Merve USLU

EPITOPES MAPPING OF RECOMBINANT CagA PROTEIN OF *Helicobacter pylori* by anti-CagA MONOCLONAL ANTIBODIES

M.S.
2014

FATİH UNIVERSITY
The Graduate School of Sciences and Engineering

Master of Science in
Biology

EPITOPES MAPPING OF RECOMBINANT CagA PROTEIN OF *Helicobacter pylori* by anti-CagA MONOCLONAL ANTIBODIES

by

Merve USLU

July 2014

EPITOPES MAPPING OF RECOMBINANT CagA PROTEIN OF *Helicobacter pylori* by anti-CagA MONOCLONAL ANTIBODIES

by

Merve USLU

A thesis submitted to

the Graduate School of Sciences and Engineering

of

Fatih University

in partial fulfillment of the requirements for the degree of

Master of Science

in

Biology

July 2014
Istanbul, Turkey

APPROVAL PAGE

This is to certify that I have read this thesis written by Merve USLU and that in my opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science in Biology.

_____

Assoc. Prof. Barık SALİH
Thesis Supervisor

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science in Biology.

_____

Assoc. Prof. Lokman ALPSOY
Head of Department

Examining Committee Members

Assoc. Prof. Barık SALİH _____

Assoc. Prof. Sevim IŞIK _____

Assoc. Prof. Lokman ALPSOY _____

It is approved that this thesis has been written in compliance with the formatting rules laid down by the Graduate School of Sciences and Engineering.

_____

Assoc. Prof. Nurullah ARSLAN
Director

July 2014

EPITOPES MAPPING OF RECOMBINANT CagA PROTEIN OF *Helicobacter pylori* by anti-CagA MONOCLONAL ANTIBODIES

Merve USLU

M.S. Thesis – Biology
July 2014

Thesis Supervisor: Assoc. Prof. Barık SALİH

ABSTRACT

The recognition of protein epitopes by specific antibodies is an important step in the antigen-antibody interactions. The detection of such epitopes can help to localize the binding sites, understand cross-reactivity between different antibodies and in the use of such antibodies in research and diagnostics.

The aims of this study were to clone the cagA 5' conserved region of *Helicobacter pylori*, to develop recombinant CagA (rCagA) proteins of different molecular weights and to identify the epitope regions of these proteins using our own anti-CagA monoclonal antibodies (Mabs) and commercial Mabs. The recombinant plasmid "pET28a(+)-rCagA" was purified from *E. coli* BL21(DE3)pLySs and the restriction cutting sites on this plasmid were designated by the SnapGene program. The plasmid vector was cut with Spe1, AvrII and Xho1 restriction enzymes at three different setups, self-ligated and transformed to *E. coli* BL21(DE3) cells. After IPTG induction, three rCagA proteins of 67 kDa, 60 kDa and 28 kDa molecular weights were obtained. A predicted number of epitopes on these proteins using the "Emini Surface Accessibility Prediction" software showed 14 epitopes in 67 kDa rCagA protein, 11 epitopes in 60 kDa rCagA protein and 6 epitopes in 28 kDa rCagA protein. The localization of these epitopes on each of these rCagA proteins were done by western blot using our own anti-rCagA Mabs (BS-53, CK-02) and three commercially available anti-CagA Mabs (A-10, B-818M, B-237H, His-probe). We found that our BS-53 Mab recognized all 3 rCagA proteins, our CK-02 Mab and A-10 Mab recognized the 67 kDa and 60 kDa rCagA proteins while the B-818M Mab and B-237H Mab recognized only the 67 kDa rCagA protein.

As a result, our BS-53 Mab recognized an epitope localized within the first 248 amino acids of all 3 rCagA proteins. While the our CK-02 Mab recognized an epitope localized in between the 248th and 469th amino acid sequence of the 67 kDa and 60 kDa rCagA proteins. The commercial Mab A-10 recognized an epitope that is localized in

the same region as our CK-02 Mab while the B-818M Mab and B-237H Mab recognized an epitope localized in between the 469th and 594th amino acid sequence of the 67kDa rCagA protein. The use of His-probe Mab in western blot further confirmed that the rCagA proteins carry the 6xHistidine tag.

This study shows that our own BS-53 Mab and CK-02 recognized epitopes localized on a different region of the rCagA protein. Thus, these Mabs can be used in research and diagnostics similar to the commercially available Mabs.

Keywords: *Helicobacter pylori,* epitope determination, recombinant CagA protein, anti-CagA monoclonal antibody

ÇEŞİTLİ anti-CagA MONOKLONAL ANTİKORLAR KULLANARAK, *Helicobacter pylori'*den ÜRETİLEN REKOMBİNANT CagA PROTEİNİNDEKİ EPİTOPLARIN HARİTALANMASI

Merve USLU

Yüksek Lisans Tezi – Biyoloji
Temmuz 2014

Tez Danışmanı: Doç. Dr. Barık SALİH

ÖZ

Protein epitoplarının spesifik antikorlar ile tanınması antikor-antijen etkilişiminin belirlenmesinde önemli bir basamaktır. Bu epitopların belirlenmesi bağlanma yerlerinin tespit edilmesinde, farklı antikorlar arasındaki karşı reaksiyonların anlaşılmasında ve bu antikorların araştırma ve tanısal alanlardaki kullanımında yardımcı olmaktadır.

Bu çalışmadaki amacımız *Helicobacter pylori'*nin cagA 5' korunmuş bölgesini klonlayarak farklı moleküler ağırlıkta rekombinant CagA (rCagA) proteinleri geliştirmek ve bu geliştirilen proteinlerdeki epitop bölgelerinin kendi ürettiğimiz anti-CagA monoklonal antikorlar ve ticari olan monoclonal antikorlar kullanarak belirlemektir. Rekombinant plazmidimiz olan "pET28a(+)-rCagA", *E.coli* BL21(DE3)pLySs hücrelerinden izole edildi ve bu plazmitteki restriksiyon kesim yerleri SnapGene program ile belirlendi. Plazmit vektörü Spe1, AvrII ve Xho1 restriksiyon enzimleri ile üç farklı yerden kesildi ve vektörün kendisine ligasyon yapılarak *E.coli* BL21(DE3) hücrelerine transformasyonu yapıldı. IPTG indüksiyon sonucu, 67 kDa, 60 kDa ve 28 kDa moleküler ağırlıktaki 3 rCagA protein elde edilerek üzerindeki tahmini epitop sayıları 67 kDa rCagA proteinde ise 14 epitop, 60kDa rCagA proteininde 11 epitop ve 28kDa rCagA proteininde 6 epitop olarak "Emini Surface Accessibility Prediction" yazılımı ile belirlendi. Bu tahmini epitoplar western blot analizinde kendi monoklonal antikorumuz olan BS-53, CK-02 ve 3 ticari anti-CagA monoklonal antikor olan A-10, B-818M, B-237H kullanılarak yapıldı. BS-53 tüm rCagA proteinlerini tanırken, CK-02 ve A-10, 60kDa ve 67 kDa rCagA proteinini tanımıştır. B-818M ve B-237H ise sadece 67 kDa rCagA proteinini tanımıştır.

Bu sonuçlara göre, BS-53 tüm rCagA proteinlerindeki ilk 248 aminoasit dizisi içerisindeki bir epitopu tanımaktadır. CK-02, 67 kDa ve 60kDa rCagA proteinlerdeki 248 ile 469'uncu aminoasit dizisi içerisindeki bir epitopu tanımaktadır. A-10 ticari

monoclonal antikoru, CK-02 ile aynı bölgedeki bir epitopu tanırken, B-818M ve B-237H monoclonal antikorları, 67 kDa rCagA proteinindeki 469 ile 594'üncü aminoasit dizisinde bulunan bir epitopu tanımaktadır. His-probe monoklonal antikorun western blot analizinde kullanılması ise rCagA proteinlerimizin 6xHistidine işaretli olduğunu bir kez daha doğrulamıştır.

Bu çalışma göstermiştir ki, kendi monoklonal antikorlarımız olan BS-53 ve CK-02, rCagA proteininin farklı bir bölgesindeki epitopu tanımaktadır. Böylece, bu monoklonal antikorlar ticari kullanılan monoklonal antikorlara benzer araştırma ve tanısal alanlarda kullanılabilecektir.

Anahtar Kelimeler: *Helicobacter pylori*, epitope belirleme, rekombinant CagA protein, anti-CagA monoklonal antikor

To my beloved parents, my brother Mehmet USLU, my grandmother Şengül TIĞ
and my uncle Abdullah TIĞ

ACKNOWLEDGEMENT

I would like to express my sincere appreciation to my major professor, Dr. Barık Salih for his guidance and suggestions throughout the entire research.

My thanks go to Assoc. Prof. Dr. Fahri AKBAŞ and Dr. Hüseyin TOMBULOĞLU for their valuable suggestions and comments.

I would like to thank Cebrail KARAKUŞ, Umidahan DJAKBAROVA, Nurdan Sena DEĞİRMENCİ and Duygu YAZICI, for their technical assistance and suggestions.

For my friends Asya NECEFOVA, Kevser YILDIRIM, Belkıs ÇAKAR, Ayşegül ÇALIŞKAN and Ayşegül KASAP, thank you for your moral support.

I would also like to express my deepest appreciation to my lovely family especially my father Hasan Hüseyin USLU and my uncle Abdullah TIĞ for their endless support, motivation and patience.

Lastly, but in no sense the least, I am thankful to all colleagues and friends who made my stay at the university a memorable and valuable experience.

TABLE OF CONTENTS

LIST OF TABLES

TABLE

LIST OF FIGURES

FIGURE

LIST OF SYSMBOLS AND ABBREVIATIONS

SYMBOL/ABBREVIATION	
Ab	Antibody
Ag	Antigen
APS	Ammonium persulfate
Bp	Base pair
BPB	Bromophenol blue
cag PAI	cag Pathogenicity Island
CBB	Comassie brilliant blue
cDNA	Complementary DNA
cagA	cytotoxin associated gene A
DMSO	Dimethyl sulfoxide
DNA	Deoxyribo Nucleic Acid
dNTP	Deoxyribo Nucleotide
E. coli	Escherichia coli
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
ESI	Electrospray ionization
EPIYA	Glu-Pro-Ile-Tyr-Ala amino acids
Fab	Fragment antigen binding
H. pylori	*Helicobacter pylori*
HPLC	High performance liquid chromatography
HRP	Horse raddish peroxidase
IEDB	Immune epitope database
IHC	Immunohistochemistry
IPTG	isopropyl-β-D-thiogalactoside
Kan	Kanamycin
kb	Kilo base
kDa	Kilo dalton
LB	Luria Broth
MALT	mucosa-associated lymphoid tissue lymphoma

Mab	Monoclonal antibody
MeOH	Methanol
MHC	Major histocompatibilty complex
MS	Mass spectroscopy
MW	Molecular weight
NMR	Nuclear magnetic resonance
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	polymerease chain reaction
PEG	Polyethylene glycol
PVDF	Polyvinylidene fluoride
rcagA	Recombinant cytotoxin associated gene A
Rpm	Revolutions per minute
SDS	Sodium dodecyl sulfide
SNP	Single nucleotide polymorphism
SPR	Surface Plasmon Resonance
T4SS	type IV secretion system
TBE	Tris-borate-EDTA
TBS	Tris buffered saline
TBS-T	TBS-Tween
TEMED	Tetramethyl etilen diamine
UV	Ultra violet
URB	Urea breath test

CHAPTER 1

INTRODUCTION

Helicobacter pylori the causative agent of various gastrodueodenal diseases such as peptic ulcers and gastric adenocarcinoma was initially discovered by Marshall and Warren in the 1984 (Marshall and Warren, 1984). More than half of world population have suffered from the infection of with *H. pylori* (Dunn et al., 1997)

H. pylori virulence factors play a crucial role in developing chronic disease owing to the persistent colonization of this bacterium inside stomach (Bauer et al., 2011). The cytotoxin-associated antigen CagA protein (120-140 kDa) is the most important virulence factor of *H. pylori* and this immunodominant protein is encoded by the cagA gene which is a part of the cag-pathogeniticity island (cagPAI) genes that produce a type IV secretion apparatus (T4SS). The CagA protein is directly injected into host cells by the T4SS a syringe-like structure and lead to cell morphologic alterations, proliferation and apoptosis (Censini et al., 1996). The cagA gene have two structures which are a 5' highly conserved region and a 3' variable region. Also, this cagA 3' variable region of *H. pylori* contains a number of repeat sequences known as EPIYA motifs that is the reason why the CagA protein showing variation in its size between 120-140 kDa. The cagA gene is also divided into Western type strains and East Asian type strains according to the type of these repeat sequences (Yamaoka et al., 2002).

Several methods have been used to detect the infection of *H. pylori* with CagA status in the diagnosis and progression of the disease. Currently, several non-invasive techniques such as stool test, serology, urea breath test (URB) and urine test are available. Among them, stool test and serology are useful for the detection of *H. pylori* infection. However these tests do not provide better information of the pathological

potential of *H. pylori* and the CagA status. The reason of this situation is related to the lack of immunoreagents such as are antigens and monoclonal antibodies (Klimovich et al., 2010).

As pointed out in some studies, the development of immunoreagents has been utilized in the infection of *H. pylori*. Li et al. (2006) developed only one Mab specific for CagA using *H. pylori* lysate as antigen. Stephens et al. (1996) utilized CagA-antigen from *H. pylori* cultures also obtained a single Mab. More recently Yasuda et al. (2009) produced Mab against a synthetic peptide of an East Asian type CagA.

Epitope mapping plays a pivotal role in a determination of antigen-antibody interaction and it is a significant step for designation epitopes that are the specific regions where an antibody binds its antigen. Identification of epitopes in an antigen makes a major contribution in an antibody production, immunodiagnostic test and epitope-driven vaccine design (EL-Manzalawy and Honavar, 2010).

Nowadays, many approaches have been developed to map epitopes of monoclonal antibodies. While X-ray crystallography, NMR and Electron microscopy are the types of structural mapping approach, the competition methods, antigen modification methods, fragmentation methods and phage display libraries or peptide panning are the types of functional mapping approaches (Morris, 2005). Besides, the use of mass spectrometry has become popular in epitope mapping because of the high sensitivity and rapid analysis time (Chambers et al., 2000). The epitope prediction programs can also be preferred as a computational approach associated with these experimental methods (Clementi et al., 2013).

Several reports have been published on the epitope mapping of the various proteins of *H. pylori* recognized by their own specific monoclonal antibodies. Fuji et al. (2004) mapped epitope for monoclonal antibodies inhibiting enzymatic activity of *H. pylori* Urease using competition method and mass spectroscopy. Li, Ning et. al (2007) developed mouse monoclonal antibodies against *H. pylori* Lpp20 protein and *H. pylori* Catalase protein and mapped the antigenic epitope of these antigens by using phage display library. Li, Mao et al. (2008) produced a neutralizing monoclonal antibody against UreB of H. *pylori* and identified the B-cell epitope of this recombinant protein recognized by neutralizing Mab using fragmentation and synthetic peptide library

methods correlated with Western Blot and immunoassay such as ELISA. Qiu, Wang et al. (2010) identified the B-cell epitopes in *H. pylori* urease B subunit using synthetic truncated peptide with ELISA and Western Blot analysis. Klimovich et al. (2010) determined the epitope region of monoclonal antibodies against the entire recombinant CagA protein fragments using serological assay development and Western Blot analysis.

In this study, we will determine the epitopes on our developed recombinant CagA protein fragments by using our own anti-cagA monoclonal antibodies and commercially available monoclonal antibodies. Our objectives were to clone the previously cloned cagA 5' conserved region of *H. pylori* into 2 different clones, to express the CagA proteins (rCagA) of the two cloned regions, to detect the predicted antigenic epitopes on each of these recombinant CagA proteins using sofware program and to identify the epitope region on each protein by using monoclonal antibodies.

CHAPTER 2

REVIEW OF LITERATURE

2.1 HELICOBACTER PYLORI

Helicobacter pylori is a spiral-shaped, gram-negative, microaerophilic bacterium that colonizes the human gastric and duodenal mucosa. Infection with *H. pylori* causes chronic active gastritis, gastric and duodenal ulceration, lymphoma of mucosa-associated lymphoid tissue (MALT) and the development of adenocarcinoma (Marshall et al., 1998) . Therefore, *H. pylori* has been classified as a group I carcinogen by the International Agency for Research on Cancer.

Almost half of the world populations are infected with *H. pylori*. Several factors such as bacterial genetic diversity, virulence factor, host constitution and environmental factors play a role in the prevalence of infection and the disease outcome.

Among the virulence factors the cytotoxic-associated gene (cagA) is the most important and plays a major role in the severity of the disease. The cagA gene encodes an immunodominant protein of 120-140 kDa molecular weight (Tummuru et al., 1993; Covacci et al., 1993). This cagA gene is carried on the cagA pathogenecity island (cagPAI) which encodes the components of a type IV secretion system (T4SS). The CagA protein is translocated into host epithelial cells where it triggers signal transduction and cell damage by this T4SS syringe like structure (Covacci et al., 2000; Higshi et al., 2002).

The structure of the cagA gene consists of a 5' highly conserved region and a 3' variable region. The 3' region of the cagA may contain variable numbers of repeat sequences known as EPIYA motifs (Yamaoka et al., 1998). These repeat motifs consist of five amino acid residues; glutamic acid-proline-isoleucine-tyrosine-alanine (Glu-Pro-

Ile-Tyr-Ala) (Furuta et al., 2011). Because of this polymorphism of the cagA gene and the types of these motifs, the CagA proteins were divided into two types as a Western type and an East Asian type (Yamaoka et al., 1998, 1999, 2000). EPIYA-A, EPIYA-B and EPIYA-C or as EPIYA-A, EPIYA-B are existed in Western isolates and EPIYA-D is existed in East Asian isolates (Hatakeyama et al., 2004).

2.2 EPITOPE MAPPING

Epitope mapping is a method for identifying the antigenic determinants or epitopes recognized by the antibodies. It also refers to the localization of the antibody-binding site (Fab) on the antigen (Ag) in antibody–antigen recognition. Determination of such epitopes is an important step in the, characterization of monoclonal antibodies (Mabs) especially those used in diagnostic and therapeutic and epitope-driven vaccine design (Saint-Remy, 1997; Nelson et al., 2000) . Such characterization reveals the specific primary structure of the target antigen that interacts with the Fab variable region of the Mab (Lu et al., 2009) .

Lots of terms such as “immunological analysis”, “fine-structure analysis”, or “epitope determination” have substituted for the term of “epitope mapping” however it has been stated that the phrase of “epitope mapping” with proper qualification is preferable in the literature (Ladner, 2007).

2.2.1 Antigenic determinants or epitopes

Since the term “epitope” has been come into use in the 1960’s by Niels Jerne, it has referred to the antigenic determinant or the place where an antibody binds its Ag in a broader sense (Onoue et al., 1965; Pomés, 2009).

The antigenic determinant or epitopes are defined as the portions of the antigen molecules that interact with the antigen-binding site of antibody also known as paratope. Immunoglobulin molecules recognize only unique conformations and spatial sites on the surface of the antigens rather than the entire antigen. The ‘epitope’ on the antigen and the ‘paratope’ on the antibody are complementary surface features that

provide an ‘induced fit’ and flexibility between them (Morris, 2001; Kulkarni-Kale et al., 2005).

2.2.2 Immune reagents

In epitope mapping, monoclonal antibodies are generally used rather than polyclonal antisera since they recognize a single epitope on a specific antigen. However, polyclonal antibodies react with a great number of epitopes of the antigen and this can make a potential difficulty while identifying one different epitope in two antigens. Also, monoclonal antibodies can be generated against the protein according to the number of epitopes in analyzing the specificity of antibodies in diseases which an immune response is an important (Saint-Remy, 1997; Kulkarni-Kale et al., 2005).

2.2.3 Cross-reactivity

Cross-reactions between antigens occur due to the dynamic surface relations between an epitope and its paratope. Cross-reactivity can be classified into two groups. In the first type of cross-reactivity which is called shared reactivity the same epitope in two related multi-determinant proteins is recognized by an antibody. However, true cross-reactivity as a second type of cross-reactivity occurs when the antibody reacts with an epitope that is different but structurally related to the epitope present in the antigen (Arnon and Van Regenmortel, 1992).

2.3 TYPES OF EPITOPE

2.3.1 Linear and Conformation

Epitopes can be generally classified into two groups namely linear epitopes and conformational epitopes. Linear epitopes are also known as ‘continuous’ or ‘sequential’ epitopes and conformational epitopes are also known as ‘discontinuous’ or ‘assembled’ epitopes. Linear epitopes are the part of contiguous amino acid sequence linked with peptide bond in the antigen but conformational epitopes are non-contiguous in a primary amino acid sequence and they are brought close together by the folding of 3-dimentional protein structure (Barlow et al., 1986; Huang and Honda, 2006).

The sequence and conformation of constituent amino acids provide specificity of linear epitopes however the spatial folding and the conformation of the contributing individual linear epitopes determine the specificity of conformational epitopes (Kulkarni-Kale et al., 2005).

Most epitopes on native proteins are conformation-dependent and when the protein is denatured or fragmented, they can be disappeared. Conformational epitopes continue to exist only in properly folded proteins or large folded fragments while linear epitopes persist after the protein is denatured or is in small peptide fragments. Western Blotting after SDS-PAGE can be used to determine whether an epitope is conformational or linear. If the antibody still binds after the protein has been denatured, the epitope is unlikely to be highly conformational however most proteins are not completely linearized by denaturation and some conformational epitopes maybe still present in the antigen after Western Blotting (Morris, 2001).

It is believed that antibody antigen interaction is often conformation dependent and conformational epitopes may also have functional advantages over linear epitopes. Conformational epitopes also provide structural and binding information correlated to the antigen and give useful information for research, new diagnostic reagent development, vaccine design and disease treatment including viral infections and cancer (Huang and Honda, 2006).

2.3.2 Protein Epitope

Carbohydrates, lipids, nucleic acids and proteins or peptides can be an antigen in the antibody recognition. Epitopes from protein or peptide consist of a minimum number of amino acid residues involved in binding to the antibody. If those residues are sequential in the primary sequence, epitopes from protein could also be a linear. If those residues are folded by its structure, they are non-linear or conformational epitopes (Lu et al., 2009). One study has showed that approximately 10% of the epitopes corresponding antibodies are cross-reactive with a linear peptide fragment of the epitope (Westhof, 1993). Determination of such linear peptide segments will be the initial step in the search for antigenic determinants in pathogenic organisms.

2.3.3 T-cell and B-cell Epitopes

Two types of lymphocytes are involved in the adaptive immune response. T cells are one type of lymphocyte that recognizes antigen and presents antigen fragment to MHC molecules within antigen processing. Another type of lymphocyte is B cells which produce antibodies that can also recognize antigen in its native form (Kindt et al., 2007). T cells recognize T cell epitopes while B cells and antibodies recognize B cell epitopes.

