ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

CLONING, HETEROLOGOUS EXPRESSION AND PURIFICATION OF PHI 29 DNA POLYMERASE VARIANTS

M.Sc. THESIS

Fatma ARAS

Department of Chemical Engineering

Chemical Engineering Programme

DECEMBER 2017

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

CLONING, HETEROLOGOUS EXPRESSION AND PURIFICATION OF PHI 29 DNA POLYMERASE VARIANTS

M.Sc. THESIS

Fatma ARAS
(506141032)

Department of Chemical Engineering

Chemical Engineering Programme

Thesis Advisor: Prof. Dr. Nuran DEVECİ AKSOY

DECEMBER 2017

ISTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

PHI 29 DNA POLİMERAZ TÜREVLERİNİN KLONLANMASI, HETEROLOG EKSPRESYONU VE SAFLAŞTIRILMASI

YÜKSEK LİSANS TEZİ

Fatma ARAS
(506141032)

Kimya Mühendisliği Anabilim Dalı

Kimya Mühendisliği Programı

Tez Danışmanı: Prof. Dr. Nuran DEVECİ AKSOY

ARALIK 2017

Fatma ARAS, a M.Sc. student of İTU Graduate School of Science Engineering and Technology student ID 506141032, successfully defended the thesis entitled "CLONING, HETEROLOGOUS EXPRESSION AND PURIFICATION OF PHI 29 DNA POLYMERASE VARIANTS", which she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor :	**Prof. Dr. Nuran DEVECİ AKSOY** Istanbul Technical University
Jury Members :	**Prof. Dr. Yüksel AVCIBAŞI GÜVENİLİR** Istanbul Technical University
	Assist.Prof. Dr. Yaşar ANDELİP AYDIN Marmara University

Date of Submission : 17 November 2017
Date of Defense : 13 December 2017

To my family,

FOREWORD

I would like to express my deepest gratitude to my supervisor Prof.Dr.Nuran DEVECI AKSOY for her encouragement and guidance.

I owe Prof.Dr.Christof M.NIEMEYER a great debt of gratitude, as he gave me an opportunity to study in Institute for Biological Interfaces 1- Biomolecular Micro- and Nanostructures, Karlsruhe Institute of Technology. I would also thank to Dr.Kersten Rabe for his help during my whole laboratory studies and great contribution to my master thesis. I wish to express my special thanks to all members of IBG-1 for their support.

Lastly, I am grateful to my family, and my friends for their support and love.

November 2017

Fatma ARAS
(Chemical Engineer)

TABLE OF CONTENTS

ABBREVIATIONS

A : Adenin
Amp : Ampicilin
App : Appendix
bp : Base Pair
C : Cytosin
kDa : kiloDalton
DNA : Deoxyribonucleic Acid
dsDNA : Double Stranded Deoxyribonucleic Acid
dNTP : Deoxyribonucleotide
DMSO : Dimethyl Sulfoxide
DTT : Dithiothreitol
E.coli : Escherichia Coli
EDTA : Ethylene Diamine-Tetra-Acetic Acid
G : Guanine
GC : 5X Phusion GC Buffer
h : Hour
HCl : Hydrochloric Acid
HF : 5X Phusion HF Buffer
IPTG : Isopropyl-β-D-1-Thiogalactopyranoside
L : Liter
LB : Lysogeny Broth
mA : Miliampere
$MgCl_2$: Magnesium Chloride
min : Minutes
mL : Mililiter
mM : Milimolar
nM : Nanomolar
NaCl : Sodium Chloride
NaOH : Sodium Hydroxide
µL : Microliter
µM : Mikromolar

OD	**:** Optical Density
PCR	**:** Polymerase Chain Reaction
RFU	**:** Relative Fluorescence Unit
rpm	**:** Rounds per Minute
s	**:** Second
ssDNA	**:** Single Stranded Deoxyribonucleic Acid
SDS	**:** Sodium Dedocyl Sulphate
SDS-PAGE	**:** Sodium Dedocyl Sulphate-Polyacrylamide Gel Electrophoresis
T	**:** Thymin
TBE	**:** Tris/Borate/EDTA Buffer
Tris	**:** 2-Amino-2-(hydroxymethyl)propane-1,3-diol
UV	**:** Ultraviolet
°C	**:** Grad Celsius

LIST OF TABLES

Page

LIST OF FIGURES

Page

CLONING, HETEROLOGOUS EXPRESSION AND PURIFICATION OF PHI 29 DNA POLYMERASE VARIANTS

SUMMARY

Today, isolation and amplification of intracellular genetic material have a vital importance for molecular biology and molecular medicine *in vitro* studies. Nucleic acid amplification techniques such as polymerization chain reaction (PCR) have been developed to synthesize very amount of DNA fragments in a short time. PCR is a technique in which DNA polymerase enzyme is used to amplify specific DNA sequences by the help of the oligonucleotide primers. DNA polymerase enzyme which synthesizes a complementary DNA plays an important role in the genome replication and in proteins synthesis.

Bacillus subtilis phage φ(Phi)29 operate a natural DNA amplification mechanism to reproduce their linear double-stranded DNA genome with φ29 DNA polymerase during the replication process. φ29 DNA polymerase, the product of viral gene 2, gained a lot of interest in the field of DNA analytics in the recent years due to its astonishing processivity and proofreading activity. However, only a limited number of genetic fusion proteins have been described.

In this study, the production and purification of φ29 DNA polymerase variants fused to a Streptavidin binding peptide to enable the immobilization of the polymerase on Streptavidin modified surfaces were carried out.

The φ29 DNA polymerase enzyme was successfully produced in the bacterial expression system thanks to the recombinant DNA technology. After protein purification in affinity chromatography system, a molecular beacon assay was performed to measure the activity of the φ29 DNA polymerase because it allows rapid and quantitative characterization of DNA polymerases activity. Both commercial and *in vitro* synthesized φ29 DNA polymerases have approximately the same performance according to the activity results. However, DNA contamination which revealed itself in each negative control samples of activity tests was a major problem caused by the DNA of *E.coli* host cells. Since φ29 DNA polymerases have a high affinity for both ssDNA and ds DNA, the genomic DNA of *E.coli* and residues of vector DNA should have immediately removed in cell lysis step, but it is not easy to get rid of these undesired DNA fragments later on affinity chromatography.

Besides combining the different functional domains of proteins in the same recombinant protein structure, multifunctional recombinant fusion proteins increase the biological activity of proteins. Another φ29 DNA polymerase labeled with a red fluorescent protein (RFP) was expressed by following the same heterologous protein expression procedures that previously applied for φ29 DNA polymerase synthesis. The purpose of the junction of RFP protein is to detect the φ29 DNA polymerase during polymerization reactions.

The φ29-RFP DNA polymerase enzyme was also purified by affinity chromatography using streptavidin binding peptide. According to the enzymatic activity results, φ29-RFP DNA polymerase also brought to a successful conclusion, but DNA contamination issue revealed itself again. Secondly, in accordance with the SDS gel images, φ29-RFP DNA polymerase is expressed at full-length and the expected size, but there were different protein bands. There may be various reasons related with these multiple bands. For instance, it may have not completely abided by the amino acid composition of Φ29-RFP DNA polymerase, or cleavage at the junction between RFP and φ29 DNA polymerase might have taken place. These are the most commonly encountered problem in fusion protein expression. Finally, both φ29 DNA polymerase and φ29-RFP DNA polymerase concentrations were detected at low level in comparison to commercial φ29 DNA polymerase concentration, but solution-orientated approaches are also given in this thesis study.

PHI 29 DNA POLİMERAZ TÜREVLERİNİN KLONLANMASI, HETEROLOG EKSPRESYONU VE SAFLAŞTIRILMASI

ÖZET

Tüm canlı organizmalar genetik bilgilerini soylarının devamlılığı için gelecek nesillere aktarırlar. Başarılı bir aktarımın sırrı ise organizmaların sahip olduğu DNA polimerazlarda yatmaktadır. DNA replikasyonundaki rolünün yanısıra, bu polimerazlar organizmaların yaşam döngülerinde önemli bir yer tutan protein sentezinde de görev almaktadırlar.

Günümüzde, hücre içi genetik materyalin izolasyonu ve çoğaltılması moleküler biyoloji ve moleküler tıptaki *in vitro* çalışmalar için hayati derecede öneme sahiptir. Nükleik asit çoğaltma tekniklerinden örneğin polimerizasyon zincir rekasiyonu (PCR) çok miktarda DNA parçalarının kısa bir sürede sentezlenmesi için geliştirilmiştir. PCR, oligonükleotid primerler yardımıyla DNA polimeraz enziminin spesifik DNA dizilimlerini çoğalttığı bir tekniktir. DNA polimeraz enzimleri farklı biyokimyasal ve fizikokimyasal özelliklerine göre sınıflandırılmalarına rağmen, polimerizasyon reaksiyon mekanizmalarını benzer şekilde yürütmektedirler.

Bacillus subtilis faj φ29, replikasyon süreci boyunca φ29 DNA polimeraz ile lineer çift DNA zincirli genomunu yeniden üretmek için doğal bir mekanizma işletmektedir. ‘’Viral gen 2’’ ürünü olan φ29 DNA polimeraz şaşırtıcı işlevselliği ve düzelticilik aktivitesi sayesinde son yıllarda DNA analiz çalışmalarında büyük bir ilgi görmektedir. φ29 DNA polimeraz enzimi yüksek kalitede hem çok kompleks DNA dizilimlerini çoğaltırken, hemde DNA yapısında meydana gelmiş mutasyonların çok kısa bir sürede tespit edilmesine imkan vermektedir. Diğer DNA polimerazlarda da bulunan 5’-3’ polimerizasyon aktivitesinin yanısıra, 3’-5’ ekzonükleaz aktivitesi sayesinde polimerizasyon reaksiyonu boyunca yanlış eşleşmiş baz çiftlerinin düzeltilmesi işlevini de gerçekleştirmektedir.

DNA çoğaltılması sürecinde aktif rol alan φ29 DNA polimeraz’ın biyokimyasal özellikleri farklı analiz çalışmalarıyla gün yüzüne çıkartılmıştır. Fakat sadece sınırlı sayıda füzyon proteinleri tanımlanmıştır. Bu çalışmada, Streptavidin yüzeyine proteinin immobilizasyonunu sağlayan Streptavidin bağlayıcı peptit (SBP etiketi) ile bağlanmış φ29 DNA polimeraz türevlerinin üretimi ve saflaştırılması rekombinant DNA teknolojisi sayesinde başarıyla gerçekleştirilmiştir.

Rekombinant proteinin üretimine ait deneysel süreç, proteinin karakteristik özellikleriyle uyumlu bir bakteriyel ekspresyon sisteminin dizaynı ile başlamaktadır. Sentetik φ29 DNA polimeraz genine sahip rekombinant ekspresyon vektörünün konak bakteri hücrelerine transferi sonrası, hücreler uygun besi ortamında geliştirilir ve bakteri kültürlerinden proteinin özütlenmesi ile devam eden süreç, proteinin afinite kolonu kullanılarak saflaştırılması ve karakterizasyon çalışmalarının yapılmasıyla tamamlanmaktadır. Deneyde kullanılan sentetik gen; yapısındaki φ29 DNA polimeraz genine ek olarak esnek bir bağlayıcıya, SBP etiketine ve yapışkan

uçlara sahiptir. Yapışkan uçlar rekombinant vektörün oluşturulmasında kullanılır iken, SBP etiketi afinite kromatografisinde proteinin kolona bağlanmasını sağlamak için yapıya ilave edilmiştir. Esnek bağlayıcı ise proteinin hareket kabiliyetini artırmak amacıyla kullanılmıştır. Ayrıca esnek bağlayıcı ile φ29 DNA polimeraz ve RFP proteini aynı yapıda buluşturularak füzyon proteinin biyolojik aktivitesi artırılmış ve proteinler arasındaki olası etkileşim en aza indirilmiştir.

Bakteriyel ekspresyon sistemi; düşük maliyeti, hızlı çoğalma özelliği ve genetik modifikasyonlara uyumluluğu ile rekombinant protein sentezinde yaygın olarak kullanılan *E.coli* kullanılarak oluşturulmuştur. pET serisi plazmid vektör ise rekombinant proteinin üretiminde görev almıştır.

Rekombinant vektörün oluşturulması ve akabinde bakteri hücrelerine transferi ile klonlama işlemi tamamlanmıştır. Seçilen bakteri kolonileri LB besi ortamında inkübe edilmiştir. Ticari Miniprep seti ile bakterilerin sahip olduğu bu rekombinant vektörler DNA sekanslama işlemi için izole edilmiştir. Sekanslama sonuçlarına göre tasarlanan yapıya sahip olan vektörler ile protein sentezine başlanmış ve elde edilen numuneler saflaştırma işlemi için hazır hale getirilirmiştir.

Protein örnekleri kolon kromatografisi öncesinde bir takım ön saflaştırma işlemlerinden geçirilmiştir. Bunlar; hem mekanik hemde kimyasal etki ile istenmeyen bakteriyel genom, tek veya çift iplik DNA kalıntıları, hücre duvarı vb. uzaklaştırılması için uygulanmıştır.

Kromatografi kolonunda, destek malzemesine sabitlenmiş özel ligandlar ile protein afinite etiketi arasındaki tersinir ilişkiye dayanarak proteinler saflaştırılır. Bu işlemin önemli noktalarından bir diğeri ise yıkama tampon çözeltisinde SBP-etiketi ile afinite açısından yarışarak proteinin kolondan ayrılmasını sağlayan ikinci bir ligand olmasıdır.

Saflaştırma sonrası protein örneklerinin beklenen moleküler ağırlığa sahip olup olmadığı ve saflaştırmanın ne derece başarılı olduğu SDS-PAGE analizi ile belirlenmiştir. SDS jellerinin görüntülenmesiyle literatürde belirtilen moleküler ağırlığına sahip φ29 DNA polimeraz sentezinin başarıyla gerçekleştirildiği kanıtlanmıştır.

In vitro çalışmalarda DNA polimerazların aktivitelerinin ölçülmesi zaruri bir ihtiyaç olarak karşımıza çıkmaktadır. Radyoizotop kullanılarak gerçekleştirilen geleneksel aktivite ölçüm metotları bu çalışmada getirdiği riskler sebebiyle tercih edilmemiştirler. Bunun yerine, yeni bir teknik olan moleküler işaret analizi yapılmıştır, çünkü bu analiz φ29 DNA polimeraz enziminin hızlı, kolay ve hassas karakterizasyonunu mümkün kılmaktadır. Polimerizasyon reaksiyonu boyunca özel DNA şablonunun yapısında bulunan florofor'a ait yayılım değerlerinin aralıklarla ölçülmesiyle aktivite testi tamamlanmıştır.

Aktivite testi sonuçlarına göre, ticari ve *in vitro* sentezlenmiş φ29 DNA polimerazlar yaklaşık aynı performansa sahiptir. Fakat aktivite testlerindeki her bir negatif kontrol örneklerindeki konakçı *E.coli* hücrelerine ait DNA tarafından kaynaklanan DNA kirliliği büyük bir problem yaratmaktadır. φ29 DNA polimeraz hem tek hemde çift zincirli DNA'ya karşı yüksek afiniteye sahip olduğu için, *E.coli* genomunun ve vektör DNA kalıntılarının hücre lizis aşamasında tamamen uzaklaştırılması gerekmektedir. Ancak afinite kromatografisi sonrası bu istenmeyen DNA parçalarından kurtulmak kolay değildir.

Çok fonksiyonlu rekombinant füzyon proteinleri farklı fonksiyonel protein bölgelerini aynı rekombinant protein yapısında bir araya getirmenin yanısıra proteinlerin biyoaktivitesini de artırmaktadır. Bu nedenle, RFP proteini ile etiketlenmiş diğer bir φ29 DNA polimeraz enzimi de önceden φ29 DNA polimeraz sentezi için uygulanmış heterolog protein ekspresyon prosedürü takip edilerek sentezlenmiştir.

φ29-RFP DNA polimeraz enzimi de yine Streptavidin bağlayıcı peptit ile afinite kromatografisi sayesinde saflaştırılmıştır. Enzimatik aktivite sonuçlarına göre, φ29-RFP DNA polimeraz enzimi de iyi bir sonuç vermektedir. İkinci olarak, SDS jel görüntülerine göre φ29-RFP DNA polimeraz enzimi bütünüyle ve istenilen moleküler ağırlıkta sentezlenmiştir, fakat jelde farklı protein bantları da görülmektedir. Bu çoklu bantların birçok farklı sebepleri olabilir. Örneğin, φ29-RFP DNA polimeraz enzimi amino asit bileşimine tamamen uymamış olabilir yada RFP ve φ29 DNA polimeraz birleşme noktasında ayrılma meydana gelmiş olabilir. Protein sentezinin yarıda kesilmiş olma ihtimali de yüksektir. Bunlar füzyon protein ekspresyonunda sıkça karşılaşılan problemlerdir. Fakat, farklı bir başlatıcıya sahip vektör kullanılarak bu sorunların üstesinden gelinilebilir.

φ29-RFP DNA polimeraz enzimi yapısında bulunan esnek bağlayıcıların füzyon proteine hareket kabiliyeti sağladığı bilinen bir gerçektir. Ancak, iki farklı proteinden oluşan bu yapının bozunmaması için yeterli dayanıklılığa sahip değildirler. Esnek bağlayıcılar yerine sabit bağlayıcıların kullanılması daha yüksek verimde istenilen proteinlerin elde edilmesine olanak sağlayacaktır.

Son olarak, UV-absorbsiyon ölçümü ile sentezlenen proteinlerin konsantrasyon değerleri belirlenmiştir. Her iki φ29 DNA polimeraz ve φ29-RFP DNA polimeraz enzimlerinin konsantrasyonları ticari φ29 DNA polimeraz ile kıyaslandığında düşük olduğu tespit edilmiştir. Deneysel süreçte meydana gelen bir takım olumsuzluklar bu düşük konsantrasyonun sebebi olabilirler. Bunların üstesinden gelmek için, indüksiyon sıcaklığı yada süresi değiştirilebilir. Bakteri kültürlerinin besi ortamı da içerik bakımından zenginleştirilirse üretimde verimliliğin artırılması muhtemeldir. Dahası, plazmid türünün yada ekspresyon sisteminin değiştirilmesi de çözüm olarak sunulabilir. Ayrıca protein örneklerinin ticari ürüne göre konsantrasyonları daha düşük olsa dahi, kayda değer büyüklükte aktivite değerlerine sahip oldukları görülmektedir.

1. INTRODUCTION

One-step whole genome amplification technologies allow for extensive genomic studies such as genome analysis of non-cultivable viruses or detection and identification of circular plasmids of some pathogens and replicative forms of viral genomes. φ29 DNA polymerase has gained ground in DNA amplification technologies thanks to its exceptional processivity, high proofreading activity and efficient DNA polymerization coupled with strand displacement activity in the absence of additional proteins or cofactors. This strand displacement activity has a great value in many amplification techniques since the displaced strand can be used as a long single strand DNA template for further DNA synthesis, without any denaturation step in polymerization [1-3].

Bacteriophage φ29 is among the smallest *Bacillus subtilis* phages and operate a natural rolling circle mechanism during the DNA replication process [4].

The linear double-stranded DNA genome is reproduced by help of φ29 DNA polymerase, the product of viral gene 2, consisting of a single peptide chain which has a molecular mass of only about 66 kDa [5-8].

The multiple enzymatic activities of φ29 DNA polymerase have been well identified by way of biochemical characterization studies based on site-directed mutagenesis, structure/function studies, and amino acid sequence alignments [5, 9-10].

Besides 5'-3' polymerisation activity like any other DNA polymerases, φ29 DNA polymerase repairs mismatched nucleotide by using its proofreading function that is provided by the 3'-5' exonuclease domain of protein [11-15]. φ29 DNA polymerase has also two distinguishable properties which contribute to complete genome replication and catalyze chain elongation: astonishing processivity and robust strand displacement activity, also called as helicase activity coupled with DNA synthesis. Throughout the initiation step of DNA replication process, φ29 DNA polymerase benefits from own DNA helicase activity to unwind double-stranded DNA. Moreover, φ29 DNA polymerase initiates the DNA synthesis by catalyzing the

formation of phosphodiester bond without an accessory processivity factors [5, 16-19].

