RECOMBINANT PRODUCTION of GLP-1 ANALOGUE USING *Escherichia coli* HOST ORGANISM

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

By
Burak Çalış
August 2024

RECOMBINANT PRODUCTION of GLP-1 ANALOGUE USING *Escherichia coli* HOST ORGANISM

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

By
Burak ÇALIŞ
August 2024

SCIENTIFIC ETHICS COMPLIANCE

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

Name-Surname: Burak ÇALIŞ
Signature :

REGULATORY COMPLIANCE

M.Sc. thesis titled Recombinant Production of GLP-1 Analogue Using *Escherichia coli* Host Organism has been prepared in accordance with the Thesis Writing Guidelines of the Abdullah Gül University, Graduate School of Engineering & Science.

Prepared By	Advisor
Burak ÇALIŞ	Asst. Prof. Özkan FİDAN
Signature	Signature

Head of the Bioengineering Program
Asst. Prof. Emel Başak GENCER AKÇOK
Signature

ACCEPTANCE AND APPROVAL

M.Sc. thesis titled Recombinant Production of GLP-1 Analogue Using *Escherichia coli* Host Organism and prepared by Burak ÇALIŞ has been accepted by the jury in the Bioengineering Graduate Program at Abdullah Gül University, Graduate School of Engineering & Science.

19 /08 / 2024

(Thesis Defense Exam Date)

JURY:

Advisor : Asst. Prof. Özkan FİDAN

Member: Asst. Prof. Emel Başak GENCER AKÇOK

Member: Asst. Prof. Rizvan İMAMOĞLU

APPROVAL:

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

…….. /…….. / ……..

(Date)

Graduate School Dean
Prof. Dr. İrfan ALAN

ABSTRACT

RECOMBINANT PRODUCTION of GLP-1 ANALOGUE USING *Escherichia coli* HOST ORGANISM

Burak Çalış
MSc. in Bioengineering
Advisor: Asst. Prof. Özkan Fidan
August 2024

Diabetes is the most serious metabolic disorder correlated with obesity, hypertension and cardiovascular conditions. High prevalence of Type II Diabetes Mellitus (T2DM) indicates the need for new medication development. In developing therapeutics, higher efficiency and fewer adverse effect features are targeted primarily. Recombinant protein-based biotechnological drug molecules have been developed and used for the treatment of T2DM. Especially, GLP-1 analogues are known by their self-limiting mechanism and insulinotropic effect. In this study, a novel GLP-1 analogue with increased stability and efficiency is produced using recombinant *E. coli*. The expression plasmid was constructed and confirmed by restriction digestion and whole plasmid sequencing. Then, itwas transformed into various *E. coli* strains followed by optimized lysis, growth and expression conditions to maximize the yield of the GLP-1 analogue. Various parameters such as pre-induction time, induction point, induction IPTG concentration and post-induction temperature were tested for the succesfull expression with maximum yield. Consequently, it was achieved that *E. coli* BL21(DE3) as strain, 0.2 mM IPTG induction at OD_{600nm} of 0.6 and 18 °C overnight post-induction growth was the most promising conditions. Under these conditions, the GLP-1 analogue was obtained in the insoluble fraction. Following protein analysis and purification, quantification was performed and the highest titer of GLP-1 analogue was measured as 626 µg/ml. As future prospect, using another host organism and changing growth conditions can provide obtaining target protein in the soluble form.
Keywords: T2DM, GLP-1 analogue, recombinant DNA technology, protein expression, E. coli

ÖZET

Escherichia coli KONAK ORGANİZMADA GLP-1 ANALOĞUNUN REKOMBİNANT ÜRETİMİ

Burak Çalış
Biyomühendislik Anabilim Dalı Yüksek Lisans
Tez Danışmanı: Asst. Prof. Özkan Fidan
Ağustos 2024

Diyabet, obezite, hipertansiyon ve kardiyovasküler hastalıklarla ilişkilendirilen en ciddi metabolik bozukluktur. Özellikle Tip II Diyabet'in yüksek prevalansı, buna karşı yeni ilaç geliştirme ihtiyacını işaret etmektedir. Geliştirilen ilaçlarda öncelikle daha yüksek verimlilik ve daha az yan etki hedeflenmektedir. Rekombinant protein temelli biyoteknolojik ilaç molekülleri bu amaç için umut vericidir. Özellikle GLP-1 analogları, kendi kendini sınırlayan mekanizmaları ve insülinotropik etkileriyle bilinmektedir. Bu çalışmada, rekombinant E. coli kullanılarak artırılmış stabilite ve verimliliğe sahip yeni bir GLP-1 analoğu üretilmiştir. Ekspresyon plazmidi, restriksiyon-ligasyon yöntemi ile oluşturulmuş ve tüm plazmit dizilimi ile doğrulanmıştır. Daha sonra, GLP-1 analoğunun verimini en üst düzeye çıkarmak için optimize edilmiş lizis, büyüme ve ekspresyon koşulları izlenerek çeşitli E. coli suşlarına transform edilmiştir. Ön indüksiyon süresi, indüksiyon noktası, indüksiyon IPTG konsantrasyonu ve indüksiyon sonrası sıcaklık gibi çeşitli parametreler, maksimum verimle başarılı ifade için test edilmiştir. Sonuç olarak, *E. coli* BL21(DE3) suşu, 0.6 OD_{600nm}'de 0.2 mM IPTG indüksiyonu ve 18 °C'de gece boyunca indüksiyon sonrası büyümenin en verimli koşullar olduğu sonucuna ulaşıldı. Bu koşullar altında hedef protein GLP-1 analoğu, çözünmeyen fraksiyon halinde elde edildi. Protein analizi ve saflaştırmanın ardından kantifikasyon yapıldı ve GLP-1 analoğunun en yüksek titresi 626 µg/ml olarak ölçüldü. Gelecekteki beklentiler olarak, başka bir konak organizma kullanmak ve büyüme koşullarını değiştirmek, hedef proteinin çözünür formda elde edilmesini sağlayabilir.

Anahtar kelimeler: T2DM, GLP-1 analoğu, rekombinant DNA teknolojisi, protein ekspresyonu, E. coli

Acknowledgements

First of all, I would like to thank Dr. Özkan Fidan, who became my advisor in the middle of the year after an unfortunate start to my Master program. I am proud to be his student and express my sincere respects because he constantly communicates healthily with me, understands me, supports me and improved my motivation. I will always be grateful to him for sharing his experiences with me and always supporting my development.

I would like to thank my lab-mates Fatma Şener and Melisa Zülal Karaman, who made my experiments enjoyable and easier every day without exception in the laboratory. It is very valuable for me that they support me even in the most difficult times and they are always by my side. I hope that together we can achieve the goals we consistently work towards. I would also like to thank our undergraduate student Ayşe Şahin for her help.

I would like to thank respectable guides TA Büşra Acar, TA Merve Şansaçar and TA Nihan Aktaş Pepe, who never spared their help and spared their time when I had difficulties, for their sincerity and I will always be grateful.

I would like to thank my dear home-mate Muhammed Hatem, who always supported me throughout this process, stood by me and made even the busiest days enjoyable. I would also like to thank my dear friend Yusuf Erdem, who always helped me when I needed in the laboratory, even though our work areas did not match.

I would like to thank TUBITAK for supporting me in this process within the scope of 2210-C Scholarship program.

Finally, I am deeply grateful to my family, who have been my support throughout my entire educational life. Their never-ending encouragement and belief in me are the most valuable factors in my success. I express my sincere love to all of my family members, who always give me strength and are the greatest luck in my life. I hope I can always be a family member that they are proud of.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

T2DM	Type II Diabetes Mellitus
SGLT2	Sodium-Glucose Transport Protein 2
GLP-1	Glucagon-like Peptide 1
DPP-IV	Dipeptidyl Peptidase 4
GPCR	G-protein-coupled Receptor
GIP	Gastric Inhibitory Peptide
PCR	Polymerase Chain Reaction
EDTA	Ethylenediaminetetraacetic Acid
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
IPTG	Isopropyl β- d-1-thiogalactopyranoside
PMSF	Phenylmethylsulfonyl Fluoride
DTT	Dithiothreitol
OD	Optical Density
Ni-NTA	Nickel-nitrilotriacetic Acid
HSA	Human Serum Albumin
PCN	Plasmid Copy Number

Chapter 1

Introduction

1.1. Diabetes Prevalence and Mechanism

Diabetes and obesity are two interconnected health problems that are becoming increasingly common in Turkey and the world. It is a chronic condition that affects the body's ability to produce or use insulin, a hormone that regulates blood sugar. There are two main types of diabetes: type I diabetes, which is an autoimmune disorder that typically develops in childhood, and type II diabetes, which is often linked to lifestyle factors such as poor diet, lack of physical activity, and obesity [1]. The incidence of diabetes in Turkey has been increasing steadily over the last few decades. Ministry of Health's 2022 data on diabetes prevalence shows that there are approximately 7 million diabetic patients in the 20-79 age range in Turkey and this corresponds to approximately 15% of the total adult population [2]. In another study, it was stated that the most common metabolic disease is type II diabetes, and approximately 5-10% of the population in developed countries has type II diabetes. In addition, 90-95% of diabetes cases worldwide are type II diabetes, 5-10% are type I diabetes, and 2-3% are other types of diabetes [3]. International Diabetes Federation stated that nearly there were 463 million diabetes patients in 2019, and this number was expected to be 700 million after 2040 [4]. Increase of T2DM prevalence is greatly correlated with the obesity and sedentary life style. Urbanization and the Western dietary habit’s, preference of high-calorie and low-nutrient foods influenced the T2DM prevalence [5]. Additionally, impact of genetic tendency on T2DM is undeniable. Because, various genetic variations are associated with insulin sensitivity and beta-cells’ functionality [6]. So, it can be said that environmental factors, life-style and genetic disposition all together have impacts on T2DM development.

Main mechanism of T2DM is caused by insulin resistance which is a condition of less sensitivity of body's cell to insulin. Thus, need of insulin is increased and secretion is enhanced to keep blood glucose levels within standard range. Pancreatic beta-cells are responsible for insulin production and they become dysfunctional over time. Eventually, these cells can't meet the body's elevated insulin requirement which is resulted by hyperglycemia [7].

T2DM is a complex condition which is not only correlated with death rate and disorder prevalence, but also it has a serious financial impact on global healthcare system. Because T2DM is closely linked with development of hypertension, obesity, hypercholesterolemia, cardiovascular conditions, neuropathy and Alzheimer's Disease as secondary complications [8]. It is reported that the average person with diabetes could spend between $3,300 and $4,600 a year on out-of-pocket costs, including prescription medications, provider visits, over-the-counter supplies, and lost wages. Patients who have chronic complications from the disease can expect to pay an additional $470 each year on average to manage their other conditions. Also, the total annual cost of diabetes in 2022 is $412.9 billion, including $306.6 billion in direct medical costs and $106.3 billion in indirect costs. So, the economic perspective of the T2DM is another factor that shows significant need of treatment [74, 75].

1.2. Current Treatments Against T2DM

The treatment of T2DM is a long period started by the use of insulin from pancreatic animal cells in 1922 by Leonard Thompson and insulin injection is accepted as primary treatment still [9]. Current treatments consist of orally administered drugs and injectable ones and they can be used alone or in combination depending on the patient status. Managing blood glucose levels, reducing body weight and easing cardiovascular risks are common targets of modern T2DM treatments [10]. However, the medications in that category usually have adverse effects with varying degrees including weight loss, hypoglycemia, gastrointestinal issues, edema, increased risk of heart failure, bone fractures, headache, pancreas inflammation, vomiting, renal failure and diarrhea [11]. Antidiabetic drugs can be divided into many classes: insulin sensitizers including biguanides and thiazolidinediones, insulin secretagogues such as sulfonylureas and meglitinides, alpha-

glucosidase inhibitors, Sodium-Glucose Transport Protein 2 (SGLT2) inhibitors and incretin-based medications including glucagon like peptide-1 (GLP-1) analogues and dipeptidyl peptidase 4 (DPP-IV) inhibitors. All these drug classes have distinct mechanisms and targets [12].

Among insulin sensitizers, metformin as an example, which is belong to biguanides class, is extracted from the *Galega officinalis* plant and helps to restrict hepatic glucose production, resulted in increased insulin sensitivity of peripheral tissue cells. Gastrointestinal side effects and lactic acidosis risks are associated with metformin drug [13]. Other class of insulin sensitizers is thiazolidinediones including pioglitazone and rosiglitazone. They help to increase insulin sensitivity and decrease blood glucose levels by activating peroxisome proliferator-activated receptor gamma (PPAR-γ) receptor. Weight gain, fluid retention and elevated fracture possibility are counted as side effects for that group of drugs [14]. Another T2DM medication class is insulin secretagogues which stimulate insulin release from β-cells in pancreas and controlling postprandial blood glucose levels. Commonly, weight gain and hypoglycemia are indicated as adverse effect for insulin secretagogues drugs. Sulfonylureas and meglitinides such as nateglinide and repaglinide are involved in that group [15-17]. Alpha-glucosidase inhibitors including acarbose, voglibose and miglitol help to delay absorption of carbohydrate in the small intestine and decreasing glucose levels in postprandial blood. As side effects, gastrointestinal problems like flatulence and diarrhea are associated with this medication group [18, 19]. SGLT2 inhibitors such as canagliflozin and depagliflozin are another alternative drug group against T2DM. Their mechanism of action is inhibiting glucose reabsorption together with improving urinary glucose excretion. However, it is known that they can cause urinary and genital infections [20]. The last class of antidiabetic drugs is incretin based ones which consist of GLP-1 receptor agonists (GLP-1 analogues) and DPP-IV inhibitors. GLP-1 analogues like exenatide, liraglutide, dulaglutide, albiglutide and lixisenatide help to elevate insulin secretion from pancreas by inhibiting glucagon release at the same time. Blood glucose reduction effect and benefits for obesity are reported by clinical studies. Also cardiovascular benefits of some GLP-1 analogues have been reported. DPP-IV inhibitors including sitagliptin, saxagliptin and vildagliptin are generally orally administered in order to prevent incretin degredation by DPP-IV protease. Both those

antidiabetic medication groups are usually well tolerated but still have risks with long-term use [21-24].

Exogenous insulin injection is the primary treatment against T2DM but it can cause many negative consequences over time. Comparing to endogenous insulin, injection from outside can lead a different insulin gradient since endogenously secreted insulin is taken by liver firstly and less than half can reach the peripheral tissue. Exogenously administered insulin cannot be taken by liver at first, it must circulate through the peripheral tissue before the liver and this situation causes peripheral hyperinsulinemia [25]. At the same time, excessively decreased blood glucose levels lead to hypoglycemia which can be life-threatening. Glucose uptake is promoted by highly concentrated insulin leading to weight gain and insulin resistance. Another negative impact of exogenous insulin uptake is increasing cardiovascular risks by promoting lipid synthesis together with inflammatory processes which is resulted in atherogenesis. Additionally, frequent injections can stimulate complications like lipodystrophy which is defined by unusual fat accumulation. Apart from physical impacts, the need of daily insulin injection may cause stress, reduced adherence and decreased quality of life which are important social/psychological impacts [26, 27].

