformation, colony PCR, induction and purification results of SQF24918.1 respectively from left to right.

25. ACD04701.1 – Amuc_0868

Akkermansia muciniphila ATCC BAA-835 / DSM 22959 / LMG 27907

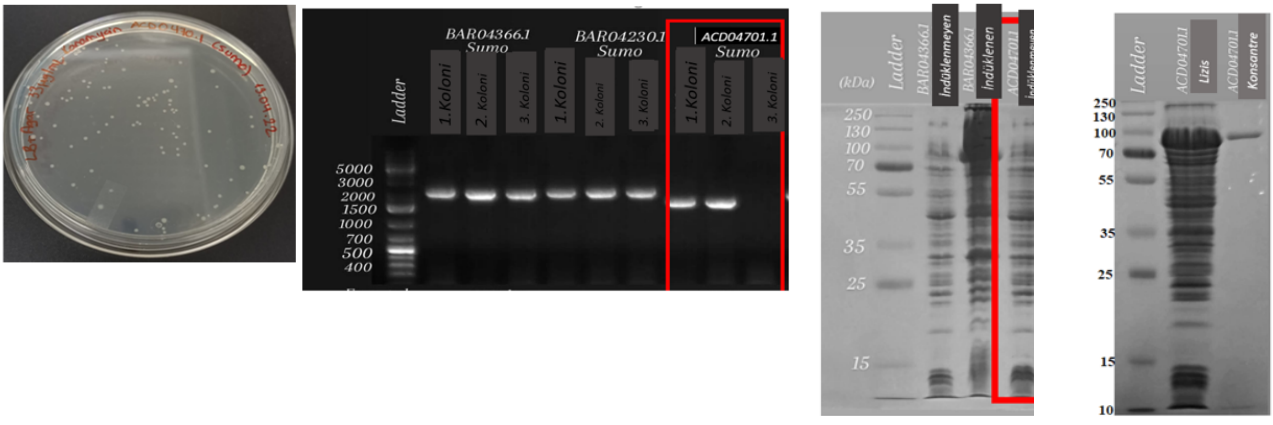


Figure 82. Transformation, colony PCR, induction and purification results of ACD04701.1 respectively from left to right.

26. ACD04208.1 –Amuc_0369

Akkermansia muciniphila ATCC BAA-835 / DSM 22959 / LMG 27907

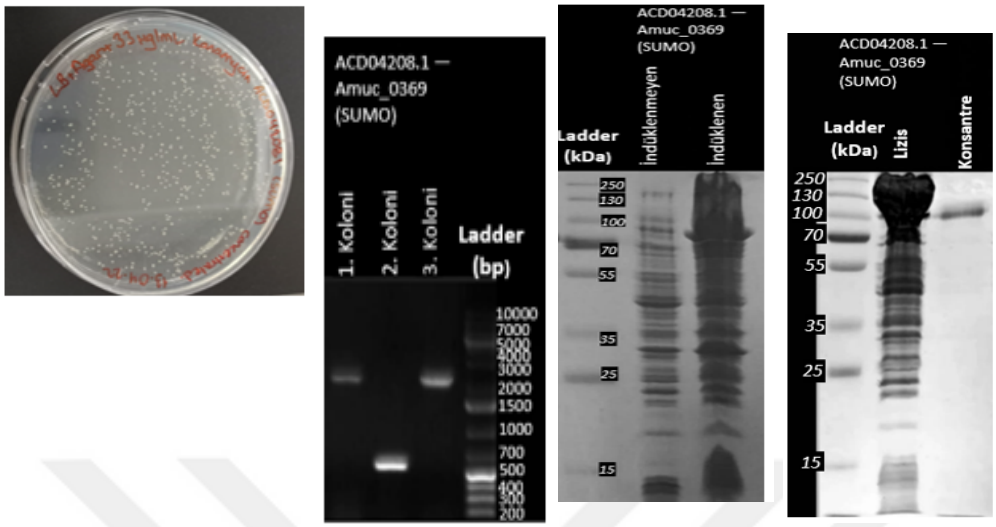


Figure 83. Transformation, colony PCR, induction and purification results of ACD04208.1 respectively from left to right.

27. BAQ97280.1 – BBBF_0073

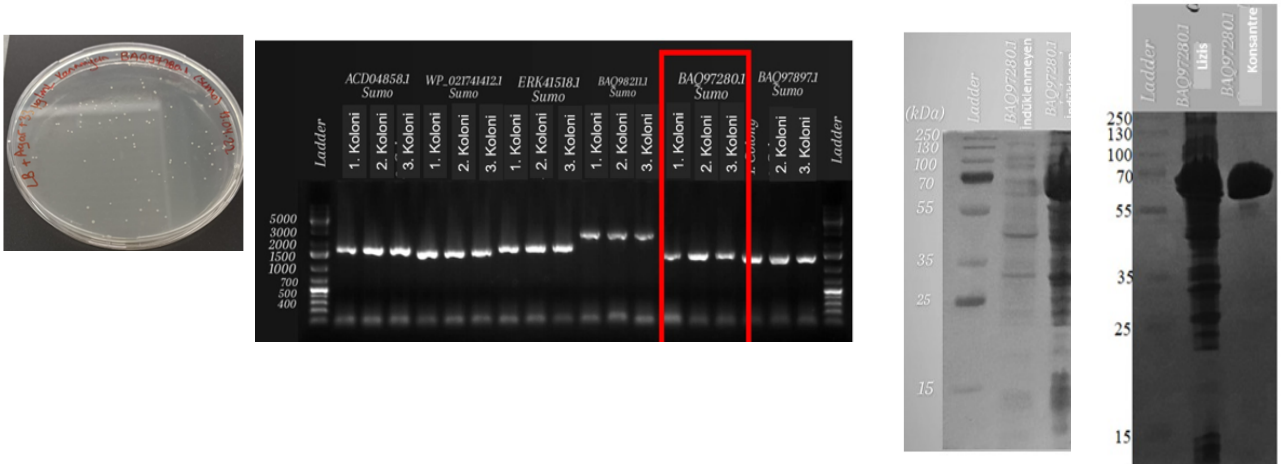


Figure 84. Transformation, colony PCR, induction and purification results of BAQ97280.1 respectively from left to right.

28. CAH09389.1 – BF9343_3608

Bacteroides fragilis ATCC 25285

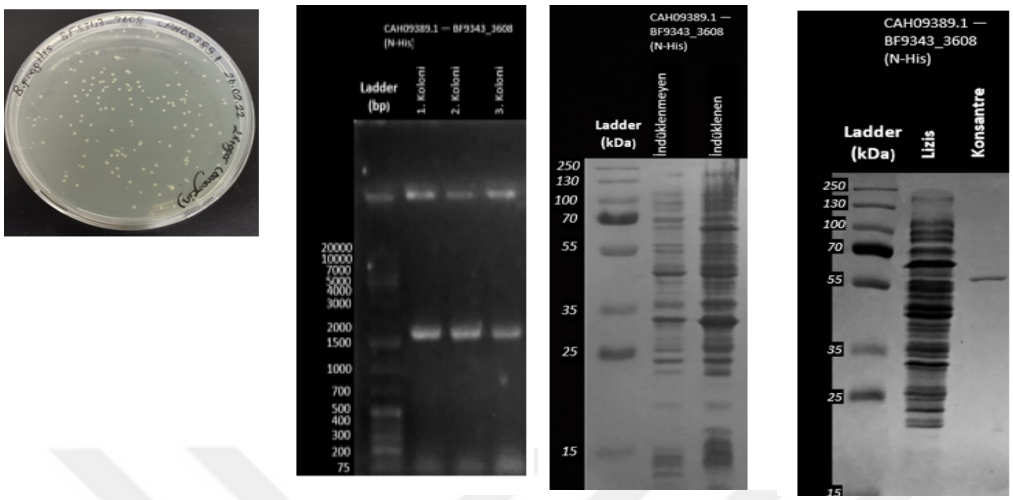


Figure 85. Transformation, colony PCR, induction and purification results of CAH09389.1 respectively from left to right.