As well as the usage of φ29 DNA polymerase in amplification reactions, the mutational analysis of φ29 DNA polymerase has contributed to the clarification of biochemical properties, both general and specific structure/function relations of DNA polymerases. However, only a limited number of genetic fusion proteins have been described [5-6, 20].

The aim of this project was the production and purification of φ29 DNA polymerase variants fused to a Streptavidin binding peptide to enable the immobilization of the polymerase on Streptavidin modified surfaces. Furthermore, a φ29 DNA polymerase labeled with a red fluorescent protein (RFP) should enable tracking of the enzyme via fluorescence detection.

2. THEORETICAL STUDY

All living organisms pass down the genetic instruction from parent to offspring in order to ensure the continuity of their lineage and maintaining life. The secret of this successful transmission is that organisms encode their own DNA polymerases which are responsible for the efficient replication and maintenance of genome. In addition to their role in replication, the DNA polymerases serve a major function in protein synthesis which is essential for the life cycle of organisms [21].

The existence of a DNA polymerase within *E.coli* cells was come to light in 1958 by Komberg and colleagues. After the discovery of novel polymerases, they were classified into families according to different physicochemical and biochemical properties. However, a chemical pathway for all DNA polymerases is the addition of deoxynucleotides into the growing end of short DNA primer strand regardless of DNA or protein synthesis processes [21].

Main replicative enzymes such as φ29 DNA polymerase belong to Family B which includes a great number of DNA polymerases originating from eukaryotic and prokaryotic cells. Additionally, many of polymerases in Family B contain exonuclease activity that ensures to act in reverse direction of DNA synthesis and repairs mismatched nucleotides such as G-A or G-T [22].

DNA-dependent DNA polymerases in Family B also referred to eukaryotic type or α-like polymerase, are indispensable for viruses since many of DNA viruses which spend their whole life in the cytoplasm are not enable to utilize from host DNA polymerases which accomplish multiple functions related to genome replication and transcription [6-7]. DNA polymerases catalyze the polymerization of nucleotides in the presence of divalent metal ions and specific primer which anneals to the template and provides a 3'- OH group. Incoming dNTPs are covalently attached by the polymerases that are only able to extend a new DNA strand in the direction of 5’ to 3’ end. The rest of complementary single-strand DNA is elongated according to A-T, G-C base pair rules. One molecule of pyrophosphate (PP_i) is released into the cytoplasm, as shown in Figure 2.1 [7, 23].

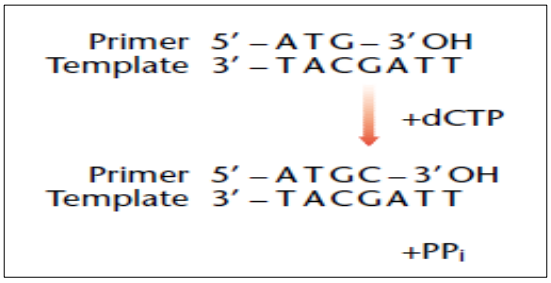

Figure 2.1 : The basic catalytic mechanism of DNA polymerisation reaction [23].

The isolation and amplification of significant DNA sequences is a fundamental necessity for molecular biology. *In vitro* nucleic acid amplification techniques such as polymerization chain reaction (PCR) and rolling circle amplification (RCA) have been developed to obtain very amount DNA fragments from limited amounts of DNA and synthesize long single-strand DNA molecules. Due to their robustness and simplicity, these techniques become crucial in a genome organization, mutation detection, clinical diagnosis etc [1, 24-27].

PCR is dependent on specific primer sequences and not convenient to amplify large DNA sequences [1, 28]. However, isothermal rolling circle amplification (RCA) offers a rapid, straightforward and inexpensive enzymatic process that amplify the considerable amount of long DNA strands by using a DNA polymerase including a suitable buffer. This method use circular nucleic acid structures and a DNA polymerase that have strand displacement ability and high polymerase processivity to generate growing repeats of the circular template [29-31].

DNA-dependent φ29 DNA polymerase gained a lot of interest in the field of DNA analytics in the recent years due to resulting sensitivity of the amplification assay [6-7, 15]. φ29 DNA polymerase is able to generate in high quality both a very complex circular DNA template and large linear DNA or cDNA in RCA reaction to detect point mutations in small amounts of genomic DNA in a few minutes. For instance, φ29 polymerase-dependent RCA using random hexamer primers generates hundreds of DNA copies by incorporating more than 70,000 nucleotides during a single binding event. Because of this amplification power, RCA attracts attention in the development of bioassays, nanobiomedicine, nanotechnology and material science [24, 32-35]. Second example; multifunctional DNA nanostructures consist of the

elongated and non-nicked DNA building blocks. Generation of long non-nicked DNA strands has been accomplished by rolling circle amplification using φ29 DNA

polymerase with high efficiency and robust strand displacement ability. These abilities of φ29 allow it to replicate a circular DNA template [36].

Lagunavicius and colleagues have presented that φ29 DNA polymerase convert target RNA into a primer for RCA thanks to its 3'-5' exoribonucleolytic activity. In the presence of this RNA-based primer that is complementary to the DNA template and dNTPs, amplification is begun [37].

On the other hand, φ29 DNA-dependent DNA polymerase has a high affinity for single-strand DNA and double strand DNA. Endogenous DNA contamination derives from expression vector used for recombinant φ29 DNA polymerase production and/or genome DNA of the host cells. This kind of contaminations easily affects the accuracy of polymerization chain reaction by resulting in non-specific DNA amplification in the reaction mixture. Furthermore, some reagent used in PCR also may include in DNA contamination.As a result, these amplifiable contaminating DNA should be removed after φ29 DNA polymerase production [38].

Blanco and colleagues have achieved the production of φ29 DNA polymerase in a highly purified form. *E.coli* NF2690 cells harboring plasmid pJLw2 were grown in LB medium. After overproduction of φ29 DNA polymerase in *E.coli* BL21DE3 competent cells, DNA fragments in the soluble extract were removed by polyethyleneimine (PEI) precipitation. Following that the obtaining PEI-free protein pellet, purification with Phosphocellulose, Blue dextran-agarose and Heparin-Sepharose columns was fulfilled in 99% purity and 33% yield [5].

Takahashi and colleagues have developed multiply-primed rolling circle amplification (MRPCA) using modified random RNA primers. However, the outcome of MPRCA has been adversely affected by DNA contamination in recombinant φ29 DNA polymerase. For this reason, they have described the preparation and purification protocols of φ29 DNA polymerase, which is free of amplifiable DNA fragments. GST-fused φ29 DNA polymerase was expressed in *E.coli* and affinity purification with glutathione sepharose 4B resin (GS4B) were performed after PEI precipitation. Then, the fusion protein was treated with ethidium monoazide (EMA) and irradiated with visible light (VL) to reduce both ssDNA and dsDNA captured by GST- φ29 DNA polymerase. After washing away the free EMA, endogenous 3'-5' exonuclease activity of the polymerase minimized the quantity of remaining oligonucleotides [38].

2.1 Heterologous Protein Expression

The production of heterologous proteins are of substantial interest for academic and commercial applications. The design of a suitable expression system relates with characteristic properties of proteins. Unfortunately, there is not any protocol or strategy covering all proteins to be expressed. Expression and purification strategies must be work out for each individual recombinant protein. Therefore, it is identified the essential stages of expression systems of desired recombinant proteins before starting to the experiments [39-41].

At theoretical level, method of expression of recombinant protein can be specified as mentioned below:

- Get your gene of interest,
- Clone it in an expression vector,
- Transform it into the host of choice and then induce for overexpression,
- Purify and characterize the obtained protein [42].

2.1.1 Design of expression system

Bacterial expression system is commonly used for recombinant protein expression due to its several advantages. However, problems arising from the nucleotide sequence of recombinant protein, protein contamination, improper folding of peptide, product degradation and protein insolubility are serious drawbacks of this system. Regardless of these drawbacks, alternative systems should be considered only after bacterial system has been tried. For example, *E.coli* as an expression host for fusion proteins is by far the most widely employed cells thanks to its rapid cell multiplication, low-cost, easily genetic modification and cultivation [40,43].

Overexpression of a gene in a foreign host resulted in some problems and one of the frequently encountered problems is that *E.coli* uses different codons from overexpressed protein and the differences in codon usage heads translation off due to lacking some tRNAs for the rarer codons in the expression host. To enhance expression of proteins that are encoded by rarely used codons by *E.coli* cells, BL21DE3 has been engineered to overcome the effects of different codon usage by supplying additional tRNAs under control of their native promoters. For routine recombinant protein expression, BL21DE3 is a convenient *E.coli* strain. A number of heterologous proteins were successfully expressed in *E.coli* BL21DE3 with T7 *lac*O

promoter system. T7 RNA polymerase that is under the control of *lac* operator elongates amino acid chains in *E.coli* cells. Additionally, ampicillin is the mostly used as antibiotic selection marker in bacterial expression system. On the other hand,there is an another way to deal with the codon utilisation. Optimisation in the original sequence of protein gene might be beneficial to avoid from troubles in translation and amino acid misincorporation [39, 42-43].

Each 20 amino acids are encoded by an average of three synonymous codons. If codons being in target protein structure are changed with synonymous ones existing in *E.coli* cells, then it will be overcome this problem in gene expression [44].

A wide array of vector systems for recombinant protein expression have been designed. pET-22b(+) expression vector consisting of an origin of replication, a transcriptional promoter, a selective marker, and a ribosome binding site (RBS) is responsible to express the recombinant protein under the control of T7 promoter and *lac* operator. After strong induction of T7 RNA polymerase with IPTG, the polymerase is produced and production of target protein is begun [39, 42, 45].

Protein affinity tags are used for high purity single-step purification of recombinant proteins expressed in *E.coli*. The streptavidin-binding peptide (SBP) engineered for optimal binding is 38 amino acids in length and used to immobilize fusion proteins on streptavidin surfaces. After The SBP tagged-recombinant protein is immobilized into streptavidin column matrix, the fusion protein is obtained with elution buffer which consists of D-desthiobiotin. The affinity of desthiobiotin for streptavidin is stronger than SBP, so the SBP-tagged fusion proteins can be efficiently eluted. In order to reutilize the same Strep-tactin column, desthiobiotin is removed by washing the column with HABA (2-[4’-hydroxy-benzenazo] benzoic acid) [46-47].

Besides combining the different functional domains of proteins in the same recombinant protein structure, flexible linkers are empirically designed with various sequence and conformations to increase the biological activity of fusion proteins by researchers. Second key feature of linkers is that it can significantly decrease unfavorable interaction between fusion proteins at the protein folding stage and derive a certain degree advantage in fusion proteins movement. Most commonly used flexible linker has $(GGGGS)_n$ sequence consisting of Glycine (Gly) and Serine (Ser) amino acids. By adjusting the length sequence, GS linker can be considerably enhanced to achieve appropriate separation between fusion protein domains [48-49].

2.1.2 Experimental procedure of protein expression

Introduction of protein gene into the host cell and the production of foreign protein in the chosen expression system are major steps of the heterologous protein expression. The experimental procedure can be specified as isolation of appropriate DNA expression vector and construction of recombinant protein vector via DNA assembly method, transforming into the desired host cell, confirmation of vector sequence and protein overexpression [43].

2.1.2.1 Isolation of DNA with polymerase chain reaction (PCR)

Polymerase chain reaction is a robust technique to synthesize DNA sequences in large quantity. In order to do a successful amplification, PCR needs several components that could extremely enhance the efficiency of polymerase chain reaction, and they are as follows;

Thermostable DNA Polymerase: Phusion® High-Fidelity DNA polymerase is a thermostable enzyme that generates long copies of DNA template with an accuracy during the high temperatures of PCR cycles as well as high reaction speed and processivity.

Primers: A pair of synthetic oligonucleotides being critical for DNA polymerase which can only prolong pre-existing single-stranded DNA pieces are designed as complementary sequences of plasmid DNA.

Template DNA: The circular pET22b(+) vector can be chosen as a template in recombinant protein construction.

Amplification Buffer: 5X Phusion HF Buffer and 5X Phusion GC Buffer including $MgCl_2$ acting as a cofactor for the enzyme are used to making a stable reaction environment.

dNTP Solution: Standard dNTP solution for PCR contains equimolar dATP,dTTP,dGTP,and dCTP nucleotides linked each other by polymerase enzyme to create a new complementary DNA strand of template DNA.

PCR additives: DMSO contributes to denaturation of template DNA with GC-rich content by breaking away Hydrogen bonds between complementary DNA strands [50-51].

2.1.2.2 DNA assembly reaction and bacterial transformation

NEBuilder® HiFi DNA Assembly Master Mix/ NEBuilder HiFi DNA Assembly Cloning Kit is performed to assemble a synthetic gene and DNA vector that has been linearised and amplified by using a high-fidelity Phusion polymerase. For this polymerization reaction, forward and reverse primers are designed to amplify DNA fragments with proper overlapping ends.

The HiFi assembly master mix includes in three different enzymes operating their special responsibilities in reaction. While the 5'-Exonuclease forms the overlap region and facilitates the annealing of DNA fragments that have overlapping ends, the DNA polymerase extends 3'ends. Lastly, DNA ligase seals nicks. By this way, fully assembled recombinant DNA was attained in one step. The DNA assembling mechanism is summarised in Figure 2.2.

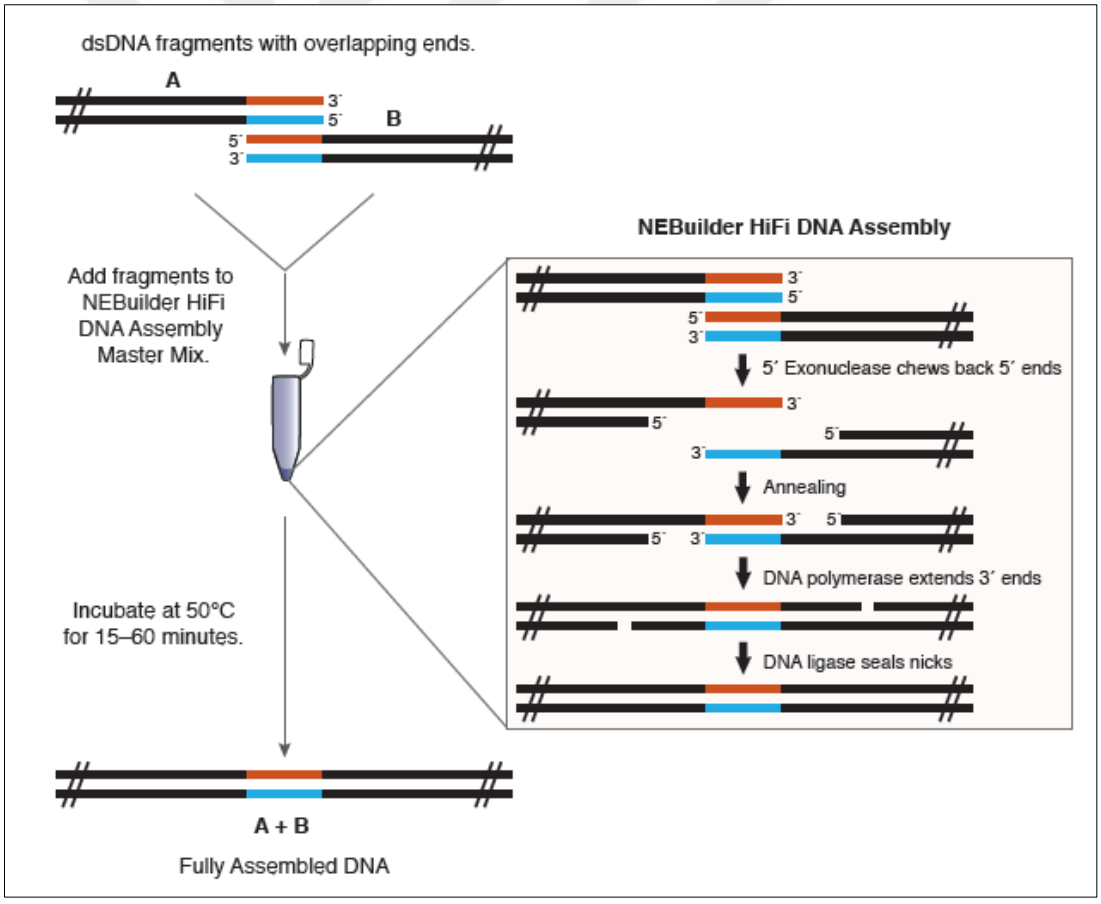

Figure 2.2 : NEBuilder® HiFi DNA assembly procedure [52].

Finally, the assembled recombinant vector is transformed into competent DH5α *E.coli* cells via Heat Schock method. DH5α cells are grown on LB (Luria-Bertani) agar plates with Ampicillin [52].

2.1.2.3 Protein purification

Preparation of the bacterial lysate is vital step to reduce protein oxidation and sample contamination with genomic DNA. Besides advantages of mechanical lysis by sonication and superspeed centrifuge, DNase and Lysozyme can accomplish to remove any particulate like cell-wall materials or viscous materials. The clarified bacterial lysate is also filtrated with vacuum-driven filtration system before loading on the affinity column [39].

Affinity chromatography is a high-tech workhorse that enables to separate the protein of interest on the basis of the specific and reversible interaction between the protein and its biospecific ligand immobilized into a support matrix. During the purification process, the biospecific ligand coupled to the column matrix covalently attaches to the target protein molecules in the sample, and all other proteins are eluted out. Competitive ligand is subsequently used to the elution of bound protein from affinity matrix column in a single step. Figure 2.3 shows the essential stages in an affinity purification [53-54]. After the affinity column is equilibrated with running buffer, target proteins reversibly bind to the ligand and unbound materials flow out. These proteins are recovered by using competitive ligand in elution buffer and collected in a purified form. Lastly, the column is re-equilibrated with running buffer.

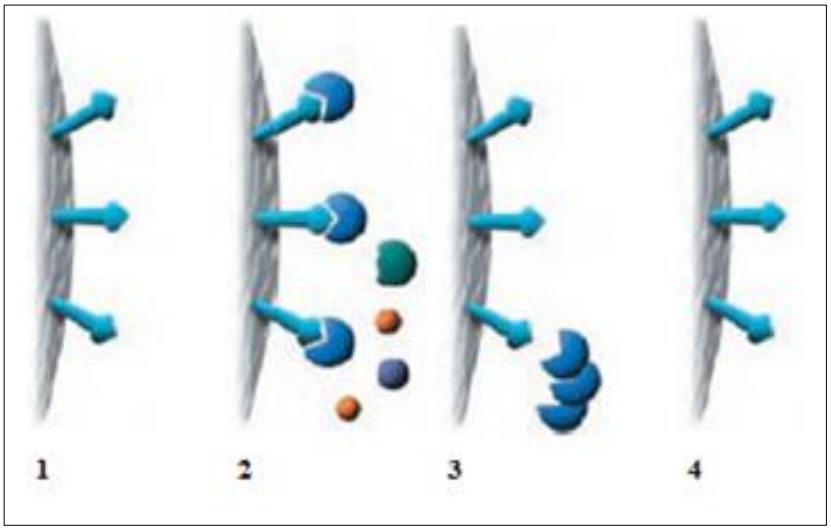

Figure 2.3 : General affinity purification steps [53].

In Figure 2.4, absorbance-column volumes graph indicates whether target proteins are captured throughout chromatography process or not. Moreover, absorbance amount is directly proportionate to the protein concentration [53].

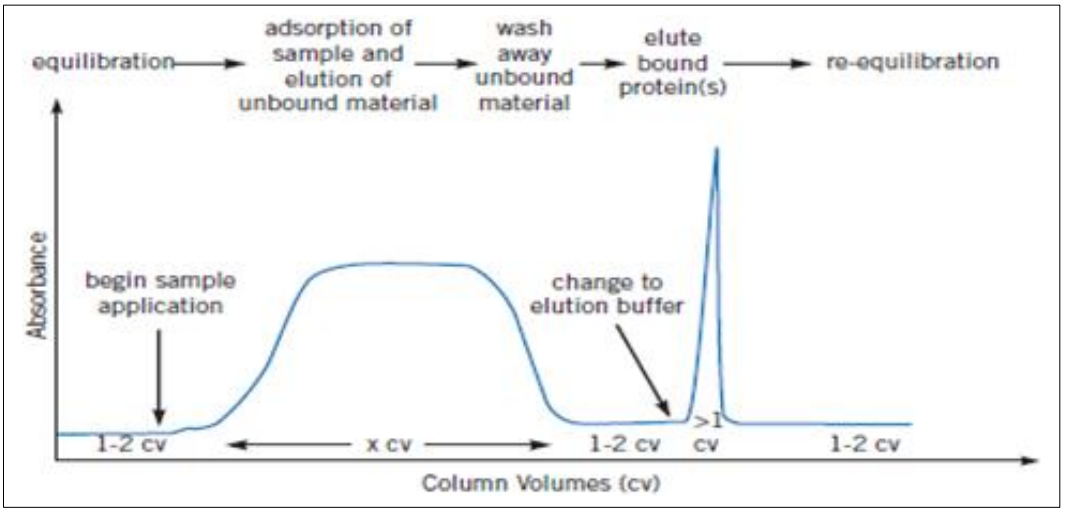

Figure 2.4 : Absorbance (A280) versus column volumes (cv) graph [53].