1.3. GLP-1 Mechanism and Structure

Glucagon-like peptide-1 (GLP-1) is an incretin hormone that primarily takes role on glucose homeostasis. Secretion of GLP-1 is carried out by enteroendocrine L cells which is located in the mucosa of the distal part of the small intestine and colon [28]. The GLP-1 receptor (GLP-1R), which belongs to G-protein-coupled receptor (GPCR) family, provides an affinity for GLP-1 molecule and its effect is exerted in response to carbohydrate and fat ingestion. GLP-1R is expressed in different tissues like pancreas, brain, heart and gastrointestinal tract whose cells can be target for GLP-1 [29].

Mechanism of action of GLP-1 is started by nutrient intake which causes release of GLP-1 into the bloodstream and its binding with GLP-1 receptors on panctreatic β cells. Upon this binding, G_s alpha subunit of GLP-1R activates adenylate cyclase, leading to elevated cyclic AMP (cAMP). Increased cAMP concentration triggers the activation of protein kinase A (PKA) resulted in raised transcription of insulin gene and insulin secretion in those cells (Figure 1.1.). The risk of hypoglycemia is minimized by this insulinotropic

effect in a glucose-dependent manner when blood glucose level is above normal level. Additionally, glucagon release from pancreatic α cells is inhibited by GLP-1 action as supportive contribution to the decreased glucose level in blood [30,31].

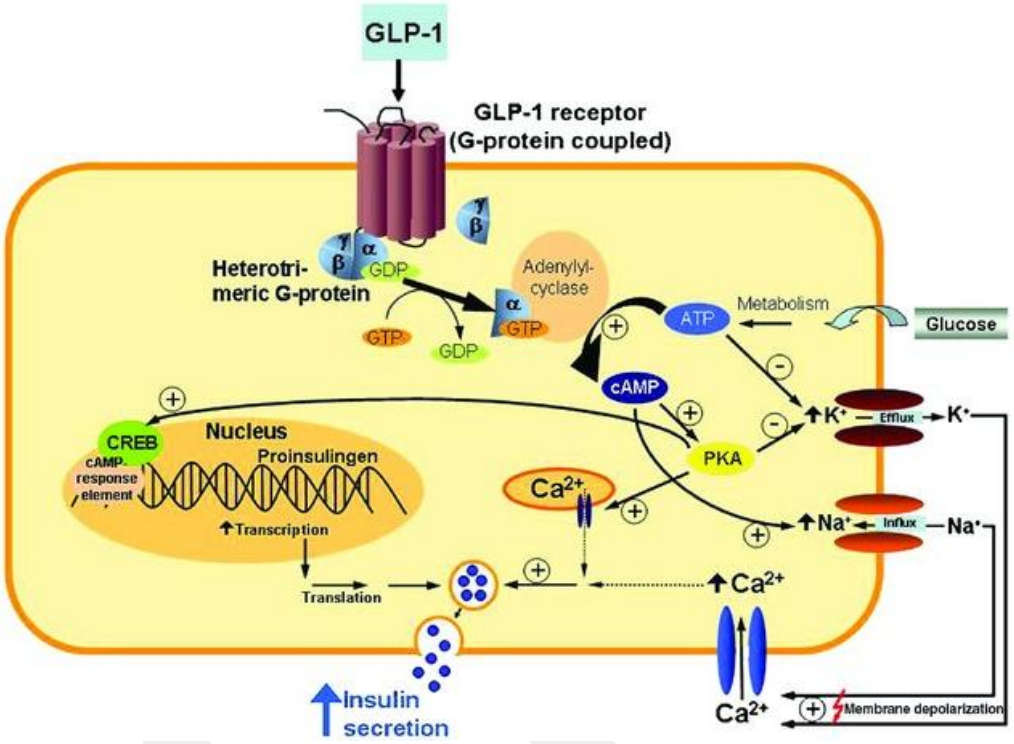

Figure 1.1. Mechanism of action of GLP-1 [31].

Structure of GLP-1 peptide is composed of 30 amino acid residue derived from the proglucagon (GCG*)*. There are two biologically active GLP-1 forms whose primary sequences are GLP-1 (7-37) and GLP-1(7-36) amide and the second one is predominant form in humans [32]. GLP-1 structure is characterized by alpha-helix arrangement which is crucial for its binding to GLP-1 receptor [33]. GLP-1R from class B GPCR family has a wide extracellular N-terminal domain which is required for ligand attachment. Intracellular signalling pathway is activated by conformational change caused by binding of GLP-1 to the receptor as it was mentioned (Figure 1.1.). The hormone-receptor complex stability is predominantly determined by the binding interface between GLP-1 and GLP-1R [34].

1.4. DPP-IV Protease Enzyme as GLP-1 Inhibitor

DPP-IV is an aminopeptidase enzyme consisting of 766 amino acids. Crystal structure of DPP-IV is composed of a tetramer and each subunit of that tetramer contains two structural N-terminal β propeller domain and C-terminal catalytic domain [35]. Primary action of DPP-IV is known as cleavage of post-proline bonds which are located downstream of the N-terminus of the proteins. Cleavage site is preferred as X-Pro or X-Ala, where X can be any residue differing from proline. Generally, it is stated that enzymatic activity requires glycosylation of an integral protein. Controversially, DPP IV is an exception in that case [36]. *In vitro* studies showed that the variety of candidate DPP-IV substrates such as growth hormone-releasing hormone, bradykinin, particular chemokines, neuropeptide Y, eotaxin, gastric inhibitory peptide (GIP), prolactin identified previously but only some of them, GLP-1 and GIP, are defined as endogenous substrates. Thus, GLP-1 is degraded by DPP-IV significantly because of being an endogenous substrate of the enzyme [37]. That's why GLP-1 analogues against DPP-IV are developed considering DPP-IV cleavage site in native GLP-1.

1.5. Production of GLP-1 Analogues Against T2DM

GLP-1 protein discovery dates back to the early 1980s and it has gained significant attention as therapeutic against diabetes [38]. Unlike other antidiabetics, the insulinotropic effect of GLP-1 has a self-limiting mechanism as it diminishes when the blood glucose concentration turns back to the normal range. Thus, hypoglycemia risk is minimized by GLP-1 therapeutics [39]. Additionally, GLP-1 blocks gastric emptying and induces satiety resulted in reduction in body weight. The cardioprotective effects of GLP-1 therapy are also reported by clinical studies [28]. All of these physiological benefits of GLP-1 made it a promising therapeutic agent against T2DM. On the other hand, the half-life ($t_{½}$) of the endogenous GLP-1 is shortened to 1-2 minutes by proteases like DPP-IV with its rapid metabolic degradative mechanism [40]. The cleavage site of DPP-IV is Ala_8-Glu_9 which is resulted in GLP-1(9–36)-NH_2 metabolite that is known having 100-fold lower binding capacity compared to native peptide to GLP-1 receptor [41].

GLP-1 analogues have a longer half-life due to the amino acid modifications and their increased resistance to DPP-IV compared to GLP-1. Additionally, increased receptor binding affinity is another significant advantage of the analogues [33]. Among the currently approved GLP-1 analogue drugs, exenatide is the first incretin analogue approved for clinical use and is also the longest studied analogue so far. Approximately 50% of the amino acid sequence of exenatide is identical to human GLP-1 and has a similar affinity for GLP-1 receptor. Since its half-life (t½) is 2-4 hours, it is administered twice a day. Early studies showed that after 30 weeks of twice daily treatment improved fasting and postprandial glucose levels [42]. Another GLP-1 analogue, lixisenatide, has a six-fold lysine sequence at the C-terminus of the peptide and differs from the natural GLP-1 sequence by removing a proline. Approximately 55% of lixisenatide is bound to receptor protein and its half-life (t½) is 2-4 hours [43]. Another GLP-1 analogue, liraglutide, was obtained by replacing the serine amino acid in GLP-1 with arginine and attaching palmitoyl fatty acid to Lys_{26}. Thanks to the conjugated fatty acid, it can bind to serum albumin, providing resistance to degradation by DPP-IV [44]. Semaglutide, another GLP-1 analogue used against type II diabetes and obesity, is 94% similar to natural GLP-1. The amino acid alanine at position 8 was replaced by Aib (2-aminobutyric acid). As in liraglutide, the amino acid lysine at position 34 in native GLP-1 was replaced by arginine in semaglutide. Additionally, the lysine amino acid at position 26 was acylated to bind with C18 fatty diacid via a hydrophilic linker "gamma-glutamyl-2xOEG (γGlu-2xOEG)" consisting of gamma-glutamic acid and two units of 8-amino-3,6-dioxaoctanoic acid (OEG). This modification increased albumin binding capacity, extending its systemic half-life and reducing renal clearance [45]. In order to overcome disadvantage of rapid clearance of peptide based drugs, fusion proteins have been developed and longer plasma circulation was provided. Albiglutide is a GLP-1 analogue harboring human serum albumin fusion protein attached to dimer of the modified GLP-1 sequence. It has gained longer plasma half-life by albumin and resistance to DPP-IV by replacing the Ala8 with Gly [76]. Lastly, tandem repeats of modified GLP-1 analogues also showed resistance against DPP-IV and enhanced stability. Positive glucoregulatory effect is exerted by oral glucose tolerance test. Thus, significant glucose lowering effect of GLP-1 tandem construct peptides was reported [46]. Another example of tandem repeated analogues of GLP-1 is that expression of fusion

protein containing 10 tandem repeated GLP-1 analogue in yeast was studied. The produced peptide was injected into diabetic rat models and lowered level of serum glucose significantly [77]. According to all these GLP-1 mimicking therapeutics, a GLP-1 analogue consisting of modified amino acid sequence in a repeated tandem form produced in the study is shown below.

Native GLP-1Amino Acid sequence:

$H^{7}AEGT^{11}FTSDV^{16}SSYLE^{21}GQAAK^{26}EFIAW^{31}LVKGR^{36}$-$NH_2$

GLP-1 Analogue-1:

$[H^{7}\mathbf{\underline{G}}EGT^{11}FTSDV^{16}SSYLE^{21}GQAAK^{26}EFIAW^{31}LV\mathbf{\underline{R}}GR^{36}\mathbf{\underline{G}}]_2$

GLP-1 Analogue-2: :

$[H^{7}\mathbf{\underline{G}}EGT^{11}FTSD\mathbf{\underline{L}}^{16}S\mathbf{\underline{KQM}}E^{21}\mathbf{\underline{EE}}A\mathbf{\underline{VR}}^{26}\mathbf{\underline{L}}FI\mathbf{\underline{E}}W^{31}L\mathbf{\underline{KD}}G\mathbf{\underline{G}}^{36}\mathbf{\underline{PSSGAPPPS}}]_2$

GLP-1 Analogue-3:

$[H^{7}\mathbf{\underline{G}}EGT^{11}FTSDV^{16}SSYLE^{21}GQAAK^{26}EFIAW^{31}L\mathbf{\underline{C}}KGR^{36}\mathbf{\underline{GGGGGGCGGGGGGGGG}}]_2$

Lixisenatide Amino Acid Sequence:

$[H^{7}\mathbf{\underline{G}}EGT^{11}FTSD\mathbf{\underline{L}}^{16}S\mathbf{\underline{KQM}}E^{21}\mathbf{\underline{EE}}A\mathbf{\underline{VR}}^{26}\mathbf{\underline{L}}FI\mathbf{\underline{E}}W^{31}L\mathbf{\underline{KN}}G\mathbf{\underline{G}}^{36}\mathbf{\underline{PSSGAPPSKKKKKK}}$-NH2

Albiglutide Amino Acid Sequence:

$[H^{7}\mathbf{\underline{G}}EGT^{11}FTSDV^{16}SSYLE^{21}GQAAK^{26}EFIAW^{31}LVKGR^{36}]_2$ - Albumin

GLP-1 Tandem Repeat [46]:

$[H^{7}\mathbf{\underline{S}}EGT^{11}FTSDV^{16}SSYLE^{21}GQAA\mathbf{\underline{Q}}^{26}EFIAW^{31}LVNGR^{36}]_6$

Native GLP-1, GLP-1 analogues to be produced in this study, commercial GLP-1 analogues and GLP-1 tandem repeat construct from the literature can be seen above. The bold-underlined residues are differing from the native human GLP-1. The analogue-1 peptide sequence is very similar to the currently marketed FDA-approved recombinant drug liraglutide. As a difference, liraglutide contains a palmitic acid group attached to the 26th amino acid in its structure. Substitution of Ala_8 for Gly, which takes place in all analogues, circumvented the DPP-IV degredative action resulted in longer half-life [47]. Substitution of Lys_{34} for Arg causes a two-fold reduction in the half maximal effective concentration (EC50) [48]. GLP-1 analogue-2 is dimeric form of antidiabetic drug exenatide peptide

sequence. It has amino acid replacements like Gly_{22} for Glu and Lys_{34} for Arg which lead to an increased affinity and activity [49, 50]. Also, it includes Tyr_{19} for Gln due to hydrogen bond interaction between GLP-1 and GLP-1R at this site and Gln is an alternative for Tyr in this manner [48]. Analogue-3 is constructed based on commercial albiglutide drug peptide sequence with slight differences. Ala_8 for Gly substitution takes place as it is seen in other analogues due to increased resistance against DPP-IV. Additionally, Val_{33} for Cys and addition of a Gly_{tail} with a Cys at the middle is a modification that is provided in purpose of forming intramolecular disulfide bonds for increased stability and longer half-life proven in animal experiments [51]. All taken together, dimeric forms of constructed 3 different GLP-1 analogues in a repeated tandem is determined as target antidiabetic peptide to be produced in this study.

1.6. *E. coli* as Expression System

Recombinant production is possible with bioproduction platforms, and some host organisms are used as platforms. It is known that the most commonly used host microorganisms are *Escherichia coli (E. coli)* as bacterial cells, *Saccharomyces cerevisiae (S. cerevisiae)* and *Pichia pastoris (P. pastoris)* as yeast cells [52]. Because they are frequently used, these host organisms have benefits such as lower cost and practicality in accessing resources, abundance of findings within the framework of recombinant protein production in the literature, successful efficiency in the products obtained, etc. [53]. Among technical advantages of using *E. coli* as expression system, high growth rate can be counted because of 20 minutes doubling time under optimum conditions allowing quick biomass accumulation and production of target molecules. When using strong promoters with an appropriate plasmid, recombinant protein can be produced in large amounts. Also, the advantages of scalability and efficient transformation offer short production time [54]. Additionally, it was reported that there are more than 100 recombinant protein products in biopharmaceutical market of US and Europe in 2012 and 34% of them are produced in *E. coli* recombinantly [55]. According to recombinant protein producing studies in the literature, *E. coli* was found to be a proper expression system for GLP-1 protein. Gao et al. stated that the expression and purification of a GLP-1 analogue was successfully performed by using *E. coli* and 1.58 mg of modified GLP-1 with the purity of up to 98% was obtained

from 1 g wet bacterial cells. Also, a significant dose-dependent glucose-lowering activity was reported in the bioactivity tests [56]. Yuliawati et al. reported that recombinant GLP-1 as a tandem multicopy of the GLP-1 encoding sequence, just like in this thesis, was expressed by optimizing codon for *E. coli.* Recombinant eight multicopy GLP-1 protein was successfully produced and purified by affinity chromatography [57].

This thesis focuses to produce a recombinant GLP-1 analogue using *E. coli* as host microorganism to develop a biotechnological therapeutic agent against T2DM. The thesis was aimed to optimize protein expression conditions for maximizing the yield as much as possible. The purpose of the whole project from wider perspective is to overcome the drawbacks of current T2DM therapeutics by biotechnologically produced GLP-1 analogue.