29. ABR38963.1 - BVU_1273

Bacteroides vulgatus 8482

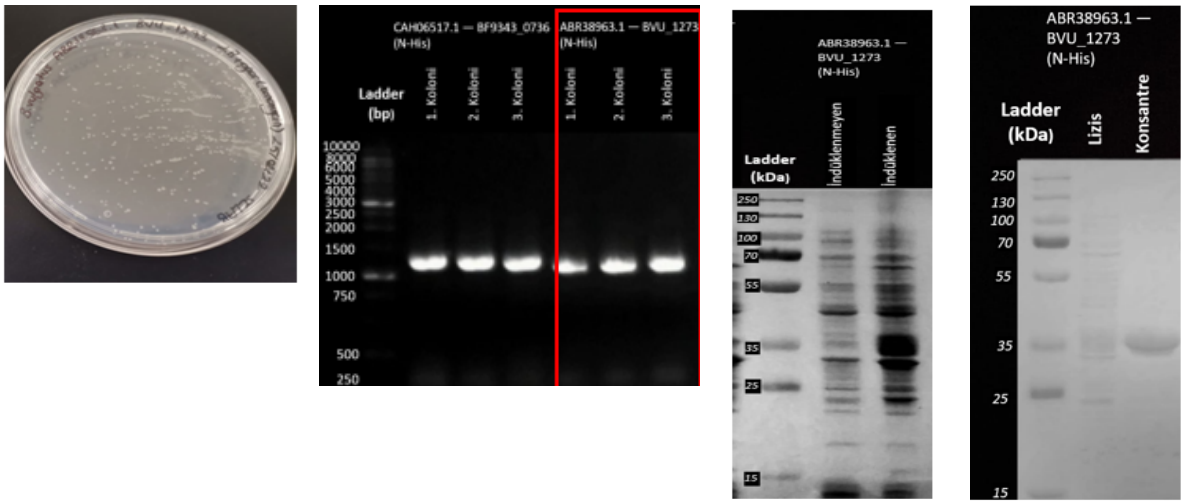


Figure 86. Transformation, colony PCR, induction and purification results of ABR38963.1 respectively from left to right.

30.AAO76145.1 – BT_1038

Bacteroides thetaiotaomicron ATCC 29148

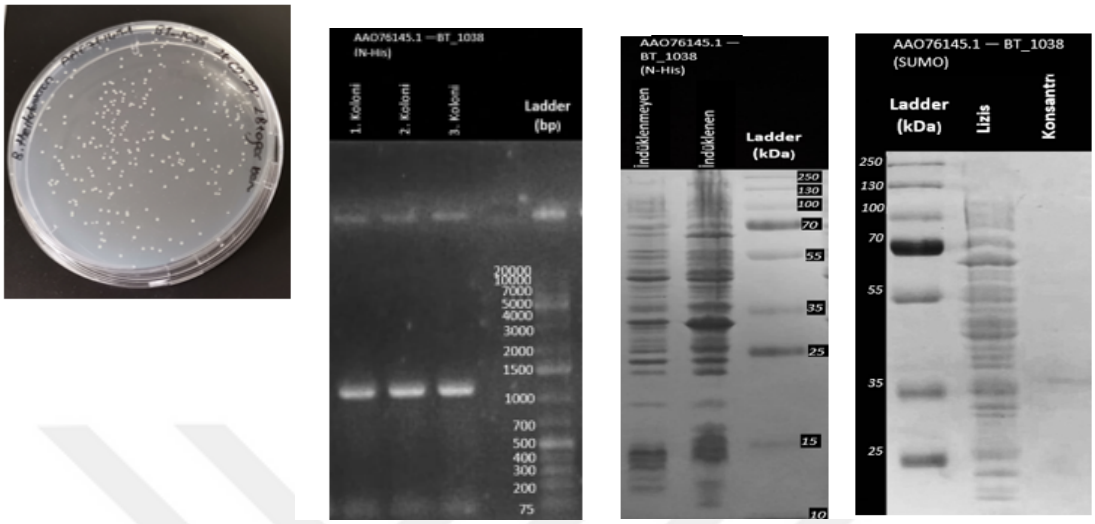


Figure 87. Transformation, colony PCR, induction and purification results of AAO76145.1 respectively from left to right.

31.ACJ51376.1 –Blon_0248

Bifidobacterium longum subsp. infantis ATCC 15697

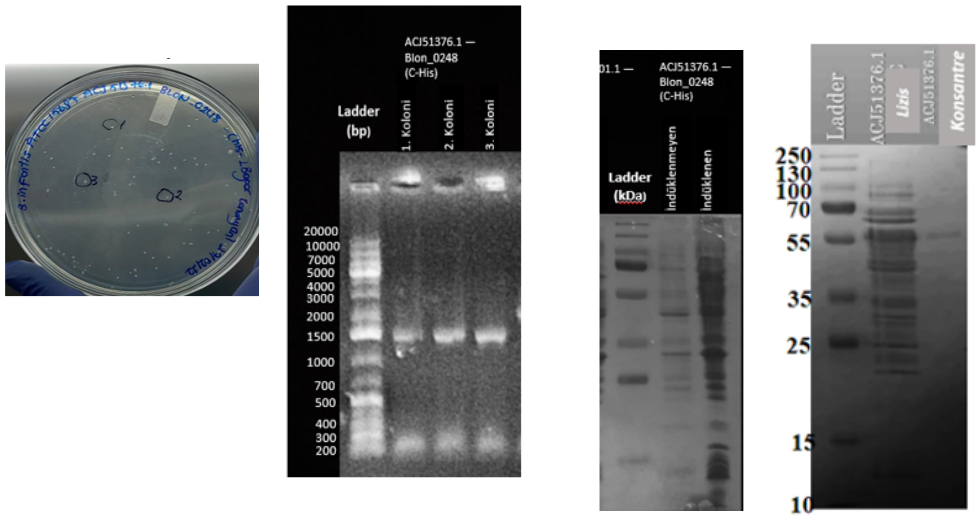


Figure 88. Transformation, colony PCR, induction and purification results of ACJ51376.1 respectively from left to right.

4.5. Measurement of Produced and Purified Enzymes' Concentration

Table 4. The concentration of produced and purified enzymes

	GeneBank ID / Accession Number **Locus tag**	**Concentration (mg/mL)**
1	ATP38112.1 CR531_08240	0.200
2	ATP36889.1 CR531_01355	0.534
3	ATP37244.1 CR531_03290	1.058
4	SQH52440.1 NCTC11324_01490	0.506
5	ATP38122.1 CR531_08290	0.406
6	ATP37586.1 CR531_05275	0.254
7	SQH51076.1 NCTC11324_00070	0.800
8	SQH51655.1 NCTC11324_00672	0.240
9	CAR86329.1 LGG_00434	0.233
10	AAO77566.1 BT_2459	0.088
11	ABR41745.1 BVU_4143	0.062
12	AAO75562.1 BT_0455	0.284
13	ACD04858.1 Amuc_1032	1.336
14	BAQ98211.1 BBBF_1004	3.580
15	BAQ97897.1	0.220

	BBBF_0690	
16	ERK41518.1 HMPREF0495_02198	0.330
17	ACJ51836.1 Blon_0732	0.800
18	ACJ53413.1 Blon_2355	0.220
19	QQA29671.1 I6G58_17045	0.364
20	BAQ30021.1 BBKW_1886	0.268
21	ABR38247.1 BVU_0537	0.213
22	SQF24907.1 NCTC12958_01101	4.260
23	SQF25661.1 NCTC12958_01892	0.776
24	SQF24918.1 NCTC12958_01112	0.860
25	ACD04701.1 Amuc_0868	1.668
26	ACD04208.1 Amuc_0369	0.208
27	BAQ97280.1 BBBF_0073	2.100
28	CAH09389.1 BF9343_3608	0.106
29	ABR38963.1 BVU_1273	3.160
30	AAO76145.1 BT_1038	0.200
31	ACJ53522.1 Blon_2468	0.650
32	ACJ51376.1 Blon_0248	0.153