2.1.2.4 Protein characterisation

Characterisation protocols are necessary to clarify whether the correct proteins have been purified, also estimate the protein concentration, and active proteins [39].

Enzyme activity

In vitro measurement of DNA polymerase activity is indispensable after protein purification process. In traditional methods, undesirable radionucleotides that have inherent risks and limitations associated with radioisotopes have been consumed. However, molecular beacon based DNA polymerase activity assay, it becomes a fast and sensitive for single-stranded DNA polymerization reaction [55-56].

The enzyme activity test plays a significant role in the quantitative screening of overexpressed enzyme function after protein purification. The method allows fast parallel monitoring of φ29 and φ29-RFP DNA polymerases activities with the use of commercially available reagents in a multi-well plate reader. Besides concentration reagents, optimum reaction conditions such as pH, temperature and ionic strength are of great importance [57-58].

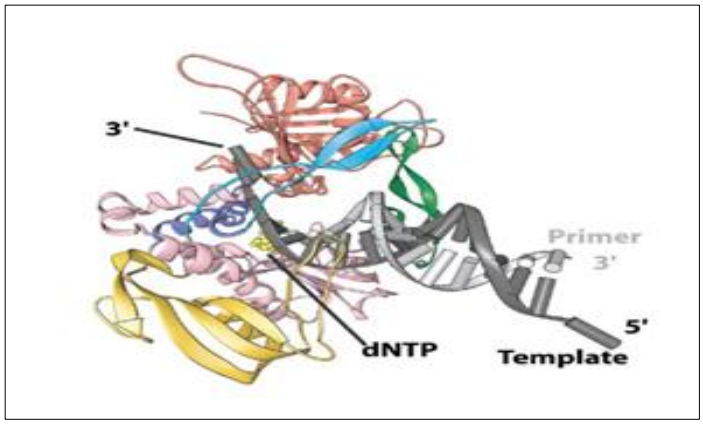

Figure 2.5 : Polymerisation reaction mechanism carried out by (colored) φ29 DNA polymerase [13].

The positions of polymerisation reaction components is demonstrated by using φ29 DNA polymerase structure in Figure 2.5 [13].

DNA polymerases extend the specific primer strand annealing to the complementary region of the template and open stem created by close combining of a fluorophore (FAM) and its quencher (TAMRA) ,shown in Figure 2.6. FAM emission values depending on the distance between fluorescence dye and acceptor are measured throughout polymerisation reaction by the multi-well plate reader [57].

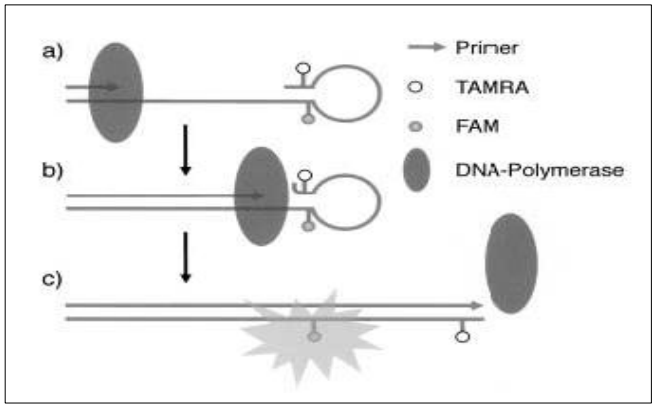

Figure 2.6 : Quantitative observation of the DNA polymerase reaction [57].

The emission values of FAM fluorescence excited at 485 nm wavelength are measured at 518nm in different time intervals. However, during the activity test of φ29-RFP DNA polymerase, RFP protein releases red fluorescence as well as the presence of FAM fluorescence in the reaction media. For the accuracy of activity tests, a multi-plate reader should only detect FAM fluorescence emission. Fortunately, emission regions of RFP protein indicated in orange are far away from

FAM emission region in accordance with the percent transmission (%T) versus wavelength (nm) graph given in Figure 2.7.

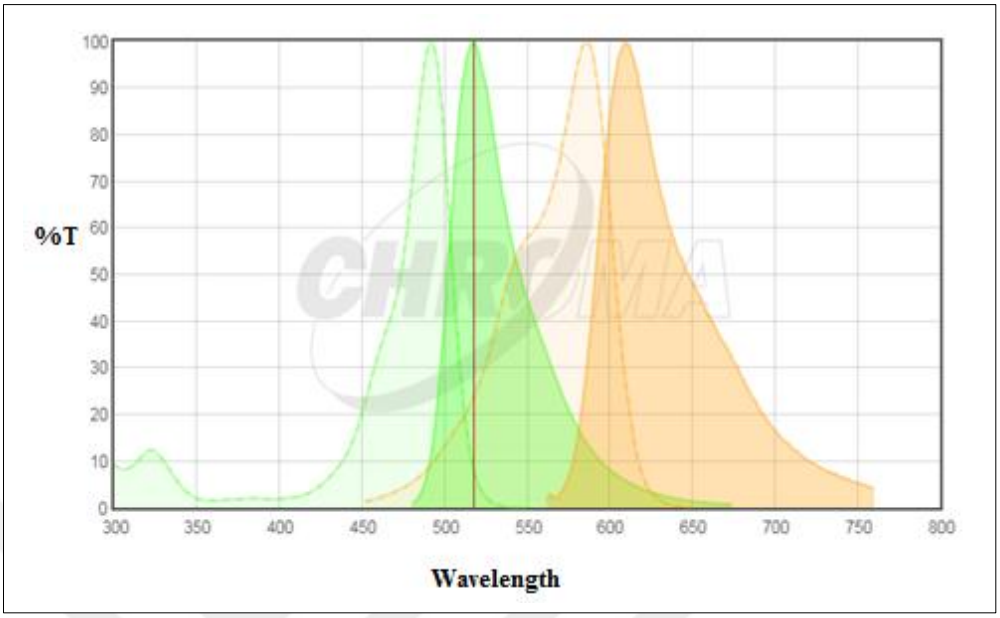

Figure 2.7 : Percent Transmission (%T) vs. Wavelength (nm) [59]. Light green region: Excitation spectra of FAM. Dark green region: Emission spectra of FAM. Light orange region: Excitation spectra of RFP. Dark orange region: Emission spectra of RFP [59].

UV absorption spectroscopy

To quantify the concentration of purified proteins, UV absorbance spectrums at A_{280} are measured and concentrations of protein samples are calculated by using Beer-Lambert Law.

$$A = \varepsilon \times d \times C \text{ [60]} \quad (2.1)$$

A= measured absorbance at 280 nm wavelength, ε= extinction coefficient (wavelength-dependent molar absorptivity coefficient with units of $M^{-1}cm^{-1}$), d = path length, cm, C= protein concentration, M [60].

SDS-PAGE analysis

After protein purification, SDS-PAGE is performed to evaluate the purity of protein samples and whether the purified proteins are of the expected molecular weights. In coomassie stained SDS gel, the intensity of the bands demonstrate the approximate amount of proteins [39].

3. EXPERIMENTAL STUDY

3.1 Materials

In order to achieve a DNA amplification with high efficiency, a PCR protocol with Phusion® High-Fidelity DNA Polymerase was performed. 5X Phusion® HF Reaction Buffer, 5X Phusion® GC Reaction Buffer and 100% DMSO were purchased from NEW ENGLAND BioLabs®Inc. 124-KSR-for, 125-KSR-rev, Phi29-SBP for, Phi29-SBP rev, 15-SG for and 16-SG rev primers purified in HPLC quality, hairpin-template, and its primer were synthesized by Sigma Aldrich.

After agarose gel electrophoresis had been carried out by using Biozym LE Agarose obtained from Biozym Scientific GmbH, Zymoclean™ Gel DNA Recovery Kit supplied from Zymo Research Corporation was applied.

NEBuilder® HiFi DNA Assembly Master Mix/ NEBuilder HiFi DNA Assembly Cloning Kit, NEW ENGLAND BioLabs®Inc. for DNA assembling and the bacterial transformation was conducted with given protocol by the company. Subsequently, ZR Plasmid Miniprep™-*Classic* commercial kit obtained from Zymo Research Corporation was implemented.

D-Desthiobiotin ($C_{10}H_{18}N_2O_3$), Strep-Tactin® Superflow Cartridge and 10X Buffer R were supplied from IBA GmbH Solutions for Life Science to use in protein purification process.

The other chemicals and consumables that were essential to the high-quality research during thesis study were supplied from Merck KGaA, Thermo Fischer Scientific, Sigma Aldrich and Panreac AppliChem GmbH.

Bacterial plasmid vectors, primers, DNA polymerases, isolated DNA fragments and other consumables used in amplification reaction were stored at -20^0C. Expressed fusion proteins and competent cells were kept at -80^0C.

3.2 Equipment

- Affinity chromatograply system, ÄKTA™ pure-laboratory scale-Chromatography System, Protein Research GE Healthcare Life Science
- Balance, ED2202S-CW Sartorius Extend Precision Balance
- Balance, Sartorius CPA64 Analytical Balance
- Box for sonication, SB1-16 Sound Protection Box, Hielscher Ultrasonic GmbH
- Benchtop micro centrifuge, Centrifuge 5415 R, Eppendorf
- Centrifuge, Sorvall LYNX 6000 Superspeed Centrifuge, Thermo Fischer Scientific
- Centrifuge, SU 1550 mini-spin centrifuge, Sun Lab
- Centrifuge, Multipurpose Centrifuge 5804 R, Eppendorf
- Electrophoresis system, Wide Mini-Sub® Cell GT Horizontal Electrophoresis System, 15x10 cm tray with PowerPac™ Basic Power Supply,Bio-Rad
- Freezer, Forma™ 900 Series -86^{0}C Upright Ultra-Low Temperature Freezer, Thermo Fischer SCIENTIFIC
- Filtration system, Steriflip® Vacuum-driven Filtration System, Merck
- Visualisation system, FluorChem M System ,ProteinSimple
- Ice machine, AF103 Ice Flaker, Scotsman
- Incubator, BD 115 Incubator with Gravity Convection, BINDER
- Microwave oven, MW 2288 CB Microwave, BOMANN
- Microplate reader, Synergy H1 Hybrid Multi-mode Reader, BioTek
- Orbital shaker, Unimax1010, HEIDOLPH
- pH meter, pH1120 Portable pH Meter, METTLER TOLEDO
- Photo scanner, CanoScan LiDE 210, Canon
- Photometer, BioPhotometer plus, Eppendorf
- Rotator, Stuart™ SB2 tube rotator , Merck
- Stirrer, Variomag MONOTHERM Electronic Stirrer
- Shaking incubator, Multitron Pro, INFORS HT
- SDS-PAGE system, Mini-PROTEAN® Tetra Cell, Bio-Rad
- Thermocycler, Eppendorf ™ Mastercycler ™ pro with vapo. protect PCR System,Thermo Scientific
- Thermostat, BTD Dry Block Heating System, Grant

- Transilluminator, Gel documentation system Super Bright, PEQLAB Biotechnologie GmbH
- Thermo shaker, PHMT-PSC18 Thermoshaker for Microtubes, Grant-Bio
- Ultrasonic homogenizer sonicator, SONOPULS HD2200 with MS73 Microtip, BANDELIN
- Vortex, Analog Vortex Mixer, VWR
- Water bath, G.HEINEMANN Ultraschall und Labortechnik

3.3 Methods

3.3.1 Cloning procedure

The *E.coli* pET22b(+) vector was transformed that it could easily produce SBP-tagged φ29 DNA polymerase variations originating from the 'Gene 2' of Bacteriophage φ29 [5].

The recombinant plasmids have been created for two different protein expression ;

pET-22b(+)-φ29 with SBP tag: As soon as the linearization of pET-22b(+) plasmid vector were completed by PCR with primers that help to open circular plasmid from the desired location of the vector, the plasmid DNA and the synthetic gene were brought together in the new recombinant DNA structure. In other words, a synthetic gene encoding φ29 DNA polymerase protein was inserted into pET-22b(+).

pET-22b(+)-φ29-RFP with SBP tag: Red Fluorescence Protein (RFP) was isolated from a previously cloned recombinant pET-22b(+) - RFP vector by PCR. [53] Meanwhile, pET-22b(+)-φ29 DNA polymerase circular vector was linearized by PCR, as well. Subsequently, linearised pET-φ29 DNA polymerase vector and amplified RFP gene were gathered in the same DNA structure by means of HiFi DNA assembly.

3.3.1.1 Polymerase chain reaction (PCR) with Phusion DNA polymerase

Two PCR mixtures which include all reaction components except DNA polymerase are assembled in the first and fifth of a sterile 0.2mL, 8 Strip- PCR tubes in a volume of 100 µL in order to decrease pipetting errors in small volumes. The mixtures were vortexed and then centrifugated by mini centrifuge. Phusion polymerase was added as a final component because its 3'-5' exonuclease activity might degrade the primers

in the lack of dNTPs [50]. It is important that Phusion HF buffer could achieve higher accuracy at nucleotide sequence of complementary DNA than GC Buffer. However, GC buffer can improve the reactions performance in case the usage of templates that have more Guanine and Cytosine nucleotides. Therefore, PCR mixtures were prepared with these two different buffers as stated in Table 3.1 [50].

Table 3.1 : PCR mixtures with Phusion DNA polymerase [50].

First PCR tube including of 1X HF Buffer		**Fifth PCR tube including of 1X GC Buffer**	
Components	Quantity	Components	Quantity
Nuclease-free H_2O	66 µL	Nuclease-free H_2O	66 µL
5X Phusion HF Buffer	20 µL	5X Phusion GC Buffer	20 µL
10µM forward primer	5 µL	10µM forward primer	5 µL
10µM reverse primer	5 µL	10µM reverse primer	5 µL
10mM dNTPs	2 µL	10mM dNTPs	2 µL
Template DNA	1 µL	Template DNA	1 µL
Phusion DNA polymerase	1 µL	Phusion DNA polymerase	1 µL

Following that these two reaction solutions were divided into each of 8-Strip PCR tubes in the volume of 23 µl, DMSO was stirred in the volumes mentioned in Table 3.2 [50];

Table 3.2 : DMSO quantities in PCR mixture [50].

PCR mixtures including of 1X HF Buffer		**PCR mixtures including of 1X GC Buffer**	
Tube Number	DMSO Quantity	Tube Number	DMSO Quantity
1	0 µL	5	0 µL
2	1 µL	6	1 µL
3	2 µL	7	2 µL
4	3 µL	8	3 µL

The concentration of DMSO directly affects melting temperature of primers during PCR. Thus, DMSO was mixed with different volumes to find optimum concentration at final reaction mixtures. Eventually, 8-strip PCR tubes were quickly centrifuged and transferred into thermocycler which was run according to special thermal conditions as described in Table 3.3 [50];

Table 3.3 : The temperature profile of PCR.

Step	Temperature	Time	Cycles
Initial Denaturation	98 ^{0}C	30 seconds	1 cycle
	98 ^{0}C (denaturation)	10 seconds	
Cycles	65 ^{0}C (annealing)	20 seconds	35 cycles
	72 ^{0}C (elongation)	3 minute 30 seconds	
Final Extension	72 ^{0}C	3 minute 30 seconds	1 cycle
Hold	4 ^{0}C	Forever	-

Firstly, the temperature of PCR process has been raised into 98 ^{0}C in order to seperate double-stranded DNA, and secondly decreased to 65 ^{0}C which allows the annealing of two forward and reverse primers. Lastly, increasing the temperature to 72 ^{0}C enables the Phusion DNA polymerase to elongate the templates. After running 35 cycles, few copies of DNA have been exponentially amplified about a billion copies [30]. Appropriate elongation and extension time were estimated according to 15-30 seconds/kb ratio [50].

The design strategies of primers

Specificity and efficiency of DNA amplification depend on the oligonucleotide primer's design. To reduce unrelated amplicons and increase of the product for each PCR cycle, annealing temperature (T_m) of primers should be taken into consideration. For example, the optimum primer length that is generally in between 18 and 24 bases and %50 G+C content provide T_m values in the range of 56-62 ^{0}C. This temperature values are adequate for efficient annealing [61].

A circular DNA molecule is created by PCR fragments that share overlapping ends. To create this DNA structure, PCR primers have to be composed of an overlap sequence and a gene-specific sequence for use in HiFi DNA assembly. While the overlap sequence is indispensable for the assembly of adjacent DNA fragments, the gene-specific sequence is necessary for the duplication of template strand during PCR [52]. Another issue is that the length of overlap sequence is dependent on the GC content within whole primer sequence. However, the length of the gene-specific sequence is determined by taking annealing temperature (Tm) of PCR into consideration [52].

The circular plasmid vector pET-22b(+) was linearized by performing polymerization chain reaction with 124-KSR forward and 125-KSR reverse primers

that open the vector from T7 expression region (App. A, Table A.1 and A.3). Therefore, single-stranded DNA sequencing was performed with T7 up and T7 term primers by LGC, Germany.

After the first polymerization reaction, amplified pET-22b(+) vector was assembled with the SBP tag-φ29 DNA polymerase synthetic gene by following HiFi DNA Assembly method. It is worthy of note that the SBP tagged-φ29 DNA polymerase synthetic gene was comprised of sticky ends, also known as overlapping ends, as stated above.

Flexible linker gene (App. A, Table A.2) shown in Figure 3.1 has been inserted between SBP tag and φ29 polymerase by Invitrogen GeneArt Gene Synthesis.

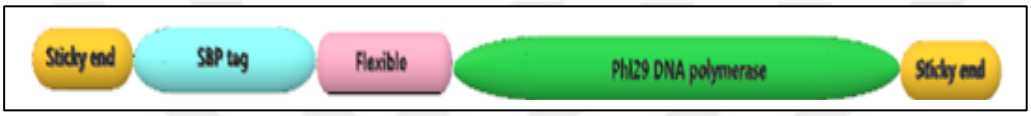

Figure 3.1 : The components of synthetic φ29 DNA polymerase gene produced by Invitrogen GeneArt Gene Synthesis.

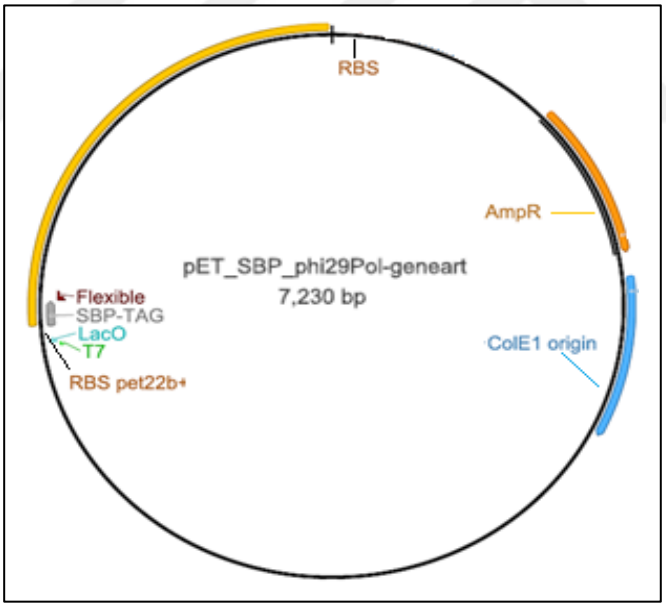

Figure 3.2 : The circular structure of pET-22b(+)-SBP tag-φ29 DNA polymerase gene, designed in Geneious 8.0.

As shown in Figure 3.2, SBP tag-φ29 DNA polymerase synthetic gene was inserted into ribosome binding site (RBS) of pET-22b(+).