Chapter 2

Materials and Method

2.1. Materials

2.1.1. Strains, Plasmids and Media

E. coli TOP10 was used for cloning purposes, while *E. coli* strains BL21(DE3)pLysS, Origami2(DE3), BL21-CodonPlus(DE3)-RIL, BL21(DE3), BL21 Star(DE3) were used for the expression of the GLP-1 analogue peptide as tandem repeat (Table 2.1.). Bacterial cells were grown in LB media composed of 1% tryptone, 0.5% east extract, 1% NaCl with the supplementation of 50 µg/mL of kanamycin to provide selectivity. pUC57 was the cloning plasmid, while pET28a (+) was used for expression (Table 2.2.).

Table 2.1. *E. coli* strains used in the study [63].

Strain	Description	Source
E.coli TOP10	F^-; *mcrA; Δ(mrr-hsdRMS-mcrBC); φ80lacZΔM15; ΔlacX74; nupG; recA1; araD139; Δ(ara-leu)7697; galE15(GalS); galK16; rpsL(*Str^R*); endA1;* λ^-	Lab Stock
E.coli BL21(DE3) pLysS	F^-; *ompT; gal; dcm;* $hsdS_B(r_B^- m_B^-)$; *λ(DE3);rne131;pLysS[T7p20* ori_{p15A}*](*Cm^R*)*	Lab Stock

E.coli Origami2(DE3)	*Δ(ara-leu)7697; ΔlacX74; ΔphoA;* PvuII*; phoR;araD139; ahpC; galE; galK; rpsL(Str^R); F'[lac$^+$ lacI^q pro];(DE3);gor522::Tn10;trx B;pLysS[T7p20 ori$_{p15A}$](Cm^R)*	Lab Stock
E.coli BL21-CodonPlus (DE3)-RIL	F$^-$; *ompT; hsdS$_B$(r$_B^-$ m$_B^-$); dcm$^+$; Tet^R; gal; (DE3); endA; Hte; [argU ileY leuW Cam^R];[argU ileY leuW Strep/Spec^R]*	Lab Stock
E.coli BL21(DE3)	F$^-$; *ompT; gal; dcm; lon; hsdS$_B$(r$_B^-$ m$_B^-$);λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]); [malB$^+$]$_{K-12}$(λ^S)*	Lab Stock
E. coli BL21 Star(DE3)	F$^-$; *ompT; gal; dcm; hsdS$_B$(r$_B^-$ m$_B^-$); rne131;* (DE3)	Lab Stock
E.coli TOP10 / pBAT1	*E. coli* TOP10 harboring pBAT1 plasmid	This study
E.coli BL21(DE3) pLysS/pBAT1	*E. coli* BL21(DE3) pLysS harboring pBAT1 plasmid to express GLP-1 Analogue	This study
E.coli Origami2(DE3)/pBAT1	*E. coli* Origami2(DE3) harboring pBAT1 plasmid to express GLP-1 Analogue	This study
E.coli BL21-CodonPlus (DE3)-RIL / pBAT1	*E. coli* BL21-CodonPlus (DE3)-RIL harboring pBAT1 plasmid to express GLP-1 Analogue 1	This study
E. coli BL21(DE3) / pBAT1	*E. coli* BL21(DE3) harboring pBAT1 plasmid to express GLP-1 Analogue 1	This study
E. coli BL21 Star(DE3) / pBAT1	*E. coli* BL21 Star(DE3) harboring pBAT1 plasmid to express GLP-1 Analogue 1	This study

Table 2.2. Plasmids used in the study

Plasmids	**Features**
pET-28a (+)	N-terminal His•Tag/thrombin/T7•Tag configuration

	plus a C-terminal His•Tag sequence [64]
pBAT1	pET-28a (+) carrying GLP-1 Tandem inserted between BamHI and HindIII sites
pUC57	A derivative of pUC19. Multiple cloning site (MCS) contains 6 restriction sites with protruding 3'-ends that resistant to *E. coli* exonuclease III [65]
pUC57-Albumin and GLP-1 Tandem	pUC57 plasmid carrying Albumin-GLP1 Tandem insert between EcoRI and BamHI

2.2. Plasmid Construction

Plasmid construction was started with double digestion of expression plasmid (pET-28a(+)) with BamHI and HindIII restriction enzymes to get backbone for following assembly. Thus, linearized vector was obtained for cloning. Restriction digestion reaction components and their amounts are listed in Table 2.3. All components were added into eppendorf tube and the tube was incubated for 6 hours in a water bath set at 37 °C. The GLP-1 Tandem Analogue coding DNA fragment was amplified by Touchdown-PCR reaction in order to be used as insert in cloning. Template used in the PCR reaction was GLP-1 Tandem (GLP-1 analogue designed for this study) carrying pUC57 plasmid ordered from Gene Universal Inc. Primers were designed to obtain PCR products with overhang regions and restriction sites for Gibson Assembly reaction (Table 2.4.) and were ordered from Oligomer Biyoteknoloji Co. PCR components and conditions are listed in Table 2.5 and Table 2.6. respectively. Since the target DNA fragment is relatively short (around 700 bp), the annealing temperature and PCR conditions were optimized to obtain the brightest target DNA bands on the gel. Optimized PCR conditions are listed in Table 2.6.

Table 2.3. Restriction digestion components

Materials	Amounts
10 X Tango Buffer	2 µl
BamHI	0.3 µl
HindIII	0.3 µl
Template (pET-28a(+) plasmid DNA)	4 µl
dH2O	13.6 µl

Table 2.4. Primer set of GLP-1 Tandem

Oligo 1 (BamHI) Forward Primer	AATGGGTCGCGGATCCCATGGTGAAGGTACATTCACATCTG
Oligo 2 (HindIII) Reverse Primer	GTGCGGCCGCAAGCTTACCACCACCACCACCACC

Table 2.5. PCR components of GLP-1 Tandem

Materials	Amounts
5 X Phusion Buffer	4 µl
Template (pUC57-GLP-1 Tandem)	0.4 µl
dNTP(10 mM)	0.4 µl
DMSO	0.4 µl
Phusion Polymerase	0.3 µl
Primers (Forward and Reverse)	0.4 µl for each
dH2O	14.1 µl

Table 2.6. PCR conditions of GLP-1 Tandem

Step	Temperature (°C)	Duration	Cycles
Denaturing	98	1 min.	1
Amplification	98	10 sec.	36
	80 to 70	30 sec.	
	72	30 sec.	
Elongation	72	10 min.	1

Both restriction digestion and PCR products (20 µl) were loaded into 1% agarose gel and it was run for 45 minutes at 100 V. The gel was screened by gel imager (BIO-RAD Molecular Imager® Gel Doc™ XR System). Gel pieces from both restriction digestion and PCR were extracted by using commercial kit (FavorPrep™ GEL/PCR Purification Kit). Instruction of the producer was followed during the protocol. Briefly, interested gel pieces were sliced and lysed, DNA binding to column was provided and washed. After that, column matrix were dried and attached DNA was eluted by elution buffer. After the gel extraction, 3 µl of each elute was loaded into 1% agarose gel and run for 45 minutes at 100 V in order to confirm gel size and purity. Purified restriction digestion and PCR products were fused by Gibson Assembly Reaction technique. DNA concentrations were measured and added into the Master mixture by following protocol provided by the producer (Takara In-Fusion® Snap Assembly Master Mix). The reaction mixture composed of 1 µl of insert DNA fragment, 8 µl of vector DNA fragment, 1 µl of 5X In-Fusion Snap Assembly Master Mix and the final mixture in a PCR tube was incubated at 50 °C for 15 minutes. Thus, the construction of the expression plasmid for GLP-1 analogue (named as pBAT1) was completed. The construct plasmid pBAT1 was transformed into *E.coli* TOP10 strain followed by isolation and the confirmation with restriction digestion as well as sequencing. Illustration of the construct plasmids are shown in Figure 2.1. T7 RNA Polymerase in the pBAT1 provides the formation of RNA from DNA followed by protein synthesis. His-tag in the pBAT1 provided the purification of target protein after production.

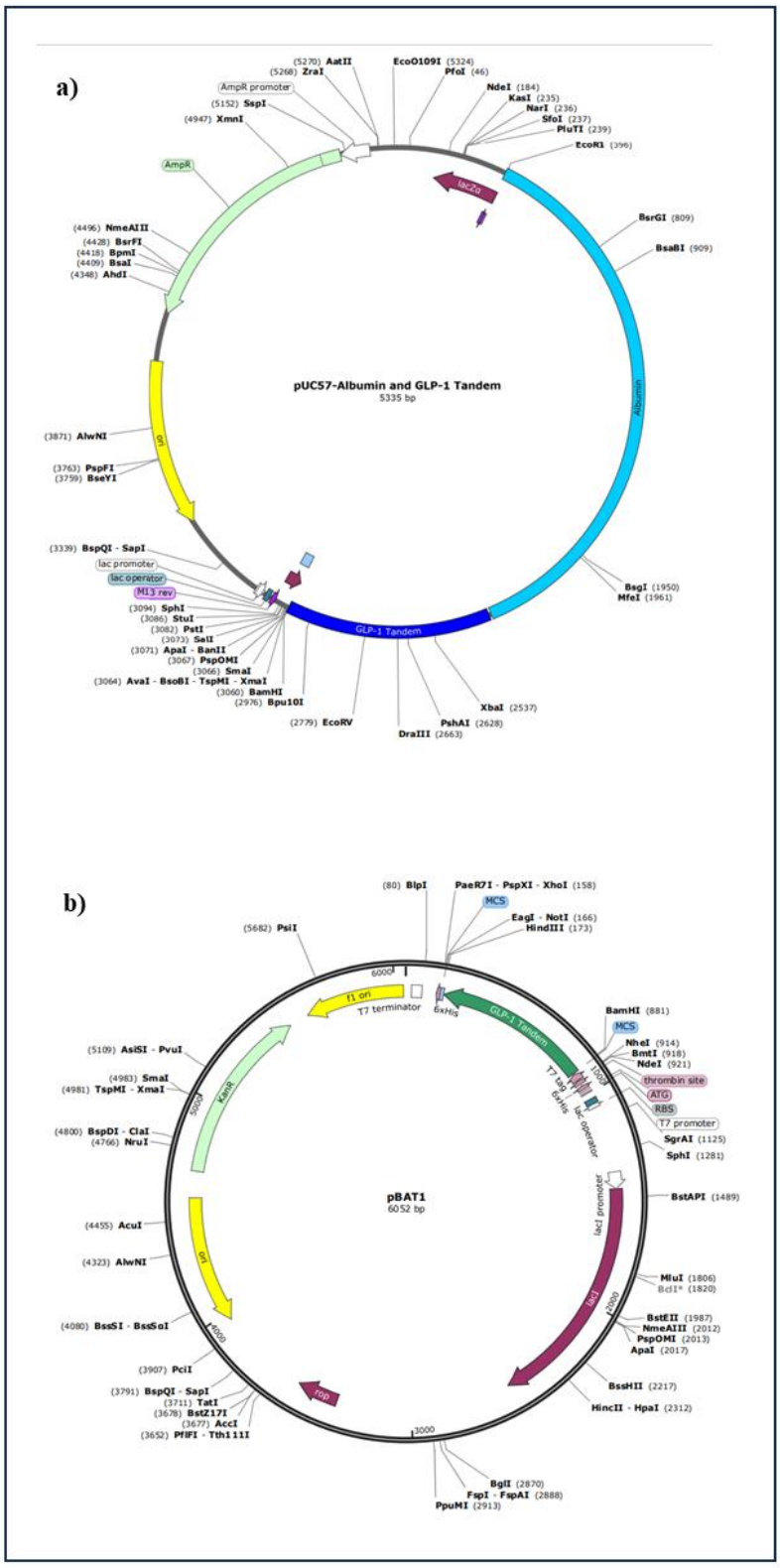

Figure 2.1. Plasmid maps visualized by SnapGene software. (a) pUC57 harboring gene coding for albumin-GLP-1 Tandem (designed and synthesized). (b) pBAT1 expression plasmid harboring the gene for GLP-1 Tandem in pET-28a(+).

2.3. Transformation

First of all, *E. coli* TOP10 strain was made competent by $CaCl_2$ treatment. The protocol was as follows: The liquid culture of *E. coli* TOP10 strain was incubated at 37 °C overnight to grow. Then, it was inoculated to 50 ml flask and incubated at 37 °C until OD_{600} reached to 0.35-0.40. The cells were collected by centrifugation at 6,000 rpm for 8 minutes at 4 °C. Cells were resuspended by ice-cold 10 ml 50 mM $CaCl_2$ and incubated on ice for 20 minutes. After incubation, cells were collected by centrifugation as in previous step. Finally, cells were resuspended by 50 mM $CaCl_2$ containing 10% glycerol and aliquoted (150 µL) to be stored at -80 °C for further use. Plasmid was up-taken into bacterial competent cells by heat-shock transformation method. Firstly, total assembly product (10 µl) was added into 100 µl competent *E. coli* TOP10 cell and stayed on ice for 30 minutes. After that, the tube was placed into water bath at 42 °C and incubated for 1 minute, then placed back to ice immediately. 1 mL LB was added onto cells and incubated at 37 °C for 30 minutes. Finally, the mixture in the tube was poured onto LB agar (containing 50 µg/ml kanamycin) plate and spreaded. It was kept until it dried and incubated at 37 °C overnight. All these steps were made in biocabinet in order to prevent contamination.

2.4. Plasmid Isolation and Confirmation

Because of the selective antibiotic in the plate, colonies that have grown were considered successful transformants. However, isolating and screening plasmids were required for the confirmation. For that purpose, manual plasmid extraction was performed without a commercial kit. This method utilizes DNA precipitation without column. Instead, isopropanol was used to precipitate DNA. The protocol was performed as follows: Firstly, 1.5 ml of culture was added to 1.5 ml eppendorf tube and centrifuged at 16,000 g for 2 minutes to pellet cells. Supernatant was discarded and this step was repeated three times because of 5 ml starting culture. After that, 250 µl of P1 buffer (50 mM Tris-HCl, 10 mM EDTA, 100 µg/ml RNAse I) was used to resuspend cells followed by addition of 250 µl of P2 buffer (8 gr of NaOH, 50 ml of 20% SDS in 950 ml dH2O) and inverted 4-6 times to

mix. Finally, 350 µl of P3 buffer was added and the tube was inverted immediately. Lysis of the cells was completed with these steps. The cloudy solution was centrifuged at 16,000 g for 10 minutes to pellet cell debris. 700 µl of the supernatant was transferred to a tube containing 700 µl of isopropanol and cooled on ice for 15 minutes followed by centrifugation at 16,000 g for 10 minutes to pellet plasmid DNA. After that, 350 µl of chilled 70% ethanol was added to resuspend pellet and spinned at 16,000 g for 5 minutes. The remaining supernatant was removed and pellet was dried via air for 20 minutes. Finally, the pellet was resuspended in 20 µl of TE buffer to be used in downstream processes. After that, the identification of isolated plasmids was provided by gel electrophoresis. 3 µl of isolates was loaded into 1% agarose gel and it was run for 45 minutes at 100 V. The gel was screened by gel imager. Secondary confirmation of the constructed plasmid is performed through the restriction digestion with EcoRV enzyme. Because, EcoRV cuts pET-28a(+) from one point (at 1573th bp position) without any manipulation. The insert DNA fragment has EcoRV recognition site (5'-GATATC-3') at 427th bp position. Thus, the restriction digestion of the correctly constructed plasmid by EcoRV must give a two DNA pieces as around 1800 bp and around 4200 bp.