4.6. Design of the novel system and test on a glycoprotein source

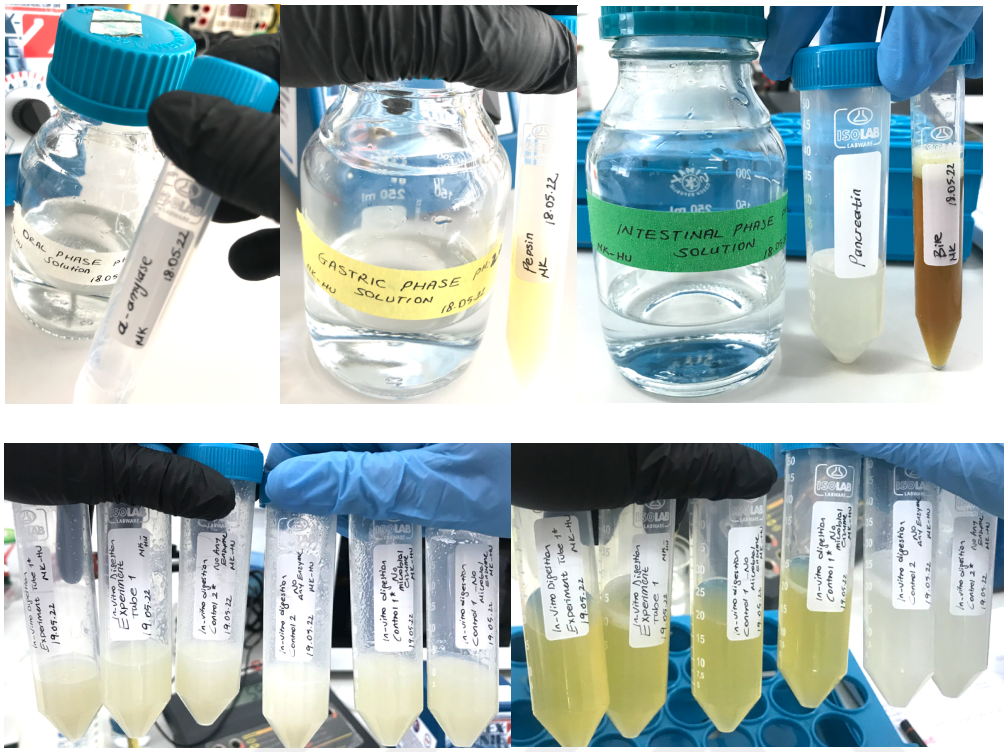


Figure 89. *In-vitro* digestion experiment and samples after digestion in each phase.

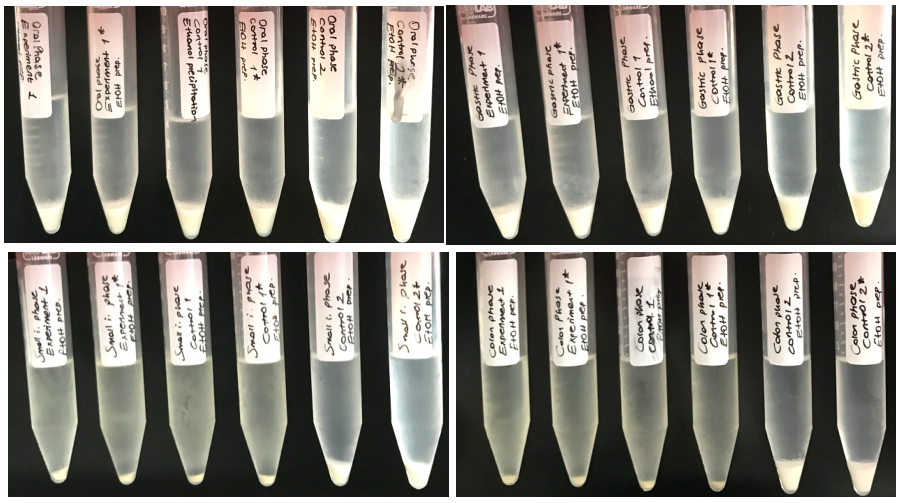
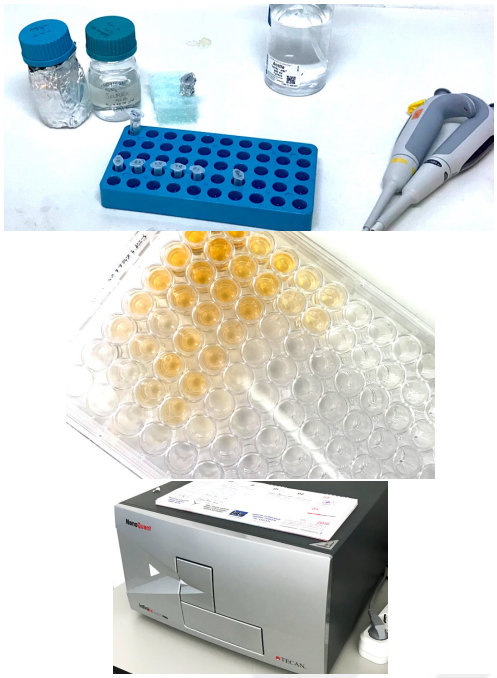
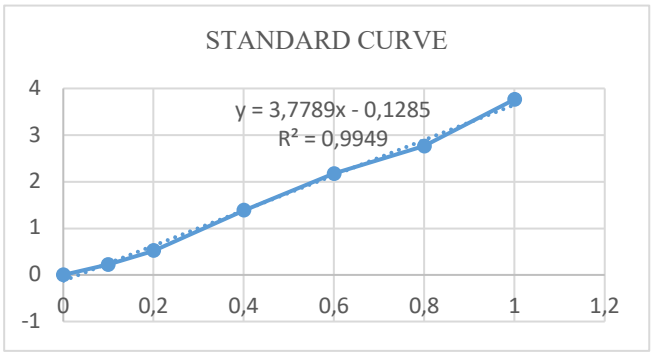


Figure 90. Digested samples after protein precipitation.

4.7. Measurement of Concentration of Released Glycans from Novel In-Vitro Digestion





OD_{490nm} Measurements (Mean Values)

	ORAL	GASTRIC	SMALL I	COLON
EXPERIMENT	2	3,91	3,4	3,8
CONTROL 1	1,51	3,7873	1,9539	2,1
CONTROL 2	1,1278	1,4	0,85	0,81

Figure 91. Measurement of carbohydrate concentration with phenol sulphuric acid assay.

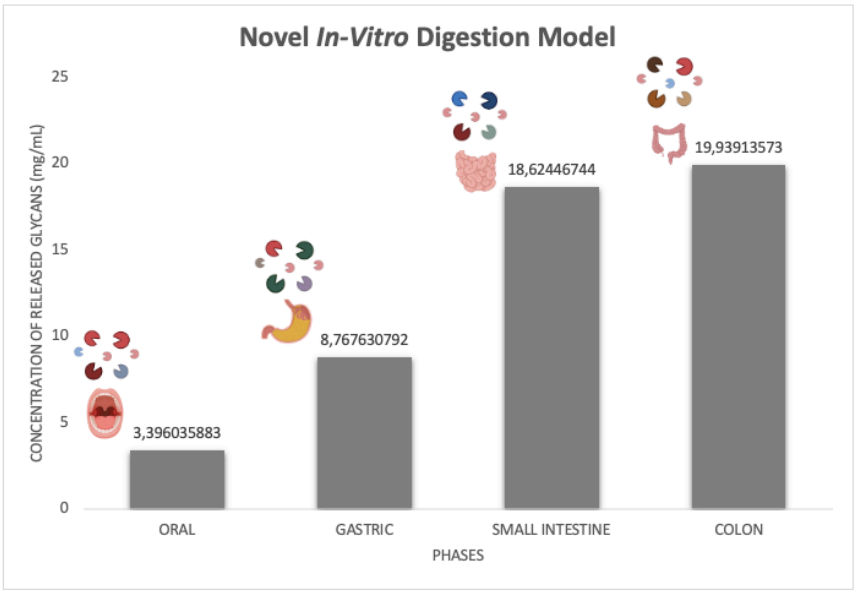


Figure 92. The concentration of released glycans from the novel *in-vitro* digestion model – A statistically significant difference between groups (except between the small intestine and colon groups) was determined by ANOVA, Tukey multiple comparison test ($p<0.05$).