In the expression of SBP tag-φ29-RFP fusion protein, another recombinant pET-RFP vector that had been previously cloned was used as an RFP gene source. As the

frame of vector has been known, PCR was used to amplify the target RFP sequence using oligonucleotide primers given in App.A, Table A.4.

As shown in Figure 3.3, SG-15 forward and SG-16 reverse primers attached to the beginning and end of RFP gene in opposite directions. By means of that, RFP gene was extracted from pET vector and produced in a large quantity by PCR.

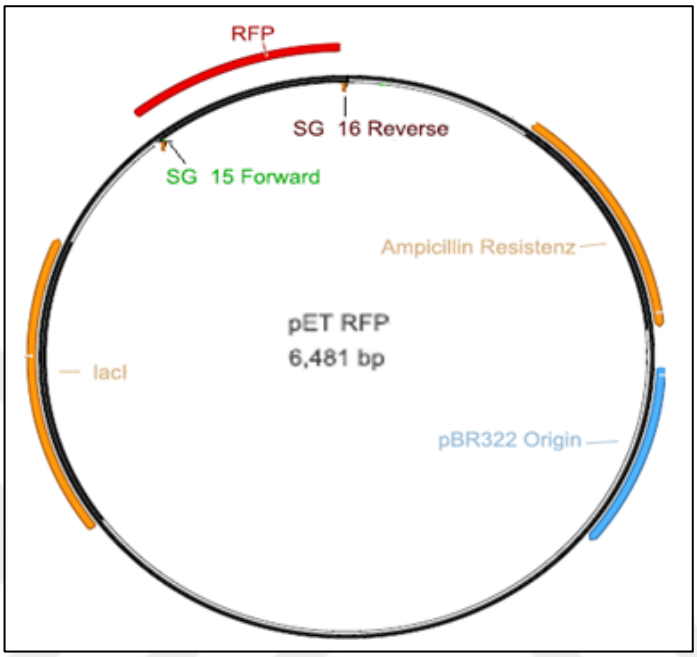

Figure 3.3 : The circular vector structure of pET-22b(+)-RFP gene, designed in Geneious 8.0.

After pET-22b(+)-SBP tag-φ29 DNA polymerase circular vector had been linearised by φ29-RFP forward and reverse primers, amplified RFP gene was inserted in between SBP-tag and flexible by performing HiFi DNA Assembly method.

Table 3.4 : Phi29-RFP forward and reverse primers, designed in Geneious 8.0.

Phi29-RFP forward primer	**Tm(C^0)**	**%GC content**	**Length (bp)**
5′-GGTCGTCATAGTACCGGTGCC GGTGGTGGTGGTAGCGGT-3′	61,6	66,7	39
Phi29-RFP reverse primer			
5′-GATCACATCTTCGCTGCTTGC CGGTTCACGCTGACCCTGCGG-3′	60,7	66,7	39

Overlap sequence of forward and reverse primers shown in red in the Table 3.4 were complementary sequences of RFP gene. Additionally, the rest of Phi29-RFP forward and reverse primers involved in the gene-specific sequences of φ29 DNA polymerase and SBP-tag, respectively.

The final structure of pET-22b(+)-SBP tag-φ29-RFP DNA polymerase vector is demonstrated in Figure 3.4.

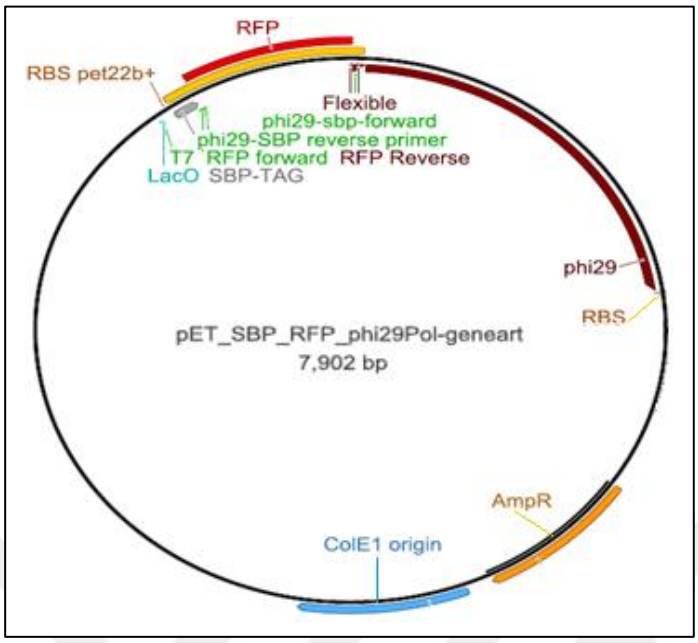

Figure 3.4 : The circular vector structure of pET-22b(+)-SBP tag-φ29-RFP DNA polymerase, designed in Geneious 8.0.

3.3.1.2 Agarose gel electrophoresis

Agarose gel electrophoresis is a favorable method to seperate the DNA fragments of varying sizes [62].

Weighed 1 g agarose powder was placed into a 250 ml-glass Erlenmeyer flask and dissolved in 100 ml 0.5X TBE buffer (App. D, Table D.2). The 1% agarose slurry was heated in the microwave until the all of the grains of agarose was completely dissolved for about 2-3 min. Molten agarose solution was cooled for a while, and 5 µl peqGreen used as nucleic acid staining solution was added to the agarose solution. After stained agarose solution had been poured into a gel casting tray with a comb that is used to make pockets for loading of DNA samples, gel solidification was completed for at least 15 min. Once the comb was removed, casting tray was placed into electrophoresis tank which was fully loaded with 0.5XTBE buffer.

Besides pipetting 5 µl Gene Ruler™ DNA Ladder Mix into an agarose gel pocket, the mixtures consisting of 1 µl 6X DNA Loading Dye and 5 µl of each PCR samples were centrifugated and slowly loaded into gel pockets in a row. The moment electrophoresis system was connected properly and switched to 120 Volt, 350 mA, 150Watt, 30 min, DNA started to migrate from negative to positive electrode thanks

to negative phosphate ions originating in the double helix structure of DNA. After seperation, the gel with DNA fragments was visualized under UV light [51,62].

Purification of DNA fragments

PCR products appearing at pretty much the same size in the first agarose gel image were picked up in the same vial to isolate linearised plasmid vector from PCR reaction mixtures by using agarose gel electrophoresis applied exactly the same as the previous method. However, the second gel was formed with relatively large pockets in order to purify all PCR samples in one run.

After the image of gel had been visualized, the gel was placed under transilluminator that supplies super bright light, and DNA bands were cut from the gel with a knife. Zymoclean ™ Gel DNA Recovery kit was used to extract pure DNA fragments from agarose gel particles. This gel purification kit was used by following given protocol by manufacturer instructions [63].

3.3.1.3 Bacterial transformation

The recombinant vector that has the fusion protein gene and ampicillin resistance were taken into commercial *E.coli* DH5α competent cells via Heat Shock method that was performed by following the HiFi Assembly Cloning kit protocol given by the manufacturer and transformed cells were grown over ampicillin-LB agar plates at 37^0C in the incubator (App. D, Table D.3). On the next day, colonies were picked from LB-agar plates with inoculation loop, transported into 5 ml Mini-cultures that comprise of ampicillin and fresh LB medium at 1:1000 ratio (App. D, Table D.3) and incubated overnight at 37^0C in incubation shaker. For efficient isolation of plasmid DNA from *E.coli* cells that grew inside Mini-cultures, ZR Plasmid Miniprep™-*Classic* commercial kit was processed and DNA concentration was measured at 260 nm wavelength in the multi-mode reader [64].

3.3.2 Procedure for the overexpression of protein

Single-stranded sequencing was carried out to confirm and identify cloned plasmids by LGC Genomics GmbH, Berlin, Germany. On the purpose of protein expression from the limited amounts of recombinant vector, plasmids that have got verified DNA sequences were introduced into chemically competent *E.coli* BL21DE3 cells via Heat Shock method and grown over ampicillin-LB agar plates o/n at 37^0C in the incubator. BL21DE3 bacterial colonies picked from LB-agar plates were transferred

into 30 mL Pre-cultures that were formed fresh LB medium supplemented with ampicillin at 1:1000 ratio and incubated o/n at 37^0C in the incubation shaker. The next day,1L Main-culture made of ampicillin and LB-medium at 1:1000 ratio was agitated with 20 ml pre-culture suspension to get in proportion as 1:50 and incubated more than 2 hours at 37^0C in the incubation shaker. During 2 hours, optical densities of samples from main cultures were measured at the 600nm wavelength in biophotometer. Once optical density values reached 0.8-0.9, 1M, 100 μL IPTG was added all main cultures to induce protein expression. Induced main cultures were incubated o/n at 25 ^{0}C [51].

To harvest bacterial cells, main cultures were centrifugated with 6x1000 LEX rotor at 10000 x g-force and 4 ^{0}C for 10 minutes in the superspeed centrifuge. The supernatants were discarded and bacterial cell pellets were resuspended in 20 ml running buffer that was used in ÄKTA chromatography system. Bacterial cell suspensions were kept at -80 ^{0}C till protein purification.

3.3.3 Protein purification

A frozen bacterial cell suspension kept at -80 ^{0}C was thawed in water-bath at 37 ^{0}C for almost 30 minutes. The moment trace amount DNase I and Lysozyme were added to the thawed bacterial cell suspension, it was rotated for 30 minutes in the rotator.

The cell suspension placed in ice was disrupted by ultrasound sonication supplied from a proper micro tip of ultrasonic homogenizer adjusted to the max amplitude 65%, 3x times for 30 seconds, and then the suspension was centrifugated at 45000 x g-force with A23-6x100 rotor of superspeed centrifuge for 1 hour. All these implementations were necessary to break down bacterial cell walls and get rid of chromosomal DNA of *E.coli*, cell debris, aggregates etc.

In the final stage of decontamination, the clear lysate was loaded into Vacuum-driven Filtration System to perform the affinity chromatography with crude protein extracts.

Before the separation of φ29-DNA polymerase protein variants from other unspecific proteins synthesized by *E.coli* was performed by ÄKTA, the synthetic φ29 gene had been tagged with a Streptavidin-binding-peptide (SBP) to enable the immobilization of the polymerase on Streptavidin-modified surfaces.

A clear lysate was loaded into a 50 ml super loop connected to ÄKTA pure lab-scale protein purification system, from GE Healthcare, operated by Unicorn 6.3 software. In the course of purification run, SBP-tagged fusion proteins were immobilized into 5 mL Strep-Tactin® column that was run with Buffer W (App. D, Table D.1), and then eluted from the column with Buffer E consisted of D-Desthiobiotin that competes with the SBP-tagged fusion protein (App. D, Table D.1). After the regeneration of column with Buffer R (App. D, Table D.1) had been processed, the column was washed with Buffer W and kept within the same buffer.Meanwhile, 20 µl samples from protein fractions and the flow through were taken for SDS-PAGE analysis.

Fusion polymerases were collected with Buffer E throughout ÄKTA process at 96 well Polypropylene microplates and storage blocks. However, this elution buffer has to be exchanged with the Storage Buffer of fusion polymerases for long-term usage (App. D, Table D.1). Because of that, a centrifugal concentrator was used for buffer exchange application in a multipurpose centrifuge. Moreover, fusion proteins were acquired at desired concentration level by use of the concentrator.

Protein Spectrum programme in multiplate reader was used to find the concentration values of proteins. The absorbance measurement of 2 µL protein samples put into Take3 plate was started from 230 nm and extended to 700 nm wavelengths in 10 seconds intervals during 10 minutes.

3.3.4 Enzyme activity test

In order to test the activity of proteins expressed from cloned genes, DNA polymerase reaction mixtures in 96 well UV microplates were prepared with hairpin-template FAM, specific primer and other reagents given in Table 3.5. FAM fluorescence was excited at 485 nm, and emission was measured by microplate reader at 518 nm wavelength in 30 seconds intervals during 30 minutes. While positive control with commercial φ29 DNA polymerase was performed to compare with expressed proteins in terms of efficiency, the negative control was carried out in the absence of primer. All reactions were performed at 37^0C [65].

Table 3.5 : Reaction mixture of enzyme activity test.

Reagents	Volume, µL
Hairpin-template (100 µmol/L)	2
Primer (10 µmol/L)	3
φ29 DNA polymerase	2
dNTPs (10 mmol/L)	2
10X φ29 DNA poly.reaction buffer	10
Nuclease-free water	81

3.3.5 SDS PAGE analysis

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was used to inspect the quality of proteins processed in ÄKTA purification system. Negatively charged protein samples by sodium dodecyl sulfate (SDS) migrate according to their specific size from negative to the positive electrode in the electrical field [51].

The SDS-PAGE separating gels containing 16% acrylamide were prepared according to the volume of components given in Table 3.6. Separating gel solution was filled between two glass plates, and isopropanol overflowed. After the completion of polymerization reaction, isopropanol was poured into the sink and collecting gel was loaded into the top of separating gel. A comb was placed into S gel to form the pocket for protein loading, and it was removed almost 20 minutes later.

Table 3.6 : SDS gel preparation.

Separating Gel(T Gel), for 6 gels	Collecting Gel(S Gel)
5.99 mL H_2O	10.2 mL H_2O
16.0 mL, 30% Acrylamide mix	2.53 mL, 30% Acrylamide mix
7.5 mL, 1.5M Tris (pH 8.8)	1.88 mL, 1.5M Tris (pH 8.8)
0.3 mL, 10%SDS	0.15 mL, 10%SDS
0.3 mL, 10%Ammonium persulfate	0.15 mL, 10%Amonium persulfate
0.012 mL TEMED	0.15 TEMED

In SDS-polyacrylamide gel electrophoresis, 1, 5, 10 µL protein and flow through samples were mixed with 14, 10, 5 µL nuclease-free water, respectively, and 5 µL 6X SDS loading dye was added to all of them. Protein and flow through samples were briefly centrifuged in the mini-spin centrifuge and denaturated in the thermostat at 95^0C for 5 minutes. SDS-page electrophoresis tank incorporated in gel cassette

were filled with 1X Maniatis buffer (App. D, Table D.2). Page Ruler™ Plus Pre-stained used as protein ladder, protein and flow through samples were loaded into the gel in a row. The SDS-gel was run at 50 mA for 30 minutes.When the electrophoresis run out, the buffer was discarded. Gel-separated from glass plates was stained with Coomassie blue staining solution on an orbital shaker for 30 minutes. After Coomassie solution had been removed, the gel was washed with distilled water for several times. Once SDS-gel decolorized, protein bands were visualized by scanning.

4. RESULTS AND DISCUSSION

The φ29 and φ29-RFP DNA polymerases were cloned into pET-22b(+) vector and expressed according to the concept of cloning and expression as mentioned earlier in Materials and Methods. It has resulted in a successful conclusion both φ29 DNA polymerase and φ29-RFP DNA polymerase production. After accomplished protein purifications were finalized, agarose gel electrophoresis, single-stranded DNA sequencing, protein concentration measurements, enzyme activity test and SDS-page were performed as characterization studies.

4.1 Cloning, Heterologous Expression of φ29 and φ29-RFP DNA Polymerases

4.1.1 SBP-tagged φ29 DNA polymerase

The length of DNA fragments obtained from the Phusion PCR amplification that was executed for SBP-tagged φ29 DNA polymerase synthesis was observed after agarose gel electrophoresis. In Figure 4.1, Gene Ruler DNA ladder mix shown in M1 and M2 columns were used to determine the approximate quantification of amplified pET-22b(+) vector. DNA bands that belong to pET-22b(+) with 5493 bp are obviously seen between 5000 and 6000 bp bands. Thus, it is concluded that Phusion PCR amplification has been worked out successfully in the reaction mixtures number 1-3-4-5-8. Additionally, slightly seen bands labeled as impurity belong to undesired short DNA fragments. Due to the existence of these impurities in reaction mixtures, second agarose gel electrophoresis has been applied to get rid of them. Finally, 2.mixture has included in a vast amount of impurities and 7-8 reaction mixtures have not possessed desired DNA bands. Therefore, they have not been preferred for further experiments.

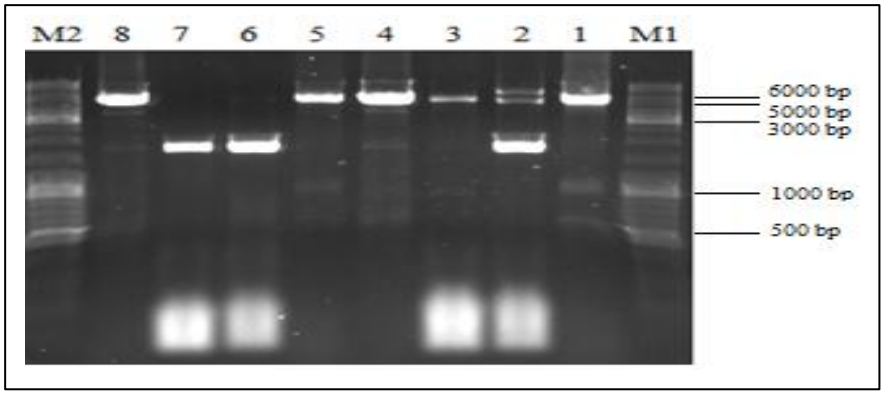

Figure 4.1 : The agarose gel image of amplified pET-22b(+) vector.

Following that PCR mixtures which have DNA template in desired length were mixed together in a single tube, second agarose gel electrophoresis was applied for the purpose of isolation and decontamination of DNA template from impurities.After gel visualization, DNA bands seen in the 5000-6000 base pair range, indicated in Figure 4.2, were cut away from agarose gel. To recover ultrapure DNA from agarose gel, Zymoclean™ Gel DNA Recovery Kit was carried out. The pET-22b(+) vector and synthetic SBP tagged-φ29 DNA polymerase gene was assembled in the same structure thanks to HiFi DNA Assembly Master Mix and then protein synthesis was started.

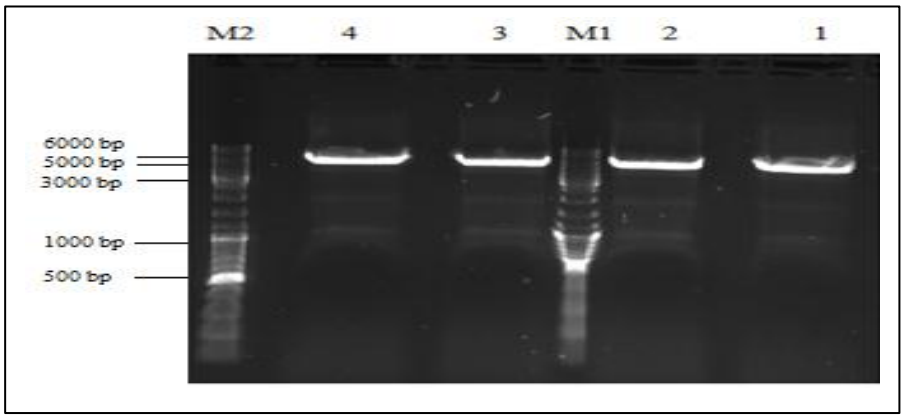

Figure 4.2 : The second agarose gel image of amplified pET-22b(+) vector.

4.1.2 SBP-tagged φ29-RFP DNA polymerase

The recombinant pET-22b(+)-SBP tag-φ29 DNA polymerase vector served as a template for the second φ29-RFP DNA polymerase protein expression. This circular recombinant vector was linearized and amplified by Phi29-RFP forward and reverse primers. It will be seen when Figure 4.3 is compared with Figure 4.1, DNA bands that belong to amplified pET-22b(+)-SBP tag-φ29 DNA polymerase vector are

longer than amplified pET-22b(+) DNA fragments. The reason for this, pET-22b(+)-SBP tag-φ29 DNA polymerase vector has DNA sequence encoding φ29 and SBP tag as well as pET-22b (+). It is clear that 2-3-6 reaction mixtures had the DNA fragments with the desired length.