As the most reliable demonstration, constructed plasmid was sent out for sequencing. Whole plasmid sequencing was performed by GetGenome Laboratory (Norwich, UK). Sequence alignment was performed to ensure the succesful insertion of DNA of interest into the expression plasmid without any mutations. This confirmed plasmid was named as pBAT1(pET28a-GLP1 Tandem), which will be used this way in following sections.

2.5. Protein Expression

Protein producing culture requires an expression strain rather than *E. coli* TOP10 which is widely used for cloning purposes. Heat-shock transformation was performed by using 3 µl of isolated pBAT1 and competent *E. coli* BL21(DE3)cells. Competent *E. coli* BL21(DE3) cells was prepared as described in section 2.3 and transformation was also performed as described in the same section. *E. coli* BL21(DE3) harboring pBAT1 was inoculated to 50 ml LB shake-flask, containing 50 µg/ml Kanamycin, as 1% and incubated at 37 °C until OD reached to 0.6-0.8 at 600 nm absorbance measured by spectrophotometer (GENESYS™ 10S UV-Visible Spectrophotometer, Thermo Scientific). After that, 0.2 mM

isopropyl β- d-1-thiogalactopyranoside (IPTG) was added to culture and incubated at 18 °C overnight. In order to check expression of the target protein, SDS-PAGE analysis was performed as follows: Firstly, cells were precipitated at 8,000 rpm for 10 minutes at 4 °C. 5 ml of lysis buffer containing 20 mM Tris-HCl, 0.5 M NaCl, 20 mM imidazole and 0.2 mM phenylmethylsulfonyl fluoride (PMSF) was added to pellet and resuspended on ice. After that, cells were disrupted by sonicator (Bandelin Sonopuls HD 4400). Sonication was performed at 60% Amp. for 10 minutes with 30 seconds breaks after each 30 seconds sonication on ice. 200 µl of the lysate was centrifuged at 16,000 rpm at 4 °C and 80 µl of supernatant was taken as soluble fraction. 200 µl of 8M urea was added to pellet and resuspended on ice followed by centrifugation at 16,000 rpm at 4 °C. This time, 80 µl of supernatant was taken as insoluble fraction and 20 µl of 5X protein loading dye containing 10% sodium dodecyl sulfate (SDS), 10 mM dithiothreitol(DTT), 20% glycerol, 0.2 M Tris-HCl, 0.05% bromophenol blue was added to both fractions. Finally, the tube containing proteins and loading dye were place in a heat rock at 95 °C for 10 minutes and stayed at room temperature for 1-2 minutes to cool down. At that point, samples were ready to be loaded into gel. SDS-PAGE resolving gel was prepared as 18% because expected light molecular weight (25 kDa) of the target protein and the stacking gel was prepared as 5%. After all the samples were loaded, the gel was run at 70 V for 35 minutes followed by 200 V for 110 minutes by using power supply (CleaverScientific, PowerPRO 300 Power Supply) and electrophoresis tank (Biocomdirect, OMNIPAGE MINI, VS10DSYS). Protein bands were imaged by using imager (Bio-Rad ChemiDoc™ MP Imaging System).

The expression of the protein of interest was tried to be maximized by changing culture conditions. Parameters including different expression strains, IPTG concentrations, induction OD values and post-induction temperatures were tested. Protein concentration was measured by comparing SDS-PAGE results considering band thickness before purification process. The most promising culture and lysis condition were selected and used for the purification of the target protein. During this period, SDS-PAGE protocol was also optimized. Because of various parameters like lysis conditions, boiling conditions, the gel concentration, running duration, running voltage etc. the proteins on the gel image were not clear in all the experiments. That's why the SDS-PAGE protocol was created as the final version after modifications.

2.6. Protein Purification and Quantification

The purification of the target protein was conducted via His-tag affinity chromatograpy and the measurement of purified protein concentration was performed by Bradford Protein Assay. Purification protocol was performed as follows: 5 ml of fresh lysate was prepared as same with previous section. Since the target protein was expressed as insoluble, 2.5 ml of insoluble fraction was taken as same with the method in previous section. At the same time, 300 µl of the resin (Thermo Scientific™ HisPur™ Ni-NTA Resin) was taken and they were transferred into an eppendorf tube. The tube was incubated on shaker (HEIDOLPH Duomax 1030 Platform Shaker) at 50 rpm at 4 °C overnight to provide attachment of His-tagged target protein with nickel resin. After the incubation, the solution was centrifuged at 700 rpm for 2 minutes at 4 °C. 80 µl of supernatant was taken as flow-through fraction in order to be loaded to gel. After removing the whole flow-though fraction carefully without disrupting the resin, 1 ml of the wash buffer (0.05 M NaH_2PO_4, 0.3 M NaCl and 5 mM imidazole) was added to resin pellet and the tube was gently inverted 8-10 times followed by centrifugation at 700 rpm for 2 minutes at 4 °C. The supernatant was taken was Wash-1 fraction for SDS-PAGE analysis. After removing the wash buffer completely, the washing step was repeated twice and totally three washing fractions were kept. Next, 1 ml of elution buffer (0.025 M NaH_2PO_4, 0.15 M NaCl and 500 mM imidazole) was added to resin pellet and the same inverting and centrifugation steps were performed. Finally, three wash; three elution fractions were obtained to load to the gel.

Protein concentration was determined by Bradford Protein Assay. The protocol was performed according to manufacturer's (Pierce™ Bradford Protein Assay Kit) instructions. The measurement was conducted with Elution-1 fraction from purification protocol. Since its imidazole concentration was 500 mM which is not compatible with Bradford reagent, it was diluted as 1:10. Also, the elution buffer was diluted with same ratio to be used as diluent in the protocol. After that, BSA standards and the target protein were prepared by given protocol and measured at 595 nm by plate reader (Varioskan LUX Multimode Microplate Reader, Thermo Scientific).

Chapter 3

Results

3.1. pBAT1 Construction

Constructing a plasmid requires linearized vector backbone at first. For that purpose, pET-28a(+) was digested with BamHI and HindIII restriction enzymes. Linearized pET-28a(+) is expected to be 5369 bp length on the gel which is shown in Figure 3.1. (a). The amplification of the DNA fragment coding for GLP-1 tandem was conducted with PCR. The target fragment is expected to be around 700 bp length. Its gel image can be seen in Figure 3.1. (b).

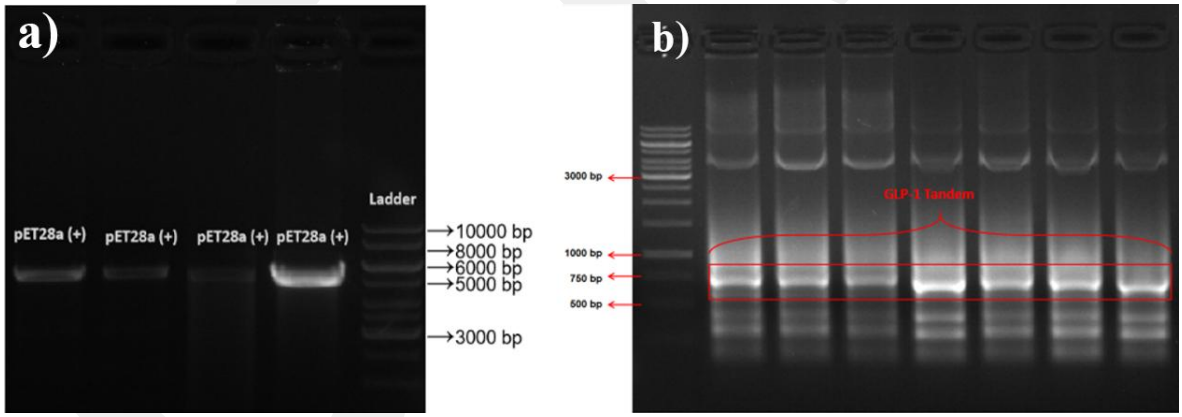

Figure 3.1. (a) Linearized pET-28a(+). (b) PCR product of GLP-1 Tandem.

Four separeted restriction digestion (pET28a(+)) tubes were prepared with exactly same components including restriction enzymes BamHI and HindIII. All of the bands are seen just above 5000 bp. The gel pieces containing linearized pET28a(+) were cut and stored for further steps. Amplification of GLP-1 tandem analogue coding DNA fragments was performed in seven separated PCR reactions with exactly same components. All of

them are seen just below the 750 bp length band of the ladder as it was expected. The gel pieces containing GLP-1 tandem were cut and stored for further steps.

Template plasmid (constructed pUC57-Albumin and GLP-1 Tandem) contains similar GLP-1 fragments as tandem. Amino acid sequence of the GLP-1 analogue in the tandem is shown in Chapter 1. Primers used in the reaction are complement to 5' of the 1st analogue fragment and 3' of the 3rd analogue fragment. However, 5' site of the analogues are very similar to each other and forward primer can bind to 5' of the 2nd and 3rd analogue. Thus, shorter PCR products than the target one can be obtained because of this similar sections. And the target longest tandem is relatively concentrated which stands on the top of among amplified fragments. So, the target GLP-1 triple tandem was clearly observed and shown in Figure 3.1. Linearized backbone and amplified insert DNA fragments were purified from the agarose gel for Gibson Assembly cloning (Figure 3.2.), in which the first line refers to purified GLP-1 Tandem DNA fragment as insert (just below 750 bp as it was expected) and the second line refers to purified linear pET28a(+) as backbone (just above 5000 bp as it was expected).

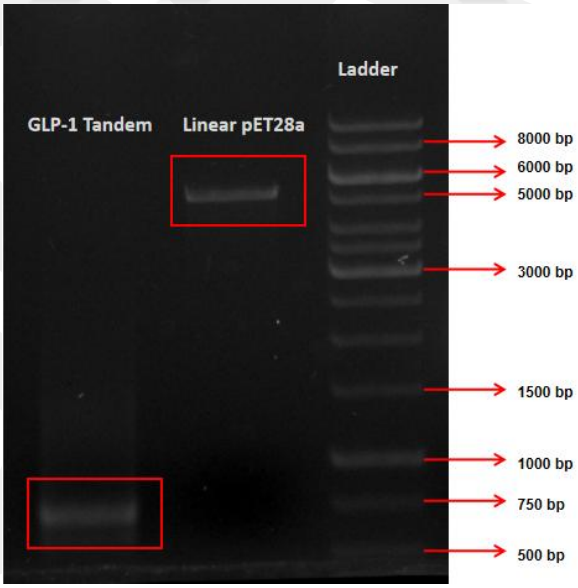

Figure 3.2. Purified pET-28a(+) vector and GLP-1 tandem insert.

Following Gibson Assembly with purified vector and insert, transformation to *E. coli* *E. coli* TOP10 competent cells took place. Colonies obtained from transformation were on a selective media as it contains Kanamycin antibiotic. The grown single colonies shown in Figure 3.3 (a) were accepted as successfully transformed cells but further confirmation was

required. Thus, plasmid isolation, confirmative restriction digestion and whole plasmid sequencing were performed. Plasmids were isolated from 5 separated colonies because of the possibility of unsuccessful transformant colony formation. Original pET28a(+) has Kanamycin resistance feature and the *E. coli* TOP10 cells took up pET28a(+) instead of pBAT1 were also able to grow on the plate. DNA bp length of the isolated plasmids are seen in Figure 3.3. (b). pBAT1 construct had linear pET28a(+) (5369 bp) as backbone and GLP-1 Tandem (678 bp) as insert. Thus, the constructed plasmid should be around 6000 bp length on the gel. However, it is seen around 4000 bp because it is in circular form instead of linear. So, it cannot be totally confirmed about DNA bp length without linearization. For that purpose, confirmative restriction digestion was performed by two restriction sites where insert DNA fragment locates between them and secondary confirmative restriction digest was performed by EcoRV only. Because original pET28a(+) has one EcoRV restriction site which should cut from that point only. However, GLP-1 Tandem DNA fragment has an EcoRV restriction site inside (between 427-432 nt in whole insert fragment) and the correct construct must be cut by EcoRV from two sites giving around a 4100 bp length fragment and around a 900 bp length fragment. These confirmative restriction digestion gel images are shown in Figure 3.4. As a final confirmation of the construct, whole plasmid sequencing result is shown in Figure 3.5. as aligned with target DNA sequence. It was demonstrated that constructed plasmid pBAT1 contains exactly the same nucleotides with designed GLP-1 Tandem coding DNA sequence.

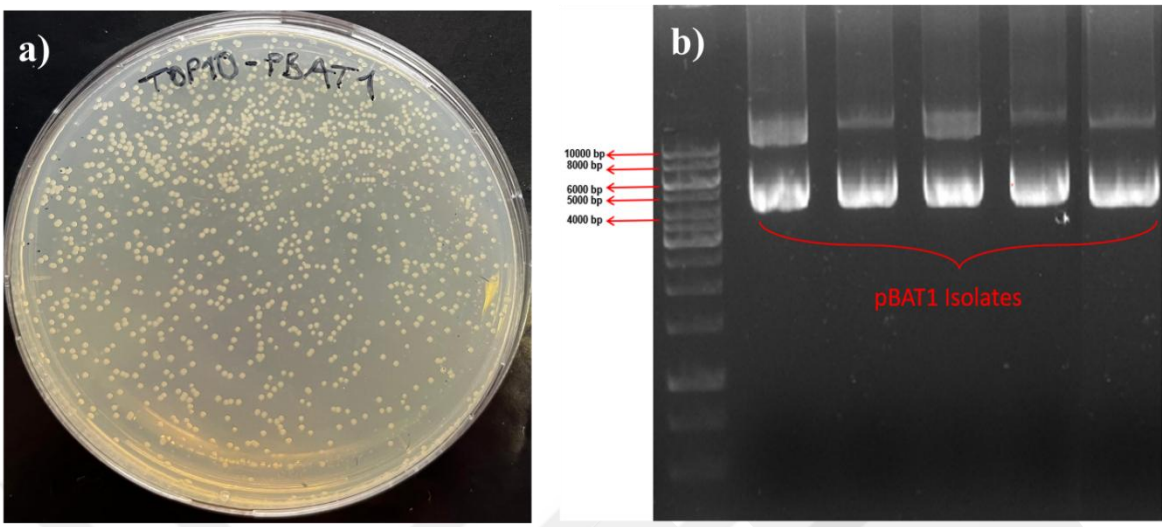

Figure 3.3. *E. coli* TOP10-pBAT1 colonies and pBAT1 isolates. (a) Grown TOP10 colonies on the Kanamycin containing LB agar were candidate to be successful transformants. There were no contamination observed considering morphology on the plate, (b) agarose gel electrophoresis imaging of isolated pBAT1 plasmids.