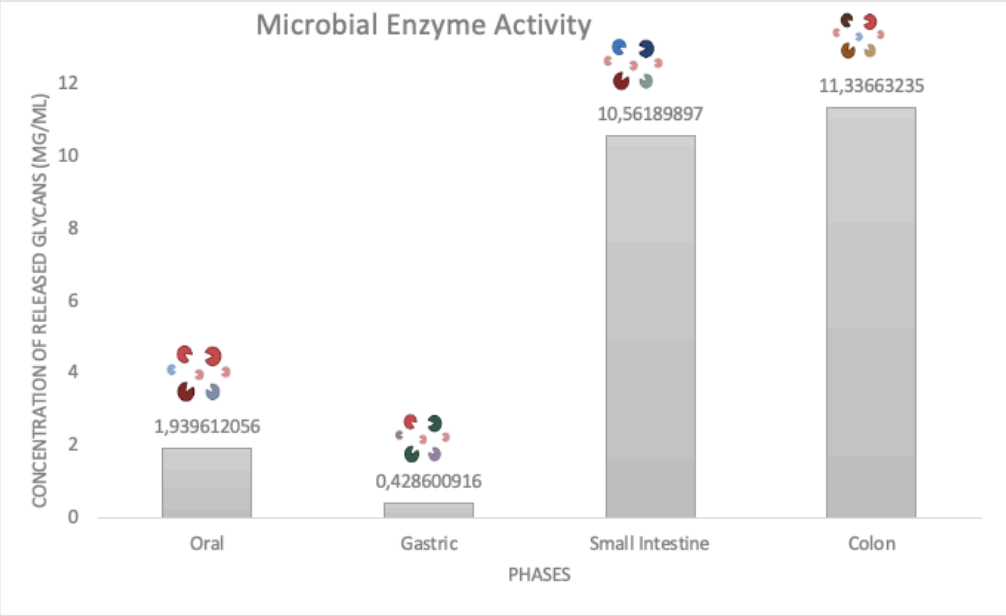


Figure 93. The concentration of released glycans by only microbial enzymes in each phase – A statistically significant difference between groups of oral – gastric and small intestine – colon was determined by ANOVA, Tukey multiple comparison test ($p<0.05$). No statistically significant difference was determined between oral and gastric groups as well as between small intestine and colon groups ($p>0.05$).

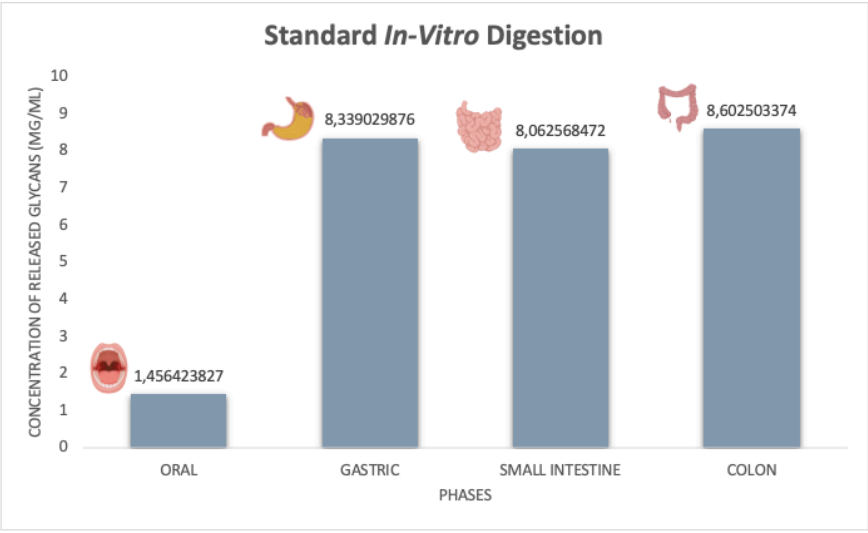


Figure 94. The concentration of released glycans from the standard *in-vitro d*igestion model - A statistically significant difference between the oral and other groups is determined by ANOVA, Tukey multiple comparison test ($p<0.05$). No statistically significant difference was determined between the groups of gastric, small intestine, and colon ($p>0.05$)

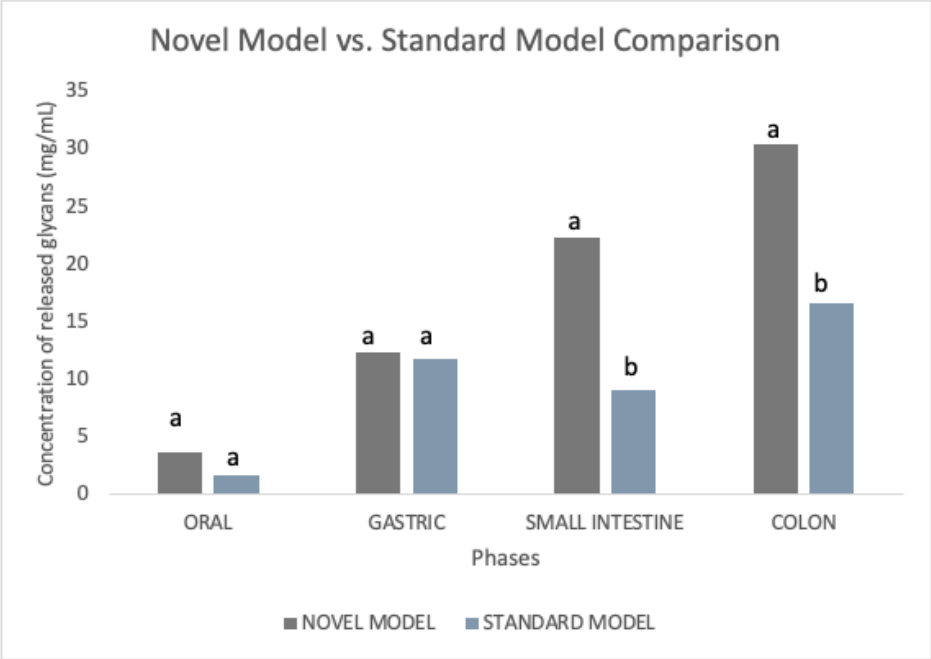


Figure 95. The comparison of released glycans from the novel and standard model - a,b: The differences between the data from the novel and standard model are significant for each phase separately ($p<0.05$), except oral and gastric phase ($p>0.05$).

CHAPTER 5

RESULTS AND RECOMMENDATIONS

The thesis mainly includes three major sections regarding experiments performed. Whilst the first one covers the bioinformatic analysis for the determination of genes, design of primers, and molecular cloning; the second section is the protein production and purification steps, and the last section includes the design of the *in-vitro* digestion model by integrating recombinant enzymes produced in the first section.