2.3.3.1 T-cell Epitope

T cell epitopes are usually linear peptides produced from protein antigens presented by the antigen presenting cells together with MHC molecules. T-cell epitopes can be used for the design of new therapeutic strategies.

2.3.3.2 B-cell Epitope

B-cell epitopes are parts of proteins or other molecules such as nucleotides, lipids, glycans or chemical compounds that antibodies produced by B-cells bind. Although it is stated that the majority of B-cell epitopes (~90%) from native proteins are conformational rather than linear (Horsfall et al., 1991), experimental determination of epitopes has focused mainly on the identification of linear B-cell epitopes. Identification of B-cell epitopes within an antigenic protein are used in the development of vaccines and in immunodiagnostics (Sela-Culang et al., 2013).

2.3.3.3 Other types of epitope

Ladner (2007) previously indicated that the epitopes can be categorized and used in many aspects such as “functional epitopes”, “structural epitopes”, “contact epitopes”, “binding epitopes, “protective epitopes”, “neutralizing epitopes”, “extracellular epitopes” and “cytoplasmic epitopes”.

2.4 EPITOPE MAPPING METHODS

In 19th century, a significant distinction between ‘structural’ epitopes and ‘functional’ epitopes were done (Van Regenmortel, 1989). Structural epitopes are in contact with the antibody but functional epitopes are amino acid residues that play an important role in binding the antibody. According to this, methods for epitope mapping can be determined by the following approaches.

2.4.1 Structural mapping approaches

X-ray crystallography is one of the structural approaches for epitope mapping. It provides the best resolution of the antigen-antibody complex for epitope recognized by a protective antibody. Although contact residues in antigen-antibody complexes are identified by X-ray crystallography, good crystals of complexes are needed. Due to this reason, highly conformational epitopes on the surface of soluble proteins has usually been preferred to use in X-ray crystallography (Morris, 2001; Clementi et al., 2013).

Another structural approach for epitope mapping is nuclear magnetic resonance (NMR). Contrary to X-ray crystallography, NMR does not need crystals of complexes because it is performed in solution. However, it may not be suitable for highly assembled epitopes since NMR has limited utility for large proteins (Morris, 2001).

Electron microscopy as a method of structural mapping can be used to identify the location of epitopes on very large antigens such as whole viruses because electron microscopy is an unable to identify contact residues therefore it is a low-resolution method for epitope mapping (Morris, 2001; Clementi et al., 2013).

2.4.2 Functional mapping approaches

Since the structural mapping approaches provide information about epitopes at the atomic level, it is not often required for most applications. Therefore, many methods have been developed to identify epitopes at the level of amino acids. These methods are essentially functional and can be separated into four groups: competition methods, antigen modification methods, fragmentation methods and peptide panning or peptide libraries.

2.4.2.1 Competition method

Competition method is used to identify whether two different monoclonal antibodies compete with each other for the same antigenic epitope as they bind to different epitopes on the antigen. This method can show that two antibodies bind to overlapping or non-overlapping epitopes but they don't bind to the same epitope. In this method, ELISA or Surface Plasmon Resonance (SPR) device can be used by immobilizing one antibody and measuring the ability to bind the captured antigen or measuring the sequential binding of antigen and a second antibody. However, this method causes a relatively low degree of mapping resolution (Ladner, 2007).

2.4.2.2 Antigen modification method

Chemical cleavage at specific amino acids is a commonly used alternative method for generating antigen fragments. Direct chemical modification may affect antigen structure and give false positive results. Protection of the antigen from acetylation by the presence of antibody can be used instead of direct chemical modification especially for determining conformational epitopes. Protein foot printing or protection from proteolytic digestion method can be used in determining epitopes. In this method, antigens are treated with proteases in the presence or absence of fairly protease-resistant antibody and differences in digestion are detected by gel electrophoresis (Morris, 2001).

If the antigen is produced from recombinant cDNA or DNA and the region of epitope is known inside the sequence, the site-directed mutagenesis can be applied as a specific mutation for determining epitopes. Also, random mutations can be done by polymerase chain reaction (PCR) in the part of the antigen and epitope-negative mutants can be detected by screening of these mutants (Morris, 2001).

Alanine scanning is the other possible strategy for the epitope mapping. In this technique, alanine mutations are introduced into antigen primary amino acid sequence and the possible alterations affecting the antibody-antigen recognition is evaluated. The reason of using alanine as a mutating amino acid is that alanine amino acid has small molecular weight and neutral charge. This technique can also be used to investigate the role of residues identified through the peptide panning or possible residues involved in the antibody-antigen interaction. It can also help to confirm the competition with a

monoclonal antibody directed against an already characterized epitope in the competitive mapping method (Clementi et al., 2013).

2.4.2.3 Simple fragmentation method

Large fragments of the antigen are commonly produced either enzymatically or recombinantly. Recombinant production can be time consuming but provides precise control of fragments. Proteolytic production is less time consuming and risky because it cleaves an antigen randomly. A simple fragmentation method contains fragmented antigens produced either by proteolytic production especially partial protease digestion or by a recombinant production following by western blotting or high-performance liquid chromatography (HPLC). Identification of these fragments that bind antibody is performed by using mass spectrometry or N-terminal micro-sequencing (Morris, 2001; Ladner, 2007).

2.4.2.4 Peptide libraries and peptide panning method

Peptide library is a way of generating overlapping antigenic fragments. Overlapping peptides can be synthesized on pins (PEPSCAN), on a cellulose membrane support (SPOTS) or on micro-arrays. Generally, this method consists of screening of commercial synthetic peptide libraries correlate with phage display methodology. Peptide panning is one possible epitope mapping approach comprising the selection of short phage-displayed peptides against the interest of monoclonal antibody. In this method, a phage population expressing a defined number of random peptides is amplified in bacteria and repeatedly ‘panned’ on the monoclonal antibody, potentially allowing the selection of peptides to bind to it. Cell-surface displayed and ribosome-displayed libraries can also be done with similar selection methods (Morris, 2001).

2.4.2.4 Mass spectroscopy (MS)

Due to its high sensitivity and rapid analysis time mass spectrometry (MS) has become a powerful tool for epitope mapping in the recent years. Many methods have been developed using mass spectrometry such as limited proteolysis and epitope foot printing, epitope excision and epitope extraction for determining linear epitopes and hydrogen/deuterium exchange and differential chemical modifications of specific amino

acid side chains for conformational and discontinuous epitopes determination (Hager-Braun and Tomer, 2005).

Recently, two MS-based epitope mapping approaches were used: epitope excision and epitope extraction approaches. In epitope excision, antibody is incubated with the antigen and the antigen-antibody complexes are then exposed to enzymatic digestion and the antigen epitopes present the sites of proteolysis are protected by the antibody binding. However, in epitope extraction, antigens are first cleaved by enzymatic digestion and then digested fragments are incubated the antibody (Lu et al., 2009).

In MS-based proteomics literature, two major approaches are indicated in the concepts of the protein biophysics and the characterization of the antibodies. One is an intact or top-down approach and the other is a bottom-up approach. Native electrospray ionization (ESI), ion-mobility measurements, and fragmentation by interaction with electron are classified as an intact or top-down approach and it provides a global view of the protein of interest. However, a bottom-up approach provides detailed information of peptide even amino-acid-residue. Protein footprinting and cross-linking can be exampled for this approach (Zhang et al., 2014).

2.5 EPITOPE PREDICTION DATABASES

The existing approaches for the determination of epitopes such as X- ray crystallography, mass spectrometry, partial proteolysis, mutagenesis analysis, expressed fragments, phage display are expensive, laborious, and time consuming. Also, many epitopes could not be identified properly by using these experimental approaches. Because of this, computational techniques are preferable for predicting epitopes and focusing experimental investigations and improving our understanding of antigen-antibody interactions (EL-Manzalawy and Honavar 2010; Sela-Culang et al., 2013).

Prediction methods are cost effective, fast and reliable methods rather than experimental methods and these prediction methods can be also applied on a genomic scale. Predicting linear B-cell epitopes would be a first step in guiding a genome wide search for B-cell antigens in pathogenic organism (Larsen et al., 2006).

Several epitope databases have been used in the computational techniques and they use the propensity values of amino acid properties, such as hydrophilicity, antigenicity, segmental mobility, flexibility and accessibility to predict antigenicity. It was stated that this accuracy lies in the range of 35–75 (Kulkarni-Kale et al., 2005).

According to a previous study, an existing epitope databases can be categorized into four main groups; T cell oriented such as MHCPEP, SYFPEITHI, FIMM, MHCBN, EPIMHC; B cell oriented such as Bcipep and Epitome; single pathogenic organism oriented such as the HIV Molecular Immunology Database and HCV Molecular Immunology Database; multifaceted database such as AntiJen and IEDB (Huang and Honda, 2006).

The immune epitope database (IEDB) is assumed that the most comprehensive database for experimentally characterized B-cell and T-cell epitopes. IEDB provides several epitope-related analysis and prediction tools for predicting linear and conformational B-cell epitopes and for visualizing the predicted conformational epitopes on the 3D structure of an antigen and for analyzing epitope data (EL-Manzalawy and Honavar, 2010).

Although most of the epitopes on protein antigens are conformational and even discontinuous, most of these methods typically identify linear epitopes. Therefore, most existing databases focus on linear rather than conformational epitopes (Sela-Culang et al., 2013).

2.6 Applications

Since epitopes play a major role in the immune recognition and the immune response, epitope mapping has several potential applications. Saint- Remy et al (1997) divided the applications of epitope mapping into several groups according to the potential application in physiology, pathology, preclinical evaluation of drugs, vaccinations, diagnosis and subtyping of micro-organism to mechanism of drug action.

Epitope mapping provides a better understanding of protein interactions between epitopes and antibodies due to its intrinsic value and generating antibody in defined specificity as research tools and helping to define the immune response to pathogenic organisms. For producing antiviral vaccine or diagnostic antigens, epitopes mapping

have been used for determining all major sites on a protein surface that can elicit an antibody response in mice or humans (Morris, 2001).

As Ladner (2007) compiled several reports which explain each of the purposes, methods and applications of epitope mapping, the epitope mapping has been used to determine the mechanism of a biological process and to identify an epitope of practical value and the binding site of a therapeutic antibody, and to connect a single nucleotide polymorphism (SNP) or other polymorphism to protein structure or antibody binding, to describe antibody binding in patents, to evaluate antibodies raised with a vaccine or design a vaccine, to understand autoimmune diseases, to qualify an antibodies for immunohistochemistry (IHC) detection. It has been also used in diagnostic use, Western Blot analysis, trans-species assays, distinguishing isoforms of antigen, allergen characterization, or therapeutic use and to discover peptide mimic of antigen, to determine structure of antigen, to discover an antigen that can differentiate between antibodies that arise from immunization and those that arise from infection.

CHAPTER 3

MATERIALS & METHODS

3.1 HELICOBACTER PYLORI

An antral biopsy specimen was taken from a 39 years old female Turkish patient with duodenal ulcer at the Istanbul Teaching and Research Hospital.

3.2 CLONING

3.2.1 Recombinant vector

A recombinant vector was constructed earlier in our laboratory by cloning the cagA 5' conserved region of *H. pylori* strain into pET-28a(+) plasmid (Ulupınar, 2012) and this recombinant plasmid was used in this study.

The size of pET-28a(+)/rcagA recombinant vector is 7022 bp containing 1785 bp rcagA gene that encoded a rCagA protein of 67 kDa molecular weight.

3.2.2 Virtual cloning analysis of recombinant vector

All virtual restriction cloning analysis was done by the "ApE plasmid editor software" and "SnapGene software". The DNA sequence of the rcagA gene containing the pET-28a(+) plasmid was uploaded into "ApE plasmid editor software" and "SnapGene software" respectively and all available restriction enzyme sites were screened on the rcagA gene. In the SnapGene software, the restriction enzymes were selected and screened. According to the available restriction enzyme cutting sites, two recombinant protein fragments were determined on the pET-28a(+)/rcagA recombinant plasmid.

3.2.3 Construction of the recombinant fragments

3.2.3.1 Construction of the first rcagA fragment

The rCagA protein developed earlier by cloning the cagA 5' conserved region cagA gene of *H. pylori* (Ulupınar, 2012) was used as a first rcagA fragment. It contains a 594 amino acids and a molecular weight of 67 kDa.

3.2.3.2 Construction of the second rcagA fragment

For a second rcagA fragment, AvrII and Xho1 restriction enzymes were selected from all restriction enzymes screened by Snapgene software. This software showed that the AvrII restriction enzyme has one restriction cutting site on our own rcagA gene and cuts at 6450 bp of pET-28a(+)/rcagA recombinant vector. The Xho1 has also one cutting site on our own rcagA gene and cuts at 6860 bp of pET-28a(+)/rcagA recombinant vector. By cutting with both AvrII and Xho1 enzymes and deleting this region, a second fragment of rcagA was constructed which contains 469 amino acids and a molecular weight of 53 kDa.

3.2.3.3 Construction of the third rcagA fragment

The output of SnapGene software showed that Spe1 restriction enzyme is a suitable for constructing a third rcagA gene fragment because it has two-cutting recognition sites on the rcagA gene located at 5743 and 6781 bp inside the pET-28a(+)/rcagA recombinant plasmid. By cutting with Spe1 restriction enzyme and deleting this site, a third rcagA gene fragment in the recombinant plasmid was constructed that contains 248 amino acids and a molecular weight of 28.1 kDa

3.3 PURIFICATION OF RECOMBINANT VECTOR

E. coli carrying rcagA gene, which is cloned from cagA 5' conserved region, inside the pET-28a(+) plasmid was cultured into LB/Kan broth and incubated at 37°C overnight. The pET-28a(+)/rcagA plasmid from *E. coli* BL21(DE3)pLySs cells was then isolated using GeneJET Plasmid Midiprep Kit (Thermo Scientific, USA) according

to the manufacturer's instructions. The concentration of this plasmid was measured using Nanodrop 2000 (Thermo Scientific, USA) as a 685 ng/µl.

3.4 CLONING OF THE FIRST rcagA GENE FRAGMENT

The isolated pET-28a(+)/rcagA recombinant plasmid (685ng/ul) was transformed into *E. coli* BL21(DE3) competent cells. 10 µl of pET28a(+)/rcagA recombinant plasmid was added into 200 µl and left on ice for 30 minutes. After 30 minutes incubation, heat shock was performed at 42°C for 90 seconds (s). 950 µl LB medium was added and the cells were incubated at 37°C for 30 minutes. After incubation, short centrifugation was done at 3000 rpm for 3 minutes. 700 µl of supernatant was discarded, remaining pellet was mixed and 100 µl of cell suspension was plated out on LB plates containing 50 µl/ml kanamycin. The plates were inverted and incubated at 37°C for overnight. For screening colonies, one colony was selected and inoculated into 2 ml LB/Kan broth. After an overnight incubation at 37°C, they were tested for IPTG induction whether or not a colony has expressed the rCagA protein.

3.5 CLONING OF THE SECOND rcagA GENE FRAGMENT

3.5.1 Cutting with AvrII and Xho1 restriction enzymes

A reaction mixture containing 25 µg pET-28a(+)/rcagA (685 ng/µl) recombinant plasmid, 5 µl AvrII restriction enzyme, 5 µl Xho1 restriction enzyme (Thermo Scientific, USA), 8 µl 10X Buffer Tango and 22 µl nuclease-free water (Thermo Scientific, USA) was prepared. After mixing gently and spinning down for a few seconds, it was incubated at 37°C for 2 hours. Thermal inactivation was done at 80°C for 20 minutes.

3.5.2 Agarose Gel Electrophoresis

3.5.2.1 Preparation of 1X TBE

1X TBE was prepared from 10X TBE stock solution (Invitrogen, USA) by diluting 100 ml 10X TBE with 900 ml distilled water.

3.5.2.2 Agarose gel preparation

0.4 g agarose powder dissolved in 40 ml of 1X TBE buffer by heating in a microwave oven (Beko MD1510) to make 1% agarose gel. Before adding 2 µl of Pronasafe Nucleic acid staining solution (Conda, Spain) into agarose solution, it was cooled down and then poured directly into the horizontal casting tray.

3.5.2.3 Running agarose gel

10 µl of cut plasmid mixed with 2 µl of 6x Loading Dye (Thermo Scientific, USA), 10 µl uncut plasmid mixed with 2 µl of 6x Loading Dye (Thermo Scientific, USA) and 5 µl Mass Ruler DNA Ladder (Thermo Scientific, USA) were loaded. The agarose gel was covered by filling with 1X TBE Buffer. Power supply (VWR Power Source 300 V) adjusted to 90 V for 40 minutes.

3.5.3 Ligation

3.5.3.1 Making a blunt-end

A reaction solution containing the 70 µl recombinant plasmid that previously cut with AvrII and Xho1 restriction enzymes was prepared inside the same tube. 5 µl Klenow Fragment (Thermo Scientific, USA), 2 µl dNTP Mix (Thermo Scientific, USA), 9 µl 10X Reaction Buffer and 4 µl nuclease-free water (Thermo Scientific, USA) were also added inside this tube. After mixing and spinning down for a few seconds, it was incubated at 37°C for 10 minutes. The reaction was stopped by heating at 75°C for 10 minutes.

3.5.3.2 Direct blunt-end self-ligation

Direct blunt end self-ligation was performed in the same tube by adding 5 µl T4 DNA Ligase (Thermo Scientific, USA), 2 µl dNTP mix (Thermo Scientific, USA), 12 µl 10X T4 DNA Ligase Buffer and 6 µl 50 % PEG 4000 (Thermo Scientific, USA). After mixing and spinning down for a few seconds, it was incubated at 22°C for 1 hour. Heat inactivation was done at 65°C for 10 minutes.

3.5.4 Transformation into competenet *E. coli* BL21(DE3) cells

3.5.4.1 $CaCl_2$ solution preparation

21.9 g of 100mM $CaCl_2.6H_2O$, 1.02 g of 5 mM $MgCl_2.6H_2O$ and 5 ml of 5 mM Tris-HCl (pH:7.6) from 1M stock Tris-HCl solution were dissolved into 995 ml distilled water and sterilized by autoclaving at 121°C for 15minutes.

3.5.4.2 Competent bacterial cells preparation

The *E. coli* BL21(DE3) stock culture was plated out on LB agar applying streak-plate technique. After incubation overnight at 37°C, a single bacterial colony was picked and subcultured into 5 ml LB broth using a sterile loop and incubated at 37°C, 120 rpm overnight. 500 µl of this overnight culture was then transferred into 50 ml LB broth in a 250 ml flask and incubated in a shaker at 37°C until an absorbance of 0.3 - 0.4 OD at 600 nm was reached. The culture was divided equally into two different 50 ml sterile tubes and centrifuged at 5000 rpm for 5 minutes at 4°C. After discarding the supernatant, the pellet was resuspended in a 25 ml of $CaCl_2$. This was incubated on ice for 30 minutes and then spun at 5000 rpm for 5 minutes at 4°C in a refrigated centrifuge. The pellet was resuspended in 1 ml of $CaCl_2$ solution and it was left on ice for 2 hours.

3.5.4.3 Transformation

The self-ligated plasmid was gently mixed with 200 µl of competent cells inside a micro-centrifuge tube and it was left on ice for 30 minutes. After 30 minutes incubation, the heat shock method was done at 42°C for 90 seconds (s) and then it was immediately transferred on ice for 2 minutes. 950 µl LB medium was added and the cells were

incubated at 37°C for 30 minutes then centrifuged at 3000 rpm for 3 minutes. 700 µl of the supernatant was discarded, the remaining pellet was mixed and 100 µl of cell suspension was plated out on LB plates containing 50 µl/ml kanamycin. The plates were inverted and incubated at 37°C overnight.

3.5.5 Screening of all bacterial colonies

All grown colonies were selected for screening and each one was inoculated into 2 ml LB/Kan broth then incubated at 37°C overnight. The bacteria were then tested for protein expression by induction with isopropyl-β-D-thiogalactoside (IPTG). After then, the bacterial lysates of each of 3 hours IPTG inductions were loaded on the SDS-PAGE gel.

3.6 CLONING OF THE THIRD rcagA GENE FRAGMENT

3.6.1 Cutting with Spe1 restriction enzyme

A reaction mixture containing 25 µg pET-28a(+)/rcagA (685 ng/µl) recombinant plasmid, 5 µl Spe1 restriction enzyme (Thermo Scientific, USA), 8 µl 10X Buffer Tango and 27 µl nuclease-free water was prepared. After mixing gently and spinning down for a few seconds, it was incubated at 37°C for 2 hours.

3.6.2 Agarose Gel Electrophoresis

10 µl of cut plasmid + 2 µl of 6x Loading Dye (Thermo Scientific, USA), 10 µl uncut plasmid + 2 µl of 6x Loading Dye (Thermo Scientific, USA) and 5 µl Mass Ruler DNA Ladder (Thermo Scientific, USA) were loaded into 1% agarose gel. This agarose gel was covered by filling with 1X TBE Buffer. Power supply (VWR Power Source 300 V) adjusted to 90 V for 40 minutes.

3.6.3 Ligation

3.6.3.1 Self-ligation of vector after gel extraction

5984 bp of cut plasmid was excised from agarose gel by using QIAquick Gel Extraction Kit (QIAGEN, Germany) according to the manufacturer's instructions. Concentration of isolated vector was determined both Nanodrop 2000 (Thermo Scientific, USA) and agarose gel.

3.6.3.2 Direct self-ligation

Since the low concentration of recombinant vector was obtained after gel extraction, direct self-ligation was performed. A ligation reaction mixture that contains the 60 µl recombinant plasmid cutting with Spe1 restriction enzyme, 10 µl 10X T4 DNA Ligase buffer, 5 µl T4 DNA Ligase (Thermo Scientific, USA) and the 25 µl Nuclease free water (Thermo Scientific, USA) was prepared. After mixing thoroughly and spinning briefly, it was incubated 20 minutes at 22°C. Thermal inactivation was done at 65°C for 10 minutes. For confirmation, 10 µl ligation mixtures with 2 µl of 6x Loading Dye (Thermo Scientific, USA) were loaded onto 1% agarose gel using 5 µl Mass Ruler DNA Ladder (Thermo Scientific, USA).

3.6.4 Transformation into competenet *E. coli* BL21(DE3) cells

The same transformation protocol described above were applied by adding all ligation mixture into 200 µl *E. coli* BL21(DE3) competent cells by a heat shock and following steps.

3.6.5 Screening of all bacterial colonies

All grown up colonies were selected and inoculated into 2 ml LB/Kan broth. After incubation overnight at 37°C, a three hour IPTG induction was done for each one of them. The bacterial lysates of each of 3h IPTG inductions were loaded on the SDS-PAGE gel whether or not each colony has expressed the rCagA protein as above.