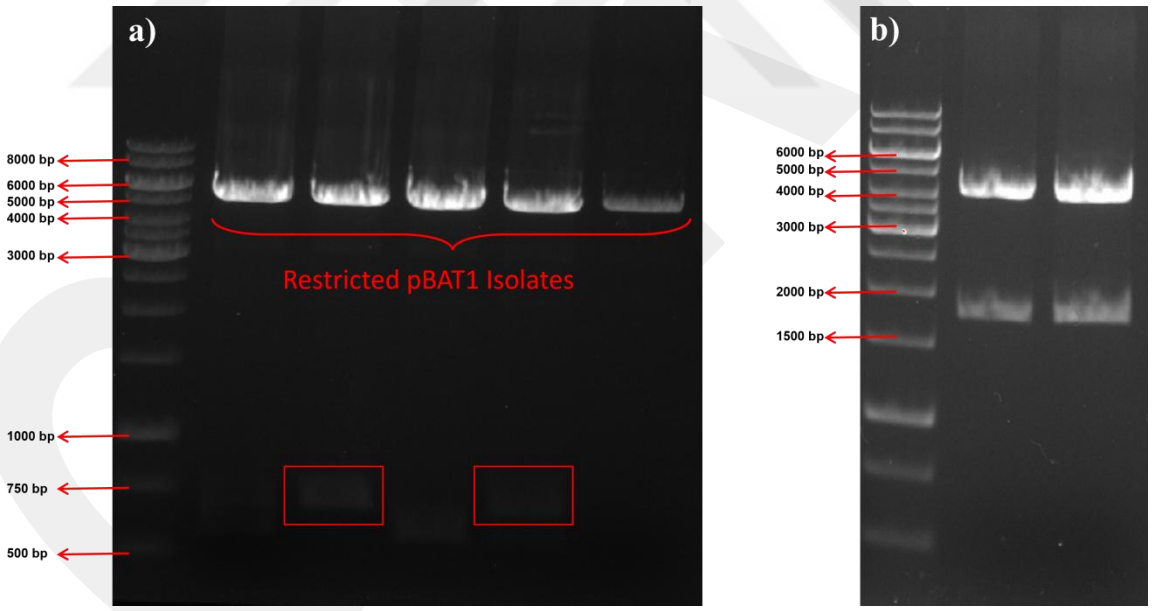

Figure 3.4. Confirmative restriction digestion of pET28a-GLP-1 Tandem isolates. (a) Correct construct's restriction products should be seen at around 5300 bp (pET28a(+)) and around 700 bp (GLP-1 Tandem) which are in red frames. (b) Further confirmation was performed on these two constructs (2nd and 4th lines) by restricting with EcoRV whose products should be seen around 4200 bp and around 1800 bp as it is in both constructs.

Figure 3.5. DNA sequence alignment. Whole plasmid sequencing was performed and the insert containing section was compared to target GLP-1 Tandem coding DNA sequence. First line refers to designed GLP-1 Tandem protein coding DNA sequence, second line refers to sequencing result of whole pBAT1 construct plasmid and matching nucleotides are shown by star at the third line.

3.2. Protein Expression and Culture Optimization

After verifying the construct pET28a-GLP1 Tandem, the isolated plasmid was transformed into *E. coli* BL21(DE3) strain for protein expression. *E. coli* BL21(DE3)-pBAT1 colonies are shown in Figure 3.6. Protein expression shake-flask (50 ml) culture induction and growth conditions were set as indicated in methodology section. Both soluble and insoluble proteins were analyzed by SDS-PAGE by comparing with control group which is *E. coli* BL21(DE3)-pET28a(+) without insert. According to molecular weight calculation by Expasy MW Tool, protein of interest should be seen at around 25 kDa on the gel. Resulting protein bands are shown in Figure 3.7. below.

Figure 3.6. *E. coli* BL21(DE3)-pBAT1 transformants. Colonies on the Kanamycin containing agar plate are bigger than *E. coli* TOP10 colonies and they grew up apart which confirm *E. coli* BL21(DE3) morphology.

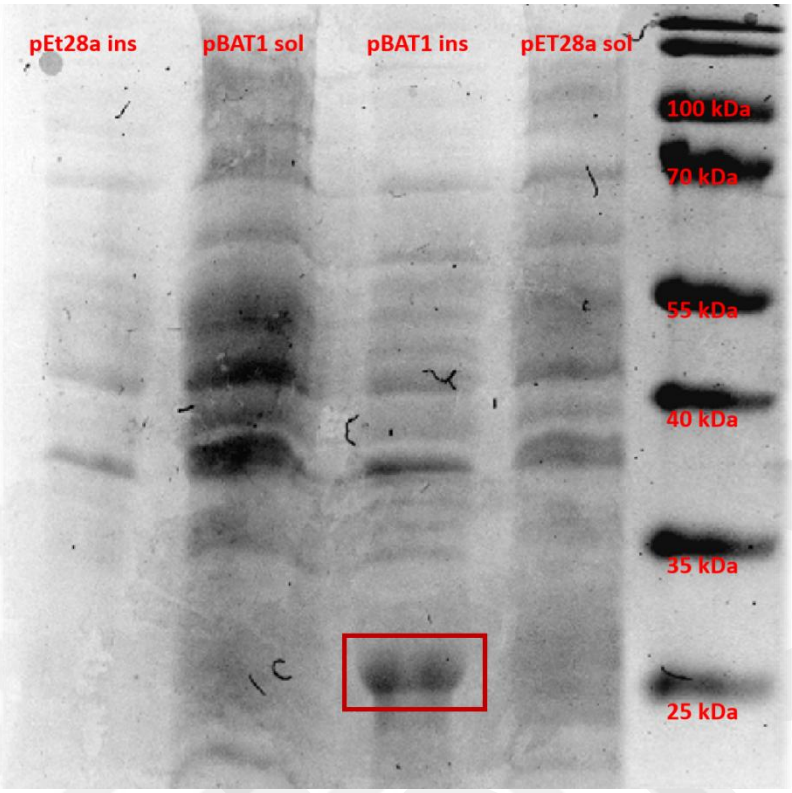

Figure 3.7. SDS-PAGE gel image of protein expression in *E. coli* BL-21(DE3). Target protein size is expected as around 25 kDa. Both soluble and insoluble fractions were analyzed, pET28a(+) bands were loaded as control. Red framed band is referred as GLP-1 Tandem protein whose control line (pET28a ins) has no any intense band at around 25 kDa.

It is the *E. coli* BL21(DE3)-pBAT1 culture which was induced with 0.2 mM IPTG and incubated at 18 °C overnight. It is clearly seen that target protein was expressed as insoluble in *E. coli* BL21(DE3)-pBAT1 culture. Also, different *E. coli* strains were tested together to check any positive effect on protein expression amount and solubility. Other parameters remained same as 0.2 mM IPTG induction when OD reached to 0.6-0.8 and 18 °C post-induction overnight growth. *E. coli* strains BL-21(DE3), Origami (tig), BL-21(DE3)RIL, BL21(DE3)STAR and BL21(DE3)pLysS were tested for better expression and their protein expression profiles are shown in Figure 3.8.

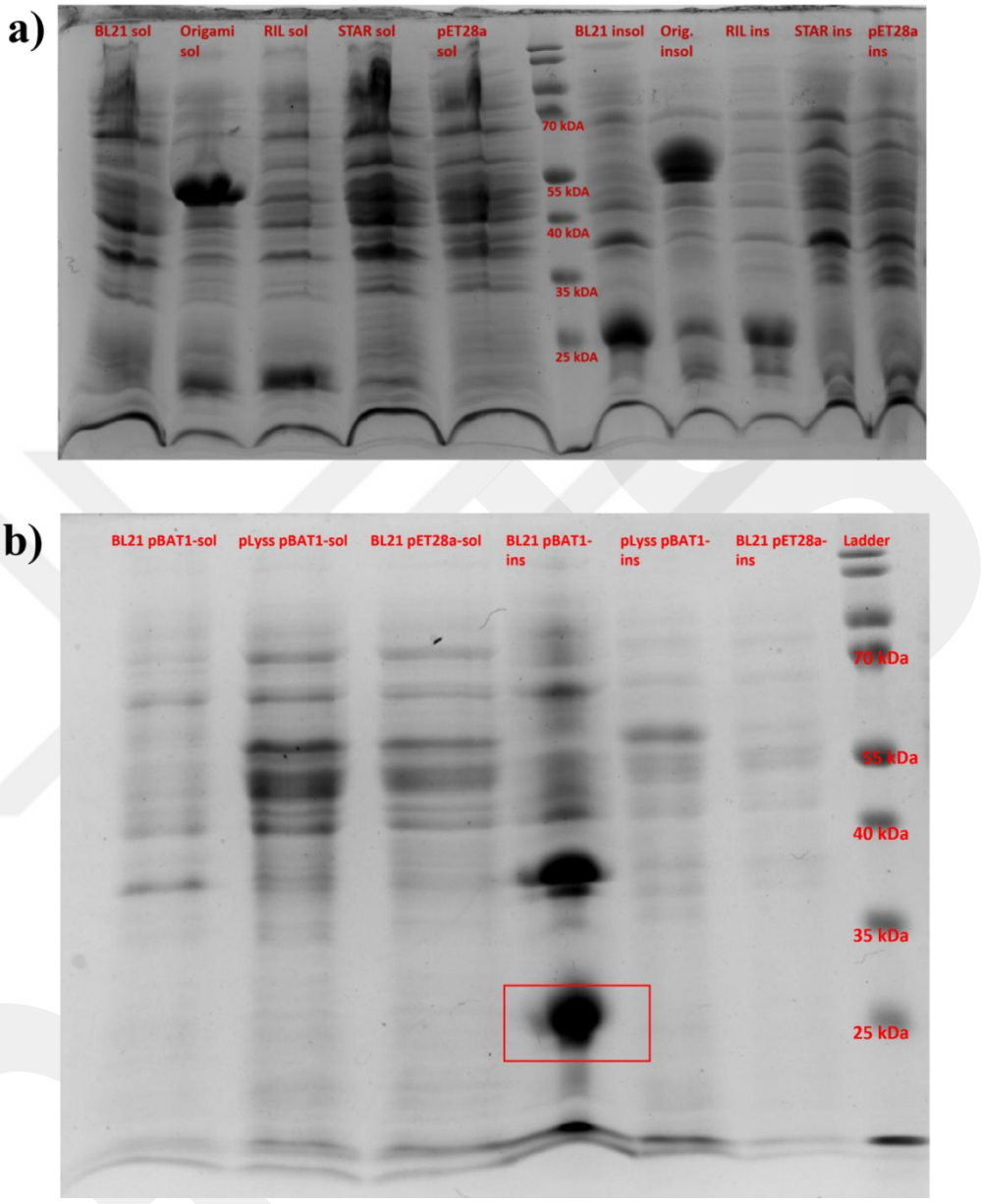

Figure 3.8. SDS-PAGE gel image of protein expression in different strains. (a) *E. coli* BL-21(DE3), *E. coli* Origami (tig), *E. coli* BL-21(DE3)RIL, *E. coli* BL21(DE3)STAR strains. (b) *E. coli* BL21(DE3)pLysS.

As it was mentioned before, protein concentration was measured by comparing band thickness of different strains carrying pBAT1. As soluble fraction, *E. coli* Origami and *E. coli* RIL strains have bands differing from control group but the size is below 25 kDa which cannot be considered as target protein. Thick bands of *E. coli* Origami(tig) strain at around 55 kDa is trigger factor(tig) protein which is expressed as chaperone by a second plasmid in *E. coli* Origami cells. That constructed strain was tested due to solubility

enhancer tig protein [58]. However, there was no any intense soluble band at around 25 kDa in *E. coli* Origami cells. Success of soluble recombinant protein expression depends on various parameters and one of them is the expression strain because host *E. coli* strains have been specifically designed by genetic modifications, shown in Table 2.1., to improve heterologous protein production. For example, *E. coli* BL-21(DE3), *E. coli* BL21(DE3)STAR strains are deficient of *ompT* and *lon* proteases which degrades T7 RNA polymerase and misfolded recombinant proteins, respectively [66]. Also, supressed *rne* gene in *E. coli* BL21(DE3)STAR encodes for RNase E which helps to prevent mRNA degradation. So, use of that strain provides enhanced mRNA stability followed by protein expression, stated as 2- to 10-fold [71]. Another host strain *E. coli* Origami has mutations in thioredoxin reductase (*trxB*) and glutathione reductase (*gor*) genes which are responsible for reducing the cytosolic environment. Reduced cytsol catalyzes the disulfide bond formation in the expressed proteins [72]. So, the formation of intramolecular disulfide bond leading to increased stability and longer half-life, in analogue-3, was supported in *E. coli* Origami strain. Considering all these features of the strains, pBAT1 was transformed to all of shown *E. coli* strains and it was obtained that the *E. coli* BL21(DE3) / pBAT1 produced the insoluble form of GLP-1 Tandem most among them. In order to check post-induction incubation temperature parameter, 6 hours at 37 °C incubation was tested in *E. coli* BL21(DE3) strain while other parameters remained same. Its protein expression profile is shown in Figure 3.9.

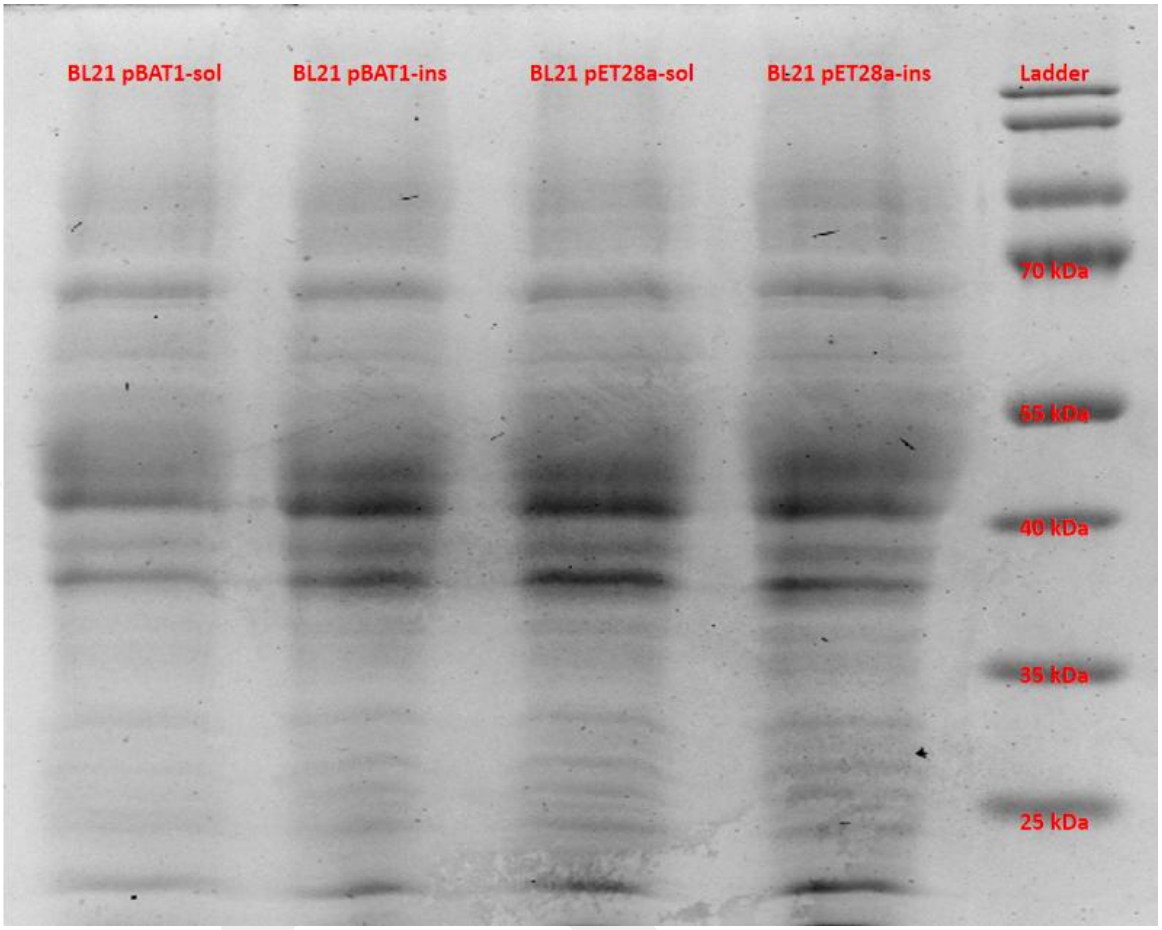

Figure 3.9. SDS-PAGE gel image of protein expression at 37 °C. Target protein size is around 25 kDa and that sized protein bands were considered while comparing.