In the first section of the thesis, target genes from microorganisms abundant in different phases of the gastrointestinal tract were selected and analyzed using bioinformatics tools. Target microorganisms vary among digestion phases, for instance, mainly Lactobacillus strains are found in the oral phase, whereas the gastric phase includes Lactobacillus and Bacteroides strains. As a variety of microorganisms inhabit in the human intestinal microbiome, diverse microorganisms were selected for small intestine and colon parts including Akkermansia, Bifidobacteria, Bacteroides, as well as Streptococcus. In addition, to design an in-vitro model with the integration of microbial enzymes phylogenetic trees based on neighbor-joining and maximum likelihood were constructed for each phase of the digestion model by appropriate model and parameters. Another notorious point is that signal peptide and transmembrane domains were searched for each gene to be cloned before the molecular cloning step, which is important to increase the efficiency of protein production. In signal peptide analysis related to Hidden Markov algorithms, 0.4 and higher results were considered signal peptides and excluded from the amino acid residues recommended by the database. As for transmembrane domain analysis, possible regions (1-1.2) were excluded from the amino acid sequence. HMMER (biosequence analysis using profile hidden Markov models) was used to determine and compare domain analysis. Regarding the results based on signal peptide and transmembrane domain analysis, identified regions were excluded from the sequence, and primers were designed based on the new gene sequence. This process was crucial to increase the yield of protein production in the following steps since transmembrane domain and signal peptides can cause the binding of recombinant protein to cells and therefore hardens the protein purification step.

In the second section of this thesis, recombinantly cloned genes were expressed using L-rhamnose which is an inducer used in protein induction step. High temperature and high inducer concentration during protein expression under strong promoter systems often result in high expression, which weakens the bacterial protein quality control. Therefore, the partially folded or misfolded protein molecules aggregate as well as form inclusion bodies. The formation of inclusion bodies poses a noticeable challenge for the large-scale recovery of proteins. To prevent this problem, the expression of recombinant genes was significantly increased by optimization studies performed before the molecular cloning step. The lower temperature, lower optical density, as well as lower inducer (L-rhamnose) concentration affect positively the protein expression than the standard expression conditions recommended by the cloning kit. By using these approaches, 32 different glycosyl hydrolase enzymes (one of them was produced in previous studies; EndoBI-1) from distinct microorganisms of phases of the digestion system were produced with 95%, which enables the mimicking digestion system *in-vitro* conditions by integrating those enzymes through the model.

In the third section of the thesis, to integrate recombinant enzymes through a conventional in-vitro digestion model, standard digestion solutions were prepared including simulated salivary fluid, simulated gastric fluid, as well as simulated intestinal fluid. All these fluids included only human-associated enzymes including amylase, pepsin, or trypsin. The point that needs to be taken into consideration is that $CaCI_2$ was lastly added to the solutions because it causes precipitation at the beginning. All appropriate chemicals, and enzymes (human-associated) were prepared for the *in-vitro* model. The temperature was adjusted to 37°C for all phases, in contrast, pH was adjusted to 7.5 for the oral and intestinal phase, 3 for the gastric phase, and 8 for the colon part. The new model designed in this thesis differs from conventional examples of *in-vitro* digestion models since it includes microbial enzymes in each phase. Moreover, conventional *in-vitro* digestion models mimic only three phases: oral, gastric, as well as intestinal for the small intestine. However, the colon part was also designed in this novel model by using only recombinant enzymes as there is no human digestion process in the colon part. The most significant point of the results of this thesis is the designing of a model with the contribution of microbial enzymes, besides, the mimicking of the colon part, which is essential to thoroughly mimicking human digestion. All reactions took place in one 50 mL falcon tube by the addition of the following enzymes and solutions based on the phase. The enzyme activities were stopped by adjusting pH while passing

through the next phases of digestion. The same experiment was also conducted without microbial enzymes as the control of the experiment. At the end of reactions of the *in-vitro* model on the model glycoprotein, released glycans at each phase were quantified by phenol sulphuric acid assay and compared with a control group to understand the contribution of microbiome-associated enzymes to human digestion. According to the findings from the novel model data, the concentration of released glycans significantly increase from the oral phase through colon, which means that microbial enzymes in intestinal and colon phase are more active on glycans. The significant difference was observed between phases, except between the small intestine and colon groups, was determined ($p<0.05$). Especially, the concentration of glycans noticeably increased from about 8 mg/mL to almost 19 mg/mL from gastric to small intestine phase, respectively. This also supports the considerable activity of microbial enzymes in intestinal and colon phase in comparison to oral and gastric phases. As for the only microbial enzyme contribution in this new model, a statistically significant difference between groups of oral – gastric and small intestine – colon was determined ($p<0.05$). No statistically significant difference was determined between oral and gastric groups as well as between small intestine and colon groups ($p>0.05$). In contrast to the novel model, the results from the concentration of released glycans from the standard *in-vitro d*igestion model has shown that a statistically significant difference between the oral and other groups was determined ($p<0.05$), however, no statistically significant difference was determined between the any groups of gastric, small intestine, and colon ($p>0.05$). The only increment from oral to gastric phase may be related to formation of peptides by pepsin enzyme which enables glycans on peptides more measurable in comparison to their form of on complex whey protein. When the data of both novel and standard model were compared, it has been shown that the differences between the data from the novel and standard models are significant for each phase separately ($p<0.05$). Even though no significant difference between oral and gastric phase data for both models, it can be clearly seen that the novel model utilizes glycans at a higher concentration rate by microbial enzymes in intestinal and colon phases in comparison to the standard model. Regarding the summary of results from this section, microbial enzymes (glycosyl hydrolases produced in this thesis) considerably affect the digestion of glycans conjugated on proteins.

In conclusion, this thesis is critical to giving a new perspective to different studies covering human digestion and glycan metabolism by microbial enzymes in a more appropriate and comprehensive way by integrating microbial enzymes into the conventional *in-vitro* digestion model. The human body carries a million microorganisms whose genes have incredible functions ranging from vitamin synthesis to digestion in the human body. As the glycans on glycoproteins are resistant to being metabolized by human-associated enzymes, they are utilized by some gut bacteria such as Bifidobacteria and Lactobacillus that use their carbohydrate-active enzymes (mainly glycosyl hydrolase enzymes) and transporters. Microorganisms and their glycosyl hydrolase enzymes which are active on prebiotics are incredibly abundant in the human host, evenly more than the number of host cells and genes. This indicates that in addition to human digestion enzymes, microorganisms and their carbohydrate-active enzymes take a noticeable part in human digestion. As the majority of microorganisms inhabit the human gut, which is related to the gut environment including pH, oxygen state, and mucus structure is suitable for the growth of many microorganisms. Therefore, the utilization of glycans massively takes place in the human gut and gut microorganisms express different carbohydrate-active enzymes to combat each other for the growth sources, glycans. As similar to the studies in the literature, the novel model in this thesis has shown that microbial enzymes from more diverse microorganisms in comparison to other phases release glycans at a higher concentration in intestinal phases. The integration of microbial enzymes is essential to combine both host and microbial enzymes in order to mimic human digestion at a higher rate and better study the digestion process in laboratory conditions. Thus, the design of a novel *in-vitro* model by integrating microbial enzymes will significantly contribute to the literature and studies related to digestion, glycans as well as the human microbiome. This novel *in-vitro* digestion model is of the utmost importance to comprehensively understand the digestion process by microbial enzymes and interactions between glycans and carbohydrate-active enzymes. A variety of scientific studies can be performed with this novel model in terms of glycans and their utilization by microbial enzymes, which would be a critical step in the glycobiology field and shed light on future studies in many other fields including the food industry, medicine, pharmacy, and even more personalized medicine.