3.7. RECOMBINANT PROTEIN INDUCTION

E. coli carrying recombinant plasmids were induced with isopropyl-β-D-thiogalactoside (IPTG) to express the CagA protein. A single colony was transferred into 5 ml of LB Broth containing 5 µl kanamycin and incubated at 37°C with shaking at 200 rpm overnight. Also, 100 µl broth culture can be stored in next day with 50% sterile glycerol containing broth at -80. After incubation 500 µl bacteria were inoculated into 10 ml of antibiotic-containing media and incubated at 37°C with shaking at 200 rpm until the optical density reached 0,60 at 600 nm. 1 ml of bacterial culture was taken as non-induced control, and protein expression was induced by 1 Mm IPTG. The culture was incubated at 37°C for 3 hours, and then 1 ml of broth was taken after 1, 2 and 4 h of incubation. The broth culture was centrifuged at 6,000 rpm for 10 minutes at 4°C, supernatant was discarded and the pellet was stored at –20°C.

3.8. SCREENING OF ALL POSITIVE COLONIES BY IPTG INDUCTION

The pellets from 3h IPTG induction of each colony were suspended with 100 µl 2X Laemmi sample buffer and heated at 5 minutes by boiling water. All pellet samples were loaded on a 12% SDS-PAGE gel. Non-induced samples were also loaded onto gel as a control

Table 3.1 2X Sample Buffer composition

2X Sample Buffer	**Amount**	**Final**
1M Tris/HCl (pH 6.8)	1,08 ml	120 mM
100% (v/v) glycerol	1,8 ml	20%
10% (w/v) SDS	3,6 ml	4%
Bromophenol Blue	0,5 ml	
β-mercaptoethanol	1 ml	10%
ddH_2O	2,52 ml	

Table 3.2 4X Sample Buffer composition

4X Sample Buffer	**Amount**	**Final**
1M Tris/HCl (pH 6.8)	2,4 ml	240 mM
100% (v/v) glycerol	4 ml	40%
SDS	0,8 g	8%
Bromophenol Blue	4 mg (500µl)	0,04%
β-mercaptoethanol	0,5 ml	5%
ddH_2O	3,1 ml (2,6 ml)	

3.9. SDS-PAGE

The vertical electrophoresis set was obtained. Glass parts, separators and comb were cleaned with 70% ethanol and dried. The vertical electrophoresis set was assembled. The separating gel was prepared in a 15 ml centrifuge tube.

Table 3.3 %12 Separating Gel composition

%12 Separating Gel (12 lanes comb)	**(5ml)**	**(5ml)**
Distilled water	1.7 ml	1700 µl
30 % Acrylamide/Bis-Acrylamide	2 ml	2000 µl
1.5 M Tris-HCl (pH 8.8)	1.3 ml	1300 µl
10% SDS	0.05 ml	50 µl
10% APS	0.05 ml	50 µl
TEMED	0.02 ml	2 µl

The front glass was labeled with marker at 1 cm below of the comb. The separating gel was poured in between two glasses of vertical electrophoresis set up to the level of the label. 1 ml of isopropanol was poured on top of separating gel in order

to make the top of the gel a straight line and the gel was let to solidify for 30 min at RT. The isopropanol was removed and the gel was washed with distilled water 4-5 times and the water was drained completely. The stacking gel was prepared in a 15 ml centrifuge tube.

Table 3.4 %5 Stacking Gel composition

%5 Stacking Gel	**(2ml)**	**(2.5 ml)**
Distilled water	1360 µl	1804 µl
30% Acrylamide/Bis-Acrylamide	332 µl	312 µl
1M Tris-HCl (pH 6.8)	252 µl	315 µl
10% SDS	20 µl	25 µl
10% APS	20 µl	25 µl
TEMED	2 µl	2.5 µl

The stacking gel was poured on top of separating gel, the comb was placed and the gel was let to solidify for 30 min. The prepared gel can be stored at 4˚C. The wells were labeled before removing the comb prior to loading the samples.

3.9.1 Running SDS-PAGE

The half of the tank was filled with 1X SDS running buffer.

Table 3.5 SDS running buffer composition

SDS running buffer	**10X**	**1X (pH 8.3)**
Tris Base	7.575 g	3 g
Glycine	36 g	14.4 g
SDS	2.5 g	1 g
dH_2O	250 ml	1000 ml

The vertical electrophoresis apparatus which was prepared before was placed in tank and the comb at the head of gel was removed. 5 µl of page ruler plus prestained

protein ladder (Thermo Scientific, USA) and 5 µl of samples that mixed with 100 µl 2X sample buffer and then heated at 5 minutes by boiling water were loaded into the well. The gel was run at 150V, 30mA for 70-120 min. Then, the power supply was powered off and the vertical electrophoresis apparatus was displaced from the tank, the apparatus was disassembled and the glasses were separated carefully. The stacking gel was removed from the separating gel and the gel was labeled from the corner of marker lane by a small cut at bottom. The separating gel was transferred to a box filled with distilled water and it was washed for 1-2 minutes. The gel was placed in Coomassie Brillant Blue (CBB) staining for 60 minutes to stain the samples.

Table 3.6 Staining solution composition

Staining solution	
0.1 % CBB	0.5 g
2 % Glacial acetic acid	10 ml
40 % MeOH (~1 hour)	200 ml
dH_2O	290 ml

The gel was destained for 2 hours and a photo was taken by scanner. Every 30 minutes the destaining solution was changed to clarify the gel completely.

Table 3.7 Destaining solution composition

Destaining Solution	
2 % Glacial acetic acid	10 ml
40 % MeOH (~1 hour)	200 ml
dH_2O	290 ml

3.10. EPITOPE DETERMINATION

The predicted epitopes on these rCagA proteins was determined by using Emini Surface Accessibility Prediction program "http://tools.immuneepitope.org/tools/".

According to this program, 6 epitopes were found in the first fragment (28 kDa), 11 epitopes in the second fragment (60 kDa) and 14 epitopes in the third fragment (67 kDa) of the rCagA proteins were predicted. The epitopes on each of these fragments were also predicted by the SnapGene program. The confirmation and localization of these predicted epitopes on each of these protein fragments were done by Western Blot using monoclonal antibodies.

3.11. WESTERN BLOT

3.11.1 Monoclonal antibodies (Mabs)

Four commercially available monoclonal antibodies (Mabs) and three of our own developed monoclonal antibodies were used as a primary antibody in western blot respectively.

- B237H Mab (Abcam, Cambridge, England) that recognizes an epitope localized within the 580 amino acids of *H. pylori* CagA C-terminal end.

- B818M Mab (Abcam, Cambridge, England) that recognizes an epitope localized between 562 and 795 amino acids of *H. pylori* CagA protein.

- A-10 Mab (Santa Cruz Biotech., CA, USA) that recognizes an epitope localized within the 1-300 amino acids of *H. pylori* CagA protein.

- His-probe Mab (H-3) (Santa Cruz Biotech., CA, USA) that recognizes a His-tagged recombinant protein.

- BS-53 Mab our own monoclonal antibody that produced in our laboratory by using hybridoma techniques.

- CK-02 Mab our own monoclonal antibody that produced in our laboratory by using hybridoma techniques.

A commercially available goat anti-mouse IgG-HRP (sc-2005) polyclonal antibody (Santa Cruz Biotech., CA, USA) was used as a secondary antibody in western blot.

3.11.2 Blotting

A 12% SDS-PAGE gel was prepared and 5 µl of page ruler plus prestained protein ladder (Thermo Scientific, USA) and 2.5 µl of prepared samples (1^{st} fragment, 2^{nd} fragment and 3^{th} fragment) were loaded into the well. The gel was run at 150V, 30mA for 70-120 min. After running gel, blotting was performed.

At first, transfer buffer containing 25 mM Tris pH=8.3, 192 mM glycine with 20% (v/v) methanol and 0.05% (w/v) SDS was prepared by mixing 3.03 g Tris, 14.4 g glycine, 0,5 g SDS and 200 ml of methanol; adding distilled deionized water (dd H2O) to 1 liter. Secondly, PVDF membrane (BioRad, USA) and filter paper (BioRad, USA) was cut to the dimensions of the gel and this gel, filter paper and fiber pads were equilibrated in transfer buffer about 15 minutes. PVDF membrane (BioRad, USA) was firstly soaked in the methanol solution (Sigma Aldrich, USA) for 5 minutes and then it was equilibrated in transfer buffer at least 5 minutes. The gel sandwich was prepared according to the manufacturer's instructions and it was placed in the module. The frozen ice cooling unit which was stored at -20 ^{o}C was added, the tank was filled with transfer buffer, a standard stir bar was added to help maintain even buffer temperature and ion distribution in the tank and the tank was placed on a magnetic stirrer. The speed was set as slow as possible to keep ion distribution even. Blotting was run for 70 min at 100 V, 300 mA. After blotting, the gel was stained with staining solution to indicate that the proteins were blotted.

3.11.4 Blocking and antibody treatment

For washing the PVDF membrane, 1X TBS-T was prepared from 10X TBS buffer solution as shown in Table 2.7.

Table 3.8 TBS Buffer and TBS-T Buffer compositions

TBS Buffer, 10 x

Reagent	Final Conc.	Volume	Mass	1x Final Conc.
Tris base (FW 121.14)	0.2 M		12.114 g	20 mM
NaCl (FW 58.44)	1.5 M		43.83 g	150 mM
H_2O		QS to 500 mL		

Adjust pH to 7.2 - 7.4 with HCl

TBS-T, 1 x

For TBS-T, add Tween 20 (Polyoxyethylene sorbitane monolaureate) to 0.1%.
For 1 L, use 100 mL 10 x TBS, 900 mL dH_2O and 1 mL Tween 20.

After blotting, the PVDF membrane was washed once with TBS-T and blocked with 5% Skim milk in TBS-T (20 mM Tris Buffered Saline containing 0.1% Tween 20 with pH 7,4) overnight at 4°C on shaker. After blocking, membrane was incubated with 10 ml of mouse anti-CagA Mab diluted in blocking solution inside the container for 1 h at room temperature (RT) on shaker. (high speed). Then, this membrane was incubated with 10 ml of HRP-linked goat anti-mouse secondary antibody (sc-2005, Santa Cruz, 1:2000, 0.1µg/ml) diluted in blocking solution for 40 min at room temperature (RT) on shaker and membrane was washed with TBS-T (5 x 5 min) on shaker (high speed).

3.11.5 Detection

The membrane was placed on the plastic sheet which was placed in a film cassette. 200 µl Luminata Crescendo Western HRP Substrate (Millipore, USA) was used to add onto the membrane and it was incubated for 1-2 minutes. At this point, the excess liquid was removed by enclosing the membrane in a sheet of plastic wrap, being careful to remove any air bubbles between the membrane and the wrap. X-ray film was placed onto the membrane and expose for 5sec, 60 sec, 1 min, 5 min and 15 minutes in dark room. Then, the X-ray film was taken develop in the X-ray film processor.

3.12 RECOMBINANT CagA PROTEIN PURIFICATION

The recombinant CagA protein was purified using the B-PER 6xHis Fusion Protein Purification Kit (Pierce, USA). The column and buffers were equilibrated to room temperature. A 100 ml of induced *E. coli* culture of bacteria (OD600 = 1.5-3.0) was centrifuged at 6000 rpm for 10 min at 4 ^{o}C and the pellet was obtained. The cell pellet was resuspended with 15 ml of 0.1% SDS in PBS solution by pipetting up and down until the cell suspension was homogenous. A 15 ml of sample was sonicated on ice for 5x15sec with 1 min intervals and then it was centrifuged at 10.000xg for 30 min at 4^{o}C and the supernatant (lysate) was transferred to a new tube. The Nickel Chelated Column was uncapped and the sodium azide storage solution was allowed to drain from the gel bed. The gel bed was prepared by adding 10 ml (2 × 5 ml) of B-PER® Reagent and allowing it to flow through the column. 15 ml of sample (2 × 5 ml) was applied up to the column and it flowed through the gel bed. Flow-through fractions was collected and saved. Column was washed by adding at least 6 ml (2 × 3 ml) of Wash Buffer 1 and allowing it to flow through the gel bed. Flow-through fractions was collected and saved. Column was washed by adding at least 9 ml (3 × 3 ml) of Wash Buffer 2 and allowing it to flow through the gel bed. Flow-through fractions was collected and saved. The 6xHis-tagged protein was eluted by adding 6 ml (2 × 3 ml) of Elution Buffer and collecting the fractions that emerge.

The concentration of all samples was monitored by measuring the absorbance at 280 nm and by using Bradford Assay. All samples including flow-through fraction, wash 1 and wash 2 fractions and elution 1 and elution 2 were mixed with 4X sample buffer with 3:1 ratio and then boiled for 5 min. After that, 5 µl marker and 7 µl of each of fraction was loaded onto 12% SDS-PAGE gel.

3.13 INDIRECT ELISA

The rCagA proteins (1st rCagA, 2nd rCagA and 3th rCagA) were confirmed with our own anti-rcagA monoclonal antibodies (BS-53, CK-02, BS-51, A-10, B818M and B-237H) for immunoreactivity using the indirect ELISA.

- 96-well ELISA plates (Grenier, Austria) were coated with 100 µl antigen solution containing 50 ng protein diluted with PBS.

PBS Buffer: Firstly, a stock of 500 mM phosphate (PO_4^-) buffer was prepared. For this, 34.8 g K_2HPO_4 was dissolved in 400 ml dH_2O and 13.6 g KH_2PO_4 was dissolved in 200 ml dH_2O. Then, KH_2PO_4 solution was added into K_2HPO_4 slowly until the pH became 7.2 and was stored at 4°C. To prepare fresh 10 mM PBS containing 150 mM NaCl, 20 ml stock buffer was mixed with 980 ml dH_2O and 8.77 g NaCl was dissolved in this solution.

- Plates were washed 3 times with PBS-T by using ELISA washer (Nunc-Immuno Wash 12).

PBS-T: PBS buffer containing 0,05% Tween 20: 0.5 ml Tween 20 (Merck) was added to 1 liter PBS buffer.

- Plates were coated with 200 µl, 1% skim milk (Fluka, Switzerland) in PBS for blocking non-specific sites. It was incubated for 1 hour at 37°C or for 2 hours at RT.
- Plates were washed 3 times with PBS-T.
- 100 µl containing 0.4 µg/ml of each anti-CagA monoclonal antibodies diluted with PBS was added and incubated for 1 hour at 37 °C.
- Plates were washed 3 times with PBS-T.
- Wells were incubated for 1 hour at 37°C with goat anti-mouse IgG conjugated with HRP (1:2000 diluted with PBS).
- Plates were washed 6 times with PBS-T.
- After the washing step, 100 µl TMB substrate buffer was added and incubated for 30 minute in the dark at room temperature (RT). The enzymatic reaction was stopped with 100 µl 1M HCl and the OD was measured at 450 nm by ELISA Reader (BioRad, USA) after 30 minutes addition of substrate.

CHAPTER 4

RESULTS

4.1 CLONING THE cagA GENE5' CONSERVED REGION

The recombinant cagA (rcagA) gene cloned from the cagA 5' conserved region was virtually analyzed in the pET-28a(+) using the SnapGene software. Figure 4.1 shows this recombinant plasmid containing rcagA gene (pink arrow).

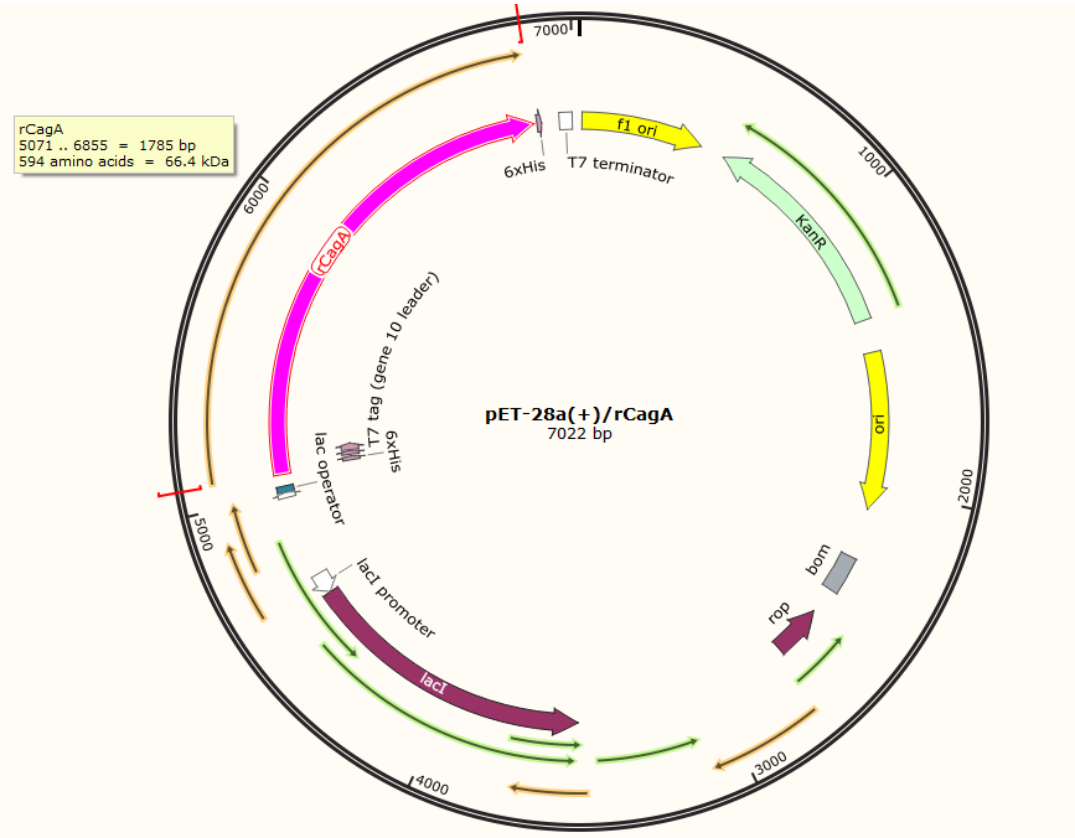


Figure 4.**1** The " pET-28a(+)/ rcagA " recombinant plasmid

4.2 Virtual cloning analysis of recombinant vector

All the restriction enzymes cutting sites were screened on our recombinant plasmid pET-28a(+)/rcagA using the SnapGene program shown in Figure 4.2 and we found three restriction enzymes that can cut at three restriction sites on our recombinant cagA gene (rcagA) shown in Figure 4.3.

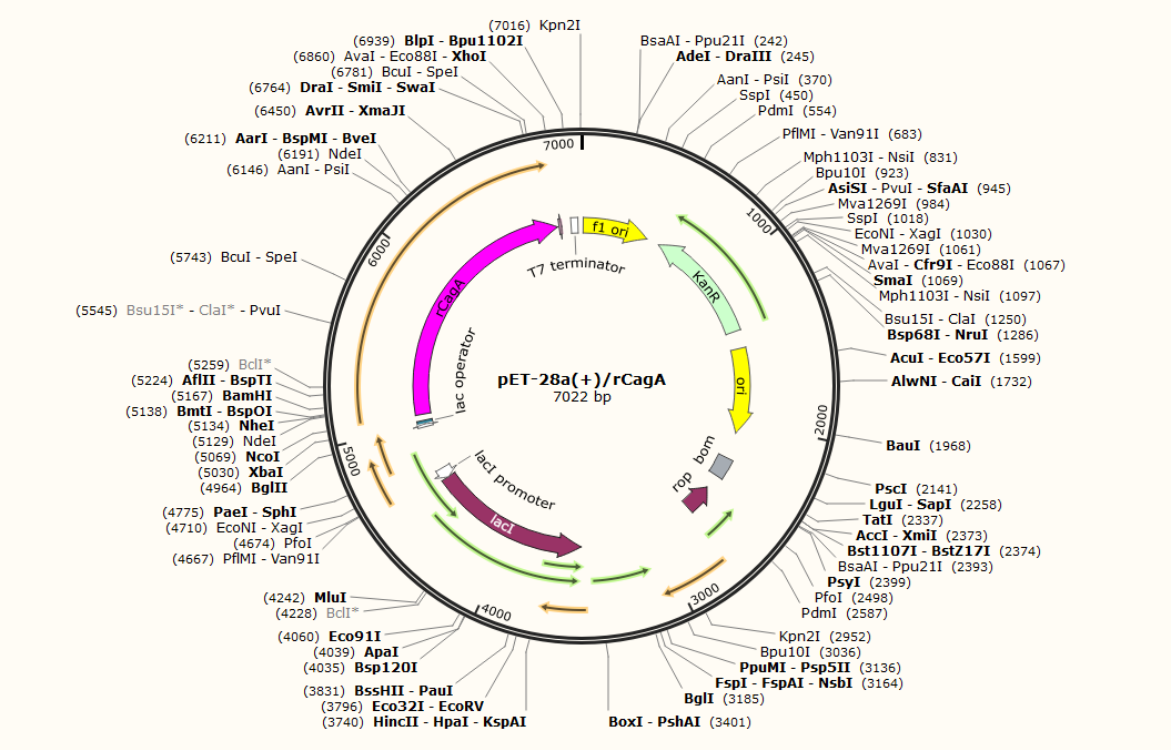


Figure 4.2 Screening of all restriction enzymes on the recombinant plasmid.

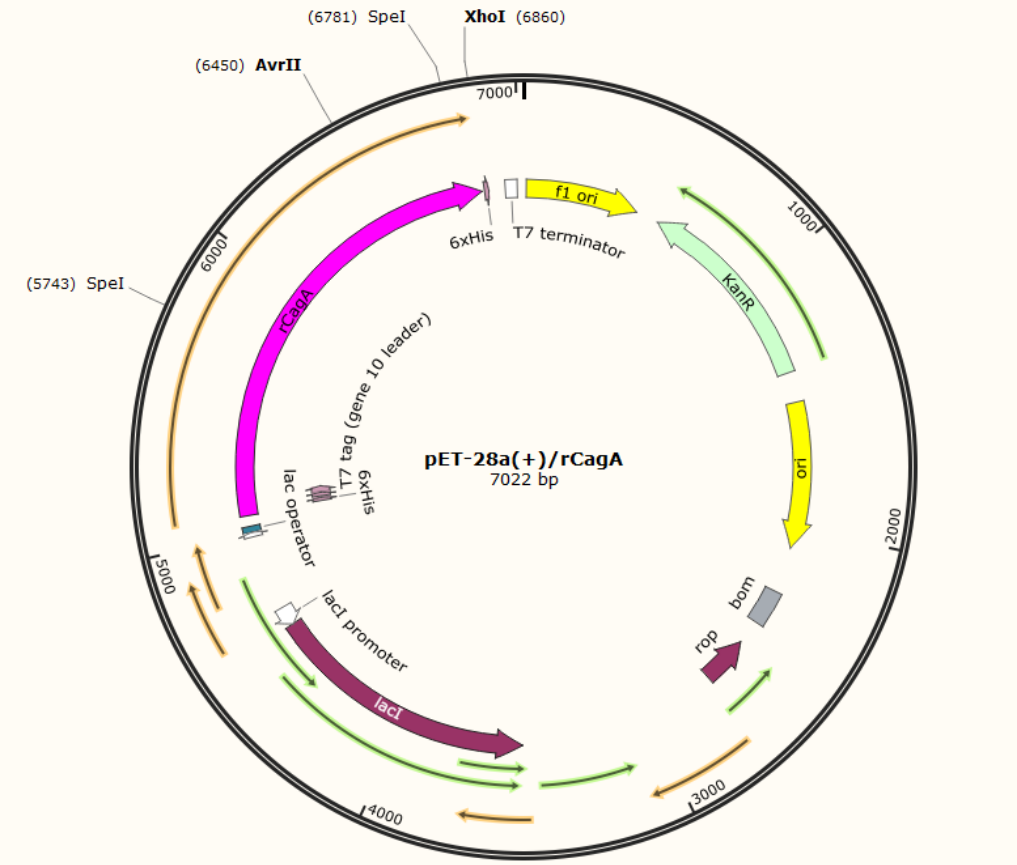


Figure 4.3 Selection of the three suitable restriction enzymes on the recombinant plasmid.