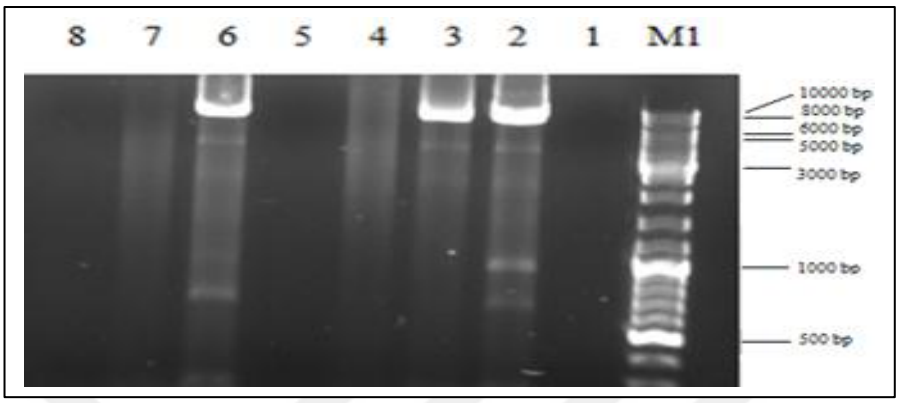

Figure 4.3 : The agarose gel image of amplified pET-22b(+)-SBP tag-φ29 DNA polymerase vector.

A second agarose gel electrophoresis was performed again to remove the impurities that present in the PCR mixtures, as it is seen in Figure 4.4, and the same method was followed to retrieve DNA fragments from agarose gel.

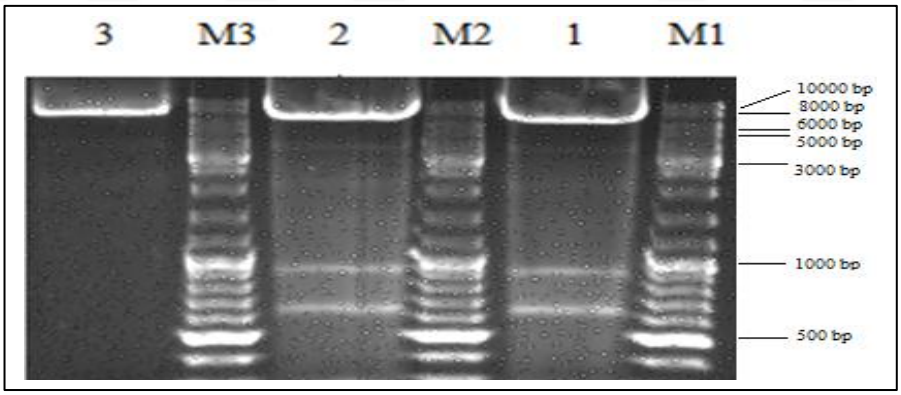

Figure 4.4 : The second agarose gel image of amplified pET-22b(+)-SBP tag-φ29 DNA polymerase vector.

4.1.3 Red flurescence protein (RFP)

The pET-22b(+)-SBP tag-φ29-RFP DNA polymerase vector was designed to generate the SBP tagged φ29-RFP DNA polymerase. In this new recombinant vector structure, a different recombinant vector, pET-22b(+)-RFP, was used as the RFP gene source which was isolated and amplified from pET-22b(+)-RFP vector with the aid of SG-15 forward and SG-16 reverse primers. The 1.-2.-6.-7. reaction mixtures

were stirred in a single tube for the next treatment, but 3.mixture was not used because it has contained impurities in high quantity, as seen in Figure 4.5. In addition, DNA bands that belong to RFP gene are situated around 700 bp, indicating that the PCR is successfully performed.

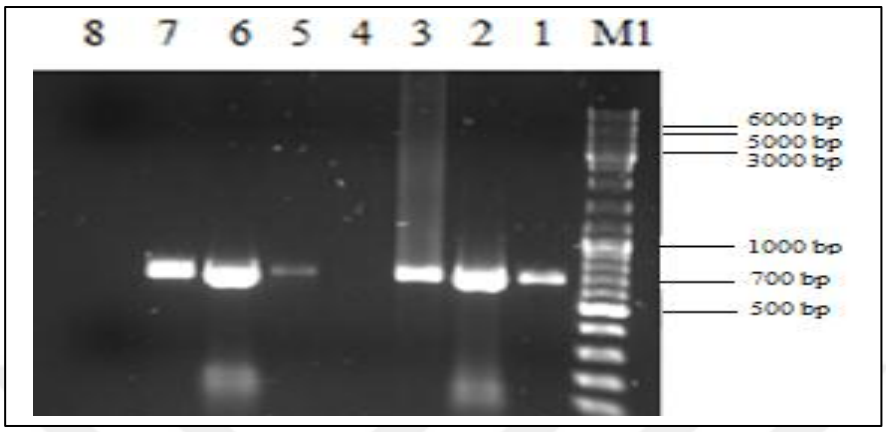

Figure 4.5 : The agarose gel image of amplified pET-22b(+)-RFP vector.

The image of agarose gel applied for DNA decontamination is shown in Figure 4.6.

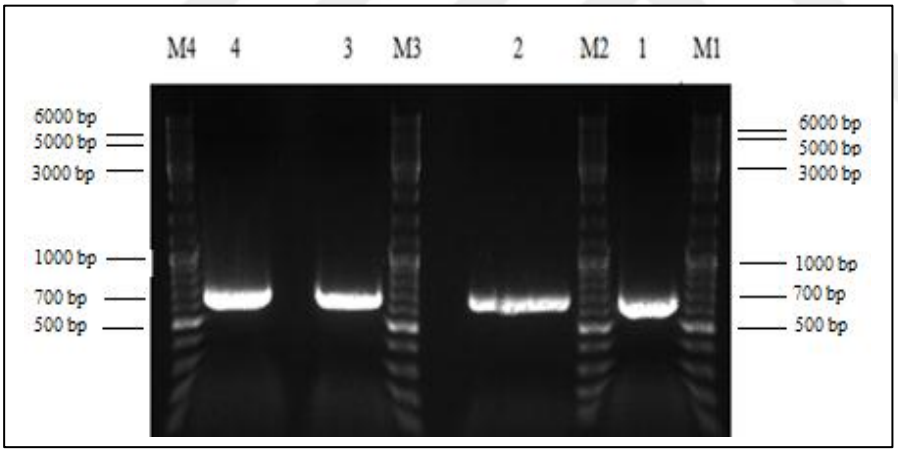

Figure 4.6 : The second agarose gel image of amplified pET-22b(+)-RFP vector.

The pET-22b(+)-SBP tag-φ29 vector and RFP gene were assembled in the same structure by applying HiFi DNA assembly and after then protein synthesis was initiated.

4.1.4 Single-stranded DNA sequencing

Both pET-22b(+)-SBPtag-φ29 and pET-22b(+)-SBPtag-φ29-RFP recombinant vectors obtained from competent cells in high purity were sent to LGC, Germany for single strand sequencing. Figure 4.7 obtained by overlapping of the sequencing

results with vector map shows that the synthetic SBP-tag-φ29 DNA polymerase gene was successfully inserted into pET-22b(+).

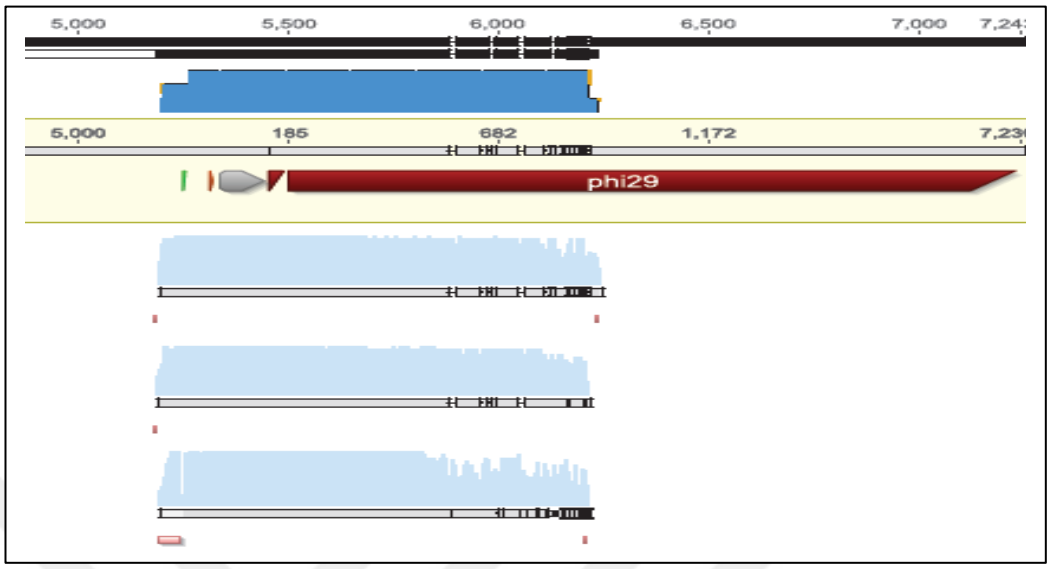

Figure 4.7 : Single strand DNA sequencing result of pET-22b(+)-SBP tag-φ29 DNA polymerase vector.

Sequencing result of the pET-22b(+)-SBP tag-φ29-RFP vector in Figure 4.8 reveals that the RFP gene was fully inserted into the pET-22b(+)-SBP tag-φ29 vector. Since the successfully cloned pET-22b (+) -SBP tag-φ29 vector was used in second protein expression, it was not necessary to check again the SBP tag-φ29 gene sequences in the recombinant vector.

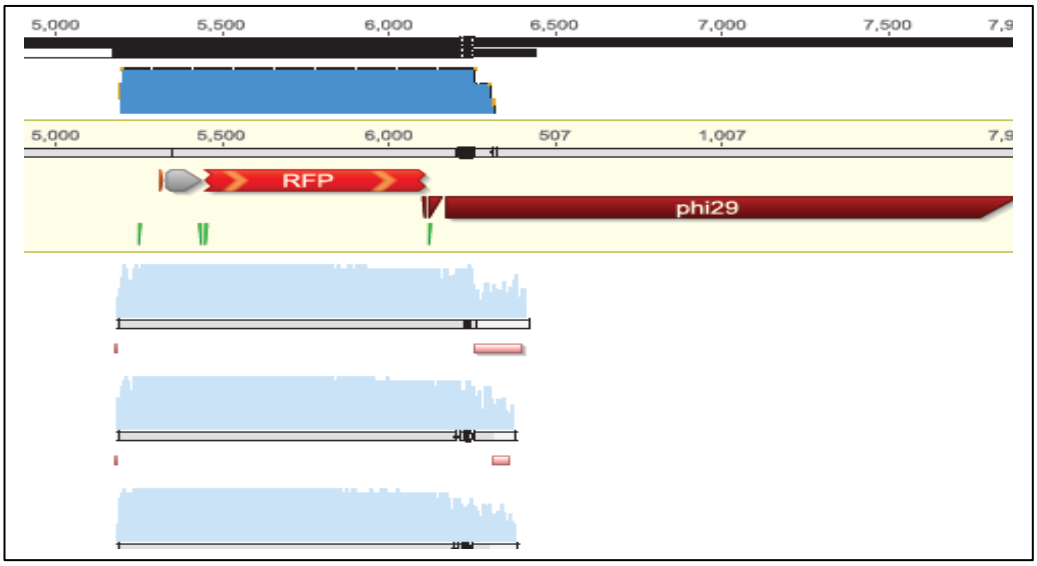

Figure 4.8 : Single strand DNA sequencing result of pET-22b(+)-SBP tag-φ29-RFP DNA polymerase vector.

4.2 Protein Purification

The AKTA affinity chromatography result which pertains to SBP tag-φ29 DNA polymerase, *in vitro* is shown in Figure 4.10. The lysate containing specific φ29 DNA polymerase and other unspecific proteins loaded into the super loop. While running buffer was flowing over throughout AKTA system, SBP tagged-φ29 DNA polymerase covalently attached to the equilibrated Streptavidin modified column surface, which possesses to immobilized SBP tag ligand. D-Desthiobiotin is a competitor included in elution buffer of SBP tag-φ29 DNA polymerase in Strep-Tactin column.Thus, SBP-tag separated from ligands by using this elution buffer, and the SBP-tag φ29 DNA polymerase protein was collected in collection tubes.

The protein peaks shown in all figures were formed by the absorbance values of them, hence a relative idea about protein concentration can be obtained.

When the results of φ29-RFP DNA polymerase are examined in Figure 4.11, B.1 and B.2 (in App. B), it can be seen that another protein peak, marked in red, originates from the RFP protein in the polymerase structure. However, there is not any peak in φ29-RFP DNA polymerase 3.sample according to Figure B.3 (in App. B), also the color differences of RFP inserted φ29 DNA polymerases can be seen in Figure 4.9.

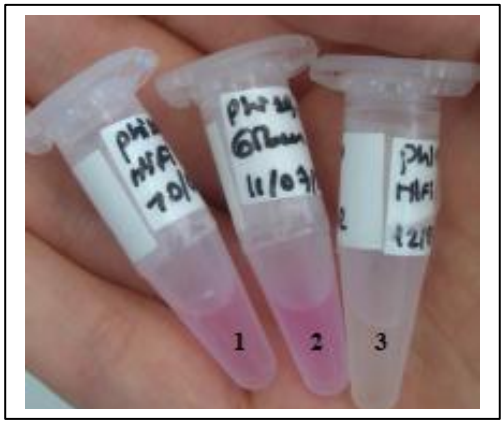

Figure 4.9 : The colour differences of φ29-RFP DNA polymerases.

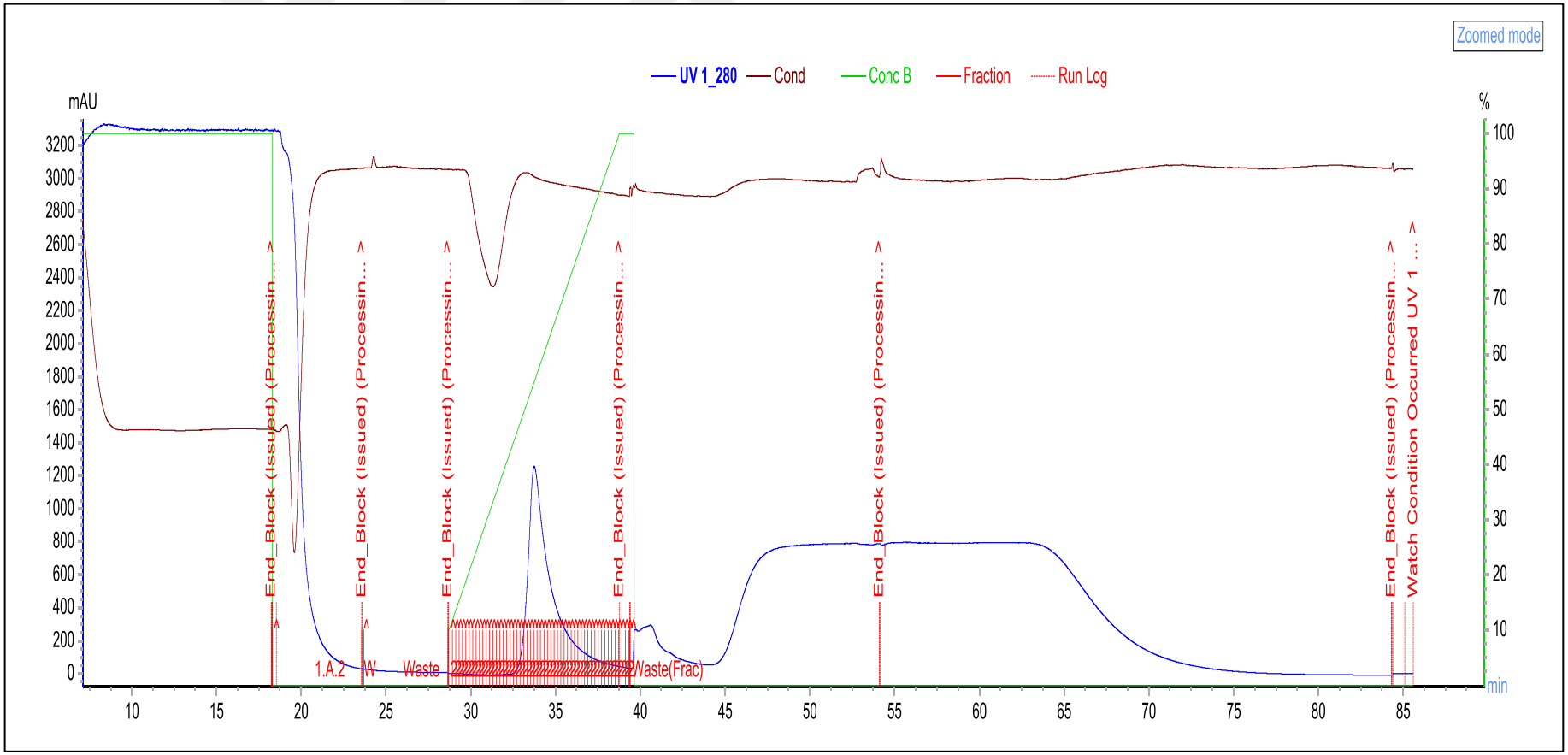

Figure 4.10 : The affinity chromatography result of SBP tag-φ29 DNA polymerase, *in vitro*.

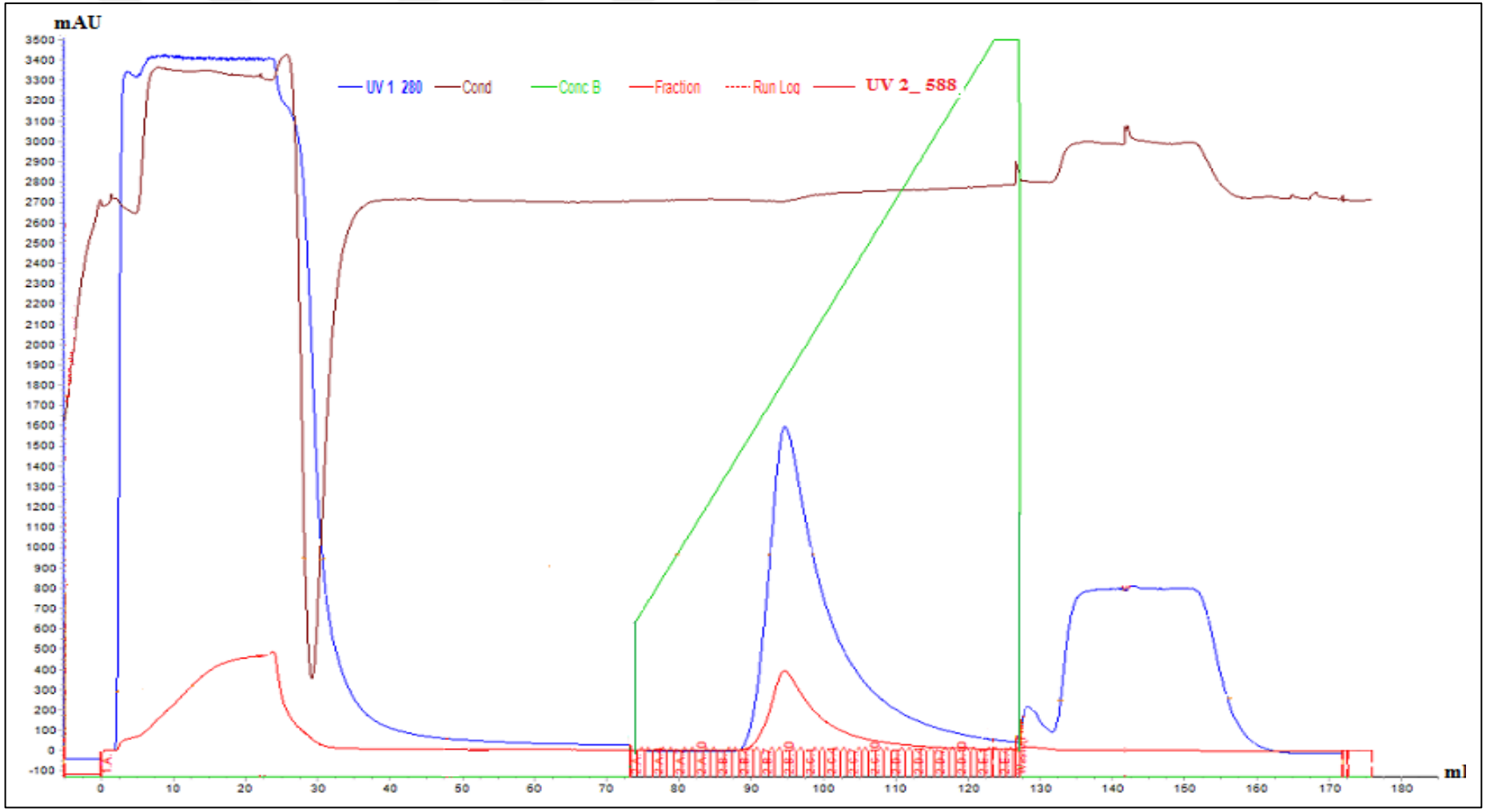

Figure 4.11 : The affinity chromatography result of SBP tag-φ29-RFP DNA polymerase 1.sample, *in vitro*.