As it can be seen in the Figure 3.9., there is no any thick band that differs from the control group at around 25 kDa. So, it can be said that higher incubation temperature did not show a positive effect on the target protein expression. Based on the previous studies, it is already stated that lower expression temperature usually enhances the amount and the solubility of target recombinant protein. Because, cell metabolism slows down at lower temperatures which leads to decreased transcription, translation and cell division. Thus, protein aggregation and activity of proteases are reduced [69]. Cultivating bacterial culture at 37 ºC until the induction point and then inducing at 15 ºC to 25 ºC is suggested due to mentioned reasons [70]. As another expression parameter, induction point was tested. For that purpose, IPTG was added when OD_{600} value reached to 0.4, 0.6, 0.8 and 1. Protein expression analysis of that group is shown in Figure 3.10.

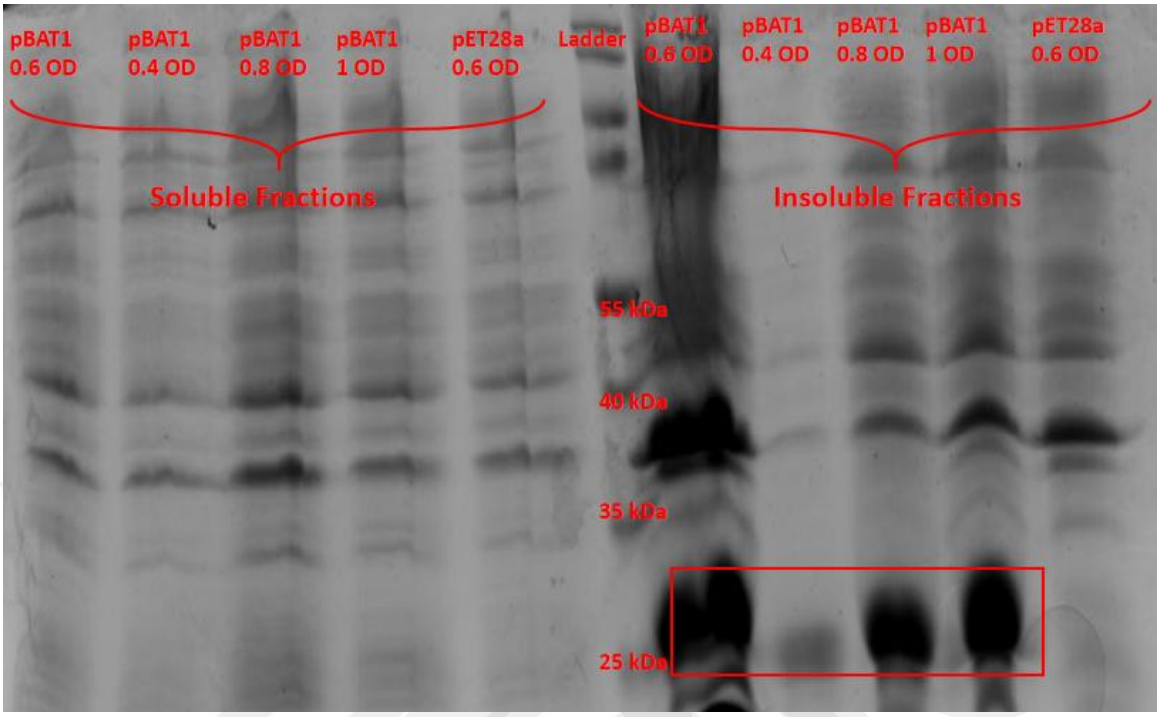

Figure 3.10. SDS-PAGE gel image of protein expression in different induction points. Both soluble and insoluble fractions were loaded and pET28a 0.6 OD was used as control. Target protein size is expected as around 25 kDa. Red framed bands are referred as GLP-1 Tandem protein whose control line (pET28a insoluble) has no any intense band at around 25 kDa.

Plasmids have various components which are required for protein expression and some of the most critical ones can be counted as origin of replication and the promoter. Plasmid copy number (PCN) is controlled by the origin of replication. High PCN elevates gene dosage and it is generally accepted as a favorable situation for high levels of target recombinant protein. However, it is not always the case [67]. It has been demonstrated that a very high gene expression is able to lead to a metabolic burden which reduces the bacterial growth rate. Thus, plasmid instability is induced followed by decreased protein synthesis [68]. Since IPTG induces *lac* operon and initiates the gene expression, different IPTG induction points were tested to decide optimum gene expression amount. Because earlier IPTG induction corresponds to greater gene expression during cultivation.

According to the obtained results from different induction points, it can be stated that relatively higher OD_{600nm} value is better since there is a clear difference between the 25 kDa band of 0.4 OD_{600nm} and others. Induction points of 0.6, 0.8 and 1 were resulted in almost same band profiles. 0.6 OD_{600nm} induction point was already supposed to be the optimum

one as a general knowledge. So, further expression cultures were induced when OD_{600} value reached to 0.6. After the induction point was determined, IPTG concentration optimization followed the experimental step. For that purpose, 0.1 mM, 0.2 mM, 0.5 mM and 1 mM IPTG concentrations were tested to observe its impact on protein expression efficiency. SDS-PAGE analysis of IPTG concentration parameter is shown in Figure 3.11.

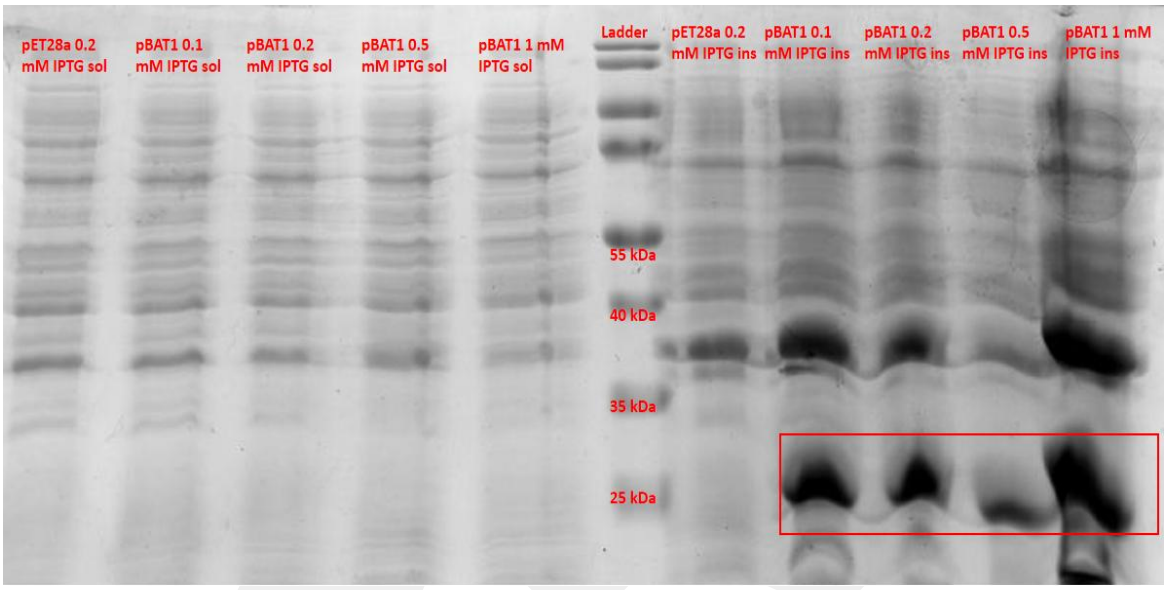

Figure 3.11. SDS-PAGE gel image of protein expression with different IPTG concentrations.

Both soluble and insoluble fractions were loaded and pET28a 0.2 mM IPTG was used as control. Target protein size is expected as around 25 kDa. Red framed bands are referred as GLP-1 Tandem protein whose control line (pET28a insoluble) has no any intense band at around 25 kDa. pBAT1 0.1 mM IPTG, 0.2 mM IPTG, 0.5 mM IPTG and 1 mM IPTG labeled cultures expressed the target protein which are seen in red frame.

Optimization of IPTG concentration revealed that 0.1 mM, 0.2 mM and 1 mM shows the similar band profiles which has relatively intense bands at around 25 kDa. However, 0.5 mM IPTG concentration did not show as intense band as them. IPTG concentration range is indicated as 0.1 mM to 1mM commonly and it is reported that reduced IPTG inducer concentration can improve the production and solubility [69]. For instance, recombinant cyclomaltodextrinase (CDase) is very sensitive to inducer concentration whose soluble active production was successful by 0.05 mM IPTG but insoluble and inactive by 0.1 mM

IPTG [73]. So, it can be said that even minor changes on inducer concentration can be resulted in exactly the opposite.

According to all performed protein expression in changed conditions, it can be clearly stated that *E. coli* BL21(DE3) as strain, 0.2 mM IPTG induction at 0.6 OD_{600nm} and 18 °C overnight post-induction growth was the most promising conditions. Under these conditions, target protein GLP-1 Tandem was obtained as insoluble fraction. That's why insoluble fraction from that culture was used in following protein purification. Soluble form of the target protein could not be produced under any of changed parameters.

3.3. Protein Purification and Quantification

Constructed pBAT1 plasmid has 6xHis-tag sequence at both N-terminal and C-terminal of the inserted protein as it was seen in the Figure 2.1.(b). Thus, purifying the target protein by using His-tag affinity of the Ni-NTA resin was the following experiment. GLP-1 Tandem protein that purified from other cellular proteins is shown and explained in the Figure 3.12.

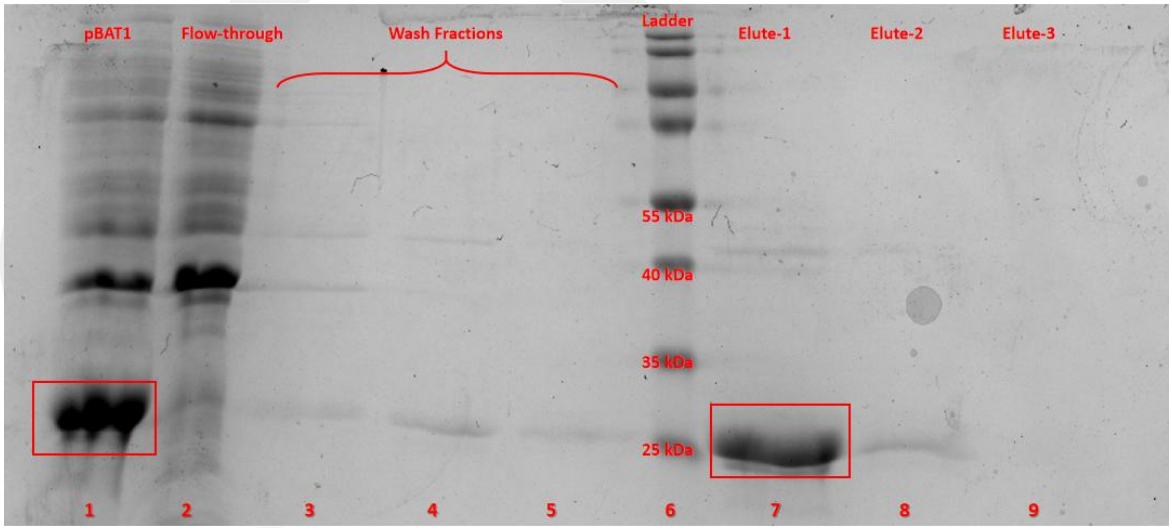

Figure 3.12. SDS-PAGE gel image of purified GLP-1 Tandem protein.

In the figure below, first line refers to insoluble fraction without performing any extra protocol after SDS-PAGE analysis. Second line refers to flow-through fraction that was took after attachment of the resin and His-tagged proteins in the sample. Thus the thick band around 25 kDa in the starting sample is not seen in the flow-through. Third, fourth

and fifth lines refer to washing fractions and almost there is no protein is seen as it is expected. Seventh line refers to first elution fraction and there is a clear intense 25 kDa band which is GLP-1 Tandem protein. Eighth and ninth lines refer to further elution fractions and almost there is no protein is seen as it is expected.

Bradford Protein Assay was performed to measure produced protein concentration. The first elute, whose band is seen in Figure 3.12., was used in quantification protocol. Albumin standard was used to obtain a standard curve seen in Figure 3.13. According to sketched graph, the equation of y = 0,0007x + 0,0078 was used to calculate protein concentration. Also, R^2 value of the curve was 0,9736 which is acceptable.

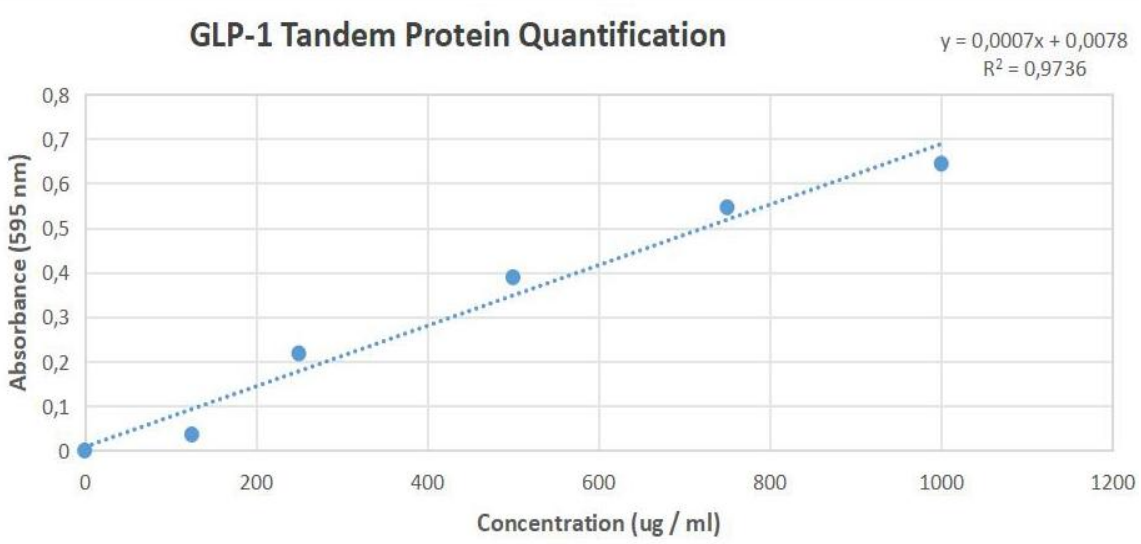

Figure 3.13. Standard curve for GLP-1 Tandem protein purification. X axis refers to protein concentration (µg/ml); y axis refers to measured absorbance value at 595 nm. The standard curve graph was created based on different concentrations of BSA that 1000 µg/ml, 750 µg/ml, 500 µg/ml, 250 µg/ml, 125 µg/ml and the blank.