4.2.1 Construction of the first rcagA fragment

The first gene fragment prepared earlier by Z. Ulupınar was used in this study. Figure 4.4 shows that this recombinant cagA (rcagA) gene that cloned from 5' conserved region of *H. pylori* inside pET-28a(+) vector in a linear form.

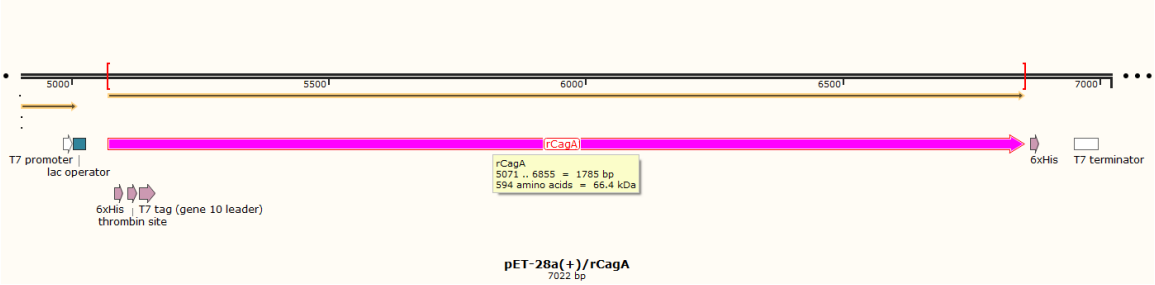


Figure 4.4 The rcagA (pink arrow) gene that cloned from *H. pylori* cagA 5' conserved region in the pET-28a(+) vector

4.2.2 Construction of the second rcagA fragment

For constructing the second rcagA gene fragment, AvrII and Xho1 restriction enzymes were selected. These enzymes cut the rcagA gene at 6450 and 6860 bp inside the pET-28a(+)/rcagA recombinant plasmid as seen in the Figure 4.5. A blunt end was performed from an uncompatible sticky ended vector by filling gaps with appropriate nucleotides and the blunt-ended self ligation of this recombinant vector after cutting AvrII and Xho1 restriction enzymes were shown in the Figure 4.6. Figure 4.7 shows the second rcagA gene fragment after cutting with AvrII and Xho1 restriction enzymes and cloning into *E.coli* BL2(DE3)pLySs cells.

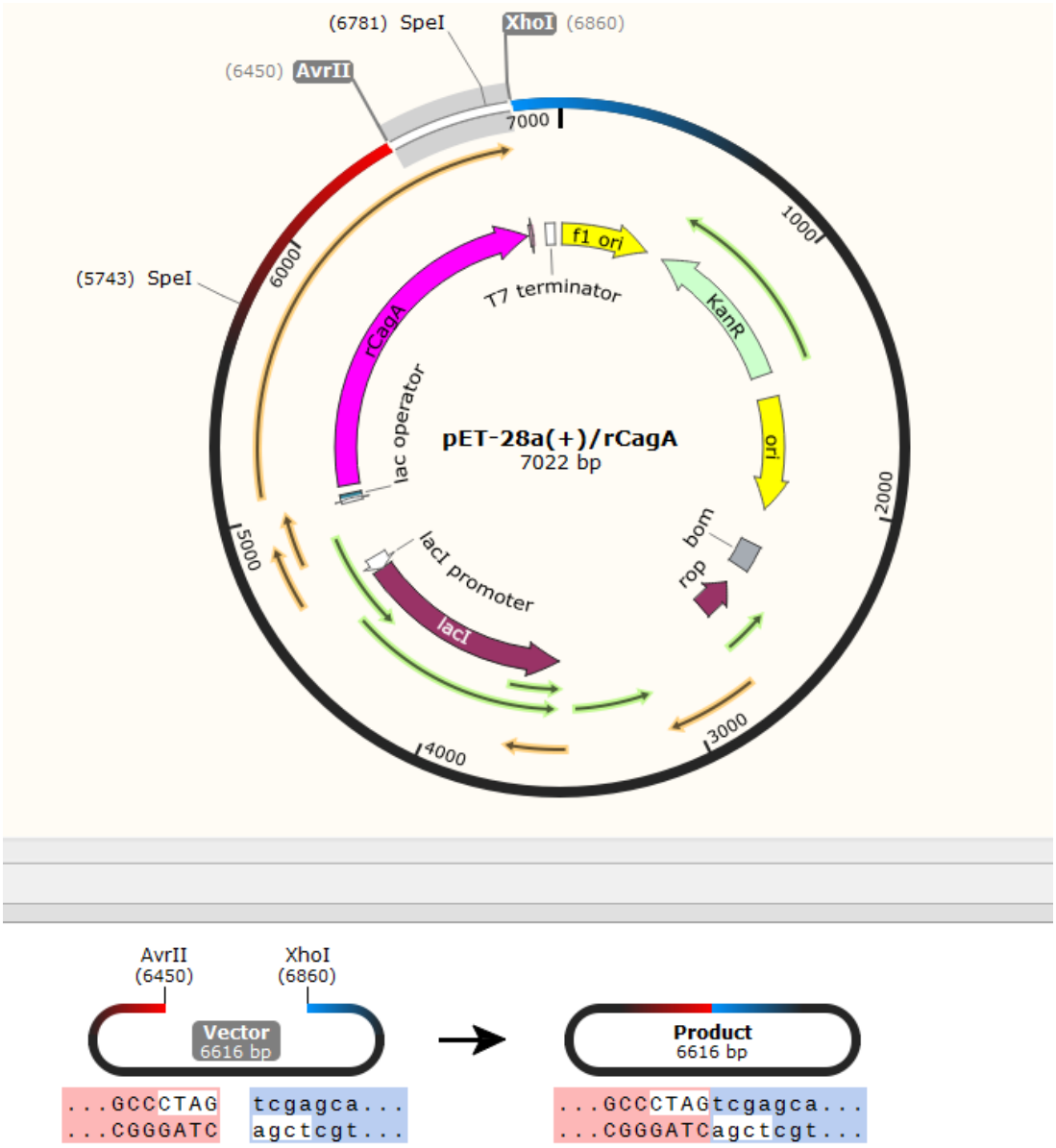


Figure 4.5 The cutting of the recombinant plasmid with AvrII and Xho1 restriction enzymes at the sites of 6450 bp and 6860 bp.

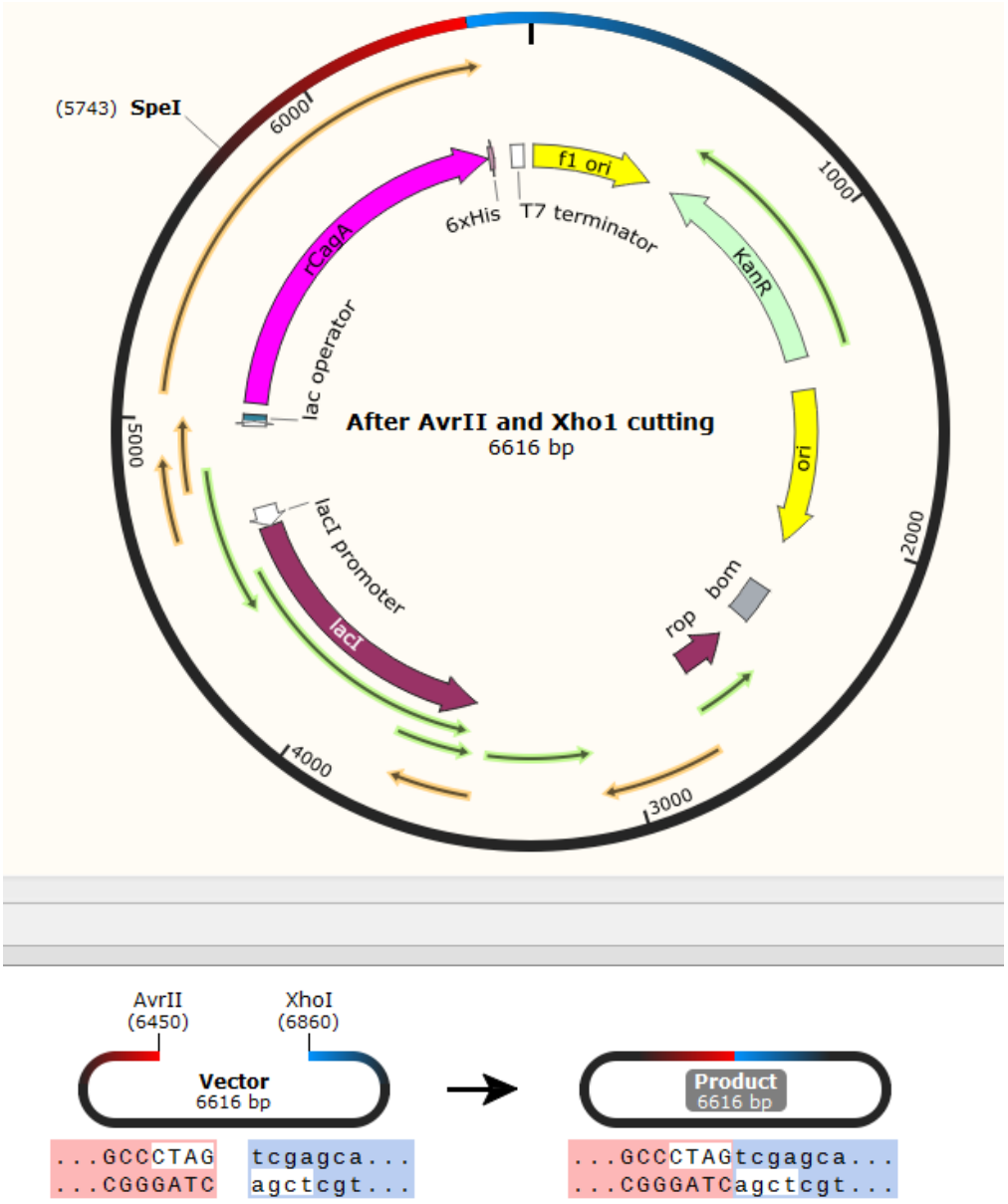


Figure 4.6 The blunt ended self-ligation after cutting of recombinant plasmid with AvrII and Xho1 restriction enzymes. A newly formed recombinant plasmid as a product is 6616 bp.

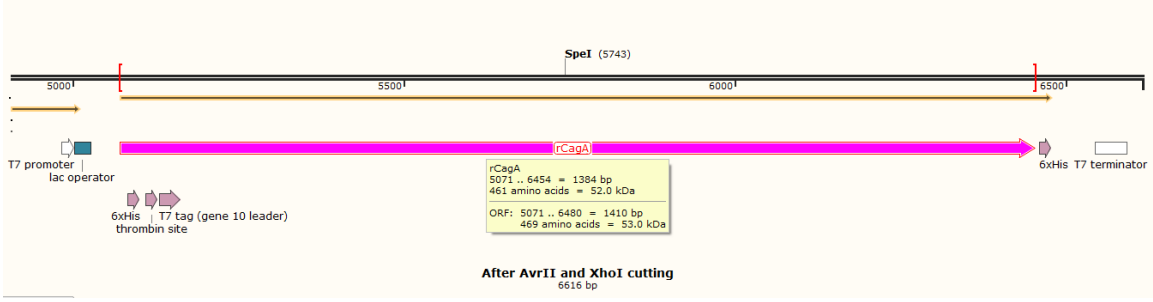


Figure 4.7 The second rcagA gene fragment after cutting with AvrII and Xho1 restriction enzymes.

4.2.3 Construction of the third rcagA fragment

For constructing the third rcagA gene fragment, Spe1 restriction enzyme was selected. It cut the rcagA gene at 5743 and 6781bp inside the pET-28a(+)/rcagA recombinant plasmid as seen in the Figure 4.8. The self ligation of this recombinant plasmid after cutting with Spe1 restriction enzyme was shown in Figure 4.9.

The third fragment of rCagA protein was constructed as a 28.1 kDa molecular weight after cutting with Spe1 restriction enzyme and cloning into *E.coli* BL2(DE3)pLySs cells as shown in Figure 4.10.

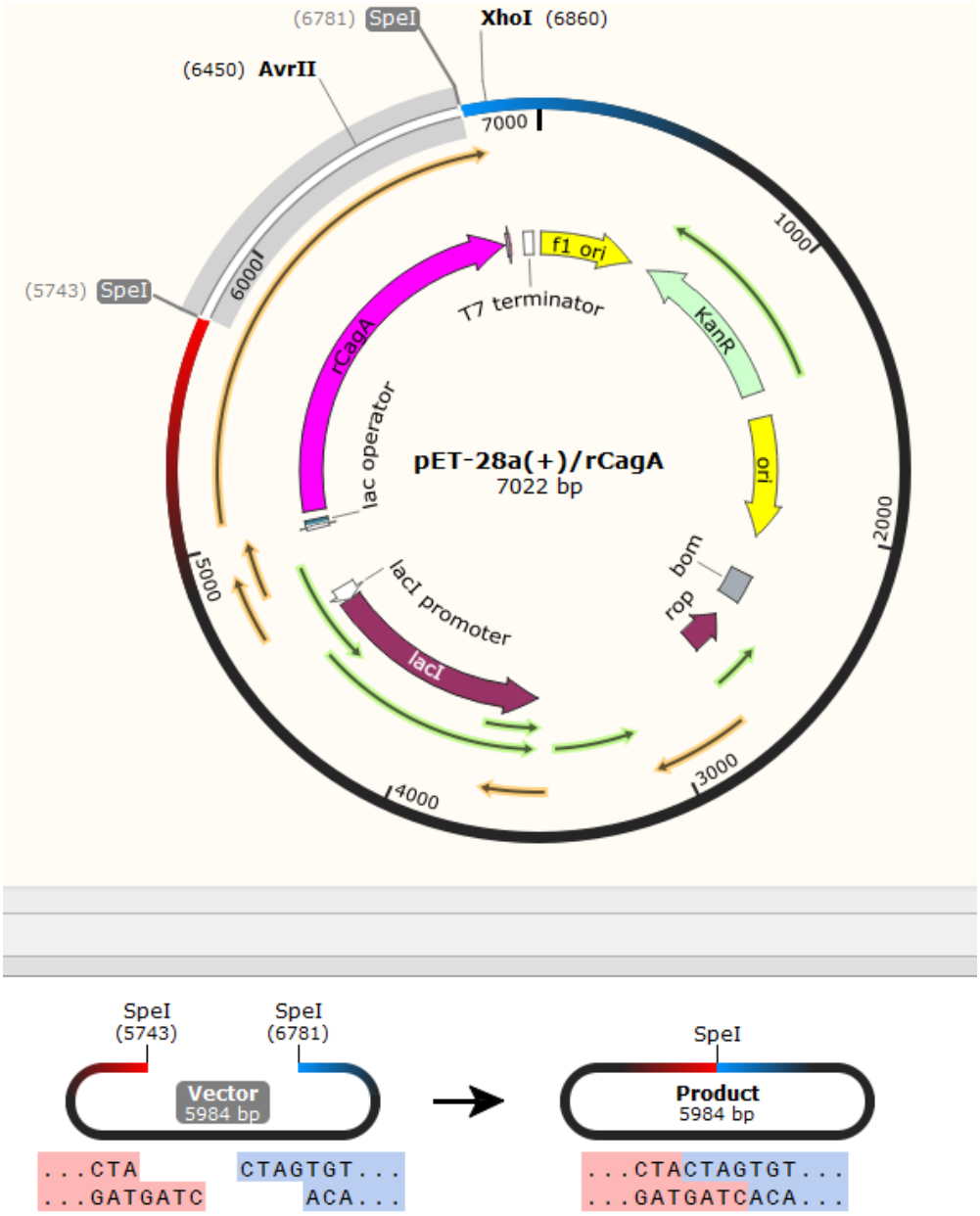


Figure 4.8 The cutting of the recombinant plasmid with Spe1 restriction enzyme at the sites of 5743 bp and 6781 bp.

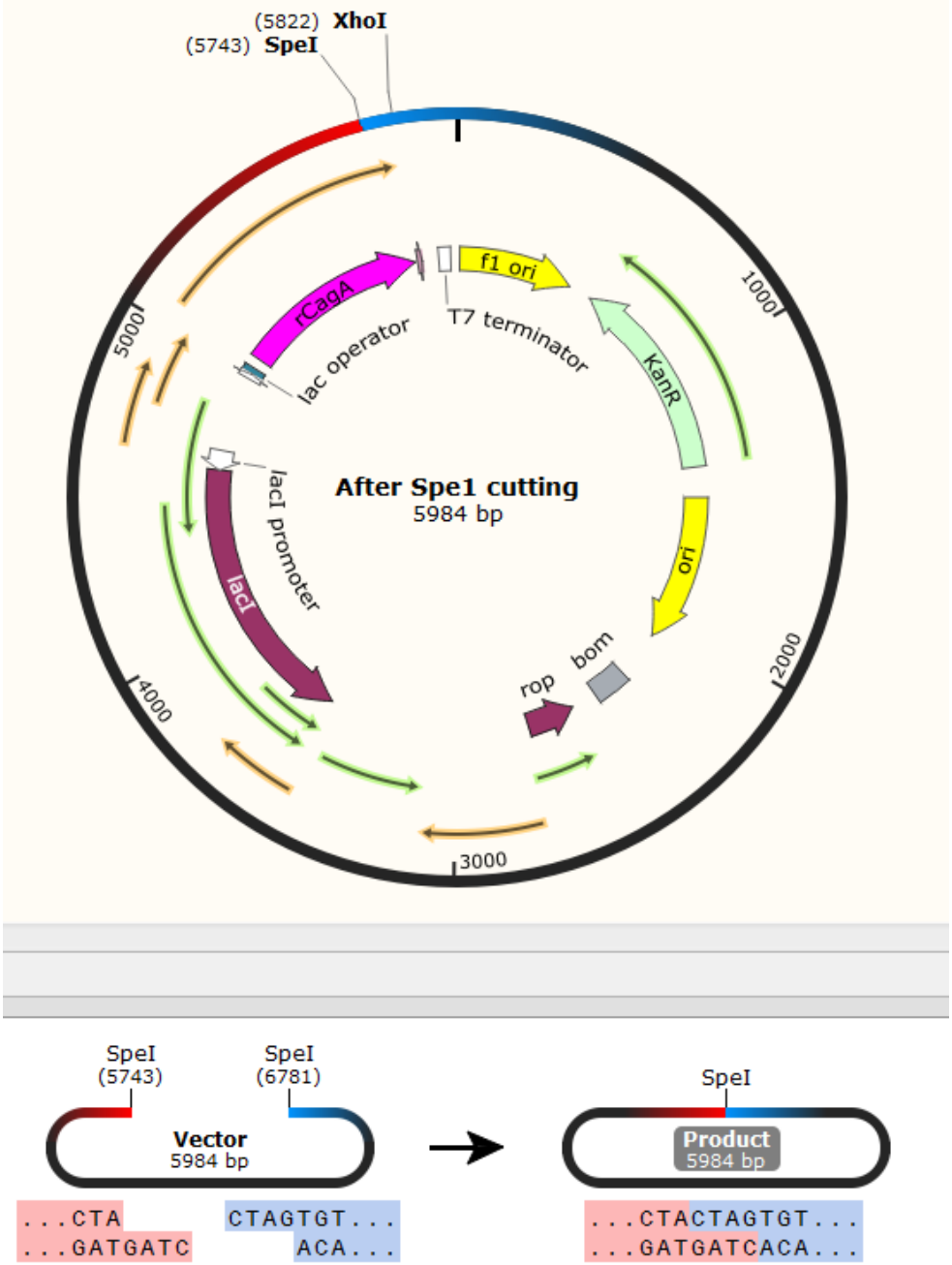


Figure 4.9: The self-ligation of this recombinant vector after cutting of the recombinant plasmid with Spe1 restriction enzyme. A newly formed recombinant plasmid as a product is 5984 bp.

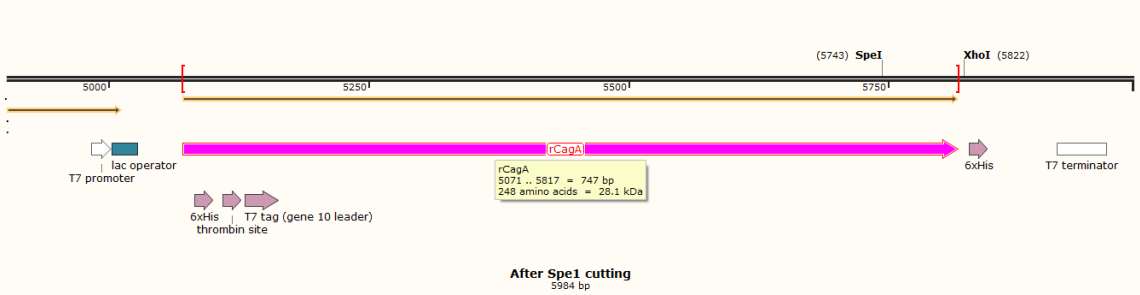


Figure 4.10 The third rcagA gene fragment after cutting with Spe1 restriction enzyme.

4.3 Purification of the recombinant vector

The " pET-28a(+)/ rcagA " recombinant plasmid was isolated and purified from *E. coli* BL21(DE3) pLySs using the GeneJET Plasmid Midiprep Kit (Thermo Scientific, USA) according to the manufacturer's instructions and it was loaded onto 1% agarose gel for confirmation as shown in Figure 4.11.