REFERENCES

Aas, J. A., Paster, B. J., Stokes, L. N., Olsen, I., & Dewhirst, F. E. (2005). "Defining the normal bacterial flora of the oral cavity". *Journal of Clinical Microbiology*, *43*(11), 5721–5732. https://doi.org/10.1128/JCM.43.11.5721-5732.2005

Alminger, M., Aura, A. M., Bohn, T., Dufour, C., El, S. N., Gomes, A., Karakaya, S., Martínez-Cuesta, M. C., Mcdougall, G. J., Requena, T., & Santos, C. N. (2014). "In Vitro Models for Studying Secondary Plant Metabolite Digestion and Bioaccessibility". *Comprehensive Reviews in Food Science and Food Safety*, *13*(4), 413–436. https://doi.org/10.1111/1541-4337.12081

Altmann, F., Schweiszer, S., & Weber, C. (1995). "Kinetic comparison of peptide: N-glycosidases F and A reveals several differences in substrate specificity". *Glycoconjugate Journal*, *12*(1), 84–93. https://doi.org/10.1007/BF00731873

Ballard, O., & Morrow, A. L. (2013). "Human Milk Composition. Nutrients and Bioactive Factors". In *Pediatric Clinics of North America*. https://doi.org/10.1016/j.pcl.2012.10.002

Bienenstock, J., Buck, R. H., Linke, H., Forsythe, P., Stanisz, A. M., & Kunze, W. A. (2013). "Fucosylated but Not Sialylated Milk Oligosaccharides Diminish Colon Motor Contractions". *PLoS ONE*, *8*(10). https://doi.org/10.1371/JOURNAL.PONE.0076236

Bik, E. M., Eckburg, P. B., Gill, S. R., Nelson, K. E., Purdom, E. A., Francois, F., Perez-Perez, G., Blaser, M. J., & Relman, D. A. (2006). "Molecular analysis of the bacterial microbiota in the human stomach". *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.0506655103

Bode, L. (2012). "Human milk oligosaccharides: Every baby needs a sugar mama". *Glycobiology*, *22*(9), 1147–1162. https://doi.org/10.1093/glycob/cws074

Bornhorst, G. M., Gouseti, O., Wickham, M. S. J., & Bakalis, S. (2016). "Engineering Digestion: Multiscale Processes of Food Digestion". *Journal of Food Science*, *81*(3), R534–R543. https://doi.org/10.1111/1750-3841.13216

Brodkorb, A., Egger, L., Alminger, M., Alvito, P., Assunção, R., Ballance, S., Bohn, T., Bourlieu-Lacanal, C., Boutrou, R., Carrière, F., Clemente, A., Corredig, M., Dupont, D., Dufour, C., Edwards, C., Golding, M., Karakaya, S., Kirkhus, B., Le Feunteun, S., ... Recio, I. (2019). "INFOGEST static in vitro simulation of gastrointestinal food digestion". *Nature Protocols*, *14*(4), 991–1014. https://doi.org/10.1038/s41596-018- 0119-1

Carlson, D. M. (1968). "Structures and immunochemical properties of oligosaccharides isolated from pig submaxillary mucins". *Journal of Biological Chemistry*.

Chaplin, A. V., Efimov, B. A., Smeianov, V. V., Kafarskaia, L. I., Pikina, A. P., & Shkoporov, A. N. (2015). "Intraspecies genomic diversity and long-term persistence of bifidobacterium longum". *PLoS ONE*, *10*(8). https://doi.org/10.1371/journal.pone.0135658

Chichlowski M., Shah N., Wampler JL., Wu SS., Vanderhoof JA. (2020). "*Bifidobacterium longum Subspecies infantis (B.infantis)* in Pediatric Nutrition: Current State of Knowledge". *Nutrients*, 12, 1581. https://doi.org/10.3390/nu12061581

Clarke, G., O'Mahony, S. M., Dinan, T. G., & Cryan, J. F. (2014). "Priming for health: Gut microbiota acquired in early life regulates physiology, brain and behaviour". *Acta Paediatrica, International Journal of Paediatrics*, *103*(8), 812–819. https://doi.org/10.1111/APA.12674

Corinaldesi, R., Stanghellini, V., Raiti, C., Rea, E., Salgemini, R., & Barbara, L. (1987). "Effect of chronic administration of cisapride on gastric emptying of a solid meal and on dyspeptic symptoms in patients with idiopathic gastroparesis". *Gut*, *28*, 300–305. https://doi.org/10.1136/gut.28.3.300

Corstens, M. N., Berton-Carabin, C. C., Schroën, K., Viau, M., & Meynier, A. (2018). "Emulsion encapsulation in calcium-alginate beads delays lipolysis during dynamic in vitro digestion". *Journal of Functional Foods*, *46*, 394–402. https://doi.org/10.1016/J.JFF.2018.05.011

Davenport, E. R., Cusanovich, D. A., Michelini, K., Barreiro, L. B., Ober, C., & Gilad, Y. (2015). "Genome-wide association studies of the human gut microbiota". *PLoS ONE*, *10*(11). https://doi.org/10.1371/JOURNAL.PONE.0140301

Debaun, R. M., & Connors, W. M. (1954). "Nutritional Assay: Relationship Between In Vitro Enzymatic Digestibility and In Vivo Protein Evaluation of Powdered Whey". *Journal of Agricultural and Food Chemistry*, *2*(10), 524–526. https://doi.org/10.1021/JF60030A007

Duar, R. M., Henrick, B. M., Casaburi, G., & Frese, S. A. (2020). "Integrating the Exosystem Services Framework to Define Dysbiosis of the Breastfed Infant Gut: The role of B. İnfantis and Human Milk Oligosaccharides". *Frontiers, Nutrition*. 7(33). https://doi:10.3389/fnut.2020.00033/bibtex

Dupuy, A. K., David, M. S., Li, L., Heider, T. N., Peterson, J. D., Montano, E. A., Dongari-Bagtzoglou, A., Diaz, P. I., & Strausbaugh, L. D. (2014). "Redefining the human oral mycobiome with improved practices in amplicon-based taxonomy: Discovery of Malassezia as a prominent commensal". *PLoS ONE*, *9*(3), 1–11. https://doi.org/10.1371/journal.pone.0090899

Dwek, R. (1993). "Analysis of Glycoprotein-Associated Oligosaccharides". *Annual Review of Biochemistry*. https://doi.org/10.1146/annurev.biochem.62.1.65

Egger, L., Ménard, O., Baumann, C., Duerr, D., Schlegel, P., Stoll, P., Vergères, G., Dupont, D., & Portmann, R. (2019). "Digestion of milk proteins: Comparing static and dynamic in vitro digestion systems with in vivo data". *Food Research International*. https://doi.org/10.1016/j.foodres.2017.12.049

Gerritsen, J., Smidt, H., Rijkers, G. T., & De Vos, W. M. (2011). "Intestinal microbiota in human health and disease: The impact of probiotics". *Genes and Nutrition*, *6*(3), 209– 240. https://doi.org/10.1007/s12263-011-0229-7

Gibson, G. R., & Roberfroid, M. B. (1995). "Dietary modulation of the human colonic microbiota: introducing the concept of prebiotics". *The Journal of Nutrition*, *125*(6), 1401–1412. https://doi.org/10.1093/JN/125.6.1401

Guerra, A., Etienne-Mesmin, L., Livrelli, V., Denis, S., Blanquet-Diot, S., & Alric, M. (2012). "Relevance and challenges in modeling human gastric and small intestinal digestion". *Trends in Biotechnology*, *30*(11), 591–600. https://doi.org/10.1016/J.TIBTECH.2012.08.001

Hawkey, C. J., Mahida, Y. R., & Hawthorne, A. B. (1992). "Therapeutic interventions in gastrointestinal disease based on an understanding of inflammatory mediators". *Agents Actions*.

Henrick, B. M., Hutton, A. A., Palumbo, M. C., Casaburi, G., Mitchell, R. D., Underwood, M. A., Smilowitz, J. T., & Frese, S. A. (2018). "Elevated Fecal pH Indicates a Profound Change in the Breastfed Infant Gut Microbiome Due to Reduction of Bifidobacterium over the Past Century". *MSphere*, *3*(2).