4.3 Characterisation of Expressed φ29 DNA Polymerase Variants

4.3.1 Protein concentration measurement

To calculate protein concentrations, their absorbtion values were measured at 280 nm and Beer-Lambert law was executed.

$$A = \varepsilon \times d \times C \text{ [60]} \quad (4.1)$$

A= measured absorbance at 280 nm wavelength, ε= extinction coefficient (wavelength-dependent molar absorptivity coefficient with units of $M^{-1}cm^{-1}$), ε for ***φ***29 DNA polymerase = 122,995 $M^{-1}cm^{-1}$ [66], d= path length (App. C, Table C.1), C=protein concentration [60].

Table 4.1 : Absorption values of φ29 DNA polymerases at 280 nm.

Protein Type	**Absorption quantity at 280nm**	**Net worth of absorption at 280 nm**	**Concentration values, µM**
Blank-Storage Buffer	0,0905	0	-
Commercial ***φ***29 DNA polymerase	0,675	0,5845	7,53
φ29 DNA polymerase, *in vitro*	0,198	0,1075	1,42

Table 4.2 : Absorption values of φ29-RFP DNA polymerase at 280 nm.

Protein Type	**Absorption quantity at 280nm**	**Net worth of absorption at 280 nm**	**Concentration values, µM**
Blank-Storage Buffer	0,042	0	-
φ29-RFP DNA polymerase, sample 1	0,4205	0,3785	4,86
φ29-RFP DNA polymerase, sample 2	0,255	0,213	2,82
φ29-RFP DNA polymerase, sample 3	0,1645	0,1225	1,63

Both φ29 DNA polymerase and φ29-RFP DNA polymerase concentrations were detected at low level in comparison to commercial φ29 DNA polymerase concentration, as shown in Table 4.1 and 4.2. Many variables may trigger to the low protein expression. For example, codon bias is the most frequently encountered problem in prokaryotic expression systems. However, codon frequency of synthetic φ29 DNA polymerase gene has been optimized in these studies by taking account of the codon usage of *E.coli* host cells before starting the expression [67].

As a solution, changing temperature or time of induction might improve the protein production yield during incubation [45]. It can be overcome the low expression of the recombinant proteins problem by altering the medium composition of bacteria cultures. LB medium is the most commonly used culture for *E.coli*. However, LB contains a little bit amounts of supplementation such as carbohydrates and divalent cations. It may lead to the inadequate nutrient content for protein production. Moreover, the type of plasmid or fusion system can be altered. In the last resort, if the yield does not increase, protein expression should be tried in other host cells [42]. These options can also enhance the yield of both φ29 DNA polymerase variants in our study.

4.3.2 Enzyme activity test

A DNA reporter molecule that forms a stable hairpin structure was used in the DNA polymerase reactions which were operated in reference to the reaction conditions stated in *Ref.65* [57,65].

It can be concluded from Figure 4.12 that φ29 DNA polymerase sample's activity is relatively higher than commercial polymerase activity. However, theoretically, it should not have occurred any polymerization reaction in negative control due to the absence of specific primer in the reaction mixture. Because of the DNA contamination, the highly processive φ29 polymerase might have performed a polymerization reaction with bacterial genomic DNA or the residual DNA fragments of a recombinant vector that brings about the contamination.

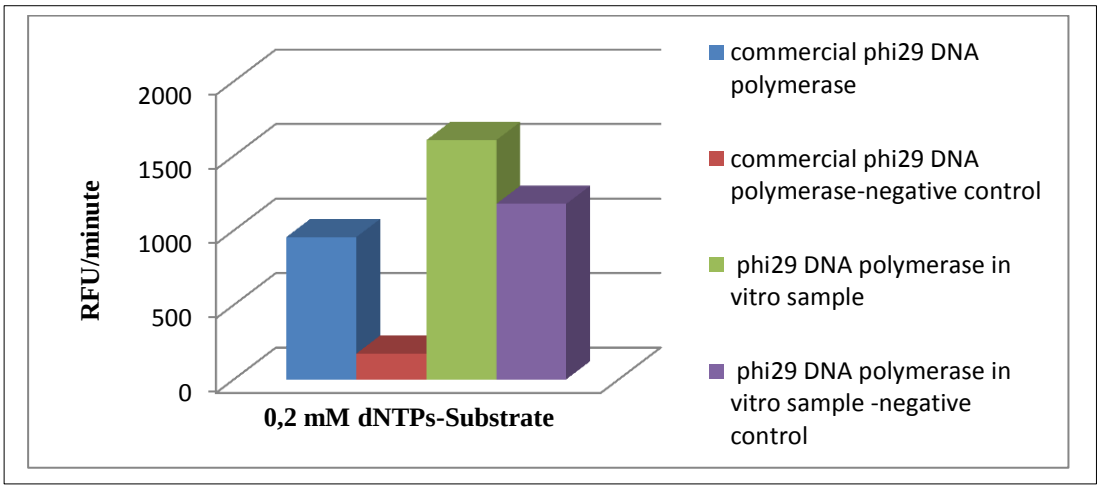

Figure 4.12 : RFU/minute values of commercial and *in vitro* synthesized φ29 DNA polymerases at constant substrate molarity.

It is obvious that φ29-RFP DNA polymerase samples also had the DNA contamination issue, according to the Figure 4.13. It can be reached to the raw activity test data in App. C.

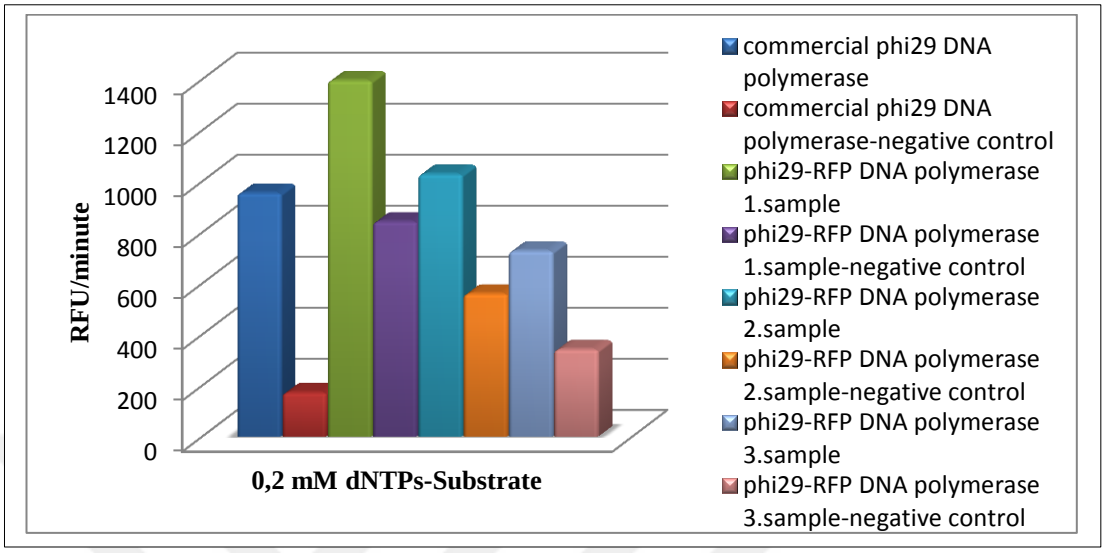

Figure 4.13 : RFU/minute values of commercial and *in vitro* synthesized φ29-RFP DNA polymerases at constant substrate molarity.

Obtaining the desired proteins after purification is not the end of the road. Their activity should be examined by making a comparison with commercial φ29 DNA polymerase. During the expression of recombinant polymerases, DNA contamination, which revealed itself in each negative control samples of activity tests was a major problem caused by the DNA of *E.coli* host cells. Since φ29 DNA polymerases have a high affinity for both single strand and double strand DNA, the genomic DNA of *E.coli* and residues of vector DNA should have immediately removed in cell lysis step. It is not easy to get rid of these undesired DNA fragments later on affinity chromatography.

4.3.3 SDS PAGE analysis

The samples from flow through and purified SBP tag-φ29 DNA polymerase were analysed in a %16 SDS gel after the completion of affinity chromatography. The visualisation of Coomassie-stained SDS gel is shown in Figure 4.14. In accordance to the gel image, the molecular mass of the purified proteins are seen about 70 kDa band which is similar value given in the literature [5].

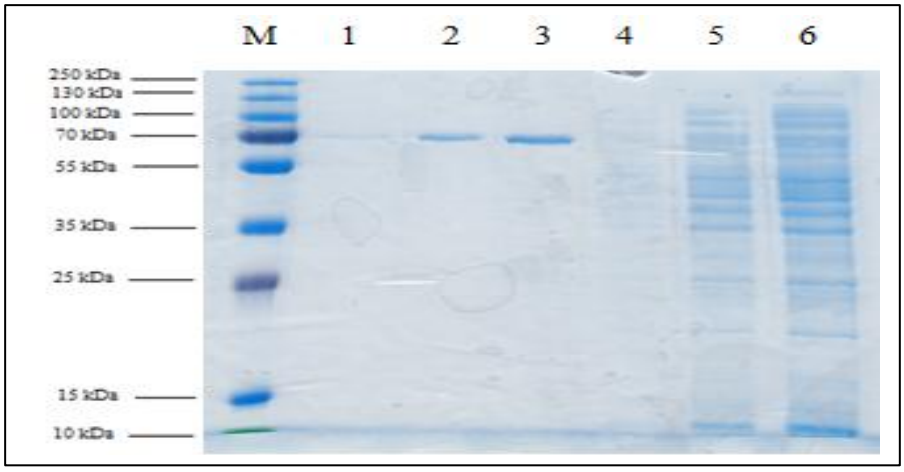

Figure 4.14 : Coomassi-stained %16 SDS gel image of φ29 DNA polymerase. M: marker -Page Ruler Plus Prestained Protein Ladder. 1, 2, 3 columns indicate SBP tag-φ29 DNA polymerase samples in different concentration values. 4, 5, 6 columns belong to flow through that includes unspecific proteins synthesides by *E.coli* cells.

φ29-RFP DNA polymerase comprises of an RFP protein with a molecular weight of 25,3 kDa and a φ29 DNA polymerase with a molecular weight of 66 kDa [68]. Therefore, it can be concluded that bands seen in the 100 kDa pertain to the successfully synthesized φ29-RFP DNA polymerase. SDS gel image shown in Figure 4.15 also indicates that φ29 and RFP proteins are seen individually in the 70 kDa and 35 kDa bands.

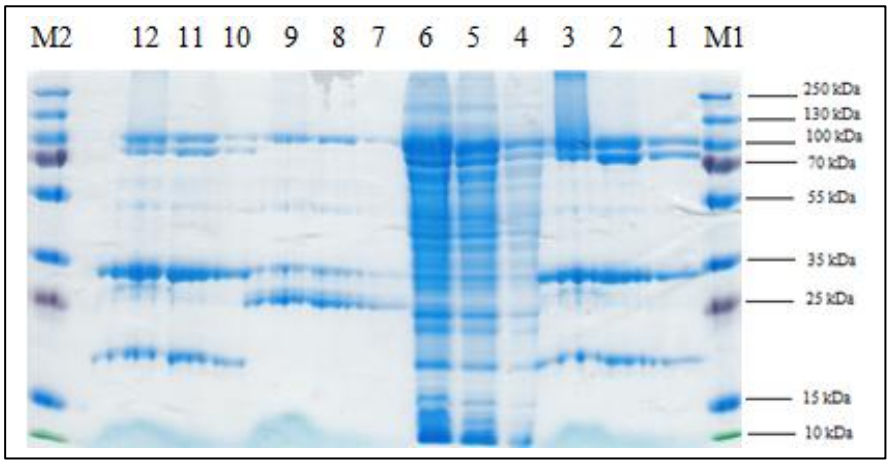

Figure 4.15 : Coomassi-stained %16 SDS gel image of φ29-RFP DNA polymerase. Loading order of samples: M1: marker 1, M2: marker 2- Page Ruler Plus Prestained Protein Ladder. 1, 2, 3 columns: φ29-RFP DNA polymerase 1.sample. 4, 5, 6 columns: Flow through. 7, 8, 9 columns: φ29-RFP DNA polymerase 3.sample. 10, 11, 12 columns: φ29-RFP DNA polymerase 2.sample.

According to the SDS-gel image visualized with Multi-colour RFB in Figure 4.16, while φ29-RFP DNA polymerase 1.and 2.samples and also RFP protein release fluorescence light in the 100 and 25 kDa bands, there is not any sign that belongs to RFP in the φ29-RFP DNA polymerase 3.sample and flow-through.

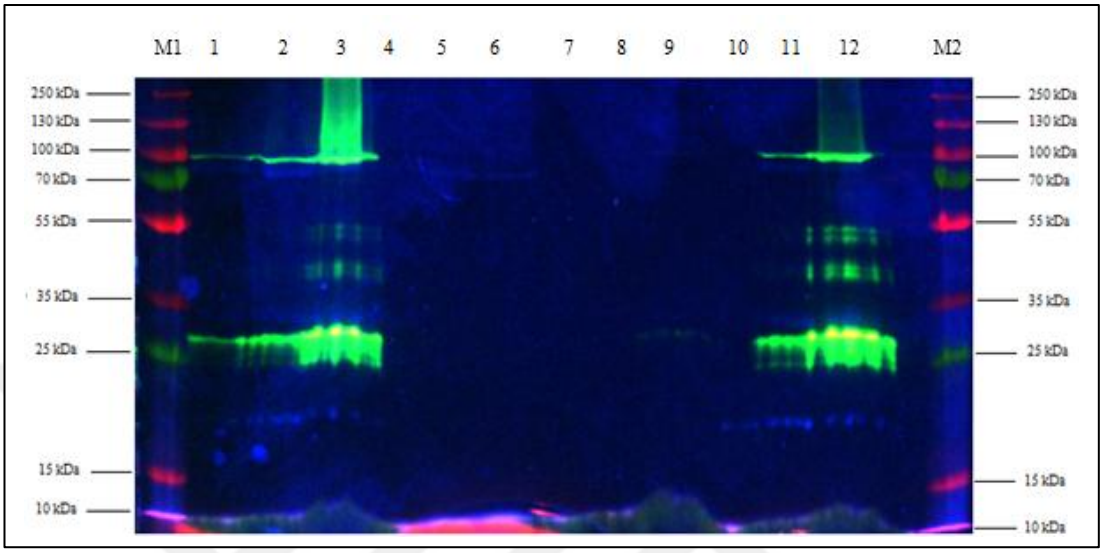

Figure 4.16 : Coomassi stained %16 SDS gel image of φ29-RFP DNA polymerase-RGB visualisation. Loading order of samples: M1: marker 1, M2: marker 2-Page Ruler Plus Prestained Protein Ladder. 1, 2, 3 columns: φ29-RFP DNA polymerase 1.sample. 4, 5, 6 columns: Flow-through. 7, 8, 9 columns: φ29-RFP DNA polymerase 3.sample. 10, 11, 12 columns: φ29-RFP DNA polymerase 2.sample.

Low protein production problem is also come out of UV spectroscopy measurement results of proteins and SDS gel bands. A slight band on SDS gel of the expected protein clarify that the cloning went as desired, but protein production yields were not desired level.

According to the Figure 4.15, φ29-RFP DNA polymerase is expressed at full-length and the expected size, but there are different protein bands in SDS gel. There may be various reasons corresponding to multiple bands. First of all, Φ29-RFP DNA polymerase may have not completely abided by the amino acid composition, so that it shows multiple bands in SDS gel. Cleavage at the junction between RFP and φ29 DNA polymerase might have taken place. As a solution, re-cloning the protein in another vector with a different promoter would be useful. However, the second possibility about multiple bands in SDS gel is that inappropriate translation termination may have arisen. Dual affinity tag purification would be a helpful method to increase the yield of full-length φ29-RFP DNA polymerase. For this purpose, a new vector that encodes two affinity tags at the C-and N-termini of the

gene should re-design. On the other hand, while the poly gly-ser linker provides an advantage to connect the functional domains both φ29 DNA polymerase and RFP in a fusion protein structure besides giving a chance certain degree of movements, lacking rigidity in this linker might lead to a limitation resulting in poor expression yields. In such a case, a rigid linker may be preferred for successfully maintaining their individual functions. Furthermore, rigid linkers permit the physically separation of covalently joined proteins, as well [49, 67].

Lastly, misfolding or trapping at an intermediate folding stage of φ29-RFP DNA polymerase cause to the multiple bands that belong to disunited RFP and φ29 DNA polymerase. Co-purification such as ion exchange would be useful for removing these undesired proteins [45].

5. CONCLUSION AND RECOMMENDATION

DNA amplification with φ29 DNA polymerase and the polymerase labeling with RFP that enables tracking of the enzyme via fluorescence detection hold promise for further application in biotechnology. In this study, SBP tag- φ29 DNA polymerase and SBP tag-φ29-RFP DNA polymerase have been successfully expressed by means of recombinant DNA technology applications.

Low production and DNA contamination were major problems encountered during protein expression and purification stages. Codon optimization, alteration in plasmid or fusion system may overcome the low production issue as earlier stated in *results and discussion*. To decontaminate the expressed polymerase variants, PEI precipitation or other chemical decontamination techniques should carry out before affinity chromatography.

Beyond these all problems, φ29 DNA polymerase variants hold promise in the amplification of very large fragments of DNA, synthesis of long single-strand DNA probes and second-strand cDNA synthesis [5].

REFERENCES

[1] **Johne, R., Müller, H., Rector, A., Van Ranst, M., Stevens, H.** (2009). Rolling-circle amplification of viral DNA genomes using phi29 polymerase, *Trends Microbiol.*, **17**(5), 205-11.

[2] **de Vega, M., Lazaro, J., Mencia, M., Blanco, L., Salas, M.** (2010). Improvement of φ29 DNA polymerase amplification performance by fusion of DNA binding motifs, *Proc.Natl.Acad.Sci.U.S.A.*, **107**(38), 16506-11.

[3] **Lizardi, P. M., Huang, X., Zhu, Z., Bray-Ward, P., Thomas, D. C., Ward, D. C.** (1998). Mutation detection and single-molecule counting using isothermal rolling-circle amplification, *Nature Genetics*, **19**(3), 225-232.

[4] **Hemphill, H. E., Whiteley, H. R.** (1975). Bacteriophages of Bacillus subtiles, *Bacteriological Reviews*, **39**(3), 257-315.

[5] **Lazaro, J. M., Blanco, L., and Salas, M.** (1995). [5] Purification of Bacteriophage ø29 DNA Polymerase, *Methods in Enzymology*, **262**, 42-49.

[6] **Rodriguez, I., Lazaro, J. M., Salas, M., de Vega, M.** (2003). φ29 DNA Polymerase Residue Phe128 of the Highly Conserved (S/T)Lx 2 h Motif is Required for a Stable and Functional Interaction with Terminal Protein, *Journal of Mecular Biology*, **325**(1), 85-97.

[7] **Choi, K. H.,** (2012). Viral Polymerases, *Adv.Exp.Med.Biol.*, **726**, 267-304.

[8] **Dufour, E., Rodriguez, I., Lazaro, J. M., de Vega, M., Salas, M.** (2003). A Conserved Insertion in Protein-primed DNA Polymerases is involved in Primer Terminus Stabilisation, *J.Mol.Biol.*, **331**(4), 781-794.

[9] **de Vega, M., Lazaro, J. M., Salas, M., Blanco, L.** (1996). Primed-terminus stabilization at the 3'-5' exonuclease active site of φ29 DNA polymerase. Involvement of two amino acid residues highly conserved in proofreading DNA polymerase, *The EMBO Journal*, **15**(5), 1182-1192.

[10] **Truniger, V., Lazaro, J. M., Salas, M., Blanco, L.** (1996). A DNA biding motif coordinating synthesis and degradation in proofreading DNA poylmerases, *The EMBO Journal*, **15**(13), 3430-3441.