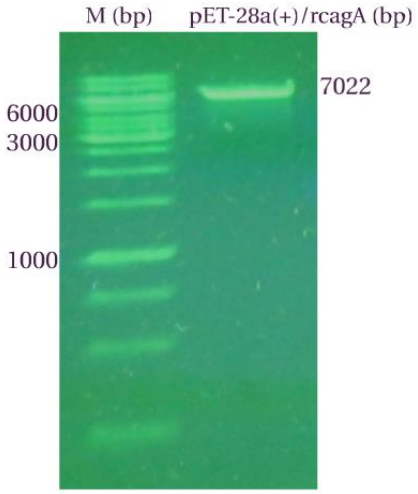


Figure 4.11 Gel electrophoresis pattern of the pET-28a(+)/rcagA recombinant plasmid M: molecular marker (1 kb ladder).

4.4 CLONING OF THE cagA 5' CONSERVED REGION GENE FRAGMENTS

4.4.1 Cutting with the restriction enzymes

A loading of agarose gel after cutting recombinant plasmid with restriction enzymes was virtually done using SnapGene software and 1 kb DNA Ladder was selected as shown in Figure 4.12. The recombinant plasmid was loaded on the 1% agarose gel as a 1.frg, the recombinant plasmid after cutting with AvrII and Xho1 restriction enzymes was loaded on the 1% agarose gel as a 2.frg and the recombinant plasmid after cutting with SpeI restriction enzyme was loaded on the 1% agarose gel as a 3.frg shown in Figure 4.13.

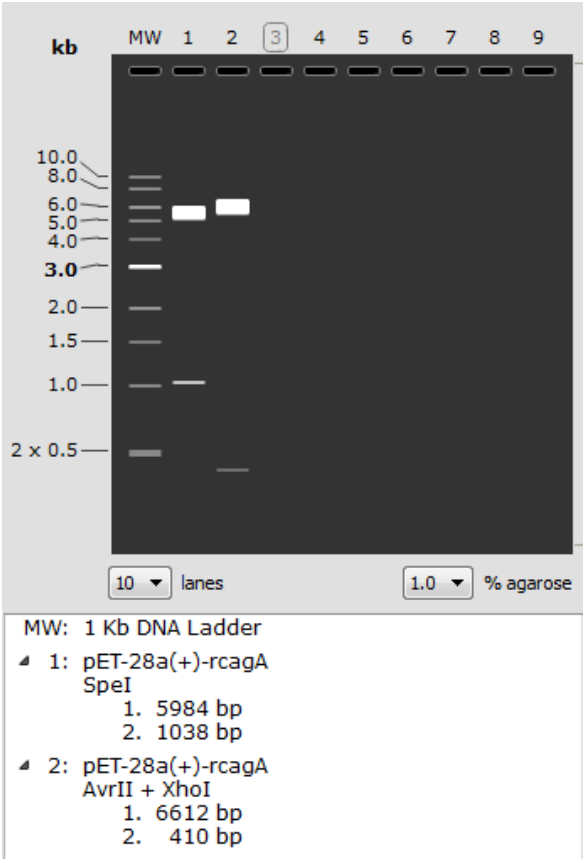


Figure 4.12 A virtual gel electrophoresis pattern of restriction cutting analysis using SnapGene software.

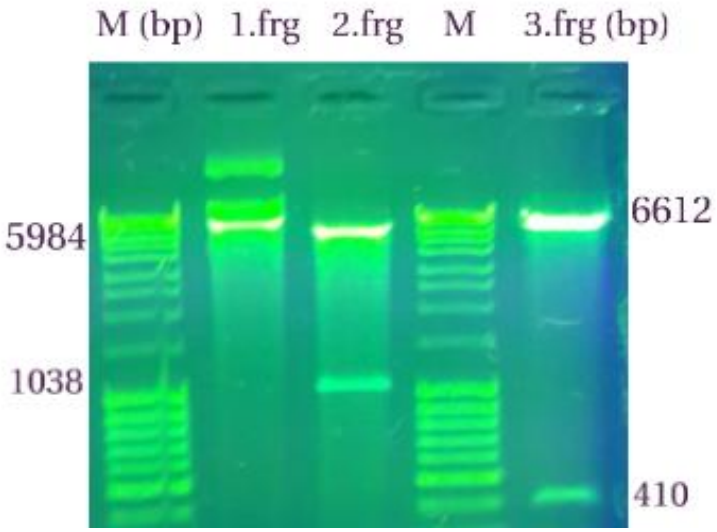


Figure 4.13 Gel electrophoresis pattern of restriction cutting analysis. M: molecular marker (1 kb ladder), 1.frg: pET-28a(+) plasmid containing rcagA gene, 2.frg: restriction cutting of pET-28a(+)/rcagA plasmid with AvrII and Xho1 restriction enzymes, 3.frg: restriction cutting of pET-28a(+)/rcagA plasmid with Spe1 restriction enzyme.

4.5 TRANSFORMATION RESULTS OF THE rcagA GENE FRAGMENTS

After cutting with restriction enzymes, the direct self ligation was done for each newly formed recombinant plasmids. A heat shock tranformation was applied and they were plated out on the LB agar containing kanamycin. Figure 4.14 shows that 6 *E.coli* BL21(DE3)pLySs colonies after transformation of the 2nd fragment and Figure 4.15 shows 23+11 *E.coli* BL21(DE3)pLySs colonies after transformation of the 3th fragment

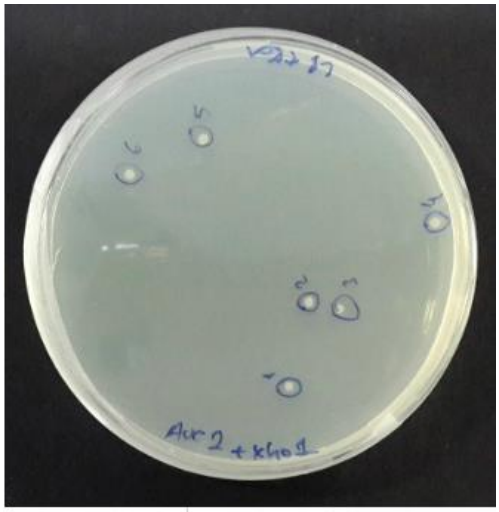


Figure 4.14 Transformation result of the second rcagA gene fragment. An agar plate showing 6 *E.coli* BL21(DE3)pLySs colonies after transformation of the 2nd fragment.

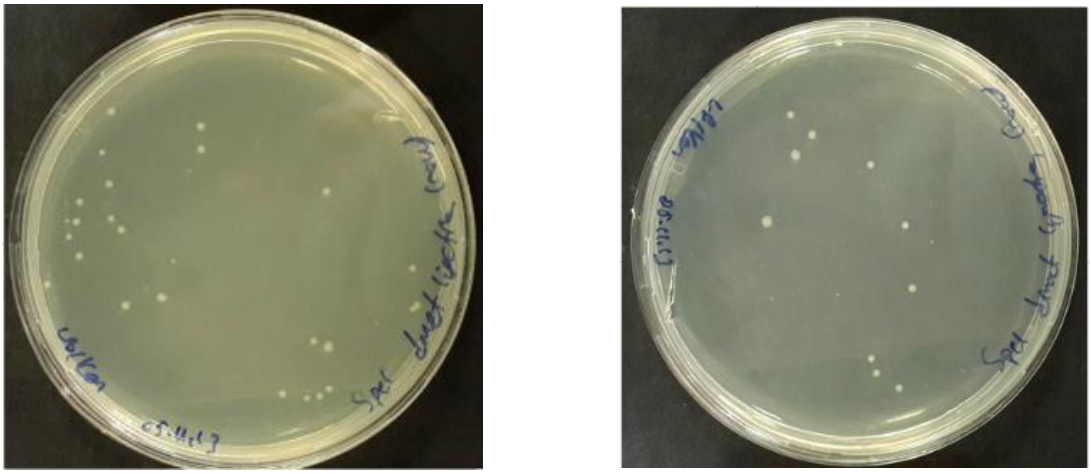


Figure 4.15 Transformation result of the third rcagA gene fragment. Two agar plates showing 23+11 *E.coli* BL21(DE3)pLySs colonies after transformation of the 3rd fragment.

4.6 RECOMBINANT PROTEIN INDUCTION

Protein expression was induced by 1 mmol/ L of isopropyl-b-d-thiogalactoside (IPTG) at culture optical density A600 = 0.60. After 3 hours of incubation cells were harvested and whole cell lysates were examined by 12% SDS–PAGE. Uninduced recombinant clones were used as negative controls.

4.7 SCREENING OF ALL COLONIES BY IPTG INDUC

All colonies that are grown on the LB agar plates containing kanamycin were observed by using both IPTG induction and SDS–PAGE whether they have correct expressed protein or not. Figure 4.16 showing a 12% SDS-PAGE gel loaded with the lysates of 3h IPTG induction for all 6 of bacterial colonies after transformation of the second rcagA gene. A non-induced (NI) bacterial colony was used as a negative control and the prestained protein ladder (10-250 kDa) was used as a marker. The bacterial colonies from the number 1 to 5 have incorrect protein expression whereas 6th bacterial colony has the correct protein induction as expected.

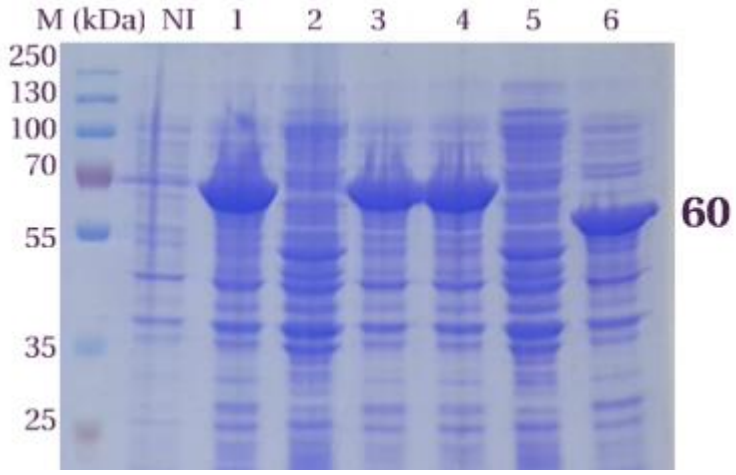


Figure 4.16 Screening of all bacterial colonies using IPTG induction and SDS-PAGE for the 2nd rCagA protein (M: prestained protein ladder; NI: Non-induced bacterial colony; 1-5: bacterial colonies that have incorrect protein induction as a negative result; 6: bacterial colony that has correct protein induction as expected.

Figure 4.17 shows the screening of the proteins from all 17 proteins bacterial colonies after transformation of the third rcagA gene. The 1st and 9th bacterial colonies have incorrect protein induction but the remaining bacterial colonies have correct protein induction as expected 28 kDa and a non-induced (NI) bacterial colony was used as a negative control and the prestained protein ladder (10-250 kDa) was used as a marker.

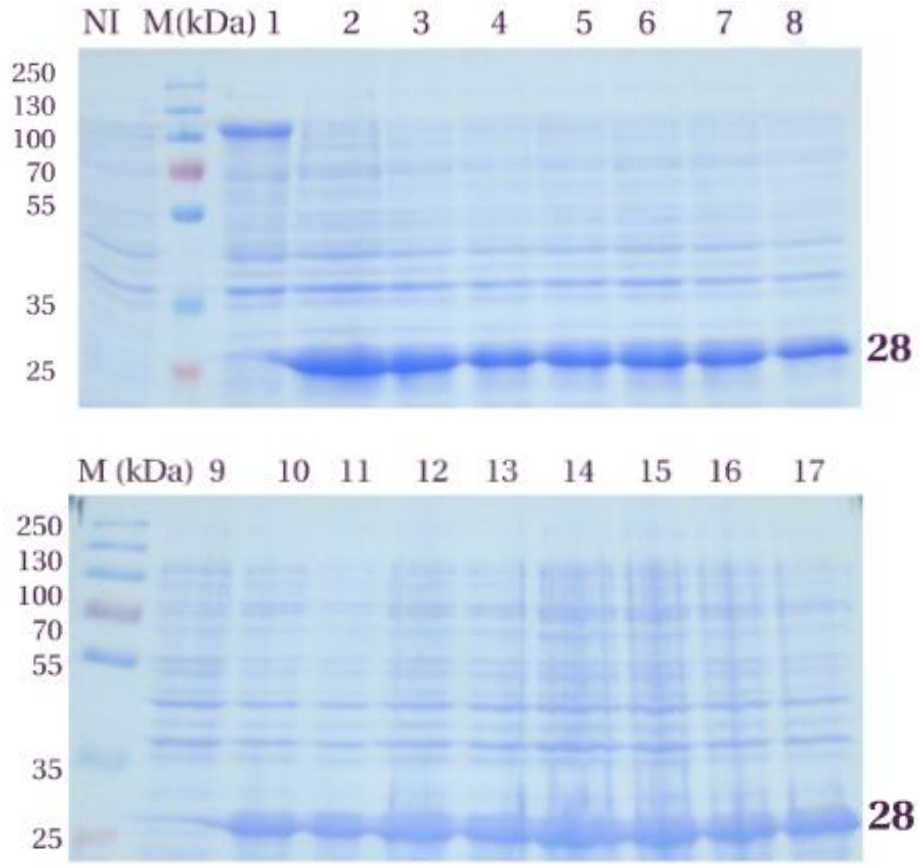


Figure 4.17 Screening of all bacterial colonies using IPTG induction and SDS-PAGE for the 3rd rCagA protein. (M: prestained protein ladder; NI: Non-induced bacterial colony; 1, 9: bacterial colonies that have incorrect protein induction as a negative result; 2-17: bacterial colonies that have correct protein induction as expected.

4.8 SDS-PAGE

One of the positive colonies was selected for each one and the lysates of 1h, 2h, 3h induction of selected positive colony were again observed onto 12% SDS–PAGE for each protein. A non-induced bacterial colony was also loaded on the gel for each of fragments. Figure 4.18 shows the three hour protein induction of selected positive colony has the 1st rCagA protein onto 12% SDS-PAGE gel. Figure 4.19 shows the three hour protein induction of selected positive colony that has the 2nd rCagA protein onto 12% SDS-PAGE gel and Figure 4.20 shows the three-hour protein induction of selected positive colony that has the 3rd rCagA protein onto 12% SDS-PAGE gel.

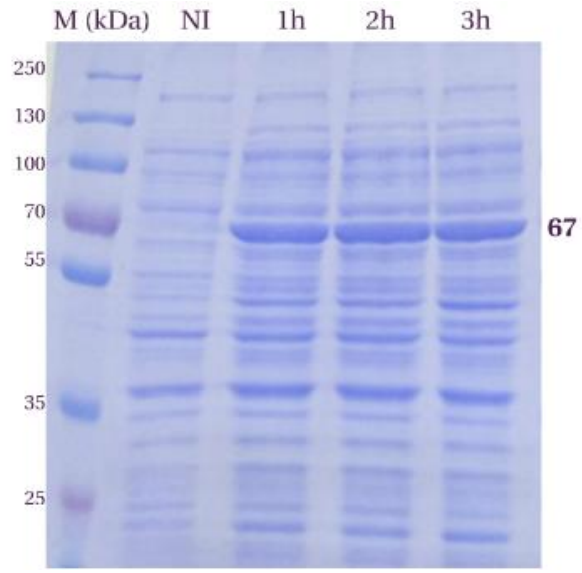


Figure 4.18 3-hour protein induction of selected positive colony for the 1st rCagA protein onto 12% SDS-PAGE gel. (M: prestained protein ladder, NI: Non-induced bacterial colony, 1h: 1 hour induction, 2h: 2 hours induction, 3h: 3 hours induction)

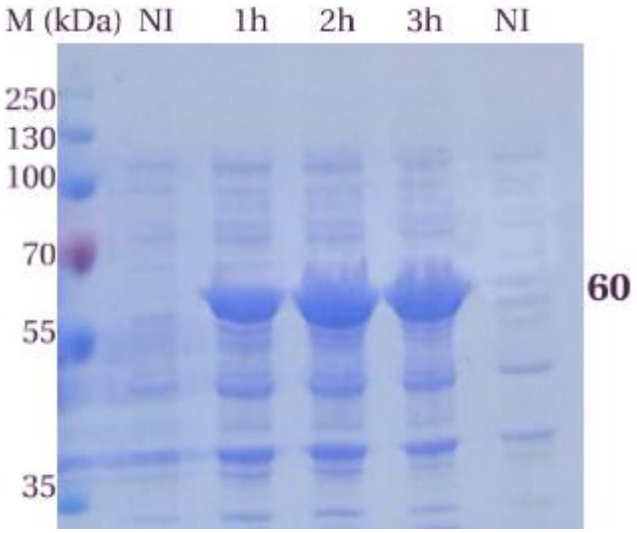


Figure 4.19 3-hour protein induction of selected positive colony for the 2nd rCagA protein onto 12% SDS-PAGE gel (M: prestained protein ladder, NI: Non-induced bacterial colony, 1h: 1 hour induction, 2h: 2 hours induction, 3h: 3 hours induction)

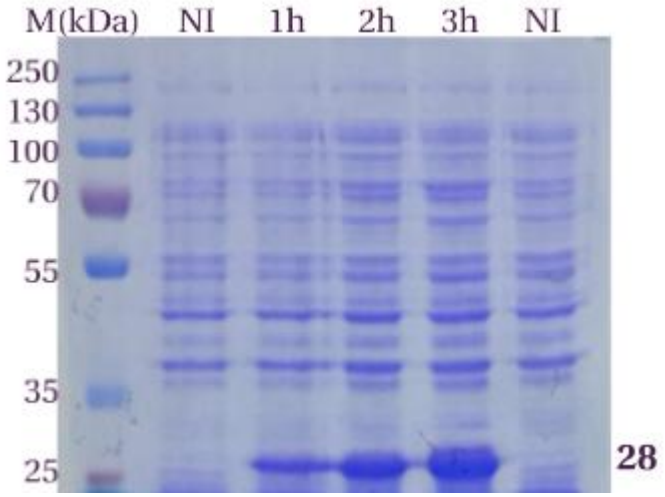


Figure 4.20 3-hour protein induction of selected positive colony for the 3rd rCagA protein onto 12% SDS-PAGE gel (M: prestained protein ladder, NI: Non-induced bacterial colony, 1h: 1 hour induction, 2h: 2 hours induction, 3h: 3 hours induction)

4.9 DNA SEQUENCE ANALYSIS OF THE rcagA FRAGMENTS

The fragments of rcagA gene were sequenced and the nucleotide sequence shown below in a red. A highlighted in yellow region indicates the 6X Histidine tag in a recombinant plasmid. The Figure 4.21 shows the DNA sequence which belongs to the first rcagA gene fragment, the Figure 4.22 shows the DNA sequence which belongs to the second rcagA gene fragment and the Figure 4.23 shows the DNA sequence which belongs to the third rcagA gene fragment.

atgggcagcagccatcatcatcatcatcacagcagcggcctggtgccgcgcggcagccatatggctag
catgactggtggacagcaaatgggtcgcggatccGAAGCGGCTTTTAACCCGCAGCAATTTATCAATA
ATCCTCAAGTGGCTTTTCTTAAGCTTGATAACGCTGTCGCTTCATTTGATCCTGATCAAAAACCAATC
GTTGATAAGAACGATAGGGATCACAGGCAAGCTTTTGAGGGAATCTCGCAATTAAGGGAAGAATACTC
CAATAAAGCGATCAAAAATCCTACCAAAAAGAATCAGTATTTTTCAGACTTTATCAATAGGAGCAACG
ATCTAATCAACAAAGACGCTCTCATTGATGTAGAATCTTCCACAAAGAGCTTTCGGAAATTTGGGGAT
CAGCGTTACCGAATTTTCGCAAGTTGGGTGTCCCATCAAAACGATCCGTCTAAAATCAACACCCGATC
GATCCGAAATTTTATGGAAAATATCATACAACCCCCTATCTCTGATGATAAAGAAAAAGCAGAGTTTT
TGAAATCTGCCAAACAATCTTTTGCAGGAATCATTATAGGGAATCAAATCCGAACGGATCAAAAGTTC
ATGGGCGTGTTTGATGAATCTTTGAAAGAGAGGCAAGAAGCAGAAAAAAATGGAGAGCCTACTAGTGG
GGATTGGTTGGATATTTTTTTATCATTTATATTTGACAAAAAACGATCTTCTGATGTCAAAGAAGCAA
TCAATCAAGAGCCAGTTCCCCATGTCCAACCAGATATAGCCACTACCACCACCGACATCCAAGGCTTA
CCGCCTGAATCTAGGGATTTGCTTGATGAAAGGGGTAATTTTTCTAAATTCACTCTTGGCGATATGGA
AATGTTAGATGTTGAGGGAGTCGCTGACATTGATCCCAATTACAAGTTCAATCAATTATTGATTCACA
ATAACGCTCTGTCTTCTGTGTTAATGGGGAGTCATAATGGCATAGAACCTGAAAAAGTTTCATTATTG
TATGCGGGCAATGGTGGTTTTGGAGCCAAGCACGATTGGAACGCCACCGTTGGTTATAAAGACCAACA
AGGTAACAATGTGGCTACAATAATTAATGTGCATATGAAAAACGGCAGTGGCTTAGTCATAGCAGGTG
GTGAGAAAGGGATTAACAACCCTAGTTTTTATCTCTACAAAGAAGACCAACTCACAGGCTCACAACGA
GCATTGAGTCAAGAAGAGATCCGAAACAAAATAGATTTCATGGAATTTCTTGCACAAAATAATGCTAA
ATTAGACAACTTGAGCGAGAAAGAGAAAGAAAAATTCCGAACTGAGATTAAGGATTTCCAAAAAGACT
CTAAGCCTTATTTAGACGCCCTAGGGAATGATCGTATTGCTCTTGTTTCTAAAAAAGACCCAAAACAT
GCAGCTTTAATTACTGAGTTTGGTAAGGGGGATTTGAGCTACACTCTCAAAGATTATGGGAAAAAAGC
AGATAAAGCTTTAGATAGGGAGAAAAATGTTACTCTTCAAGGTAGCCTAAAACATGATGGCGTGATGT
TTGTTGATTATTCTAATTTCAAATACACCAACGCCTCCAAGAATCCCAATAAGGGTGTAGGCGTTACG
AATGGCGTTTCCCATTTAGACGCAGGCTTTAGCGAGGTAGCTGTCTTTAATTTGCCTGATTTAAATAA
TCTCGCTATCACTAGTGTCGTAAGGCAGGATTTAGAAGATAAATTGTGGGCTAAAGGATTGTCCTCAC
AAGAAGCATCTTGCTGA

Figure 4.21 DNA sequence result of the 1st rcagA gene. Highlighted in yellow part indicates the 6X Histidine tag, written in red part indicates the 1st rcagA gene.

atgggcagcagccatcatcatcatcatcacagcagcggcctggtgccgcgcggcagccatatggctag
catgactggtggacagcaaatgggtcgcggatccGAAGCGGCTTTTAACCCGCAGCAATTTATCAATA
ATCCTCAAGTGGCTTTTCTTAAGCTTGATAACGCTGTCGCTTCATTTGATCCTGATCAAAAACCAATC
GTTGATAAGAACGATAGGGATCACAGGCAAGCTTTTGAGGGAATCTCGCAATTAAGGGAAGAATACTC
CAATAAAGCGATCAAAAATCCTACCAAAAAGAATCAGTATTTTTCAGACTTTATCAATAGGAGCAACG
ATCTAATCAACAAAGACGCTCTCATTGATGTAGAATCTTCCACAAAGAGCTTTCGGAAATTTGGGGAT
CAGCGTTACCGAATTTTCGCAAGTTGGGTGTCCCATCAAAACGATCCGTCTAAAATCAACACCCGATC
GATCCGAAATTTTATGGAAAATATCATACAACCCCCTATCTCTGATGATAAAGAAAAAGCAGAGTTTT
TGAAATCTGCCAAACAATCTTTTGCAGGAATCATTATAGGGAATCAAATCCGAACGGATCAAAAGTTC
ATGGGCGTGTTTGATGAATCTTTGAAAGAGAGGCAAGAAGCAGAAAAAAATGGAGAGCCTACTAGTGG
GGATTGGTTGGATATTTTTTTATCATTTATATTTGACAAAAAACGATCTTCTGATGTCAAAGAAGCAA
TCAATCAAGAGCCAGTTCCCCATGTCCAACCAGATATAGCCACTACCACCACCGACATCCAAGGCTTA
CCGCCTGAATCTAGGGATTTGCTTGATGAAAGGGGTAATTTTTCTAAATTCACTCTTGGCGATATGGA
AATGTTAGATGTTGAGGGAGTCGCTGACATTGATCCCAATTACAAGTTCAATCAATTATTGATTCACA
ATAACGCTCTGTCTTCTGTGTTAATGGGGAGTCATAATGGCATAGAACCTGAAAAAGTTTCATTATTG
TATGCGGGCAATGGTGGTTTTGGAGCCAAGCACGATTGGAACGCCACCGTTGGTTATAAAGACCAACA
AGGTAACAATGTGGCTACAATAATTAATGTGCATATGAAAAACGGCAGTGGCTTAGTCATAGCAGGTG
GTGAGAAAGGGATTAACAACCCTAGTTTTTATCTCTACAAAGAAGACCAACTCACAGGCTCACAACGA
GCATTGAGTCAAGAAGAGATCCGAAACAAAATAGATTTCATGGAATTTCTTGCACAAAATAATGCTAA
ATTAGACAACTTGAGCGAGAAAGAGAAAGAAAAATTCCGAACTGAGATTAAGGATTTCCAAAAAGACT
CTAAGCCTTATTTAGACGCCCTAGtcgagcaccaccaccaccaccactga

Figure 4.22 DNA sequence result of the 2nd rcagA gene fragment. Highlighted in yellow parts indicate the 6X Histidine tag, written in red part indicates the 2nd rcagA gene fragment in the plasmid.

ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAG
CATGACTGGTGGACAGCAAATGGGTCGCGGATCCGAAGCGGCTTTTAACCCGCAGCAATTTATCAATA
ATCCTCAAGTGGCTTTTCTTAAGCTTGATAACGCTGTCGCTTCATTTGATCCTGATCAAAAACCAATC
GTTGATAAGAACGATAGGGATCACAGGCAAGCTTTTGAGGGAATCTCGCAATTAAGGGAAGAATACTC
CAATAAAGCGATCAAAAATCCTACCAAAAAGAATCAGTATTTTTCAGACTTTATCAATAGGAGCAACG
ATCTAATCAACAAAGACGCTCTCATTGATGTAGAATCTTCCACAAAGAGCTTTCGGAAATTTGGGGAT
CAGCGTTACCGAATTTTCGCAAGTTGGGTGTCCCATCAAAACGATCCGTCTAAAATCAACACCCGATC
GATCCGAAATTTTATGGAAAATATCATACAACCCCCTATCTCTGATGATAAAGAAAAAGCAGAGTTTT
TGAAATCTGCCAAACAATCTTTTGCAGGAATCATTATAGGGAATCAAATCCGAACGGATCAAAAGTTC
ATGGGCGTGTTTGATGAATCTTTGAAAGAGAGGCAAGAAGCAGAAAAAAATGGAGAGCCTACTAGTTT
CGTAAGGCGGAATTTAGAGGACAAACTAACCGCTAAAGGATTGTCCTCACAAGAAGCATCTTGCTGA

Figure 4.23 DNA sequence result of the 3rd rcagA gene fragment. Highlighted in yellow part indicates the 6X Histidine tag, written in red part indicates the 3th rcagA gene fragment in the plasmid.

4.10 MULTIPLE SEQUENCE ALIGNMENTS

All rcagA gene fargments were aligned between each other by using ClustalW software shown below in the Figure 4.25. Multiple sequence alignments showed a homology between the first rcagA gene fragment and the second rcagA gene fragment of %99.72, the first rcagA gene fragment and third rcagA gene fragment of %99.46, the second rcagA gene fragment and the third rcagA gene fragment of %98.93 as shown in the multiple sequence alignments below.

SeqA	Name	Length	SeqB	Name	Length	Score
1	1.Fragment	1785	2	2.Fragment	1410	99.72
1	1.Fragment	1785	3	3.Fragment	747	99.46
2	2.Fragment	1410	3	3.Fragment	747	98.93

Figure 4.24 Scores of multiple alignment results between the first rcagA gene fragment (1.Fragment), the second rcagA gene fragment (2.Fragment) and the third rcagA gene fragment (3.Fragment).

```
CLUSTAL 2.1 multiple sequence alignment


1.Fragment      ATGGGCAGCAGCCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCAT 60
2.Fragment      ATGGGCAGCAGCCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCAT 60
3.Fragment      ATGGGCAGCAGCCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCAT 60
                ************************************************************

1.Fragment      ATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGCGGATCCGAAGCGGCTTTTAACCCG 120
2.Fragment      ATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGCGGATCCGAAGCGGCTTTTAACCCG 120
3.Fragment      ATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGCGGATCCGAAGCGGCTTTTAACCCG 120
                ************************************************************

1.Fragment      CAGCAATTTATCAATAATCCTCAAGTGGCTTTTCTTAAGCTTGATAACGCTGTCGCTTCA 180
2.Fragment      CAGCAATTTATCAATAATCCTCAAGTGGCTTTTCTTAAGCTTGATAACGCTGTCGCTTCA 180
3.Fragment      CAGCAATTTATCAATAATCCTCAAGTGGCTTTTCTTAAGCTTGATAACGCTGTCGCTTCA 180
                ************************************************************

1.Fragment      TTTGATCCTGATCAAAAACCAATCGTTGATAAGAACGATAGGGATCACAGGCAAGCTTTT 240
2.Fragment      TTTGATCCTGATCAAAAACCAATCGTTGATAAGAACGATAGGGATCACAGGCAAGCTTTT 240
3.Fragment      TTTGATCCTGATCAAAAACCAATCGTTGATAAGAACGATAGGGATCACAGGCAAGCTTTT 240
                ************************************************************

1.Fragment      GAGGGAATCTCGCAATTAAGGGAAGAATACTCCAATAAAGCGATCAAAAATCCTACCAAA 300
2.Fragment      GAGGGAATCTCGCAATTAAGGGAAGAATACTCCAATAAAGCGATCAAAAATCCTACCAAA 300
3.Fragment      GAGGGAATCTCGCAATTAAGGGAAGAATACTCCAATAAAGCGATCAAAAATCCTACCAAA 300
                ************************************************************

1.Fragment      AAGAATCAGTATTTTTCAGACTTTATCAATAGGAGCAACGATCTAATCAACAAAGACGCT 360
2.Fragment      AAGAATCAGTATTTTTCAGACTTTATCAATAGGAGCAACGATCTAATCAACAAAGACGCT 360
3.Fragment      AAGAATCAGTATTTTTCAGACTTTATCAATAGGAGCAACGATCTAATCAACAAAGACGCT 360
                ************************************************************

1.Fragment      CTCATTGATGTAGAATCTTCCACAAAGAGCTTTCGGAAATTTGGGGATCAGCGTTACCGA 420
2.Fragment      CTCATTGATGTAGAATCTTCCACAAAGAGCTTTCGGAAATTTGGGGATCAGCGTTACCGA 420
3.Fragment      CTCATTGATGTAGAATCTTCCACAAAGAGCTTTCGGAAATTTGGGGATCAGCGTTACCGA 420
                ************************************************************

1.Fragment      ATTTTCGCAAGTTGGGTGTCCCATCAAAACGATCCGTCTAAAATCAACACCCGATCGATC 480
2.Fragment      ATTTTCGCAAGTTGGGTGTCCCATCAAAACGATCCGTCTAAAATCAACACCCGATCGATC 480
3.Fragment      ATTTTCGCAAGTTGGGTGTCCCATCAAAACGATCCGTCTAAAATCAACACCCGATCGATC 480
                ************************************************************

1.Fragment      CGAAATTTTATGGAAAATATCATACAACCCCCTATCTCTGATGATAAAGAAAAAGCAGAG 540
2.Fragment      CGAAATTTTATGGAAAATATCATACAACCCCCTATCTCTGATGATAAAGAAAAAGCAGAG 540
3.Fragment      CGAAATTTTATGGAAAATATCATACAACCCCCTATCTCTGATGATAAAGAAAAAGCAGAG 540
                ************************************************************

1.Fragment      TTTTTGAAATCTGCCAAACAATCTTTTGCAGGAATCATTATAGGGAATCAAATCCGAACG 600
2.Fragment      TTTTTGAAATCTGCCAAACAATCTTTTGCAGGAATCATTATAGGGAATCAAATCCGAACG 600
3.Fragment      TTTTTGAAATCTGCCAAACAATCTTTTGCAGGAATCATTATAGGGAATCAAATCCGAACG 600
                ************************************************************
```

Figure 4.25 Multiple sequence alignment was done with ClustalW2 software between the first rcagA gene fragment (1.Fragment), the second rcagA gene fragment (2.Fragment) and the third rcagA gene fragment (3.Fragment)

```
1.Fragment      GATCAAAAGTTCATGGGCGTGTTTGATGAATCTTTGAAAGAGAGGCAAGAAGCAGAAAAA 660
2.Fragment      GATCAAAAGTTCATGGGCGTGTTTGATGAATCTTTGAAAGAGAGGCAAGAAGCAGAAAAA 660
3.Fragment      GATCAAAAGTTCATGGGCGTGTTTGATGAATCTTTGAAAGAGAGGCAAGAAGCAGAAAAA 660
                ************************************************************

1.Fragment      AATGGAGAGCCTACTAGTGGGGATTGGTTGGATATTTTTTTATCATTTATATTTGACAAA 720
2.Fragment      AATGGAGAGCCTACTAGTGGGGATTGGTTGGATATTTTTTTATCATTTATATTTGACAAA 720
3.Fragment      AATGGAGAGCCTACTAGT-----TTCGT------------------------------A 684
                ******************     ** **                             *

1.Fragment      AAACGATCTTCTGATGTCAAAGAAGCAATCAATCAAGAGCCAGTTCCCCATGTCCAACCA 780
2.Fragment      AAACGATCTTCTGATGTCAAAGAAGCAATCAATCAAGAGCCAGTTCCCCATGTCCAACCA 780
3.Fragment      AGGCG------------------------------------------------------- 689
                *  **

1.Fragment      GATATAGCCACTACCACCACCGACATCCAAGGCTTACCGCCTGAATCTAGGGATTTGCTT 840
2.Fragment      GATATAGCCACTACCACCACCGACATCCAAGGCTTACCGCCTGAATCTAGGGATTTGCTT 840
3.Fragment      ----------------------------------------GAATTTAGAG-------- 699
                                                       **** *** *

1.Fragment      GATGAAAGGGGTAATTTTTCTAAATTCACTCTTGGCGATATGGAAATGTTAGATGTTGAG 900
2.Fragment      GATGAAAGGGGTAATTTTTCTAAATTCACTCTTGGCGATATGGAAATGTTAGATGTTGAG 900
3.Fragment      ------------------------------------------------------------


1.Fragment      GGAGTCGCTGACATTGATCCCAATTACAAGTTCAATCAATTATTGATTCACAATAACGCT 960
2.Fragment      GGAGTCGCTGACATTGATCCCAATTACAAGTTCAATCAATTATTGATTCACAATAACGCT 960
3.Fragment      ---------GACA--------AACTA---------------------------------- 708
                         ****        ** **

1.Fragment      CTGTCTTCTGTGTTAATGGGGAGTCATAATGGCATAGAACCTGAAAAAGTTTCATTATTG 1020
2.Fragment      CTGTCTTCTGTGTTAATGGGGAGTCATAATGGCATAGAACCTGAAAAAGTTTCATTATTG 1020
3.Fragment      ------------------------------------------------------------


1.Fragment      TATGCGGGCAATGGTGGTTTTGGAGCCAAGCACGATTGGAACGCCACCGTTGGTTATAAA 1080
2.Fragment      TATGCGGGCAATGGTGGTTTTGGAGCCAAGCACGATTGGAACGCCACCGTTGGTTATAAA 1080
3.Fragment      ----------------------------------------------ACCGCT--------- 714
                                                             **** *

1.Fragment      GACCAACAAGGTAACAATGTGGCTACAATAATTAATGTGCATATGAAAAACGGCAGTGGC 1140
2.Fragment      GACCAACAAGGTAACAATGTGGCTACAATAATTAATGTGCATATGAAAAACGGCAGTGGC 1140
3.Fragment      ------------------------------------------------------------


1.Fragment      TTAGTCATAGCAGGTGGTGAGAAAGGGATTAACAACCCTAGTTTTTATCTCTACAAAGAA 1200
2.Fragment      TTAGTCATAGCAGGTGGTGAGAAAGGGATTAACAACCCTAGTTTTTATCTCTACAAAGAA 1200
3.Fragment      ----------------------AAAGGATTGTC--------------------------- 725
                                     ** *****  *
```

Figure 4.25. (cont.)

```
1.Fragment      GACCAACTCACAGGCTCACAACGAGCATTGAGTCAAGAAGAGATCCGAAACAAAATAGAT 1260
2.Fragment      GACCAACTCACAGGCTCACAACGAGCATTGAGTCAAGAAGAGATCCGAAACAAAATAGAT 1260
3.Fragment      --------------CTCACAAGAAGCAT------------------------------ 739
                              *******  *****

1.Fragment      TTCATGGAATTTCTTGCACAAAATAATGCTAAATTAGACAACTTGAGCGAGAAAGAGAAA 1320
2.Fragment      TTCATGGAATTTCTTGCACAAAATAATGCTAAATTAGACAACTTGAGCGAGAAAGAGAAA 1320
3.Fragment      ------------CTTGC-------------------------TGA-------------- 747
                            *****                         ***

1.Fragment      GAAAAATTCCGAACTGAGATTAAGGATTTCCAAAAAGACTCTAAGCCTTATTTAGACGCC 1380
2.Fragment      GAAAAATTCCGAACTGAGATTAAGGATTTCCAAAAAGACTCTAAGCCTTATTTAGACGCC 1380
3.Fragment      ------------------------------------------------------------


1.Fragment      CTAGGGAATGATCGTATTGCTCTTGTTTCTAAAAAAGACCCAAAACATGCAGCTTTAATT 1440
2.Fragment      CTAG-------------------------------------------------------- 1384
3.Fragment      ------------------------------------------------------------


1.Fragment      ACTGAGTTTGGTAAGGGGGATTTGAGCTACACTCTCAAAGATTATGGGAAAAAAGCAGAT 1500
2.Fragment      --------------------TCGAGC--------------------------------- 1390
3.Fragment      ------------------------------------------------------------


1.Fragment      AAAGCTTTAGATAGGGAGAAAAATGTTACTCTTCAAGGTAGCCTAAAACATGATGGCGTG 1560
2.Fragment      ------------------------------------------------------------
3.Fragment      ------------------------------------------------------------


1.Fragment      ATGTTTGTTGATTATTCTAATTTCAAATACACCAACGCCTCCAAGAATCCCAATAAGGGT 1620
2.Fragment      ----------------------------ACCACCACCACCA------CCACTGA---- 1410
3.Fragment      ------------------------------------------------------------


1.Fragment      GTAGGCGTTACGAATGGCGTTTCCCATTTAGACGCAGGCTTTAGCGAGGTAGCTGTCTTT 1680
2.Fragment      ------------------------------------------------------------
3.Fragment      ------------------------------------------------------------


1.Fragment      AATTTGCCTGATTTAAATAATCTCGCTATCACTAGTGTCGTAAGGCAGGATTTAGAAGAT 1740
2.Fragment      ------------------------------------------------------------
3.Fragment      ------------------------------------------------------------


1.Fragment      AAATTGTGGGCTAAAGGATTGTCCTCACAAGAAGCATCTTGCTGA 1785
2.Fragment      ---------------------------------------------
3.Fragment      ---------------------------------------------
```

Figure 4. 25 (Cont.)

```
1.Fragment      GATCAAAAGTTCATGGGCGTGTTTGATGAATCTTTGAAAGAGAGGCAAGAAGCAGAAAAA 660
2.Fragment      GATCAAAAGTTCATGGGCGTGTTTGATGAATCTTTGAAAGAGAGGCAAGAAGCAGAAAAA 660
3.Fragment      GATCAAAAGTTCATGGGCGTGTTTGATGAATCTTTGAAAGAGAGGCAAGAAGCAGAAAAA 660
                ************************************************************

1.Fragment      AATGGAGAGCCTACTAGTGGGGATTGGTTGGATATTTTTTTATCATTTATATTTGACAAA 720
2.Fragment      AATGGAGAGCCTACTAGTGGGGATTGGTTGGATATTTTTTTATCATTTATATTTGACAAA 720
3.Fragment      AATGGAGAGCCTACTAGT-----TTCGT------------------------------A 684
                ******************     ** **                             *

1.Fragment      AAACGATCTTCTGATGTCAAAGAAGCAATCAATCAAGAGCCAGTTCCCCATGTCCAACCA 780
2.Fragment      AAACGATCTTCTGATGTCAAAGAAGCAATCAATCAAGAGCCAGTTCCCCATGTCCAACCA 780
3.Fragment      AGGCG------------------------------------------------------- 689
                *  **

1.Fragment      GATATAGCCACTACCACCACCGACATCCAAGGCTTACCGCCTGAATCTAGGGATTTGCTT 840
2.Fragment      GATATAGCCACTACCACCACCGACATCCAAGGCTTACCGCCTGAATCTAGGGATTTGCTT 840
3.Fragment      ----------------------------------------GAATTTAGAG-------- 699
                                                       **** *** *

1.Fragment      GATGAAAGGGGTAATTTTTCTAAATTCACTCTTGGCGATATGGAAATGTTAGATGTTGAG 900
2.Fragment      GATGAAAGGGGTAATTTTTCTAAATTCACTCTTGGCGATATGGAAATGTTAGATGTTGAG 900
3.Fragment      ------------------------------------------------------------

1.Fragment      GGAGTCGCTGACATTGATCCCAATTACAAGTTCAATCAATTATTGATTCACAATAACGCT 960
2.Fragment      GGAGTCGCTGACATTGATCCCAATTACAAGTTCAATCAATTATTGATTCACAATAACGCT 960
3.Fragment      ---------GACA--------AACTA---------------------------------- 708
                         ****        ** **

1.Fragment      CTGTCTTCTGTGTTAATGGGGAGTCATAATGGCATAGAACCTGAAAAAGTTTCATTATTG 1020
2.Fragment      CTGTCTTCTGTGTTAATGGGGAGTCATAATGGCATAGAACCTGAAAAAGTTTCATTATTG 1020
3.Fragment      ------------------------------------------------------------

1.Fragment      TATGCGGGCAATGGTGGTTTTGGAGCCAAGCACGATTGGAACGCCACCGTTGGTTATAAA 1080
2.Fragment      TATGCGGGCAATGGTGGTTTTGGAGCCAAGCACGATTGGAACGCCACCGTTGGTTATAAA 1080
3.Fragment      ----------------------------------------------ACCGCT--------- 714
                                                             **** *

1.Fragment      GACCAACAAGGTAACAATGTGGCTACAATAATTAATGTGCATATGAAAAACGGCAGTGGC 1140
2.Fragment      GACCAACAAGGTAACAATGTGGCTACAATAATTAATGTGCATATGAAAAACGGCAGTGGC 1140
3.Fragment      ------------------------------------------------------------

1.Fragment      TTAGTCATAGCAGGTGGTGAGAAAGGGATTAACAACCCTAGTTTTTATCTCTACAAAGAA 1200
2.Fragment      TTAGTCATAGCAGGTGGTGAGAAAGGGATTAACAACCCTAGTTTTTATCTCTACAAAGAA 1200
3.Fragment      ----------------------AAAGGATTGTC--------------------------- 725
                                     ** *****  *
```

Figure 4.25 (cont.)

4.11 AMINO ACID SEQUENCE ANALYSIS OF THE rCagA PROTEIN FRAGMENTS

The amino acid sequence of all three recombinant CagA proteins were deduced from the nucleotide sequence by using ExPasy translate tool as shown below.

MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSEAAFNPQQFINNPQVAFLKLDNAVAS
FDPDQKPIVDKNDRDHRQAFEGISQLREEYSNKAIKNPTKKNQYFSDFINRSNDLINKDA
LIDVESSTKSFRKFGDQRYRIFASWVSHQNDPSKINTRSIRNFMENIIQPPISDDKEKAE
FLKSAKQSFAGIIIGNQIRTDQKFMGVFDESLKERQEAEKNGEPTSGDWLDIFLSFIFDK
KRSSDVKEAINQEPVPHVQPDIATTTTDIQGLPPESRDLLDERGNFSKFTLGDMEMLDVE
GVADIDPNYKFNQLLIHNNALSSVLMGSHNGIEPEKVSLLYAGNGGFGAKHDWNATVGYK
DQQGNNVATIINVHMKNGSGLVIAGGEKGINNPSFYLYKEDQLTGSQRALSQEEIRNKID
FMEFLAQNNAKLDNLSEKEKEKFRTEIKDFQKDSKPYLDALGNDRIALVSKKDPKHAALI
TEFGKGDLSYTLKDYGKKADKALDREKNVTLQGSLKHDGVMFVDYSNFKYTNASKNPNKG
VGVTNGVSHLDAGFSEVAVFNLPDLNNLAITSVVRQDLEDKLWAKGLSSQEASC

Figure 4.26 Amino acid sequence analysis of the 1st rCagA protein (67kDa)

MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSEAAFNPQQFINNPQVAFLKLDNAVAS
FDPDQKPIVDKNDRDHRQAFEGISQLREEYSNKAIKNPTKKNQYFSDFINRSNDLINKDA
LIDVESSTKSFRKFGDQRYRIFASWVSHQNDPSKINTRSIRNFMENIIQPPISDDKEKAE
FLKSAKQSFAGIIIGNQIRTDQKFMGVFDESLKERQEAEKNGEPTSGDWLDIFLSFIFDK
KRSSDVKEAINQEPVPHVQPDIATTTTDIQGLPPESRDLLDERGNFSKFTLGDMEMLDVE
GVADIDPNYKFNQLLIHNNALSSVLMGSHNGIEPEKVSLLYAGNGGFGAKHDWNATVGYK
DQQGNNVATIINVHMKNGSGLVIAGGEKGINNPSFYLYKEDQLTGSQRALSQEEIRNKID
FMEFLAQNNAKLDNLSEKEKEKFRTEIKDFQKDSKPYLDALVEHHHHHH

Figure 4.27 Amino acid sequence analysis of the 2nd rCagA protein (60kDa)

MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRGSEAAFNPQQFINNPQVAFLKLDNAVAS
FDPDQKPIVDKNDRDHRQAFEGISQLREEYSNKAIKNPTKKNQYFSDFINRSNDLINKDA
LIDVESSTKSFRKFGDQRYRIFASWVSHQNDPSKINTRSIRNFMENIIQPPISDDKEKAE
FLKSAKQSFAGIIIGNQIRTDQKFMGVFDESLKERQEAEKNGEPTSFVRRNLEDKLTAKG
LSSQEASC

Figure 4.28 Amino acid sequence analysis of the 3rd rCagA protein (28kDa)

4.12 MULTIPLE PROTEIN SEQUENCE ALIGNMET

The three rCagA protein fargments were aligned between each other by using ClustalW2 software. Multiple protein sequence alignments showed a homology between the 1st rCagA protein and 2nd rCagA protein showed 98.29%, the 1st rCagA protein and 3rd rCagA protein showed 98.39%, the 2nd rCagA protein and 3rd rCagA protein showed 90.73% in between as shown below.

SeqA	Name	Length	SeqB	Name	Length	Score
1	1.rCagA	594	2	2.rCagA	468	98.29
1	1.rCagA	594	3	3.rCagA	248	98.39
2	2.rCagA	468	3	3.rCagA	248	90.73

Figure 4.29 Scores of multiple protein alignment results between the first rCagA protein (1st rcagA), the second rCagA protein (2nd rCagA) and the third rCagA protein (3rd rCagA) done with ClustalW2 software.

Figure 4.30 Multiple protein alignment was done with ClustalW2 software between the first rCagA protein (1st rcagA), the second rCagA protein (2nd rCagA) and the third rCagA protein (3rd rCagA) done with ClustalW2 software

4.13 EPITOPE DETERMINATION

Epitope analysis of each of these recombinant proteins was predicted by using Emini Surface Accessibility Prediction program "http://tools.immuneepitope.org/tools/" and these predicted epitopes were located on the each of these rCagA proteins using SnapGene software program as shown below.

Emini Surface Accessibility Prediction

Input Sequences

```
  1 MGSSHHHHHH SSGLVPRGSH MASMTGGQQM GRGSEAAFNP QQFINNPQVA FLKLDNAVAS
 61 FDPDQKPIVD KNDRDHRQAF EGISQLREEY SNKAIKNPTK KNQYFSDFIN RSNDLINKDA
121 LIDVESSTKS FRKFGDQRYR IFASWVSHQN DPSKINTRSI RNFMENIIQP PISDDKEKAE
181 FLKSAKQSFA GIIIGNQIRT DQKFMGVFDE SLKERQEAEK NGEPTSGDWL DIFLSFIFDK
241 KRSSDVKEAI NQEPVPHVQP DIATTTTDIQ GLPPESRDLL DERGNFSKFT LGDMEMLDVE
301 GVADIDPNYK FNQLLIHNNA LSSVLMGSHN GIEPEKVSLL YAGNGGFGAK HDWNATVGYK
361 DQQGNNVATI INVHMKNGSG LVIAGGEKGI NNPSFYLYKE DQLTGSQRAL SQEEIRNKID
421 FMEFLAQNNA KLDNLSEKEK EKFRTEIKDF QKDSKPYLDA LGNDRIALVS KKDPKHAALI
481 TEFGKGDLSY TLKDYGKKAD KALDREKNVT LQGSLKHDGV MFVDYSNFKY TNASKNPNKG
541 VGVTNGVSHL DAGFSEVAVF NLPDLNNLAI TSVVRQDLED KLWAKGLSSQ EASC
```

Predicted peptides:

No.	Start Position	End Position	Peptide	Peptide Length
1	71	78	KNDRDHRQ	8
2	86	91	LREEYS	6
3	97	104	NPTKKNQY	8
4	134	139	FGDQRY	6
5	148	156	HQNDPSKIN	9
6	171	178	PISDDKEK	8
7	211	223	SLKERQEAEKNGE	13
8	239	245	DKKRSSD	7
9	306	311	DPNYKF	6
10	435	443	LSEKEKEKF	9
11	448	457	KDFQKDSKPY	10
12	491	499	TLKDYGKKA	9
13	502	507	ALDREK	6
14	526	537	SNFKYTNASKNP	12

Figure 4.31 The 14 predicted epitopes of the 1st rCagA protein given by a program.

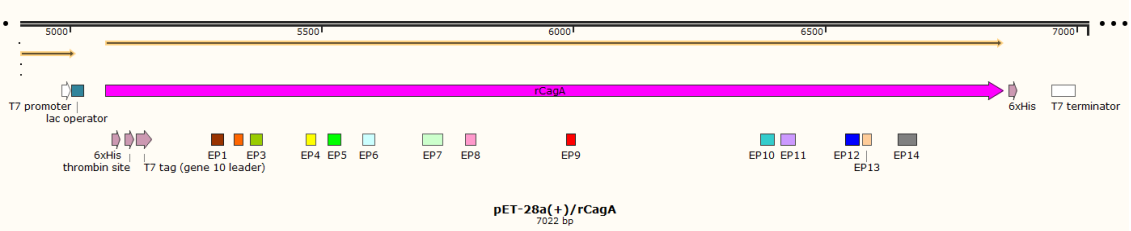


Figure 4.32 The location of these predicted epitopes on the 1st rCagA protein.

Emini Surface Accessibility Prediction

Sequence:

```
  1 MGSSHHHHHH SSGLVPRGSH MASMTGGQQM GRGSEAAFNP QQFINNPQVA FLKLDNAVAS
 61 FDPDQKPIVD KNDRDHRQAF EGISQLREEY SNKAIKNPTK KNQYFSDFIN RSNDLINKDA
121 LIDVESSTKS FRKFGDQRYR IFASWVSHQN DPSKINTRSI RNFMENIIQP PISDDKEKAE
181 FLKSAKQSFA GIIIGNQIRT DQKFMGVFDE SLKERQEAEK NGEPTSGDWL DIFLSFIFDK
241 KRSSDVKEAI NQEPVPHVQP DIATTTTDIQ GLPPESRDLL DERGNFSKFT LGDMEMLDVE
301 GVADIDPNYK FNQLLIHNNA LSSVLMGSHN GIEPEKVSLL YAGNGGFGAK HDWNATVGYK
361 DQQGNNVATI INVHMKNGSG LVIAGGEKGI NNPSFYLYKE DQLTGSQRAL SQEEIRNKID
421 FMEFLAQNNA KLDNLSEKEK EKFRTEIKDF QKDSKPYLDA LVEHHHHHH
```

Predicted peptides:

No.	Start Position	End Position	Peptide	Peptide Length
1	71	78	KNDRDHRQ	8
2	86	91	LREEYS	6
3	97	103	NPTKKQY	7
4	133	138	FGDQRY	6
5	147	155	HQNDPSKIN	9
6	170	177	PISDDKEK	8
7	210	222	SLKERQEAEKNGE	13
8	238	244	DKKRSSD	7
9	305	310	DPNYKF	6
10	434	442	LSEKEKEKF	9
11	448	456	DFQKDSKPY	9

Figure 4.33 The 11 predicted epitopes of the 2nd rCagA protein given by a program.

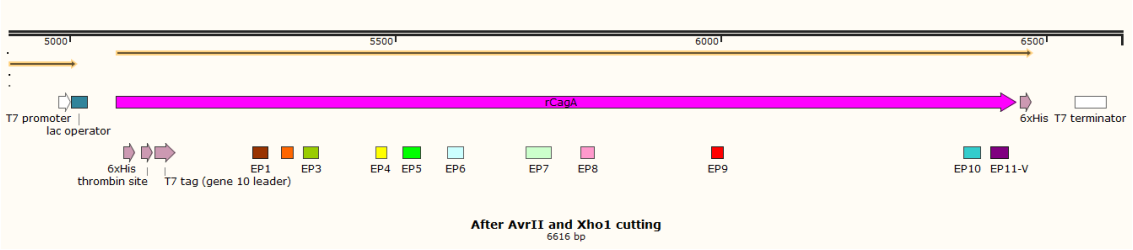


Figure 4.34 The location of these predicted epitopes on the 2nd rCagA protein.

Emini Surface Accessibility Prediction

Input Sequences

```
  1 MGSSHHHHHH SSGLVPRGSH MASMTGGQQM GRGSEAAFNP QQFINNPQVA FLKLDNAVAS
 61 FDPDQKPIVD KNDRDHRQAF EGISQLREEY SNKAIKNPTK KNQYFSDFIN RSNDLINKDA
121 LIDVESSTKS FRKFGDQRYR IFASWVSHQN DPSKINTRSI RNFMENIIQP PISDDKEKAE
181 FLKSAKQSFA GIIIGNQIRT DQKFMGVFDE SLKERQEAEK NGEPTSFVRR NLEDKLTAKG
241 LSSQEASC
```

Predicted peptides:

No.	Start Position	End Position	Peptide	Peptide Length
1	71	78	KNDRDHRQ	8
2	86	91	LREEYS	6
3	97	104	NPTKKNQY	8
4	149	156	QNDPSKIN	8
5	173	178	SDDKEK	6
6	211	223	SLKERQEAEKNGE	13

Figure 4.35 The 6 predicted epitopes of the 3th rCagA protein given by a program.

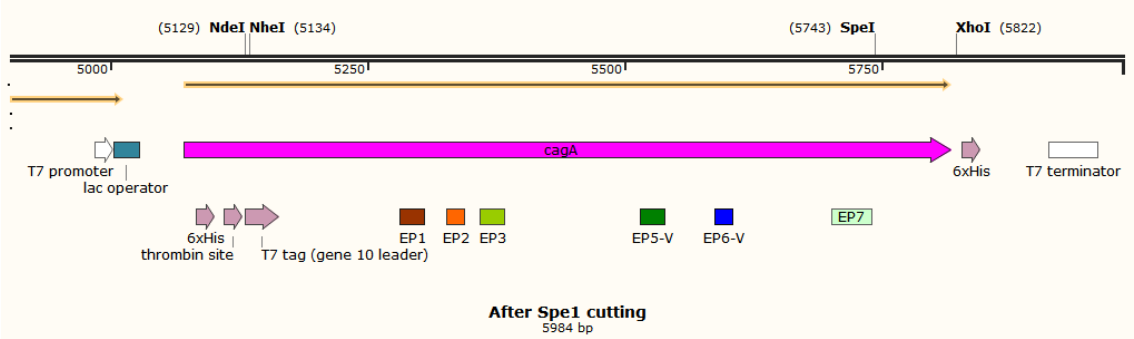


Figure 4.36 The location of these predicted epitopes on the 3rd rCagA protein.

4.14 WESTERN BLOT

The localization of these predicted epitopes on each of these proteins was done using Western Blot. All three recombinant proteins were loaded onto 12% SDS-PAGE prior to Western Blot as shown in the Figure 4.37. In Western Blot analysis, our own monoclonal antibodies (BS-53 and CK-02) and the four commercially available monoclonal antibodies (A-10, B-818, HP-317, His-probe) were used as a primary antibody respectively and the GTxMS-HRP was used as a secondary antibody for all rCagA proteins as shown below in the Figure 4.38 to Figure 4.42.

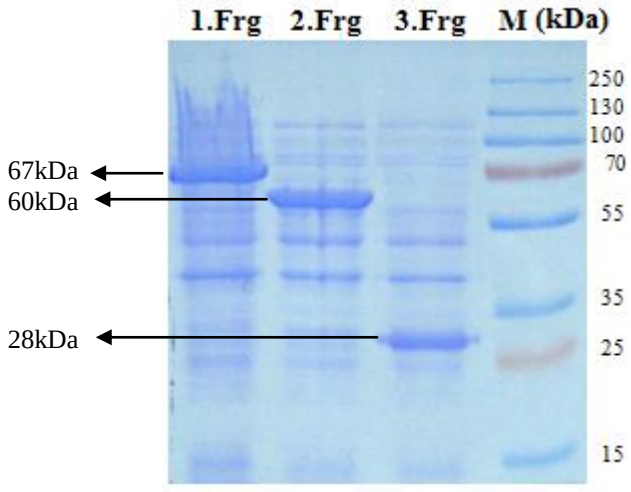


Figure 4.37 All three recombinant proteins were loaded onto 12% SDS-PAGE prior to Western Blot showing the 67 kDa of 1st rCagA protein, the 60 kDa of 2nd rCagA protein and the 28 kDa of 3rd rCagA protein. M: prestained protein ladder.

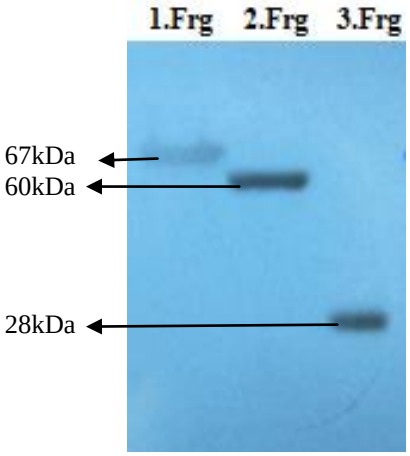


Figure 4.38 Western Blot result using His-probe (H-3) monoclonal antibody showing the 67 kDa of 1st rCagA protein, the 60 kDa of 2nd rCagA protein and the 28 kDa of 3rd rCagA protein.

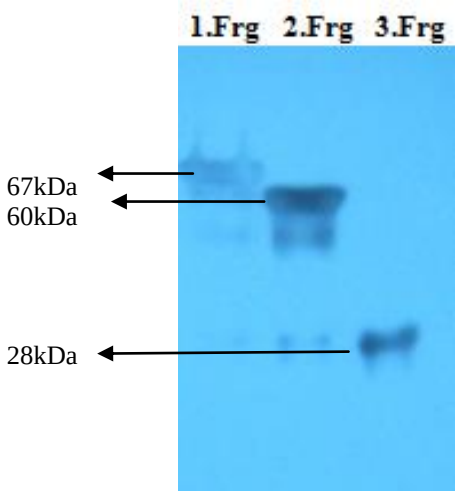


Figure 4.39 Western Blot result using BS-53 monoclonal antibody showing the 67 kDa of 1st rcagA protein, the 60 kDa of 2nd rcagA protein and the 28 kDa of 3rd rcagA protein.

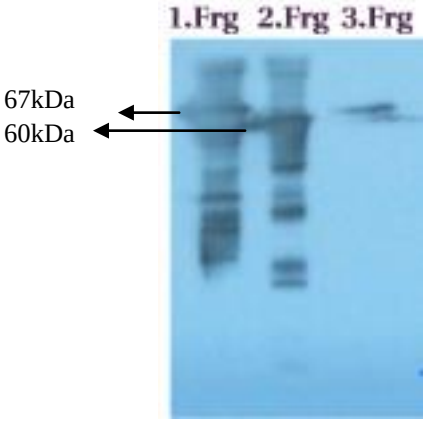


Figure 4.40 Western Blot result using CK-02 monoclonal antibody showing the 67 kDa of 1st rcagA protein, the 60 kDa of 2nd rcagA protein.

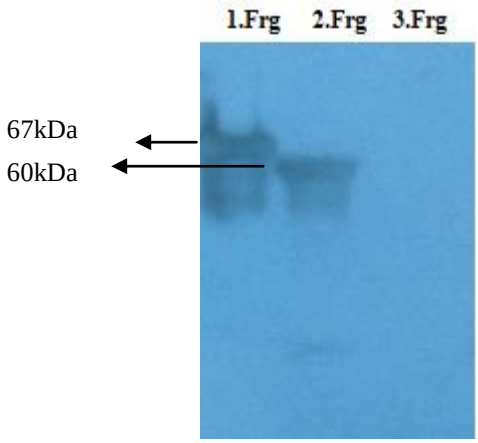
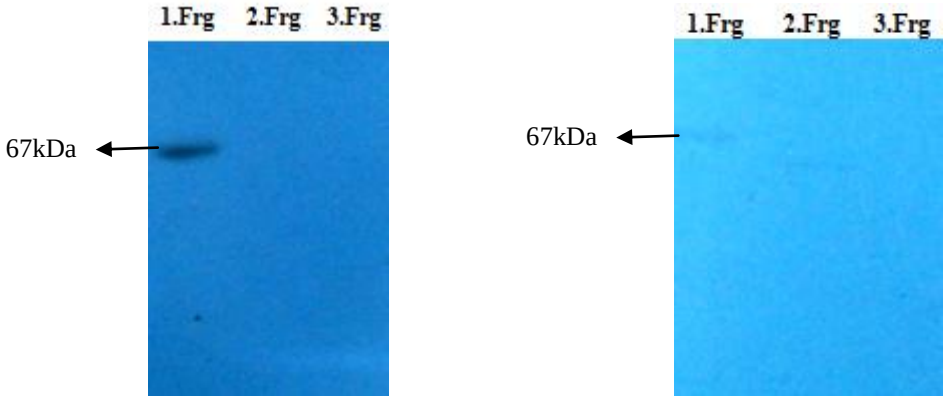


Figure 4.41 Western Blot result using A-10 monoclonal antibody showing the 67 kDa of 1st rcagA protein, the 60 kDa of 2nd rcagA protein.



(a) (b)

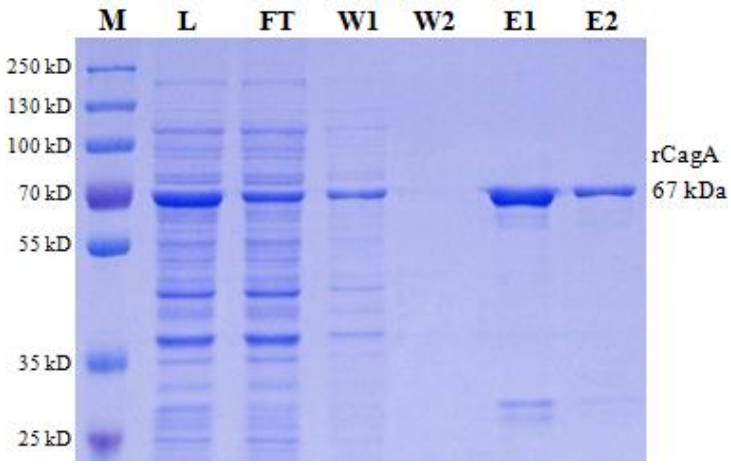
Figure 4.42 Western Blot result using B-818M (a) and B-237H (b) monoclonal antibodies respectively showing the 67 kDa of 1st rcagA protein.

4.15 RECOMBINANT PROTEIN PURIFICATION

By using the B-PER 6xHis Fusion Protein Purification Kit (Pierce, USA), the recombinant CagA proteins (1st rCagA, 2nd rcagA and 3th rcagA) were purified from a 100 ml of IPTG induced *E. coli* culture of bacteria respectively. While using a protocol of this kit, 0.1% SDS in PBS solution was used instead of the B-PER® Reagent to avoid inclusion bodies and to get soluble proteins from insoluble form. The samples from each fraction (Flow-through, Wash 1, Wash 2, Elution 1 and Elution 2) were loaded

onto 12% SDS-PAGE gel and the protein concentrations of samples were measured by Bradford assay as shown below.

(a)

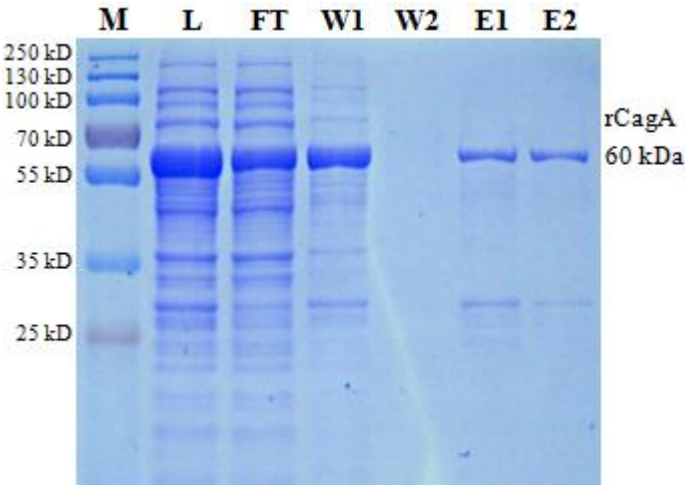


(b)

Samples	Bradford Assay (µg/ml)
L (Lysate)	586
FT (Flow Through)	488
W1 (Wash 1)	398
W2 (Wash 2)	-
E1 (Elution 1)	565
E2 (Elution 2)	52

Figure 4.43 (a) SDS-PAGE of purified the 1st rCagA protein (67 kDa). M: molecular marker, L: Lysate; FT: Flow-Through; W1: Wash1; W2: Wash2; E1: Elution 1; E2: Elution 2. (b) Measurement of protein concentration of each fraction using Bradford Assay

(a)

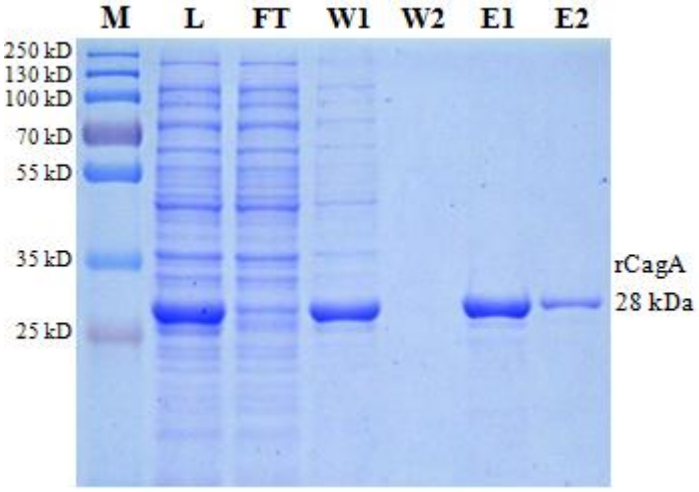


(b)

Samples	Bradford Assay (µg/ml)
L (Lysate)	410
FT (Flow Through)	480
W1 (Wash 1)	320
W2 (Wash 2)	-
E1 (Elution 1)	84
E2 (Elution 2)	86

Figure 4.44 (a) SDS-PAGE of purified the 2nd rCagA protein (60 kDa). M: molecular marker, L: Lysate; FT: Flow-Through; W1: Wash1; W2: Wash2; E1: Elution 1; E2: Elution 2. (b) Measurement of protein concentration of each fraction using Bradford Assay.