Based on the obtained equation (y = 0,0007x + 0,0078), target protein concentration was determined. Absorbance value of the target protein was measured as 0,446 and it was used in the equation to calculate concentration. The calculation gives the protein concentration as 62.6 µg/ml. The measured protein was 1:10 diluted form in order to provide imidazole and Bradford reagent compatibility. Because the producer states that higher imidazole concentration like 500 mM is not compatible with the reagent and it needs to be decreased at the beginning of the experiment. Thus, concentration of target protein becomes 626 µg/ml (626 mg/L).

Chapter 4

Conclusions and Future Prospects

4.1. Conclusions and Discussion

T2DM is the most serious metabolic disease with its high correlation with development of hypertension, obesity, hypercholesterolemia, cardiovascular conditions, neuropathy and Alzheimer's Disease as secondary complications. Millions of people suffer from this disorder worldwide that indicates need of novel medication development. That was the main motivation of this thesis. In that manner, the study mainly purposed to produce GLP-1 analogue using *E. coli* as host organism. Basis of this purpose arised from the benefits of GLP-1 compared to alternative approaches against T2DM. Primary benefits are the lower risk of hypoglycemia and weight loss which make GLP-1 analogues safer and more preferable for long-term use. The adverse effects of usual treatments like different insulin gradient, hypoglycemia, cardiovascular risks, lipodystrophy and lowered life quality were aimed to be eliminated by recombinantly producing a GLP-1 analogue as GLP-1 analogues eliminate these adverse effects. The results of the study in various contexts proved that the target protein was successfully produced. But it was in insoluble form which makes it non-functional without further process. In the field of recombinant protein, obtaining high yields of soluble protein can be a troublesome process. The experiment can be delayed due to required complex steps to obtain native protein in that case [60]. Because inclusion body formation is a well-known consequence of using *E. coli* expression system and there are some methods to enhance solubility. For instance, using mutant strains,

decreasing post-induction temperature, changing IPTG concentration, cell lysis conditions etc. [58]. In purpose of enhancing both expression and solubility, mentioned parameters were optimized in this study. Optimum conditions were provided for expression amount which can be seen in result section. However, solubility could not be provided in applied optimization conditions. As an example from the literature, Yuliawati et al. produced eight multicopy of GLP-1 analogue by using *E. coli* as inclusion body form [57]. On the other hand, Sharma and Chaudhuri reported a study using almost same expression system and conditions as our system which resulted in soluble recombinant human serum albumin (HSA) production [58]. In order to facilitate expression and purification of target protein, fusion proteins like Thioredoxin (Trx) and Glutathione S-transferase (GST) can also be constructed [56]. Also, solubility of the protein is exactly correlated with amino acid content of the protein because of chemical interactions and physiological properties that determines folding [59]. That's why same recombinant systems with same expression conditions may not show same results for a different protein. In conclusion, it can be said that GLP-1 Tandem analogue can be produced (626 µg/ml) in *E. coli* BL21(DE3) strain under the conditions of 0.2 mM IPTG induction at 0.6 OD_{600} and overnight post-induction growth as most promising parameters. All taken together, it can be unquestionably stated that various parameters have significant impact on efficiency of recombinant protein production.

4.2. Societal Impact and Contribution to Global Sustainability

Producing high-yield recombinant GLP-1 analogue using microorganism provide an important advancement in the field of biotechnology which includes societal and sustainable implications. T2DM and obesity are two correlated disorders that stand as a global health concern. Using *E. coli* to produce recombinant protein is an approach that offers cost-effective and scalable method to meet the elevating demand for T2DM therapeutics. Thus, making them accessible and affordable for patients worldwide is an important basis of this study. As societal perspective, successful GLP-1 analogue therapeutics are able to reduce the prevalence of T2DM and obesity-related diseases. In that manner, healthcare costs can be decreased resulted in increased life quality for all patients

around the world. In addition to its direct impact on health expenditures as an economic benefit, indirect positive effects can also be mentioned. Because in healthier societies, productivity increases, absenteeism decreases, and as a result, economic growth and stability occur. The mentioned results can be considered as social benefit and contribution to global sustainability. Additionally, the use of *E. coli* host expression system meets with global sustainability principles because it is a well-known, rapid growing microorganism which can be cultured in renewable resources. Thus, environmental footprint of the production process is diminished. Moreover, it will also contribute to the sustainable pharmaceutical manufacturing since recombinant protein production significantly reduces the harms of traditional chemical synthesis. Because, recombinant GLP-1 production does nor require using toxic materials and its production process does not generate hazardous chemical waste. Also, as it was mentioned before, using *E. coli* expression system has advantages like being not expensive and easier to obtain required related materials. The drugs currently used against type 2 diabetes and their side effects have been mentioned before. It was also mentioned that glp-1 analogues are able to eliminate some of these side effects. In other words, health expenses may be higher than expected during the treatment of type 2 diabetes in conventional methods. Based on all these, the positive effects of recombinant glp-1 analogues on social and economic sustainability in many aspects cannot be ignored.

4.3. Future Prospects

The study mainly focused on producing target GLP-1 Tandem protein under optimum conditions. Although expression amount was good enough, formation of inclusion bodies is the main challenge that needs to be considered in further researches. In the case of using *E. coli* expression system, fusion proteins can be integrated to construct, amino acid content can be modified, induction parameters can be changed in a wider range, additive molecules like osmolyte or co-expressing chaperones like DnaK, DnaJ and GroEL/ES can be used to favor soluble protein production [61]. In the case of obtaining insoluble form of the target protein like in the study, refolding process is an alternative approach. However, it is still a labor-intensive, low-yield and complex task in current field of biotechnology [62]. Lastly, culture growth steps were performed in 50 mL shake-flask which was not large-scale

production. So, large-scale production of the functional GLP-1 analogue can be provided in bioreactor by using another host organism like yeast. Our following project is designed to produce novel high-yield GLP-1 analogues with biological activity. As a final future prospect, it can be said that some of the approved recombinant GLP-1 analogue drugs mentioned in section 1.2. have been able to overcome drawbacks of other types of therapeutics. That's why analyzing clinical consequences of similar drugs is a valuable guide for planning further studies.

BIBLIOGRAPHY

[1] Bhupathiraju, S. N., & Hu, F. B. (2016). Epidemiology of obesity and diabetes and their cardiovascular complications. *Circulation Research, 118*(11), 1723–1735. https://doi.org/10.1161/CIRCRESAHA.115.306825

[2] Sağlığın Geliştirilmesi Genel Müdürlüğü. (n.d.). *Dünya Diyabet Günü 2020*. Retrieved June 30, 2024, from https://sggm.saglik.gov.tr/TR-76887/dunya-diyabet-gunu-2020.html

[3] Hekim, M. (2015). Tip II diyabet, hipertansiyon ve obezitenin önlenmesinde fiziksel aktivitenin önemi. *Journal of International Social Research, 8*(38), 1081–1086. https://doi.org/10.17719/jisr.20153813713

[4] Saeedi, P., Petersohn, I., Salpea, P., Malanda, B., Karuranga, S., Unwin, N., Colagiuri, S., Guariguata, L., Motala, A. A., Ogurtsova, K., Shaw, J. E., Bright, D., & Williams, R. (2019). Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Research and Clinical Practice, 157*, 107843. https://doi.org/10.1016/j.diabres.2019.107843

[5] The Lancet. (2023). Diabetes: a defining disease of the 21st century. *The Lancet, 401*(10394), 2087–2087. https://doi.org/10.1016/s0140-6736(23)01296-5

[6] Kreienkamp, R. J., Voight, B. F., Gloyn, A. L., & Udler, M. S. (2023). Genetics of type 2 diabetes (J. M. Lawrence, S. S. Casagrande, W. H. Herman, D. J. Wexler, & W. T. Cefalu, Eds.). National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). https://www.ncbi.nlm.nih.gov/books/NBK597726/

[7] Reed, J., Bain, S., & Kanamarlapudi, V. (2021). A review of current trends with type 2 diabetes epidemiology, aetiology, pathogenesis, treatments and future perspectives. *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy, 14*(1), 3567–3602. https://doi.org/10.2147/dmso.s319895

[8] Patel, R., & Keyes, D. (2023). Lifestyle modification for diabetes and heart disease prevention. StatPearls Publishing. https://pubmed.ncbi.nlm.nih.gov/36251831/

[9] Deepthi, B., Sowjanya, K., Lidiya, B., Bhargavi, R. S., & Babu, P. S. (2018). A modern review of diabetes mellitus: An annihilatory metabolic disorder. *Journal of in Silico & in Vitro Pharmacology, 3*(1). https://doi.org/10.21767/2469-6692.100014

[10] Rendell, M. (2004). The role of sulphonylureas in the management of type 2 diabetes mellitus. *Drugs, 64*(12), 1339–1358. https://doi.org/10.2165/00003495-200464120-00006

[11] Raptis, S., & Dimitriadis, G. (2001). Oral hypoglycemic agents: Insulin secretagogues, α-glucosidase inhibitors and insulin sensitizers. *Experimental and Clinical Endocrinology & Diabetes, 109*(Suppl 2), S265–S287. https://doi.org/10.1055/s-2001-18588

[12] van Baar, M. J. B., van Ruiten, C. C., Muskiet, M. H. A., van Bloemendaal, L., van IJzerman, R. G., & van Raalte, D. H. (2018). SGLT2 inhibitors in combination therapy: From mechanisms to clinical considerations in type 2 diabetes management. *Diabetes Care, 41*(8), 1543–1556. https://doi.org/10.2337/dc18-0588

[13] Foretz, M., Guigas, B., Bertrand, L., Pollak, M., & Viollet, B. (2014). Metformin: From mechanisms of action to therapies. *Cell Metabolism, 20*(6), 953–966. https://doi.org/10.1016/j.cmet.2014.09.018

[14] Diamant, M., & Heine, R. J. (2003). Thiazolidinediones in type 2 diabetes mellitus. *Drugs, 63*(13), 1373–1405. https://doi.org/10.2165/00003495-200363130-00004

[15] Korytkowski, M. T. (2004). Sulfonylurea treatment of type 2 diabetes mellitus: Focus on glimepiride. *Pharmacotherapy, 24*(5), 606–620. https://doi.org/10.1592/phco.24.6.606.34752

[16] Fuhlendorff, J., Rorsman, P., Kofod, H., Brand, C. L., Rolin, B., MacKay, P., Shymko, R., & Carr, R. D. (1998). Stimulation of insulin release by repaglinide and glibenclamide involves both common and distinct processes. *Diabetes, 47*(3), 345–351. https://doi.org/10.2337/diabetes.47.3.345

[17] Landgraf, R. (2000). Meglitinide analogues in the treatment of type 2 diabetes mellitus. *Drugs & Aging, 17*(5), 411–425. https://doi.org/10.2165/00002512-200017050-00007

[18] Meneilly, G. S., Ryan, E. A., Radziuk, J., Lau, D. C., Yale, J. F., Morais, J., Chiasson, J. L., Rabasa-Lhoret, R., Maheux, P., Tessier, D., Wolever, T., Josse, R. G., & Elahi, D. (2000). Effect of acarbose on insulin sensitivity in elderly patients with diabetes. *Diabetes Care, 23*(8), 1162–1167. https://doi.org/10.2337/diacare.23.8.1162

[19] van de Laar, F. A., Lucassen, P. L., Akkermans, R. P., van de Lisdonk, E. H., Rutten, G. E., & van Weel, C. (2004). α-Glucosidase inhibitors for patients with type 2 diabetes: Results from a Cochrane systematic review and meta-analysis. *Diabetes Care, 28*(1), 154–163. https://doi.org/10.2337/diacare.28.1.154

[20] Moradi-Marjaneh, R., Paseban, M., & Sahebkar, A. (2019). Natural products with SGLT2 inhibitory activity: Possibilities of application for the treatment of diabetes. *Phytotherapy Research, 33*(10), 2518–2530. https://doi.org/10.1002/ptr.6421

[21] Yang, L., Yuan, J., & Zhou, Z. (2014). Emerging roles of dipeptidyl peptidase 4 inhibitors: Anti-inflammatory and immunomodulatory effect and its application in diabetes mellitus. *Canadian Journal of Diabetes, 38*(6), 473–479. https://doi.org/10.1016/j.jcjd.2014.01.008

[22] Drucker, D. J., Sherman, S. I., Gorelick, F. S., Bergenstal, R. M., Sherwin, R. S., & Buse, J. B. (2010). Incretin-based therapies for the treatment of type 2 diabetes: Evaluation of the risks and benefits. *Diabetes Care, 33*(2), 428–433. https://doi.org/10.2337/dc09-1499

[23] Kendall, D. A., Cuddihy, R., & Bergenstal, R. M. (2009). Clinical application of incretin-based therapy: Therapeutic potential, patient selection and clinical use. *The American Journal of Medicine, 122*(6), S37–S50. https://doi.org/10.1016/j.amjmed.2009.03.015

[24] Blahova, J., Martiniakova, M., Babikova, M., Kovacova, V., Mondockova, V., & Omelka, R. (2021). Pharmaceutical production of therapeutic proteins by genetically modified bacteria. *Ecopersia, 9*(1), 1113–1123.