[11] **Saturno, J., Lazaro, J. M., Estaban, F. J., Blanco, L., Salas, M.** (1997). φ29 DNA Polymerase Residue Lys383, Invariant at Motif B of DNA-dependent Polymerases, is Involved in dNTP Binding, *J.Mol.Biol.*, **269**(3), 313-325.

[12] **de Vega, M., Lazaro, J. M., Salas, M., Blanco, L.** (1998). Mutational Analysis of φ29 DNA Polymerase Residues Acting as ssDNA Ligands for 3'-5' Exonucleolysis, *Journal of Molecular Biology*, **279**(4), 807-822.

[13] **Kamtekar, S., Berman, A. J, Wang, J., Lazaro, J. M., de Vega, M., Blanco, M., …, Steitz, T.A,** (2004). Insights into Strand Displacement and Processivity from the Crystal Structure of the Protein-Primed DNA Polymerase of Bacteriophage φ29, *Molecular Cell*, **16**(4), 609-618. Doi: 10.1016/j.molcel.2004.10.019.

[14] **Lagunavicius, A., Kiveryte, Z., Zimbaite-Ruskuliene, V., Radzvilavicius, T., Janulaitis, A.** (2008). Duality of Polynucleotide Substrates for Phi29 DNA Polymerase: 3'-5' RNase Activity of the Enzyme, *RNA*, **14**(3), 503-513.

[15] **Bernard, A, Lazaro, J. M., Salas, M., Blanco, L.** (1990). The highly conserved amino acid sequence motif Try-Gly-Asp-Thr-Asp-Ser in α-like DNA polymerases is required by phage φ29 DNA polymerase for protein-primed initiation and polymerization, *Proc.Natl.Acad.Sci.U.S.A.*, **87**(12), 4610-4614.

[16] **Mendez, J., Blanco, L., Salas, M.** (1997). Protein-primed DNA replication: a transition between two modes of priming by a unique DNA polymerase, *The EMBO Journal*, **16**(9), 2519-2527.

[17] **Eisenbrandt, R., Lazaro, J. M., Salas, M., de Vega, M.** (2002). φ29 DNA polymerase residues Tyr59, His61 and Phe69 of the highly conserved Exoll motif are essential for interaction with the terminal protein, *Nucleic Acids Research*, **30**(6), 1379-1386.

[18] **Rodriguez, I., Lazaro, J. M., Blanco, L., Kamtekar, S., Berman, A., Wang, J.,…, de Vega, M.** (2005). A specific subdomain in φ29 DNA polymerase confers both processivity and strand-displacement capacity, *Proc.Natl.Acad.Sci.U.S.A.*, **102**(18), 6407-12. Doi: 10.1073/pnas.0500597102

[19] **Blanco, L., Salas, M.** (1985). Replication of phage φ29 DNA with purified terminal protein in DNA polymerase: Synthesis of full-length φ29 DNA, *Proc.Natl.Acad.Sci.U.S.A.*, **82**(19), 6404-6408.

[20] **Rector, A., Tachezy, R., Ranst, M.,** (2004). A sequence-Independent Strategy for Detection and Cloning of Circular DNA Virus Genomes by Using Multiply Primed Rolling-Circle Amplification, *Journal of Virology*, **78**(10), 4993-4998.

[21] **Joyce, C. M., Steitz, T. A.** (1994). Function and Structure Relationships in DNA Polymerases, *Annu.Rev.Biochem.*, **63**(1), 777-822.

[22] **Garcia-Diaz, M., Bebenek, K.** (2007). Multiple Functions of DNA Polymerases, *CRC Crit.Rev.Plant.Sci.*, **26**(2), 105-122.

[23] **Cotterill, S., Kearsey, S.** (2009). Eukaryotic DNA Polymerases, *Encyclopedia of Life Sciences (ELS)*, John Wiley and Sons Ltd., Chichester.

[24] **Dean, F. B., Nelson, J. R., Giesler, T. L., Lasken R. S.** (2001). Rapid Amplification of Plasmid and Phage DNA Using Phi29 DNA Polymerase and Multiply-Primer Rolling Circle Amplification, *Genome research*, **11**(6), 1095-1099.

[25] **Lazaro, J. M., Blanco, L., and Salas, M.** (1995). [5] Purification of Bacteriophage ø29 DNA Polymerase, *Methods in Enzymology*, **262**, 42-49.

[26] **Deng, H., Gao, Z.** (2015). Bioanalytical applications of isothermal nucleic acid amplification techniques, *Anal.Chim.Acta.*, 853, 30-45.

[27] **Demidov, V. V.** (2002). Rolling-circle amplification in DNA diagnostics: the power of simplicity, *Expert Review of Molecular Diagnostics*, **2**(6), 542-548.

[28] **Takahashi, H., Yamamoto, K., Ohtani, T., Sugiyama, S.** (2009). Cell-free cloning using multiply-primed rolling circle amplification with modified RNA primers, *Bio techniques*, **47**(1), 609-615.

[29] **Kim, E., Zwi-Dantsis, L., Reznikov, N., Hansel, C., Agarwal, S., Stevens, M. M.** (2017). One-pot synthesis of multiple protein-encapsulated DNA flowers and their application in intracellular protein delivery, *Adv.Mater.*, **29**(26), 10.1002.

[30] **Li, J., Macdonald, J.** (2015). Advances isothermal amplification: novel strategies inspired by biological process, *Biosensors and Bioelectronics*, **64**, 196-211.

[31] **Deng, H., Gao, Z.** (2015). Bioanaytical applications of isothermal nucleic acid amplification techniques, *Analytica Chimica Acta*, **853**, 30-45.

[32] **Zhao, Y., Chen, F., Li, Q., Wang, L., Fan, C.** (2015). Isothermal amplification of nucleic acids, *Chemical Reviews*, **115**(22), 12491-545.

[33] **Ali, M. M., Li, F., Zhang, Z., Zhang, K., Kang, D., Ankrum, J., Le, …, Zhao, W.** (2014). Rolling circle amplification: a versatile tool for chemical biology, material science and medicine, *Chem.Soc.Rev.*, **43**(10), 3324-3341.

[34] **Gill, P., Ghaemi, A.** (2008). Nucleic acid isothermal amplification Technologies-a review, Nucleosides, *Nucleotides and Nucleicacids*, **27**(3), 224-243.

[35] **Deadman, R., Jones, K.** (2004). DNA purification through amplification: Use of Phi29 DNA polymerase to prepare DNA for genomic analysis, *Life Science News*, **18**, 14-15.

[36] **Zhu, G., Hu, R., Zhao, Z., Chen, Z., Zhang, X., Tan, W.** (2013). Noncanonical self-assembly of multifunctional DNA nanoflowers for biomedical application, *J.Am.Chem.Soc.*, **135**(44), 16438-16445.

[37] **Lagunavicius, A., Merkiene, E., Kiveryte, Z., Savaneviciute, A., Zimbaite-Ruskuliene, V., Radzvilavicius, T., Janulaitis, A.** (2009), Novel application of phi29 DNA polymerase: RNA detection and analysis *in vitro* and in situ by target RNA-primed RCA, *RNA*, **15**(5), 765-771. Doi: 10.1261/rna.1279909.

[38] **Takahashi, H., Yamazaki, H., Akanuma, S., Kanahara, H., Saito, T., Chimuro, T., …, Kobori, T.** (2014). Preparation of Phi29 DNA Polymerase Free of Amplifiable DNA Using Ethidium Monoazide, an Ultraviolet-Free Light-Emitting Diode Lamp and Trehalose. *PLoS ONE*, **9**(2), e82624. doi: 10.1371/journal.pone.0082624.

[39] **Gräslund, S., Nordlund, P., Weigelt, J., Hallberg, B. M., Bray, J., Gileadi, O., …, Ming, J.** (2008). Protein production and purification, *Nature Methods*, **5**(2), 135-146. Doi: 10.1038/nmeth.f.202.

[40] **Gomes, A.R., Byregowda, S. M., Veeregowda, B. M., Balamurugan, V.** (2016). An overview of heterologous expression host systems for the production of recombinant proteins, *Adv.Anim.Vet.Sci*., **4**(7), 346-356.

[41] **Madzak, C., Gaillardin, C., Beckerich, J. M.** (2004). Heterologous protein expression and secretion in the non-conventional yeast *Yarrowia lipolytica*: a review, *J.Biotechnol*, **109**(1), 63-81.

[42] **Rosano, G. L., Ceccarelli, E. A.** (2014). Recombinant Protein Expression in Escherichia Coli: Advances and Challenges. *Frontiers in Microbiology*, **5**, 172.

[43] **Rai, M., Padh, H.** (2001). Expression systems for production of heterologous proteins, *Current Science*, **80**(9), 1121-1128.

[44] **Gustafsson, C., Govindarajan, S., Minshull, J.** (2004). Codon bias and heterologous protein expression, *Trends in biotechnology*, **22**(7), 346-53.

[45] **Gerstein, A. S.,** (2004). Molecular biology problem solver: a laboratory guide. John Wiley and Sons Inc.

[46] **Keefe, A. D., Wilson, D. S., Seelig, B., Szostak, J. W.** (2001). One-step purification of recombinant proteins using a nanomolar-affinity streptavidin-binding peptide, the SBP-tag, *Protein Expr.Purif*., **23**(3), 440-446.

[47] **Kimple, M. E., Brill, A. L., Pasker, R. L.** (2013). Overview of affinity tags for protein purification, *Curr.Protoc.Protein Sci.*, 9.9.

[48] **Yang, G. G., Xu, X. Y., Ding, Y.,Cui, Q. Q., Wang, Z., Zhang, Q. Y., …, Zhang, R.G.** (2015) Linker length affects expression and bioactivity of the onconase fusion protein in Pichia pastoris, *Genet.Mol.Res.*, **14**(4),19360-70. Doi: 10.4238/2015.December.29.46.

[49] **Chen, X., Zaro, J. L., Shen, W. C.** (2013). Fusion protein linker: property, design and functionality, *Adv.Drug.Deliv.Rev.*, **65**(10), 1357-1369.

[50] **Url-1 <** https://goo.gl/8LcsH8 >, data retrieved 19.10.2017.

[51] **Andersen, J., Spikes, D. A.** (2011). Techniques in molecular and cellular biology, Kendall Hunt publishing company, U.S.A.

[52] **Url-2 <** https://goo.gl/Peu8iG >, data retrieved 19.10.2017.

[53] **Affinity Chromatography Principles and Methods,** 18-1022-29 AE, GE Healthcare Life Science (2007), <http://proteins.gelifesciences.com/knowledge-library/protein-handbooks/ > data retrieved 19.10.2017.

[54] **Mallik, B., Chakravarti, B., Chakravarti, D. N.** (2016). Principles of chromatography, *Curr.Protoc.Essential Lab.Tech*, **13**(6), 1.1-6.1.23.

[55] **Zweitzig, D., Riccardello, N. M., Sodowich, B. I., O'Hara, S. M.** (2012). Characterization of a novel DNA polymerase activity assay enabling sensitive, quantitative and universal detection of viable microbes, *Nucleic Acids research*, **40**(14), e109.

[56] **Driscoll, M. D., Rentergent, J., Hay, S.** (2014). A quantitative fluorescence-based steady-state assay of DNA polymerase, *The FEBS Journal*, **281**(8), 2042-2050.

[57] **Summerer, D., Marx, A.** (2002). A molecular beacon for quantitative monitoring of the DNA polymerase reaction in real-time, *Angew.Chem.Int.Ed.Engl.*, **41**(19), 3620-2,3516.

[58] **Bisswanger, H.,** (2014), Enzyme assays. *Perspectives in Science*, 1(1), 41-55.

[59] **Url-3** < https://goo.gl/tES894 >, data retrieved 19.10.2017.

[60] **Url-4** < https://goo.gl/v7cn61 >, data retrieved 19.10.2017.

[61] **Dieffenbach, C. W., Lowe, T. M., Dveksler, G. S.** (1993). General concepts for PCR primer design, *PCR Methods and Applications*, **3**(3), 530-537.

[62] **Lee, P. Y., Costumbrado, J., Hsu, C. Y., Kim, Y. H.** (2012). Agarose gel electrophoresis for seperation of DNA fragments, *Journal of Visualized Experiments*, **62**, 3923.

[63] **Url-5** < https://goo.gl/Htu44G >, data retrieved 19.10.2017.

[64] **Url-6 <** https://goo.gl/mM8Rzw >, data retrieved 19.10.2017.

[65] **Summerer, D., Marx, A.** (2002). Supporting Information for *Angew.Chem.Int. Ed.* Z19502, Wiley-VCH.

[66] **Url-7 <** https://goo.gl/62uyHT >, data retrieved 19.10.2017.

[67] **Trinh, R., Gurbaxani, B., Morrison, S. L., Seyfzadeh, M.** (2004). Optimisation of codon pair use within the $(GGGGS)_3$ linker sequence results in enhanced protein expression, *Molecular Immunology*, **40**(10), 717-722.

[68] **Url-8 <** https://goo.gl/Cwpi2T >, data retrieved 19.10.2017.

APPENDICES

APPENDIX A: Primer sequences

APPENDIX B: AEKTA purification results

APPENDIX C: Enzyme kinetics data

APPENDIX D: Buffers and solutions

APPENDIX A

Table A.1 : KSR forward and reverse primers.

124-KSR forward primer	125-KSR reverse primer
5’-GATCCGGCTGCTAACAA AGCCCG-3’	5’-ATGTATATCTCCTTCTTAA AGTTAAACAAAATTATTTCT AGAGGGG-3’

Table A.2 : Flexible linker.

Linker
5’-GGTGGTGGTGGTAGCGGTGGTG GTGTTCAGGTGGTGGTGGTAGTGGT-3’

Table A.3 : T7 forward and reverse primers.

T7 forward primer	T7 reverse primer
5’-TAATACGACTCACTATAGGG-3’	5’-GCTAGTTATTGCTCAGCGG-3’

Table A.4 : SG-15 forward and SG-16 reverse primers designed in Geneious 8.0, Primer3.

SG-15 forward primer	Tm(C^0)	%GC content	Length(bp)
5´-GCAAGCAGCGAAGATGTGATC-3’	60,3	52,4	21
SG-16 reverse primer			
5´-GGCACCGGTACTATGACGAC-3´	60,2	60,0	20

APPENDIX B

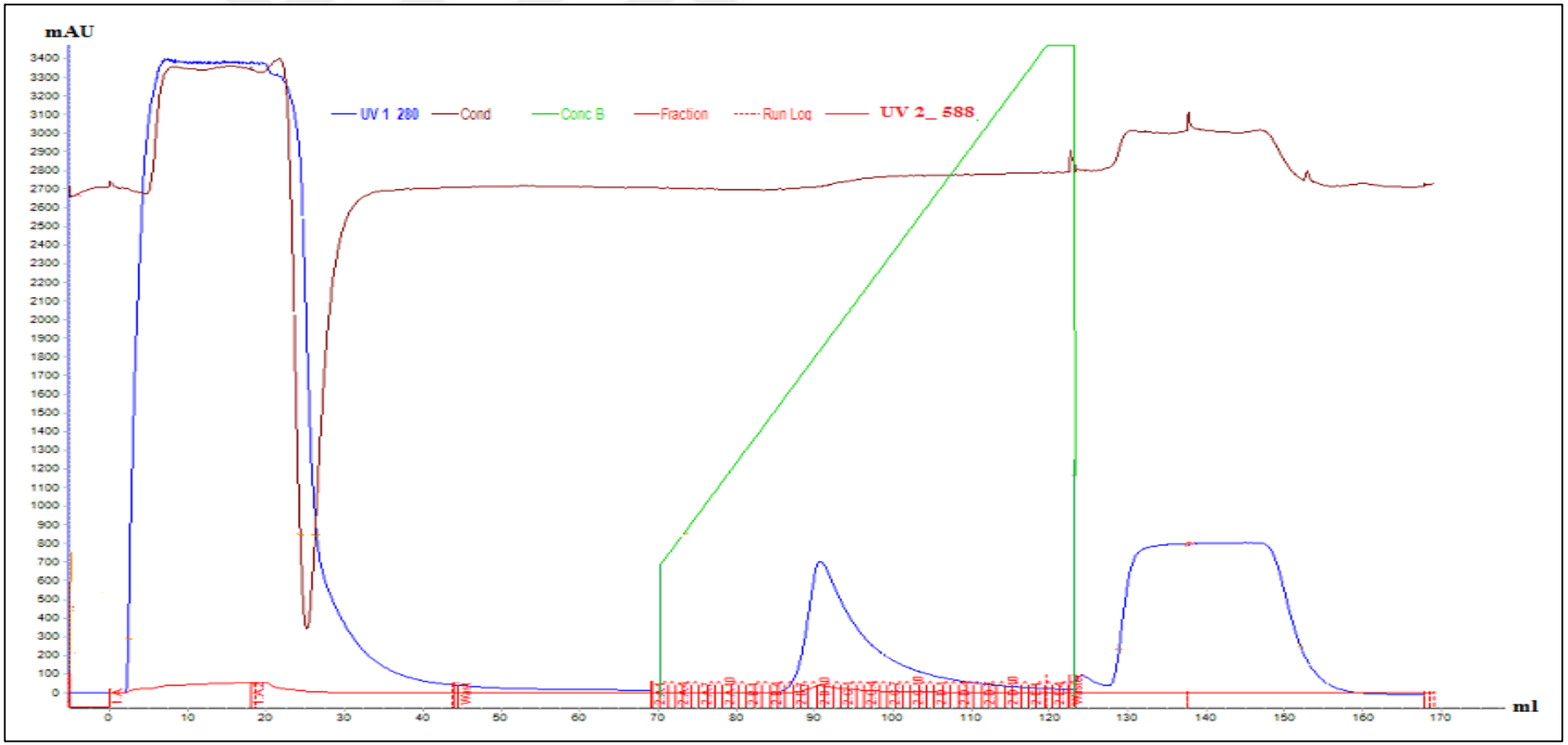

Figure B.1: The affinity chromatography result of SBP tag-φ29-RFP DNA polymerase 2.sample, *in vitro*.

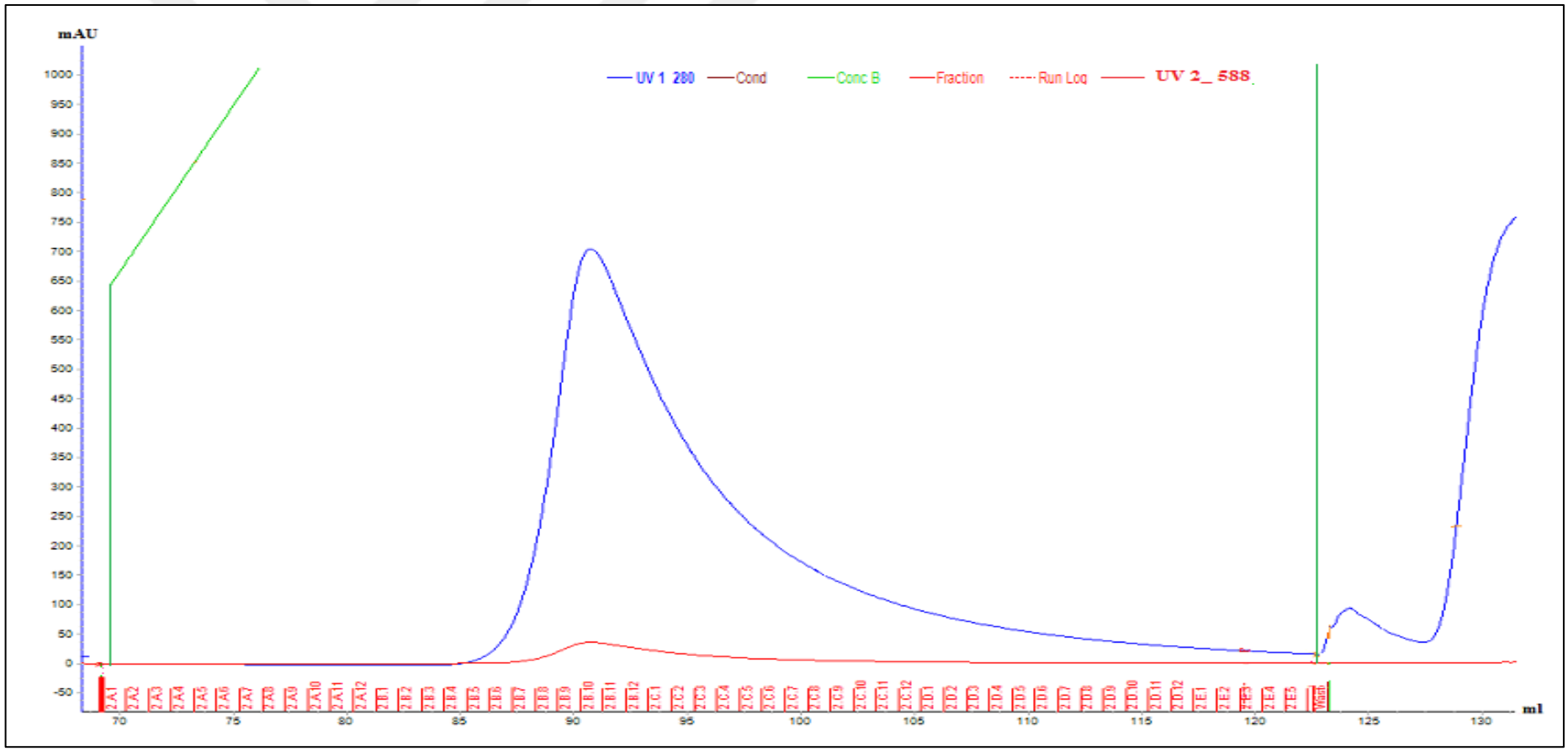

Figure B.2: The affinity chromatography result of SBP tag-φ29-RFP DNA polymerase 2.sample, *in vitro* (zoomed in the protein peak).