Hillman, E. T., Lu, H., Yao, T., & Nakatsu, C. H. (2017). "Microbial ecology along the gastrointestinal tract". *Microbes and Environments*, *32*(4), 300–313. https://doi.org/10.1264/jsme2.ME17017

Hutkins, R. W., Krumbeck, J. A., Bindels, L. B., Cani, P. D., Fahey, G., Goh, Y. J., Hamaker, B., Martens, E. C., Mills, D. A., Rastal, R. A., Vaughan, E., & Sanders, M. E. (2016). "Prebiotics: Why definitions matter". *Current Opinion in Biotechnology*, *37*, 1–7. https://doi.org/10.1016/J.COPBIO.2015.09.001

Ji, H., Hu, J., Zuo, S., Zhang, S., Li, M., & Nie, S. (2021)." In vitro gastrointestinal digestion and fermentation models and their applications in food carbohydrates". https://doi.org/10.1080/10408398.2021.1884841.

Karav, S., Bell, J. M. L. N. D. M., Parc, A. Le, Liu, Y., Mills, D. A., Block, D. E., & Barile, D. (2015). "Characterizing the release of bioactive N- glycans from dairy products by a novel endo-β-N-acetylglucosaminidase". *Biotechnology Progress*, *31*(5), 1331– 1339. https://doi.org/10.1002/btpr.2135

Karav, S., 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, S., Casaburi, G., & Frese, S. A. (2018). "Reduced colonic mucin degradation in breastfed infants colonized by Bifidobacterium longum subsp.infantis EVC001". *FEBS Open Bio*, *8*(10), 1649–1657. https://doi.org/10.1002/2211-5463.12516

Karav, S., German, J. B., Rouquié, C., Le Parc, A., & Barile, D. (2017). "Studying lactoferrin N-glycosylation". In *International Journal of Molecular Sciences*. https://doi.org/10.3390/ijms18040870

Karav, S., Le Parc, A., de Moura, J. M. L. N., 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*, AEM. 00547-16.

Karav, S., Parc, A. Le, Moura Bell, J. M. L. N. de, Rouqui, 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*(0), 46–53. https://doi.org/http://dx.doi.org/10.1016/j.enzmictec.2015.05.007

Koropatkin, N. M., Cameron, E. A., & Martens, E. C. (2012). "How glycan metabolism shapes the human gut microbiota". In *Nature Reviews Microbiology*. https://doi.org/10.1038/nrmicro2746

Kunz, C., Rudloff, S., Baier, W., Klein, N., & Strobel, S. (2000). "Oligosaccharides in human milk: structural, functional, and metabolic aspects". *Annual Review of Nutrition*, *20*(1), 699–722.

Leimena, M. M., Ramiro-Garcia, J., Davids, M., van den Bogert, B., Smidt, H., Smid, E. J., Boekhorst, J., Zoetendal, E. G., Schaap, P. J., & Kleerebezem, M. (2013). "A comprehensive metatranscriptome analysis pipeline and its validation using human small intestine microbiota datasets". *BMC Genomics*. https://doi.org/10.1186/1471- 2164-14-530

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*. https://doi.org/10.1111/j.1751-7915.2008.00072.x

Marcano, J., Hernando, I., & Fiszman, S. (2015). "Invitro measurements of intragastric rheological properties and their relationships with the potential satiating capacity of cheese

pies with konjac glucomannan". *Food Hydrocolloids*, *51*, 16–22. https://doi.org/10.1016/j.foodhyd.2015.04.028

Ménard, O., Cattenoz, T., Guillemin, H., Souchon, I., Deglaire, A., Dupont, D., & Picque, D. (2014). "Validation of a new in vitro dynamic system to simulate infant digestion". *Food Chemistry*, *145*, 1039–1045. https://doi.org/10.1016/j.foodchem.2013.09.036

Minekus, M., Alminger, M., Alvito, P., Ballance, S., Bohn, T., Bourlieu, C., Carrière, F., Boutrou, R., Corredig, M., Dupont, D., Dufour, C., Egger, L., Golding, M., Karakaya, S., Kirkhus, B., Le Feunteun, S., Lesmes, U., Macierzanka, A., Mackie, A., ... Brodkorb, A. (2014). "A standardised static in vitro digestion method suitable for food – an international consensus". *Food Funct.*, *5*(6), 1113–1124. https://doi.org/10.1039/C3FO60702J

Moremen, K. W., Tiemeyer, M., & Nairn, A. V. (2012). "Vertebrate protein glycosylation: diversity, synthesis and function". *Nature Reviews Molecular Cell Biology*, *13*(7), 448– 462.

Morgan, B. L. G., & Winick, M. (1980). "Effects of Administration of N-Acetylneuraminic Acid (NANA) on Brain NANA Content and Behavior". *The Journal of Nutrition*, *110*(3), 416–424. https://doi.org/10.1093/jn/110.3.416

O'Hara, A. M., & Shanahan, F. (2006). "The gut flora as a forgotten organ". *EMBO Reports*, *7*(7), 688–693. https://doi.org/10.1038/sj.embor.7400731

Olin, A., Henckel, E., Chen, Y., Lakshmikanth, T., Pou, C., Mikes, J., Gustafsson, A., Bernhardsson, A. K., Zhang, C., Bohlin, K., & Brodin, P. (2018). "Stereotypic Immune System Development in Newborn Children". *Cell*, *174*(5), 1277-1292.e14. https://doi.org/10.1016/J.CELL.2018.06.045

Parc, A. Le, Karav, S., Bell, J. M. L. N. D. M., Frese, S. A., Liu, Y., Mills, D. A., Block, D. E., & Barile, D. (2015). "A novel endo-β-N-acetylglucosaminidase releases specific N-glycans depending on different reaction conditions". *Biotechnology Progress*, *31*(5), 1323–1330. https://doi.org/10.1002/btpr.2133

Qazi, W. M., Ballance, S., Kousoulaki, K., Uhlen, A. K., Kleinegris, D. M. M., Skjånes, K., & Rieder, A. (2021). "*Protein Enrichment of Wheat Bread with Microalgae".*

Sek, L., Porter, C. J. H., & Charman, W. N. (2001). "Characterisation and quantification of medium chain and long chain triglycerides and their in vitro digestion products, by HPTLC coupled with in situ densitometric analysis". *Journal of Pharmaceutical and Biomedical Analysis*. https://doi.org/10.1016/S0731-7085(00)00528-8

Sela, D. 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*. https://doi.org/10.1128/AEM.06762-11

Shafquat, A., Joice, R., Simmons, S. L., & Huttenhower, C. (2014). "Functional and phylogenetic assembly of microbial communities in the human microbiome". In *Trends in Microbiology*. https://doi.org/10.1016/j.tim.2014.01.011

Sojar, H. T., & Bahl, O. P. (1987). "Chemical Deglycosylation of Glycoproteins". *Methods in Enzymology*. https://doi.org/10.1016/0076-6879(87)38029-2

Tannock, G. W., & Savage, D. C. (1974). "Influences of dietary and environmental stress on microbial populations in the murine gastrointestinal tract". *Infection and Immunity*, *9*(3), 591–598. https://doi.org/10.1128/iai.9.3.591-598.1974

Townsend, S. D., & Moore, R. E. (2019). "*Temporal development of the infant gut microbiome"*. https://doi.org/10.1098/rsob.190128

Tretter, V., Altmann, F., & März, L. (1991). "Peptide-N4-(N-acetyl-β- glucosaminyl) asparagine amidase F cannot release glycans with fucose attached α1 → 3 to the asparagine-linked N-acetylglucosamine residue". *European Journal of Biochemistry*, *199*(3), 647–652. https://doi.org/10.1111/j.1432-1033.1991.tb16166.x

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". *Journal of Biological Chemistry*.