(a)



(b)

Samples	Bradford Assay (µg/ml)
L (Lysate)	260
FT (Flow Through)	300
W1 (Wash 1)	170
W2 (Wash 2)	-
E1 (Elution 1)	610
E2 (Elution 2)	73

Figure 4.45 (a) SDS-PAGE of purified the 3rd rCagA protein (28 kDa). M: molecular marker, L: Lysate; FT: Flow-Through; W1: Wash1; W2: Wash2; E1: Elution 1; E2: Elution 2. (b) Measurement of protein concentration of each fraction using Bradford Assay.

4.16 ELISA

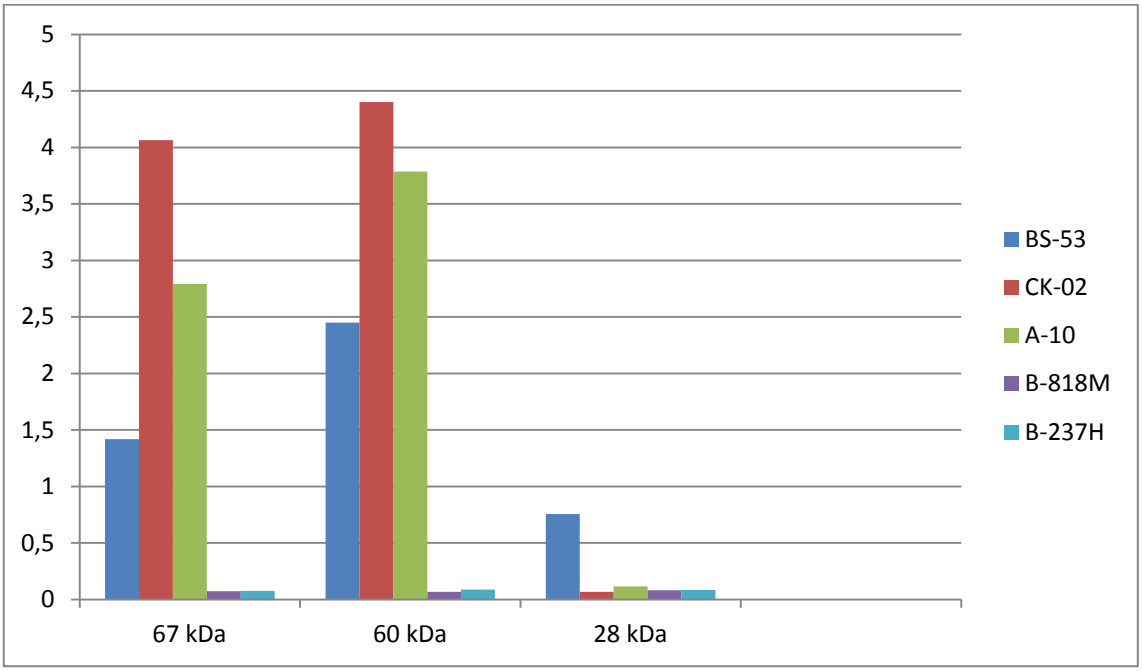


Figure 4.46 ELISA results of all three rCagA proteins (67kDa, 60kDa, 28kDa) that reacted with both our monoclonal antibodies (BS-53, CK-02) and the commercially available antibodies (A-10, B818M, B237H).

According to the Figure 4.40, BS-53 Mab highly reacted with all three rcagA proteins whereas CK-02 Mab reacted with 67 kDa rcagA protein and 60 kDa rcagA protein. The commercial available monoclonal antibodies (A-10, B-818M and B-237H) were used as positive controls to compare our own anti-cagA monoclonal antibodies.

CHAPTER 5

DISCUSSION & CONCLUSION

Helicobacter pylori an important bacterial pathogen that cause chronic active gastritis and peptic ulceration. Since it also plays an important role in the development of gastric cancer and gastric mucosal associated-lymphoid tissue (MALT) lymphoma, it has been classified as a group I carcinogen by the International Agency for Research on Cancer (Gisbert 2008; Sugimoto and Yamaoka 2011).

The most important virulent factor of *H. pylori* is the cytotoxin-associated gene A (cagA). This gene encodes a highly immunodominant protein (CagA) that is injected into the gastric epithelial cell via type IV secretion system (Höcker et al., 2003).

The cagA gene of *H. pylori* is divided into two regions: a cagA 5' conserved region and a 3' variable region (Yamaoka et al., 2010). This cagA 3' variable region contains EPIYA motifs that are the different numbers of repeat sequences. The reason for selecting only the cagA 5' conserved region is that the 3' region shows great variability that can affect the outcome of the results.

The cagA 5' conserved region of *H. pylori* was firstly cloned into the pET-28a(+) vector (pET-28a(+)/rcagA) by Z. Ulupınar. Two different clones were developed from this first recombinant cagA (rcagA) gene using restriction cloning techniques. Multiple alignments of the first rcagA gene, the second rcagA gene fragment and the third rcagA gene fragment showed 99.72%, 99.46% and 98.93% homology in between eachother

The first rCagA protein (67 kDa) was recombinantly fragmented into two overlapped proteins which are 60 kDa and 28 kDa molecular weights to determine the epitope region of the rCagA protein against our own anti-cagA monoclonal antibodies. Multiple protein alignments of the first rCagA protein and the second rCagA protein showed 98.29% homology, the first rCagA protein and the third rCagA protein 98.39% homology and 90.73% homology in between of the second rCagA protein and the third rCagA protein.

Epitope mapping is determination of which part of the antigen is recognized by the antibody and identify the antibody-binding site on the antigen (Morris, 2001). Identification of the epitopes is usually done with monoclonal antibodies rather than polyclonal antibodies and this identification is a pivotal step in the characterization of monoclonal antibodies and epitope-driven diagnostics (Nelson et al., 2000).

Many methods have been developed for mapping protein epitopes of monoclonal antibodies and these can be classified into the structural mapping approach and functional mapping approach. Both of these two approaches have advantages and disadvantages in their own applications. Although the structural mapping approach gives information about epitopes at the atomic level, the functional mapping method gives information at the amino acids level (Clementi et al., 2013).

Klimovich et al. (2010) produced the four overlapped N-terminally His_6-tagged recombinant CagA protein fragments from the cloning of the entire cagA gene and the four monoclonal antibodies using hibridoma production and they mapped the distinct linear epitopes inside conserved regions of the cytotoxin and the epitope in the variable area of CagA.

In this study, we have selected the fragmentation method as a functional mapping approach by making overlapping recombinant cagA gene fragments and Western Blot analysis to determine epitope of these rCagA proteins against the anti-cagA monoclonal antibodies.

As the epitope prediction programs provide a better and rapid understanding of the antibody binding site on an antigen, they are preferable for predicting epitopes (EL-Manzalawy and Honavar, 2010). In this study, the number of predicted epitopes on each

of these recombinant proteins where 14 epitopes in the 67 kDa rCagA protein, 11 epitopes in the 60 kDa rCagA protein and 6 epitopes in the 28 kDa rCagA protein were determined using “Emini Surface Accessibility Prediction” software in the IEDB Analysis Resource program.

The localization of these epitopes on each of these rCagA proteins were done by Western Blot analysis using both our monoclonal antibodies (BS-53, CK-02) produced by hybridoma production in our laboratory and the commercially available monoclonal antibodies (A-10, B818M and B237H). The reason of using His-probe Mab in western blot is to confirm whether the rCagA proteins carry the 6xHistidine tag or not. ELISA was done to confirm the reactivity of these monoclonal antibodies against these rCagA proteins.

Our first developed BS-53 Mab recognized an epitope localized within the first 248 amino acids of all 3 rCagA proteins but our developed CK-02 Mab recognized an epitope localized in between the 248th and 469th amino acid sequence of the 67 kDa rCagA protein and 60 kDa rCagA protein. The commercially available monoclonal antibody, the A-10 Mab recognized an epitope localized in between the 248th and 469th amino acid sequence of the 67 kDa rCagA protein and 60 kDa rCagA protein. While the B-818M Mab and the B-237H Mab recognized an epitope localized in between the 469th and 594th amino acid sequence of the 67kDa rCagA protein.

On the basis of this study, our own BS-53 Mab recognizes all three rCagA proteins developed from cloning of the cagA 5’ region of *H. pylori*. While comparing the other commercial available Mabs, BS-53 is the only monoclonal antibody that recognizes an epitope localized on a different region of the rCagA protein that was not reacted with them. The CK-02 Mab which is our second developed antibody recognizes an epitope localized between the 248th and 469th amino acid sequence of the first rCagA protein and the second rCagA protein. Both of these two Mabs that were developed in our laboratory have recognized the recombinant cagA proteins cloned from the 5’ conserved region as the commercially available Mabs did. In other words, these two Mabs are working in a similar way to the commercially available Mabs.

In conclusion, cloning the cagA 5’ gene of *H. pylori* to produce different recombinant CagA proteins was done. Also, the epitope determination of the

monoclonal antibodies against our developed rCagA protein was successfully done and these monoclonal antibodies can be commercially used in today's marketing and can be further investigated in the future of *H. pylori* epitope-based treatment or vaccine. In addition, our rCagA proteins can aslo be utilized for dignostic and research puposes since it shares over 80% homology with strains from other parts of the world.

In summary, our developed rCagA proteins and monoclonal antibodies provide perspectives for immunodiagnostics and research in the field of *H. pylori* related diseases.

REFERENCES

Arnon, R. and Van Regenmortel, M., "Structural Basis of antigenic specificity and design of new vaccines.", *The FASEB journal*, Vol. 6, No. 14, pp. 3265-3274, 1992.

Baerga-Ortiz, A., Hughes, C. A., et al., "Epitope mapping of a monoclonal antibody against human thrombin by H/D-exchange mass spectrometry reveals selection of a diverse sequence in a highly conserved protein.", *Protein Science*, Vol. 11, No. 6, pp. 1300-1308, 2002.

Barlow, D., Edwards, M., et al., "Continuous and discontinuous protein antigenic determinants.", 1986.

Burkovitz, A., Leiderman, O., et al., "Computational Identification of Antigen-Binding Antibody Fragments.", *The Journal of Immunology*, Vol. 190, No. 5, pp. 2327-2334, 2013.

Chambers, G., Lawrie, L., et al., "Proteomics: a new approach to the study of disease.", *The Journal of pathology*, Vol. 192, No. 3, pp. 280-288, 2000.

Censini, S., Lange, C., et al., "cag, a pathogenicity island of *Helicobacter pylori*, encodes type I-specific and disease-associated virulence factors.", *Proceedings of the National Academy of Sciences*, Vol. 93, No. 25, pp. 14648-14653, 1996.

Clementi, N., Mancini, N., Castelli, M., Clementi, M., Burioni, R., "Characterization of epitopes recognized by monoclonal antibodies: experimental approaches supported by freely accessible bioinformatic tools.", *Drug discovery today*, Vol. 18, No. 9, pp. 464-471, 2013.

Covacci, A., Censini, S., et al., "Molecular characterization of the 128-kDa immunodominant antigen of *Helicobacter pylori* associated with cytotoxicity and duodenal ulcer.", *Proceedings of the National Academy of Sciences*, Vol. 90, No. 12, pp. 5791-5795, 2013.

Covacci, A., "*Helicobacter pylori* pathogenicity: the bacterium and the host.", *Eur J Gastroenterol Hepatol*, Vol. 12, No. 9, pp. 1050-2, 2000.

Dall'Antonia, F., Pavkov-Keller, T., et al., "Structure of allergens and structure based epitope predictions.", *Methods*, Vol. 66, No.1, pp. 3-21, 2014.

Dunn, B. E., Cohen, H., et al., "*Helicobacter pylori*." *Clinical Microbiology Reviews*, Vol. 10, No. 4, pp.720-741, 1997.

EL-Manzalawy, Y. and V. Honavar., "Recent advances in B-cell epitope prediction methods.", *Immunome research*, Vol. 6, No. Suppl 2: S2, 2010.

Fujii, R., Morihara, F., et al., "Epitope mapping and features of the epitope for monoclonal antibodies inhibiting enzymatic activity of *Helicobacter pylori* urease.", *Biotechnology and bioengineering*, Vol. 86, No.4, pp. 434-444, 2004.

González, L., Marrero, K., et al., "Cloning and expression of a recombinant CagA-gene fragment of *Helicobacter pylori* and its preliminary evaluation in serodiagnosis", *Biomédica*, Vol. 4, pp. 546-553, 2013.

Furuta, Y., Yahara, K., et al., "Evolution of cagA oncogene of *Helicobacter pylori* through recombination.", *PLoS One*, Vol. 6, No. 8, pp. 234-99, 2011.

Hager-Braun, C. and Tomer, K. B., "Determination of protein-derived epitopes by mass spectrometry.", *Expert Review of Proteomics*, Vol. 2, No. 5, pp.745-756, 2005.

Hasanzadeh, L., Ghaznavi-Rad, E., et al., "Expression and Antigenic Evaluation of VacA Antigenic Fragment of *Helicobacter pylori*.", Iranian journal of basic medical sciences, Vol. 16, No. 7, pp. 835, 2013.

Hatakeyama, M., " Oncogenic mechanisms of the *Helicobacter pylori* CagA protein", *Nat. Rev. Cancer*, Vol. 4, pp. 688–694, 2004.

Hatakeyama, M., "*Helicobacter pylori* and gastric carcinogenesis.", *J Gastroenterol*, Vol. 44, No. 4, pp. 239-48, 2009.

Higashi, H., Tsutsumi, R., et al., " SHP-2 tyrosine phosphatase as an intracellular target of *Helicobacter pylori* CagA protein", *Science*, Vol. 295, pp. 683–686, 2002.

Hirota, K., Nagata, K., et al., "Identification of an Antigenic Epitope in *Helicobacter pylori* Urease That Induces Neutralizing Antibody Production.", *Infection and Immunity*, Vol. 69, No. 11, pp. 6597-6603, 2001.

Horsfall, A. C., Hay, F. C., et al., "Epitope mapping.", *Immunology today*, Vol. 12, No. 7, pp. 211-213, 1991.

Höcker M, Hohenberger P., "*Helicobacter pylori* virulence factors-one part of a big picture.", *Lancet*, Vol. 362, No. 9391, pp. 1231-1233, 2003.

Huang, J. and Honda W., "CED: a conformational epitope database.", *BMC immunology*, Vol. 7, No. 1, pp. 7, 2006.

Hua, R. H., Liu, L. K., et al., "Comprehensive mapping of a novel NS1 epitope conserved in flaviviruses within the Japanese encephalitis virus serocomplex.", *Virus research*, Vol. 185, pp. 103-109, 2014.

Hua, R. H., Liu L. K., Chen Z. S., Li,Y. N., Bu, Z. G., "Comprehensive Mapping Antigenic Epitopes of NS1 Protein of Japanese Encephalitis Virus with Monoclonal Antibodies.", *PLoS ONE*, Vol. 8, No. 6, pp. 67553, 2013.

Kindt, T. J., Goldsby,R. A., et al., *Kuby immunology*, Macmillan, 2007.

Klimovich, A. V., Samoylovich, M. P., et al., "Development of Immunoreagents for Diagnostics of CagA-Positive *Helicobacter pylori* Infections.", *Helicobacter*, Vol. 15, No. 3, pp. 193-200, 2010.

Kulkarni-Kale, U., Bhosle, S., et al., "CEP: a conformational epitope prediction server.", *Nucleic acids research*, Vol. 33, No. suppl 2, pp. 168-171, 2005.

Ladner, R. C., "Mapping the epitopes of antibodies.", *Biotechnology and Genetic, Engineering Reviews*, Vol. 24, No. 1, pp. 1-30, 2007.

Larsen, J. E., O. Lund, et al., "Improved method for predicting linear B-cell epitopes.", *Immunome research*, Vol. 2, No.1, pp. 2, 2006.

Li, H. X., Mao, X. H., et al., "Screening and identification of a novel B-cell neutralizing epitope from *Helicobacter pylori* UreB.", *Vaccine*, Vol. 26, No. 52, pp. 6945-6949, 2008.

Li, Y., Ning, Y., et al., "Production of mouse monoclonal antibodies against *Helicobacter pylori* Catalase and mapping the antigenic epitope by phage display library.", *Vaccine*, Vol. 26, No. 9, pp. 1263-1269, 2008.

Li, Y., Ning, Y., et al., "Production of mouse monoclonal antibodies against *Helicobacter pylori* Lpp20 and mapping the antigenic epitope by phage display library.", *Journal of Immunological Methods*, Vol. 325, No. 1–2, pp. 1-8, 2007.

Li, Y., Ning, Y., et al., "Mimotopes selected with a neutralizing antibody against urease B from Helicobacter pylori induce enzyme inhibitory antibodies in mice upon vaccination.", *BMC biotechnology*, Vol. 10, No. 1, pp. 84, 2010.

Lu, X., DeFelippis, M. R., et al., "Linear epitope mapping by native mass spectrometry.", *Analytical biochemistry*, Vol. 395, No. 1, pp. 100-107, 2009.

Marshall, D. G., Dundon, W. G., et al., "*Helicobacter pylori* - a conundrum of genetic diversity.", *Microbiology*, Vol. 144, No. 11, pp. 2925-2939, 1998.

Marshall, B. and Warren, J. R., "Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration.", *The Lancet*, Vol. 323, No. 8390, pp.1311-1315, 1984.

Menezes-Souza, D., Oliveira Mendes, T. A. De, et al., "Mapping B-Cell Epitopes for the Peroxidoxin of Leishmania (Viannia) braziliensis and Its Potential for the Clinical Diagnosis of Tegumentary and Visceral Leishmaniasis.", PLoS One, Vol. 9, No. 6, pp. 99216, 2014.

Morris, G. E., *Epitope Mapping: B-cell Epitopes*. eLS, John Wiley & Sons, Ltd., 2001.

Morris, G. E., ”Epitope mapping: Immunochemical Protocol”, *Springer***,** pp. 255-267, 2005.

Nelson, P. N., Reynolds, G. M., et al., "Demystified …: Monoclonal antibodies." , *Molecular Pathology*, Vol. 53, No. 3, pp. 111-117, 2000.

Nagai, H., Suzukawa, M., et al., "Immunological Responses and Epitope Mapping by Tuberculosis-Associated Antigens within the RD1 Region in Japanese Patients.", *Journal of Immunology Research*, Vol. 2014, pp. 8, 2014.

Mitić, N. S., Pavlović, M. D., Jandrlić, D. R., "Epitope distribution in ordered and disordered protein regions - part A. T-cell.", *J Immunol Methods*, Vol. 406, No. 83, pp. 103, 2014.

Onoue, K., Yagi, Y., et al., "Number of binding sites of rabbit macroglobulin antibody and its subunits.", Immunochemistry, Vol. 2, No. 4, pp. 401-415, 1965.

Qiu, Y., Wang, Y. C., et al., "Identification of B-cell epitopes in urease B subunit of *Helicobacter pylori* bound by neutralizing antibodies.", Vaccine, Vol. 28, No. 32, pp. 5220-5227, 2010.

Pavlović, M. D., Jandrlić, D. R., et al., "Epitope distribution in ordered and disordered protein regions. Part B — Ordered regions and disordered binding sites are targets of T- and B-cell immunity.", *Journal of Immunological Methods*, Vol. 407, pp. 90-107, 2014.

Westhof E., et al., "PREDITOP: A program for antigenicity prediction.", *Journal of Molecular Graphics*, Vol. 11, No. 3, pp. 204-210, 1993.

Pomés, A., "Relevant B cell epitopes in allergic disease.", *International archives of allergy and immunology*, Vol. 152, No. 1, pp. 1-11, 2009.

Roncarati, D., Spohn, G., et al., "Expression, purification and characterization of the membrane-associated HrcA repressor protein of *Helicobacter pylori*.", *Protein Expr Purif*, Vol. 51, No. 2, pp. 267-275, 2007.

Saint-Remy, J. M. R., "Epitope mapping: a new method for biological evaluation and immunotoxicology.", *Toxicology*, Vol. 119, No. 1, pp.77-81, 1997.

Sela-Culang, I., Kunik, V., Ofran Y., "The structural basis of antibody-antigen recognition." , *Frontiers in immunology*, Vol. 4, pp. 302, 2013.

Sugimoto, M., Yamaoka, Y., et al., "Review of *Helicobacter pylori* Infection and Chronic Renal Failure", *Therapeutic Apheresis and Dialysis*, Vol. 15, No. 1, pp. 1–9, 2010.

Tummuru, M. K., Cover, T. L., et al., "Cloning and expression of a high-molecular-mass major antigen of *Helicobacter pylori*: evidence of linkage to cytotoxin production.", *Infection and Immunity*, Vol. 61, No. 5, pp. 1799-1809, 1993.

Turner, K., Zabetakis, D., et al., "Isolation and Epitope Mapping of Staphylococcal Enterotoxin B Single-Domain Antibodies.", *Sensors*, Vol. 14, No. 6, pp. 10846-10863, 2014.

Ulupınar, Z., *Cloning the cagA gene of Helicobacter pylori for the production of recombinant CagA protein*, M.S. Thesis, Fatih University, 2012.

Van Regenmortel, M., "Structural and functional approaches to the study of protein antigenicity.", *Immunology today*, Vol. 10, No. 8, pp. 266-272, 1989.

Yamaoka, Y., et al., "Molecular epidemiology of *Helicobacter pylori*: separation of *H. pylori* from East Asian and non-Asian countries", "*Epidemiol Infect.*", Vol. 124, pp. 91–96, 2000.

Yamaoka, Y., Kodama, T., et al., "Variants of the 3' region of the cagA gene in *Helicobacter pylori* isolates from patients with different *H. pylori*-associated diseases", *J Clin Microbiol*, Vol. 36, pp. 2258-2263, 1998.

Yamaoka, Y., El-Zimaity, H. M., et al., "Relationship between the cagA 3' repeat region of *Helicobacter pylori*, gastric histology, and susceptibility to low pH.", *Gastroenterology*, Vol. 117, No. 2, pp. 342-349, 1999.

Yamada, K., Sugiyama, T., et al., "Fragmented CagA protein is highly immunoreactive in Japanese patients.", *Helicobacter*, Vol. 17, No. 3, pp. 187-192, 2012.

Zhang, Z., Pan, H., et al., "Mass spectrometry for structural characterization of therapeutic antibodies.", *Mass Spectrometry Reviews*, Vol. 28, No.1, pp.147-176, 2009.