[25] Huang, C. J., Lin, H., & Yang, X. (2012). Industrial production of recombinant therapeutics in Escherichia coli and its recent advancements. *Journal of Industrial Microbiology & Biotechnology, 39*(3), 383–399. https://doi.org/10.1007/s10295-011-1036-1

[26] Werner, R. G., Noé, W., Kopp, K., & Schluter, M. (2007). Appropriate mammalian expression systems for biopharmaceuticals. *Drugs in R&D, 8*(5), 285–293. https://doi.org/10.2165/00126839-200708050-00001

[27] Borse, S. P., Chhipa, A. S., Sharma, V., Singh, D. P., & Nivsarkar, M. (2020). Management of type 2 diabetes: Current strategies, unfocused aspects, challenges and alternatives. *Medical Principles and Practice, 30*(2). https://doi.org/10.1159/000511002

[28] Holst, J. J. (2007). The physiology of glucagon-like peptide 1. *Physiological Reviews, 87*(4), 1409–1439. https://doi.org/10.1152/physrev.00034.2006

[29] Calanna, S., Christensen, M., Holst, J. J., Laferrère, B., Gluud, L. L., Vilsbøll, T., & Knop, F. K. (2013). Secretion of glucagon-like peptide-1 in patients with type 2 diabetes mellitus: Systematic review and meta-analyses of clinical studies. *Diabetologia, 56*(5), 965–972. https://doi.org/10.1007/s00125-013-2841-0

[30] Sheikh, A. (2013). Direct cardiovascular effects of glucagon like peptide-1. *Diabetology & Metabolic Syndrome, 5*(1). https://doi.org/10.1186/1758-5996-5-47

[31] Verspohl, E. J. (2009). Novel therapeutics for type 2 diabetes: Incretin hormone mimetics (glucagon-like peptide-1 receptor agonists) and dipeptidyl peptidase-4 inhibitors. *Pharmacology & Therapeutics, 124*(1), 113–138. https://doi.org/10.1016/j.pharmthera.2009.06.002

[32] Jin, T. (2008). Mechanisms underlying proglucagon gene expression. *Journal of Endocrinology, 198*(1), 17–28. https://doi.org/10.1677/joe-08-0085

[33] Donnelly, D. (2012). The structure and function of the glucagon-like peptide-1 receptor and its ligands. *British Journal of Pharmacology, 166*(1), 27–41. https://doi.org/10.1111/j.1476-5381.2011.01687.x

[34] Alavi, S. E., Cabot, P. J., & Moyle, P. M. (2019). Glucagon-like peptide-1 receptor agonists and strategies to improve their efficiency. *Molecular Pharmaceutics, 16*(6), 2278–2295. https://doi.org/10.1021/acs.molpharmaceut.9b00308

[35] Engel, M., Hoffmann, T., Wagner, L., Wermann, M., Heiser, U., Kiefersauer, R., Huber, R., Bode, W., Demuth, H.-U., & Brandstetter, H. (2003). The crystal structure of dipeptidyl peptidase IV (CD26) reveals its functional regulation and enzymatic mechanism. *Proceedings of the National Academy of Sciences, 100*(9), 5063–5068. https://doi.org/10.1073/pnas.0230620100

[36] Gupta, R., Walunj, S., Tokala, R., Parsa, K., Singh, S., & Pal, M. (2009). Emerging drug candidates of dipeptidyl peptidase IV (DPP IV) inhibitor class for the treatment of

type 2 diabetes. *Current Drug Targets, 10*(1), 71–87. https://doi.org/10.2174/138945009787122860

[37] Green, B. D., Flatt, P. R., & Bailey, C. J. (2006). Dipeptidyl peptidase IV (DPP IV) inhibitors: A newly emerging drug class for the treatment of type 2 diabetes. *Diabetes and Vascular Disease Research, 3*(3), 159–165. https://doi.org/10.3132/dvdr.2006.024

[38] Bell, G. I., Santerre, R. F., & Mullenbach, G. T. (1983). Hamster preproglucagon contains the sequence of glucagon and two related peptides. *Nature, 302*(5910), 716–718. https://doi.org/10.1038/302716a0

[39] Perry, T., & Greig, N. H. (2003). The glucagon-like peptides: A double-edged therapeutic sword? *Trends in Pharmacological Sciences, 24*(7), 377–383. https://doi.org/10.1016/s0165-6147(03)00160-3

[40] Hupe-Sodmann, K., Göke, R., Göke, B., Thole, H. H., Zimmermann, B., Voigt, K., & McGregor, G. P. (1997). Endoproteolysis of glucagon-like peptide (GLP)-1(7–36) amide by ectopeptidases in RINm5F cells. *Peptides, 18*(5), 625–632. https://doi.org/10.1016/s0196-9781(97)00123-x

[41] Montrose-Rafizadeh, C., Yang, H., Rodgers, B. D., Beday, A., Pritchette, L. A., & Eng, J. (1997). High potency antagonists of the pancreatic glucagon-like peptide-1 receptor. *The Journal of Biological Chemistry, 272*(34), 21201–21206. https://doi.org/10.1074/jbc.272.34.21201

[42] Kyriacou, A., & Ahmed, A. B. (2010). Exenatide use in the management of type 2 diabetes mellitus. *Pharmaceuticals, 3*(8), 2554–2567. https://doi.org/10.3390/ph3082554

[43] Barnett, A. (2011). Lixisenatide: Evidence for its potential use in the treatment of type 2 diabetes. *Core Evidence, 6*, 67. https://doi.org/10.2147/ce.s15525

[44] Klen, J., & Dolžan, V. (2022). Glucagon-like peptide-1 receptor agonists in the management of type 2 diabetes mellitus and obesity: The impact of pharmacological properties and genetic factors. *International Journal of Molecular Sciences, 23*(7), 3451. https://doi.org/10.3390/ijms23073451

[45] Mahapatra, M. K., Karuppasamy, M., & Sahoo, B. M. (2022). Semaglutide, a glucagon-like peptide-1 receptor agonist with cardiovascular benefits for management of type 2 diabetes. *Reviews in Endocrine and Metabolic Disorders, 23*(3). https://doi.org/10.1007/s11154-021-09699-1

[46] Xu, F., Wang, K. Y., Wang, N., Li, G., & Liu, D. (2017). Modified human glucagon-like peptide-1 (GLP-1) produced in E. coli has a long-acting therapeutic effect in type 2 diabetic mice. *PLOS ONE, 12*(7), e0181939. https://doi.org/10.1371/journal.pone.0181939

[47] Burcelin, R., Dolci, W., & Thorens, B. (1999). Long-lasting antidiabetic effect of a dipeptidyl peptidase IV-resistant analog of glucagon-like peptide-1. *Metabolism, 48*(2), 252–258. https://doi.org/10.1016/s0026-0495(99)90043-4

[48] Wang, L. (2022). Designing a dual GLP-1R/GIPR agonist from tirzepatide: Comparing residues between tirzepatide, GLP-1, and GIP. *Drug Design, Development and Therapy, 16*, 1547–1559. https://doi.org/10.2147/dddt.s358989

[49] Miranda, L. P., Winters, K. A., Gegg, C. V., Patel, A., Aral, J., Long, J., Zhang, J., Diamond, S., Guido, M., Stanislaus, S., Ma, M., Li, H., Rose, M. J., Poppe, L., & Véniant, M. M. (2008). Design and synthesis of conformationally constrained glucagon-like peptide-1 derivatives with increased plasma stability and prolonged in vivo activity. *Journal of Medicinal Chemistry, 51*(9), 2758–2765. https://doi.org/10.1021/jm701522b

[50] Knudsen, L. B., Nielsen, P. F., Huusfeldt, P. O., Johansen, N. L., Madsen, K., Pedersen, F. Z., Thøgersen, H., Wilken, M., & Agersø, H. (2000). Potent derivatives of glucagon-like peptide-1 with pharmacokinetic properties suitable for once daily administration. *Journal of Medicinal Chemistry, 43*(9), 1664–1669. https://doi.org/10.1021/jm9909645

[51] Li, Y., Zheng, X., Tang, L., Xu, W., & Gong, M. (2011). GLP-1 analogs containing disulfide bond exhibited prolonged half-life in vivo than GLP-1. *Peptides, 32*(6), 1303–1312. https://doi.org/10.1016/j.peptides.2011.04.001

[52] Fidan, O., & Zhan, J. (2015). Recent advances in engineering yeast for pharmaceutical protein production. *RSC Advances, 5*(105), 86665–86674. https://doi.org/10.1039/c5ra13003d

[53] Lozano Terol, G., Gallego-Jara, J., Sola Martínez, R. A., Martínez Vivancos, A., Cánovas Díaz, M., & de Diego Puente, T. (2021). Impact of the expression system on recombinant protein production in *Escherichia coli* BL21. *Frontiers in Microbiology, 12*. https://doi.org/10.3389/fmicb.2021.682001

[54] Pontrelli, S., Chiu, T.-Y., Lan, E. I., Chen, F. Y.-H., Chang, P., & Liao, J. C. (2018). *Escherichia coli* as a host for metabolic engineering. *Metabolic Engineering, 50*, 16–46. https://doi.org/10.1016/j.ymben.2018.04.008

[55] Meyer, H.-P., & Schmidhalter, D. R. (2012). Microbial expression systems and manufacturing from a market and economic perspective. *InTech EBooks*. https://doi.org/10.5772/29417

[56] Gao, M., Ma, C., Liu, W., Zhu, J., Tian, H., Gao, X., & Yao, W. (2010). Production and purification of an analog of glucagon-like peptide-1 by auto-induction and on-column cleavage in *Escherichia coli*. *World Journal of Microbiology & Biotechnology*, 26(9), 1675–1682. https://doi.org/10.1007/s11274-010-0345-3

[57] Yuliawati, N., Fathurahman, A. T., Kusharyoto, W., & Ningrum, R. A. (2023). Construction of multicopy glucagon-like peptide-1 (GLP-1) open reading frame and its expression in *Escherichia coli*. *AIP Conference Proceedings*. https://doi.org/10.1063/5.0118377

[58] Sharma, A., & Chaudhuri, T. K. (2017). Revisiting *Escherichia coli* as microbial factory for enhanced production of human serum albumin. *Microbial Cell Factories, 16*. https://doi.org/10.1186/s12934-017-0784-8

[59] Trevino, S. R., Scholtz, J. M., & Pace, C. N. (2007). Amino acid contribution to protein solubility: Asp, Glu, and Ser contribute more favorably than the other hydrophilic amino acids in RNase Sa. *Journal of Molecular Biology, 366*(2), 449–460. https://doi.org/10.1016/j.jmb.2006.10.026

[60] Gutiérrez-González, M., Farías, C., Tello, S., Pérez-Etcheverry, D., Romero, A., Zúñiga, R., Ribeiro, C. H., Lorenzo-Ferreiro, C., & Molina, M. C. (2019). Optimization of culture conditions for the expression of three different insoluble proteins in *Escherichia coli*. *Scientific Reports, 9*(1), 16850. https://doi.org/10.1038/s41598-019-53200-7

[61] Atroshenko, D. L., Sergeev, E. P., Golovina, D. I., & Pometun, A. A. (2024). Additivities for soluble recombinant protein expression in cytoplasm of *Escherichia coli*. *Fermentation, 10*(3), 120. https://doi.org/10.3390/fermentation10030120

[62] Yamaguchi, S., Yamamoto, E., Mannen, T., Nagamune, T., & Nagamune, T. (2012). Protein refolding using chemical refolding additives. *Biotechnology Journal, 8*(1), 17–31. https://doi.org/10.1002/biot.201200025

[63] E. coli genotypes. (n.d.). *Www.novoprolabs.com*. https://www.novoprolabs.com/tools/e-coli-genotypes

[64] Merck | Yaşam Bilimleri | Endüstriyel ve Laboratuvar Kimyasalları | eMağaza. (2024). *Merckmillipore.com*. https://www.merckmillipore.com/TR/tr

[65] GenScript - Make Research Easy - The leader in molecular cloning and gene synthesis, peptide synthesis, protein and antibody engineering. (n.d.). *Www.genscript.com*. https://www.genscript.com/

[66] Grodberg, J., & Dunn, J. J. (1988). ompT encodes the *Escherichia coli* outer membrane protease that cleaves T7 RNA polymerase during purification. *Journal of Bacteriology, 170*(3), 1245–1253. https://doi.org/10.1128/jb.170.3.1245-1253.1988

[67] Mergulhão, F. J., Monteiro, G. A., Cabral, J. M. S., & Taipa, M. A. (2004). Design of bacterial vector systems for the production of recombinant proteins in *Escherichia coli*. *Journal of Microbiology and Biotechnology, 14*(1), 1–14.

[68] Gomes, L., Monteiro, G., & Mergulhão, F. (2020). The impact of IPTG induction on plasmid stability and heterologous protein expression by *Escherichia coli* biofilms. *International Journal of Molecular Sciences, 21*(2), 576. https://doi.org/10.3390/ijms21020576

[69] Francis, D. M., & Page, R. (2010). Strategies to optimize protein expression in *E. coli*. *Current Protocols in Protein Science, Chapter 5*, Unit 5.24.1-29. https://doi.org/10.1002/0471140864.ps0524s61

[70] Peti, W., & Page, R. (2007). Strategies to maximize heterologous protein expression in *Escherichia coli* with minimal cost. *Protein Expression and Purification, 51*(1), 1–10. https://doi.org/10.1016/j.pep.2006.06.024

[71] Lopez, P. J., Marchand, I., Joyce, S. A., & Dreyfus, M. (1999). The C-terminal half of RNase E, which organizes the *Escherichia coli* degradosome, participates in mRNA degradation but not rRNA processing in vivo. *Molecular Microbiology, 33*(1), 188–199. https://doi.org/10.1046/j.1365-2958.1999.01465.x

[72] Prinz, W. A., Åslund, F., Holmgren, A., & Beckwith, J. (1997). The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the *Escherichia coli* cytoplasm. *Journal of Biological Chemistry, 272*(25), 15661–15667. https://doi.org/10.1074/jbc.272.25.15661

[73] Turner, P., Holst, O., & Karlsson, E. N. (2005). Optimized expression of soluble cyclomaltodextrinase of thermophilic origin in *Escherichia coli* by using a soluble fusion-tag and by tuning of inducer concentration. *Protein Expression and Purification, 39*(1), 54–60. https://doi.org/10.1016/j.pep.2004.09.012

[74] Nguyen, A. (2020, January 17). The True Cost of Diabetes. GoodRx; GoodRx. https://www.goodrx.com/conditions/diabetes/true-cost-of-diabetes

[75] American Diabetes Association. (2023, November 1). New American Diabetes Association Report Finds Annual Costs of Diabetes to be $412.9 Billion | ADA. Diabetes.org. https://diabetes.org/newsroom/press-releases/new-american-diabetes-association-report-finds-annual-costs-diabetes-be

[76] Glaesner, W., Mark Vick, A., Millican, R., Ellis, B., Tschang, S.-H., Tian, Y., Bokvist, K., Brenner, M., Koester, A., Porksen, N., Etgen, G., & Bumol, T. (2010). Engineering and characterization of the long-acting glucagon-like peptide-1 analogue LY2189265, an Fc fusion protein. *Diabetes/Metabolism Research and Reviews*, *26*(4), 287–296. https://doi.org/10.1002/dmrr.1080

[77] Hou, J., Yan, R., Yang, L., Wu, Z., Wang, C., Ding, D., Li, N., Ma, C., & Li, M. (2007). High-Level Expression of Fusion Protein Containing 10 Tandem Repeated GLP-1 Analogs in Yeast Pichia pastoris and Its Biological Activity in a Diabetic Rat Model. *Bioscience Biotechnology and Biochemistry*, *71*(6), 1462–1469. https://doi.org/10.1271/bbb.60694

CURRICULUM VITAE

2023 - Present	**Fidan Lab, Abdullah Gül University, Kayseri, Turkey** Bioengineering master student working on recombinant production of peptide analogues
10/2022-12/2022	**Ortakçı Lab, Abdullah Gül University, Kayseri, Turkey** Bioengineering master student working on recombinant microbial protein production
06/2021-08/2021	**DETA-GEN Genetic Diseases Diagnostic Center, Kayseri, Turkey** Intern student in cytogenetic and molecular genetic departments
2018-2022	**Aydın Lab, Abdullah Gül University, Kayseri, Turkey** Research student working on isolation and identification of lactic acid bacteria strains from fermented Turkish food and investigation of probiotic features of isolated lactic acid bacteria.
2017-2022	**B.Sc., Molecular Biology and Genetics, Abdullah Gül University, Kayseri, Turkey**

PUBLICATIONS AND PRESENTATIONS

J1) Cebeci, A., Polat, M.F., Çalış, B. Isolation of Lactic Acid Bacteria from Tarhana published in COMU Journal of Agriculture Faculty (July 2020).

C1) Cebeci, A, Polat, M, Çalış, B. (2019, 28-29 June). Isolation of Lactic Acid Bacteria from Tarhana Poster Presentation in 2nd International Eurasian Conference on Biological and Chemical Sciences, Ankara, Turkey (June 2019).