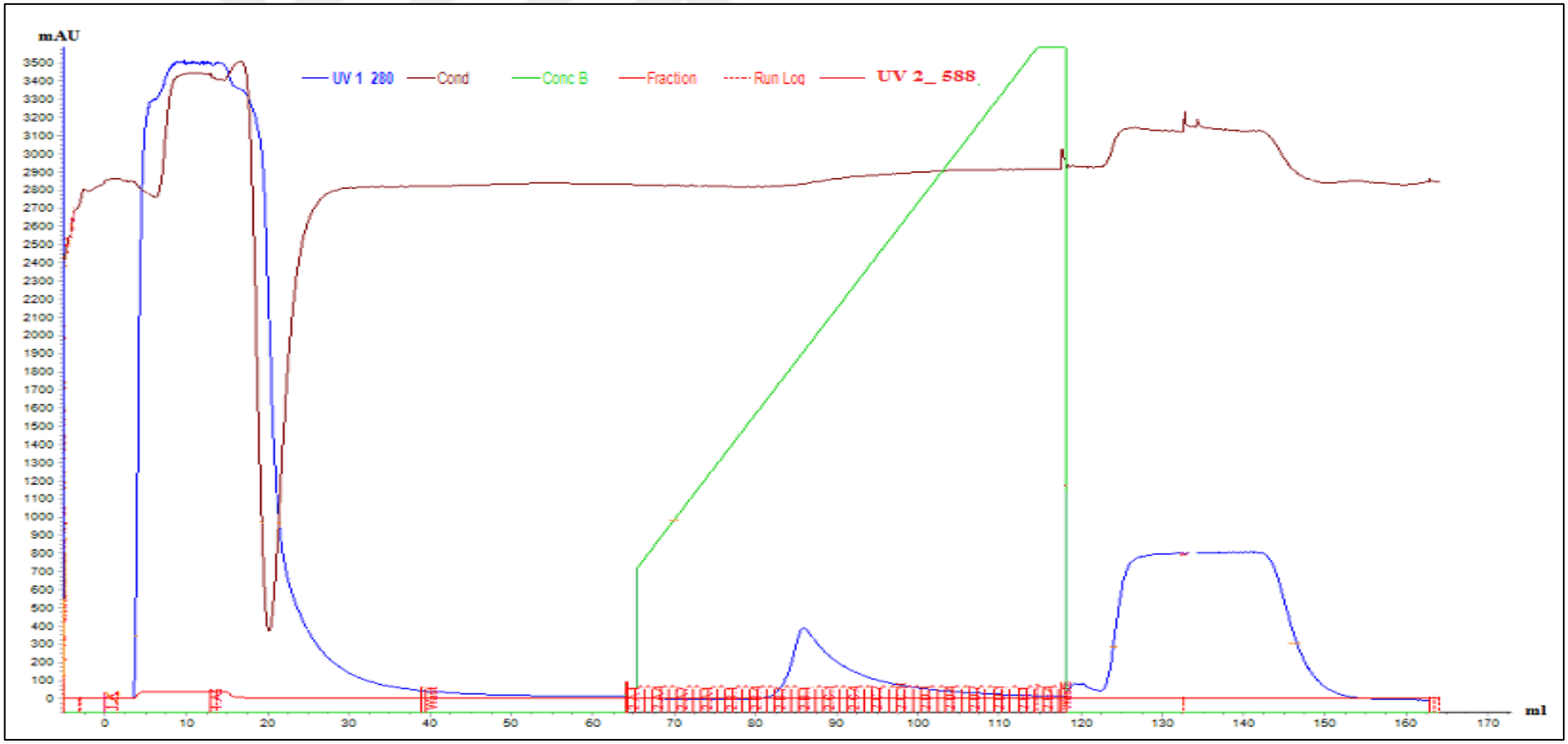

Figure B.3: The affinity chromatography result of SBP tag-φ29-RFP DNA polymerase 3.sample, *in vitro*.

APPENDIX C

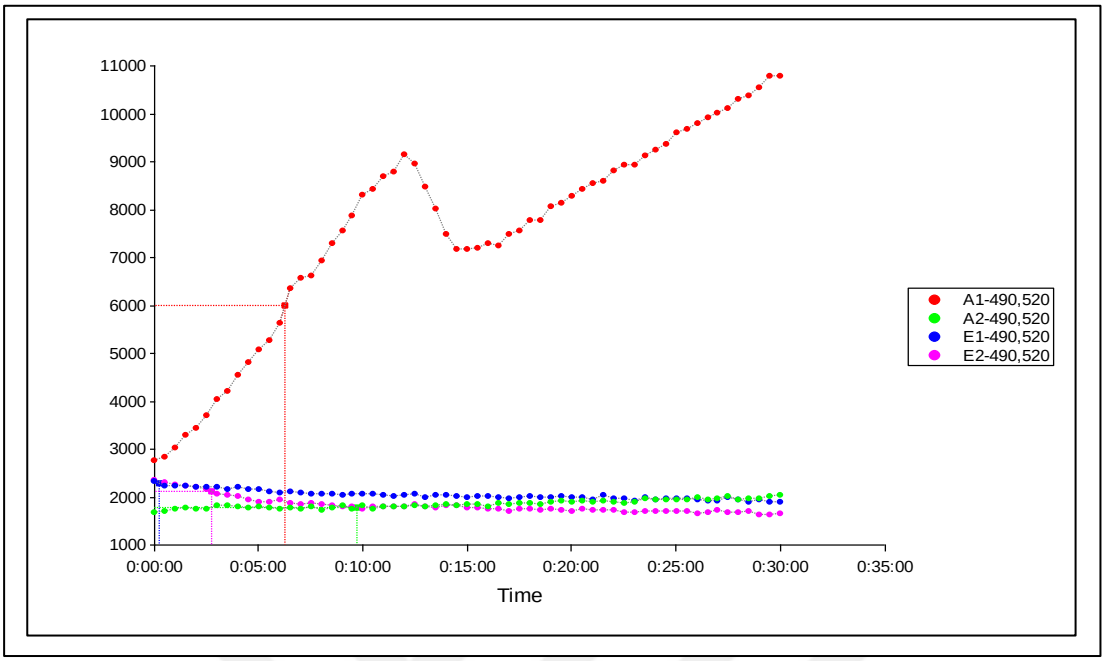

Figure C.1: RFU-time graph of commercial φ29 DNA polymerase. A1: commercial φ29 DNA polymerase, A2: commercial φ29 DNA polymerase-negative control, E1: control sample without polymerase, E2: control sample without polymerase-negative control.

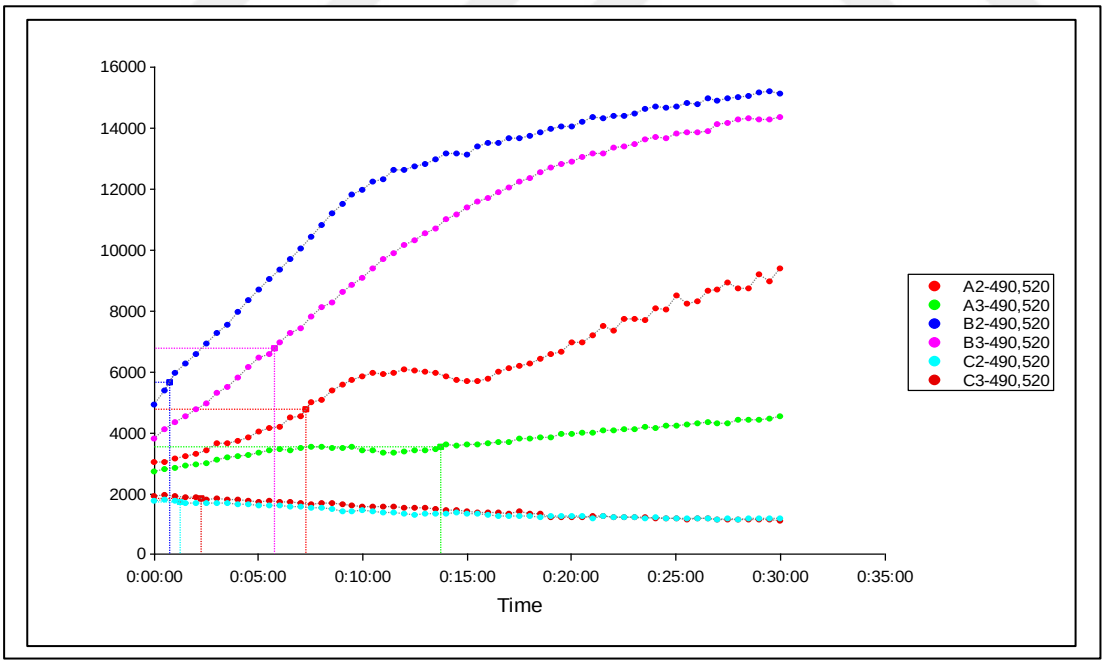

Figure C.2: RFU-time graph of φ29 DNA polymerase, *in vitro*. A2: commercial φ29 DNA polymerase, A3: commercial φ29 DNA polymerase-negative control, B2: φ29 DNA polymerase *in vitro*, B3: φ29 DNA polymerase *in vitro*-negative control, E1: control sample without polymerase, E2: control sample without polymerase-negative control.

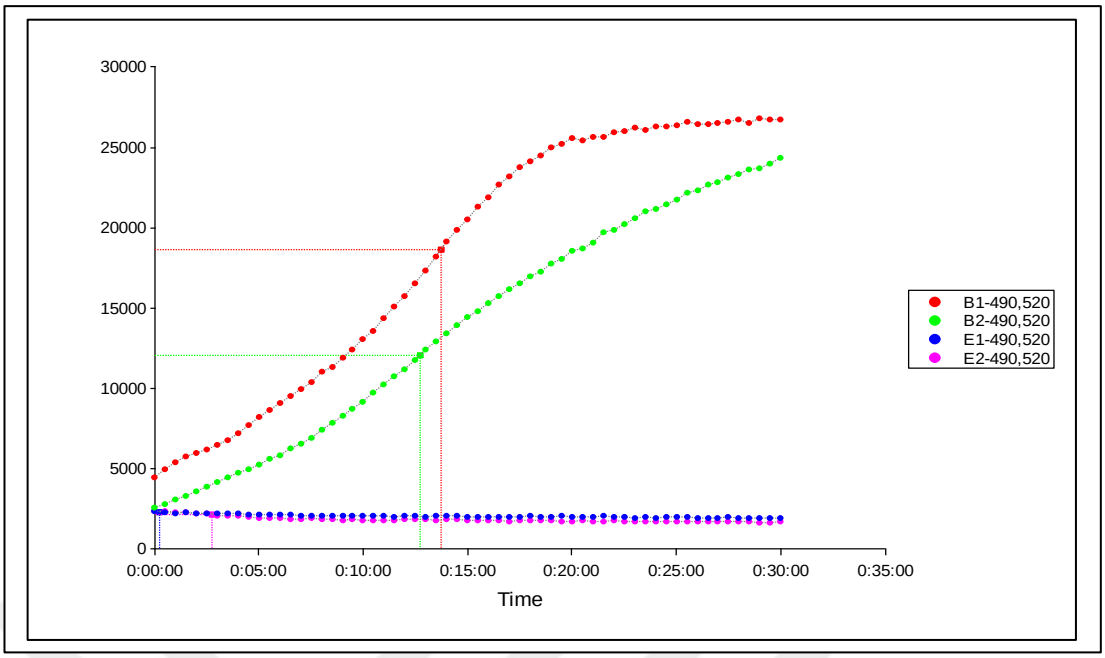

Figure C.3: RFU-time graph of φ29-RFP DNA polymerase 1.sample. B1:φ29-RFP DNA polymerase *in vitro*, B2: φ29-RFP DNA polymerase *in vitro*-negative control, E1: control sample without polymerase, E2: control sample without polymerase-negative control.

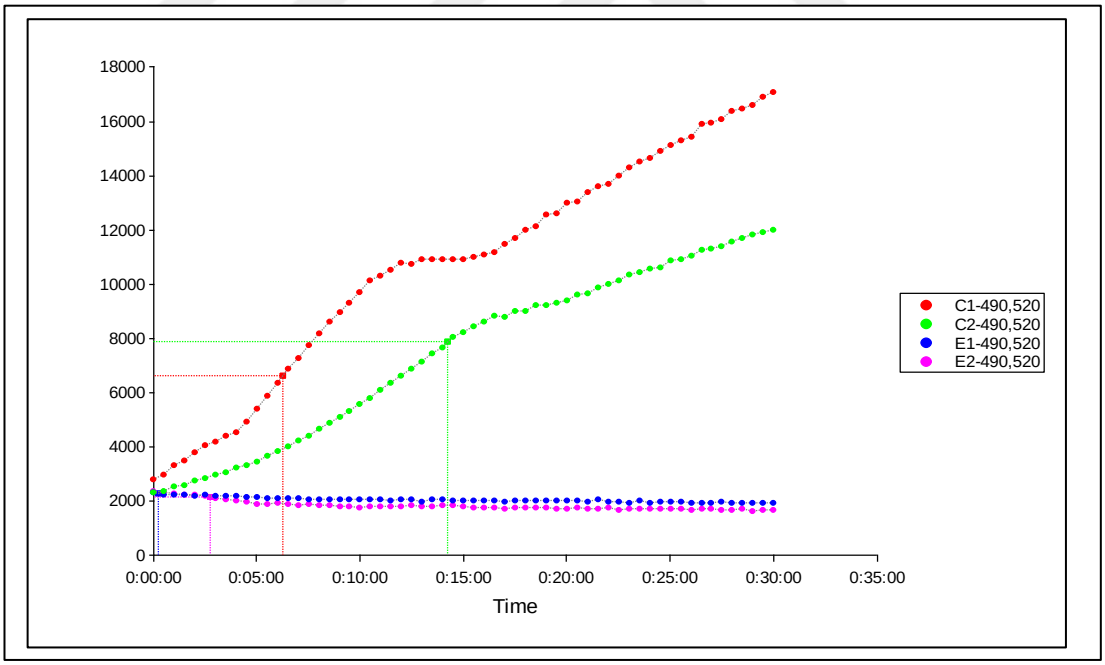

Figure C.4: RFU-time graph of φ29-RFP DNA polymerase 2.sample. C1:φ29-RFP DNA polymerase *in vitro*, C2: φ29-RFP DNA polymerase *in vitro*-negative control, E1: control sample without polymerase, E2: control sample without polymerase-negative control.

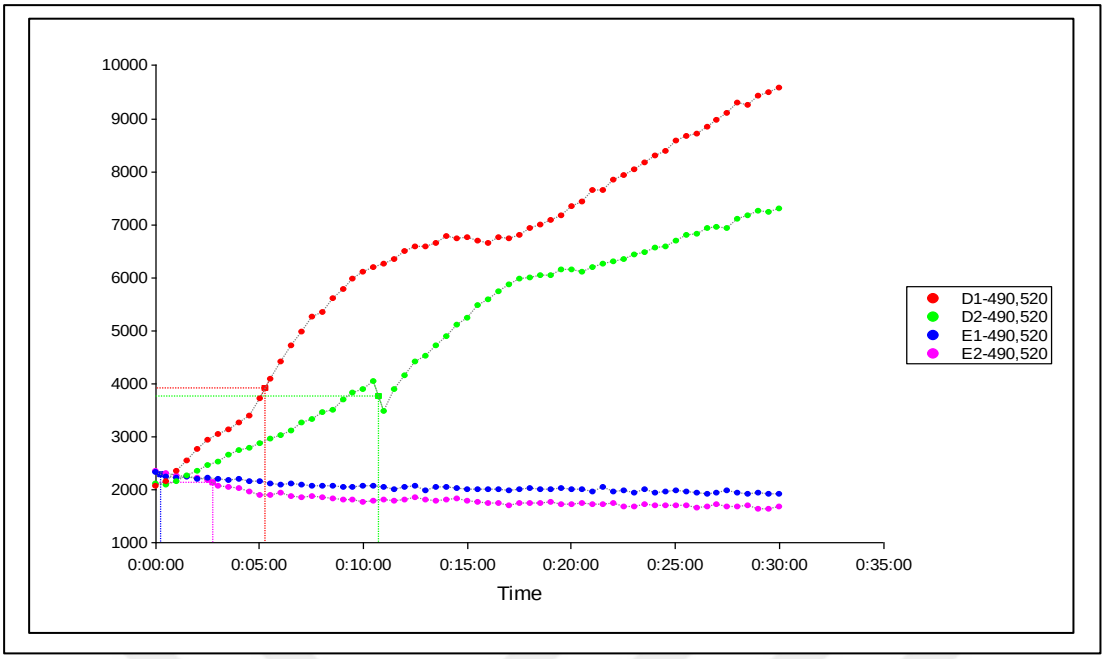

Figure C.5: RFU-time graph of φ29-RFP DNA polymerase 3.sample. D1:φ29-RFP DNA polymerase *in vitro*, D2: φ29-RFP DNA polymerase *in vitro*-negative control, E1: control sample without polymerase, E2: control sample without polymerase-negative control.

Table C.1: The length of Synergy-Take 3 plate wells.

Well type	d, length of well
A2	0,6462
A3	0,6378
A4	0,6398
A5	0,6241
A6	0,6198
A7	0,6133
A8	0,6162
A9	0,6057

APPENDIX D

Table D.1: Buffers used in affinity chromatography system.

Buffer Name	Composition
Buffer W, pH:8	100 mM Tris-Base 140 mM NaCl
Buffer E, pH:8	100 mM Tris-Base 140 mM NaCl 2,5 mM D-Desthiobiotin
Buffer R (HABA)	Diluted from 10X stock Buffer R solution
Storage Buffer for φ29 DNA polymerase, pH: 7,4	10 mM Tris-HCl 100 mM KCl 1 mM DTT 0,1 mM EDTA 50% Glycerol 0.5% Tween®20

Table D.2: Buffers used in electrophoresis.

Buffer Name	Composition
10X TBE, pH:8,3	890 mM Tris-Base 890 mM Boric Acid 20 mM EDTA
1X Maniatis	25 mM Tris-HCl 200 mM Glycine 0.1%(w/v) SDS

Table D.3: Medium compositions.

Medium	Composition
Lysogeny broth (LB), pH:7.0	10 g/L Trypton 5 g/L Yeast extract 10 g/L Sodium chloride
Lysogeny broth (LB) agar, pH:7.0	10 g/L Trypton 5 g/L Yeast extract 10 g/L Sodium chloride 16 g/L Agar powder

CURRICULUM VITAE

Name Surname **:** Fatma ARAS

Place and Date of Birth **:** Kırşehir/Turkey -10/06/1992

E-Mail **:** fatmaaras537@gmail.com

EDUCATION :

B.Sc. **:** 2014, Gazi University, Engineering Faculty, Chemical Engineering Department