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*. https://doi.org/10.1016/j.ab.2013.10.019

Ursell, L. K., Metcalf, J. L., Parfrey, L. W., & Knight, R. (2012). "Defining the human microbiome". *Nutrition Reviews*, *70*(SUPPL. 1). https://doi.org/10.1111/j.1753-4887.2012.00493.x

van Berkel, P. H. C., Geerts, M. E. J., van Veen, H. A., Kooiman, P. M., Pieper, F. R., de Boer, H. A., & Nuijens, J. H. (1995). "Glycosylated and unglycosylated human lactoferrins both bind iron and show identical affinities towards human lysozyme and bacterial lipopolysaccharide, but differ in their susceptibilities towards tryptic proteolysis". *Biochemical Journal*, *312*(1), 107–114. https://doi.org/10.1042/bj3120107

Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W., & Etzler, M. E. (2009). "Essentials of Glycobiology". *Cold Spring Harbor (NY*, *039*, 2015–2017. https://www.ncbi.nlm.nih.gov/books/NBK1908/

Wada, Y., & Lönnerdal, B. (2015). "Bioactive peptides released by in vitro digestion of standard and hydrolyzed infant formulas". *Peptides*. https://doi.org/10.1016/j.peptides.2015.09.005

Walsh, C., Lane, J. A., van Sinderen, D., & Hickey, R. M. (2020). "Human milk oligosaccharides: Shaping the infant gut microbiota and supporting health". *Journal of Functional Foods*, *72*. https://doi.org/10.1016/j.jff.2020.104074

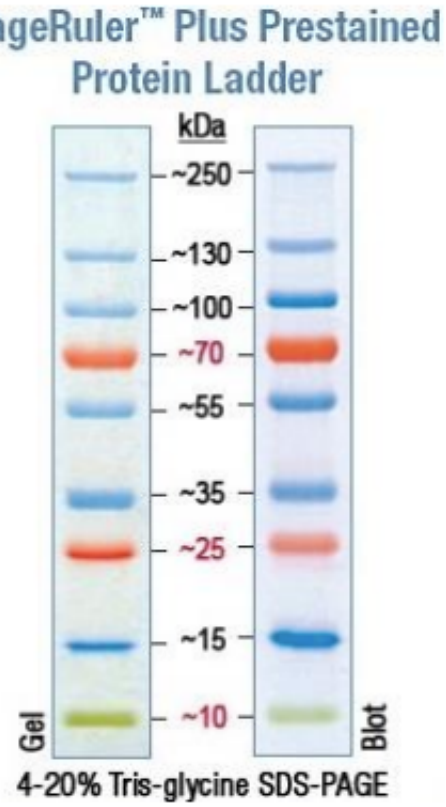
Wiciński, M., Sawicka, E., Gębalski, J., Kubiak, K., & Malinowski, B. (2020). "Human Milk Oligosaccharides: Health Benefits, Potential Applications in Infant Formulas, and Pharmacology". *Nutrients*, *12*(1). https://doi.org/10.3390/NU12010266

Wopereis, H., Oozeer, R., Knipping, K., Belzer, C., & Knol, J. (2014). "The first thousand days - intestinal microbiology of early life: Establishing a symbiosis". *Pediatric Allergy and Immunology*, *25*(5), 428–438. https://doi.org/10.1111/pai.12232

APPENDICES

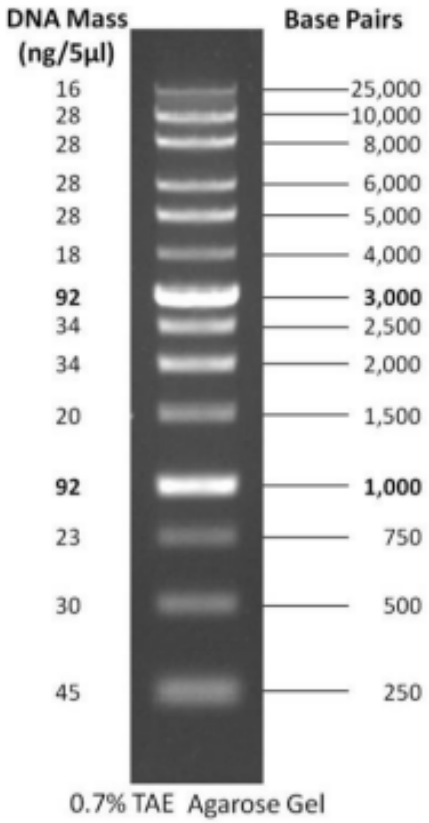
APPENDIX 1

PROTEIN LADDER USED IN SDS-PAGE



APPENDIX 2

DNA LADDER USED IN AGAROSE GEL ELECTROPHORESIS



APPENDIX 3

BUFFERS USED IN THE THESIS

	Lysis Buffer pH 7.5-8	**Equilibration Buffer pH 7.4**	**Wash Buffer pH 7.4**	**Elution Buffer pH 7.4**	**Tris-HCl pH 6.8**	**Tris-HCl pH 8.8**	**Na_2HPO_4 0.2 M, pH 5**
NaCl	200 mM	300 mM	300 mM	300 mM	-	-	-
Tris	50 mM	-	-	-	0.5 M	1.5 M	-
Imidazole	1 mM	10 mM	25 mM	250 mM	-	-	-
NaH_2PO_4	-	20 mM	20 mM	20 mM	-	-	-
SDS	1%	-	-	-	-	-	-
$Na_2HPO_4 \cdot 2H_2O$	-	-	-	-	-	-	1 M
$NaH_2PO_4 \cdot H_2O$	-	-	-	-	-	-	1 M

Chemical	**mol/L (M)**	**SSF (pH:7)**	**SGF (pH:3)**	**SIF (pH:7)**
KCl	0.5	15.1 mL	6.9 mL	6.8 mL
KH_2PO_4	0.5	3.7 mL	0.9 mL	0.8 mL
$NaHCO_3$	1	6.8 mL	12.5 mL	42.5 mL
NaCl	2	-	11.8 mL	9.6 mL
$MgCl_2(H_2O)_6$	0.15	0.5 mL	0.4 mL	1.1 mL
$(NH_4)_2CO_3$	0.5	0.06 mL	0.5 mL	-
HCl	6	0.09 mL	1.3 mL	0.7 mL
	0.3			

APPENDIX 4

ORAL PRESENTATION


4th International Eurasian Conference on Biological and Chemical Sciences (EurasianBioChem 2021)

Certificate OF ATTENDANCE

This Certificate Is Proudly Presented To

Merve KAPLAN

in recognition of participation in the

"4th International Eurasian Conference on Biological and Chemical Sciences

4th International Eurasian Conference on Biological and Chemical Sciences (EurasianBioChem 2021) November 24-26, 2021. www.EurasianBioChem.org

➢ ***ORAL PRESENTATION***

Novel *In-Vitro* Digestion Model Designed by Integration of Microbiome Associated Enzymes

Merve KAPLAN[1], Sercan KARAV[2*]

[1]Canakkale Onsekiz Mart University, School of Graduate Studies, Molecular Biology and Genetics, Canakkale, Turkey.
[2*] Canakkale Onsekiz Mart University, Faculty of Science and Literature, Molecular Biology and Genetics, Canakkale, Turkey.

*Corresponding author s

Abstract

The human microbiome includes around 100 trillion bacterial cells, which is 10 times more than human cells. Therefore, it is inevitable that bacteria play important roles in our body. One of the most important functions of bacteria is taken place in the digestion process in the human body. Human milk oligosaccharides (HMOs) and conjugated glycans, for instance, cannot be digested by humans due to the absence of specific glycosidase enzymes. Therefore, they reach the colon where they are used as a carbon source by some bacteria such as Bifidobacteria. Such compounds are called prebiotics which enhance the growth of beneficial bacteria. To better understand the digestion process, *in-vitro* digestion systems are used in many laboratories with their extensive advantages. However, current *in-vitro* digestion models are not available for significant glycan studies because of microbiome-based enzymes specificity. So, the design of novel models including both host and microbiome-based enzymes is critical to pave the way for glycan research. For this purpose, we studied to generate a novel model with recombinant microbial enzymes which have a role in digestion. Firstly, target microorganisms that predominate in the human GI system (four phases; oral, gastric, small intestine, and colon) were determined and specific glycosidases of these microorganisms were identified using bioinformatic methods. Then, these enzymes were cloned, produced, and integrated through a conventional *in-vitro* digestion model to study the prebiotic properties of a variety of sources.

Keywords: Microbiome, Glycans, *In-vitro* Digestion Model