DOKUZ EYLÜL UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

MOLECULAR CLONING AND RECOMBINANT PRODUCTION OF HUMAN (EpCAM) EXTRACELLULAR DOMAIN

by

Elsayed Nabil Elsayed Hassan ZAABOUT

March, 2019

İZMİR

MOLECULAR CLONING AND RECOMBINANT PRODUCTION OF HUMAN (EpCAM) EXTRACELLULAR DOMAIN

A Thesis Submitted to the
Graduate School of Natural and Applied Sciences of Dokuz Eylül University
In Partial Fulfillment of the Requirements for the Degree of Master of Science
in Biotechnology Department

by

Elsayed Nabil Elsayed Hassan ZAABOUT

March, 2019

İZMİR

M.Sc THESIS EXAMINATION RESULT FORM

We have read the thesis entitled **"MOLECULAR CLONING AND RECOMBINANT PRODUCTION OF HUMAN (EpCAM) EXTRACELLULAR DOMAIN"** completed by **ELSAYED NABIL EL. ZAABOUT** under supervision of **PROF. DR. HÜLYA AYAR KAYALI** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Prof. Dr. Hülya AYAR KAYALI

Supervisor

Prof.Dr.Sinan AKGÖL

(Jury Member)

Prof.Dr. M.Nalan TÜZNEN

(Jury Member)

Prof.Dr. Kadriye ERTEKİN

Director
Graduate School of Natural and Applied Sciences

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my advisor Prof. Dr. Hülya AYAR KAYALI for the continuous support of my M.Sc. study and research, for her motivation, enthusiasm, and assistance. I could not have imagined having a better advisor and mentor for my M.Sc study.

I would like to express my deepest acknowledge to Assoc. Prof. Kasım DİRİL for his patience, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis.

I am particularly grateful for the assistance given by all lab mates.

Last but not least, I would like to thank my parents and my siblings (Samar, Ahmed, and Hanin) for enduring my alienation.

My lovely thanks go to my aunty Dilek KOKTAS for her support.

I would like to thank my dearest friends Eren Egemen NARİN and Ali UZUN-YAYLA for believing in me.

Elsayed Nabil Elsayed Hassan ZAABOUT

MOLECULAR CLONING AND RECOMBINANT PRODUCTION OF HUMAN (EpCAM) EXTRACELLULAR DOMAIN

ABSTRACT

Epithelial cell adhesion molecule (EpCAM) is a type I transmembrane glycoprotein of approximately 40 kDa size that facilitates Ca^{2+}-independent homotypic cell and cell adhesion in epithelia. EpCAM consist of two domains; extracellular domain (EpEX) and intracellular domain (EpICD), separated by a transmembrane helix. EpCAM's roles include cellular signaling, and migration. The overexpression of EpCAM in epithelial tumors and downregulation in normal epithelia is used as a diagnostic tumor marker. Previous studies demonstrated that EpCAM can serve as an attractive anti-tumor target such as, immunotherapy, RNA aptamers, and vaccination. Therefore, we aimed to express and purify recombinant EpEX molecules in an *(*E. coli) based expression system in order to use it in downstream applications. RNA isolated from two different human cancer lines highly expressing EpCAM was used to amplify the cDNA coding for EpEX. PCR amplified EpEX cDNA fragments were ligated into pGEX-6P1, a plasmid vector for bacterial expression and purification of recombinant proteins. The correct pGEX-6P1-EpEX clone was isolated and verified by colony PCR screening, restriction analysis, and subsequent Sanger sequencing. The plasmid was transferred to BL21 strain for expression of the recombinant protein. In order to obtain the highest yield and purity for recombinant EpEX molecule, induction, expression and purification conditions were optimized. GST-EpEX was purified by affinity chromatography utilizing glutathione-agarose beads. HRV3C protease was used to cleave and release the recombinant EpEX molecules from the GST tag that remains bound to the beads. And finally, the results were analyzed by Coomassie Blue staining of SDS-PAGE separated protein samples.

Keywords: Epithelial cell adhesion molecules (EpCAM), Recombinant protein production, molecular cloning

İNSAN EpCAM'IN EKSTRASELLÜLER BÖLGESİNİN MOLEKÜLER KLONLANMASI VE REKOMBİNANT ÜRETİMİ

ÖZ

Epitelyal hücre adezyon molekülü (EpCAM, CD326), epitel hücrelerde Ca^{+2} 'a bağımlı homotipik hücre-hücre adezyonuna aracılık eden ve yaklaşık 40 kDa boyutunda, tip I transmembran glikoproteinidir. EpCAM iki alt domainden oluşmaktadır; bir hücre dışı domain (EpEX) ve bir hücre içi domain (EpICD), bu iki domain bir transmembran sarmal ile ayrılmıştır. Sağlıklı dokularda, EpCAM hücreler arası sinyal iletimi ve hücre göçü gibi bir çok temel biyolojik proseste görev almaktadır. EpCAM'in epiteliyal tümör dokularında yüksek seviyede ekspresyonu ve sağlıklı epitel hücrelerinde düşükseviyede eksprese olması kanser tanısında bir biyolojik belirteç olarak kullanılmaktadır. Daha önce yapılan çalışmalar EpCAM'in iyi bir anti-tümör hedefi olarak görev yapabileceğini göstermiştir. Bu teknikler başlıca, immünoterapötik stratejiler, aşılama, RNA aptamerleri. Bu çalışmada, sonraki uygulamalarda kullanılması amacıyla (*E. coli*) tabanlı bir ekspresyon sisteminde rekombinant EpEX moleküllerini ifade etmeyi ve saflaştırmayı amaçlanmıştır. Bu amaçla, EpEX'i yüksek seviyede ifade eden iki farklı insan kanser hattından toplam fragmentleri, bakteriyel ekspresyon ve rekombinant proteinlerin saflaştırılması için bir plazmit vektörü olan pGEX-6P1'e entegre edildi. Doğru pGEX-6P1-EpEX klonunlarını içeren kolonilerin tespiti için öncelikle PCR taraması ve sonrasında restriksiyon enzimleri ile kesilen plasmidlere Sanger dizi analizi yapılarak doğru klonlar seçildi. Seçilen klonlar recombinant protein üretimi amacıyla *E. coli* BL21 strainine transfekte edildi. Rekombinant EpEX molekülü için en yüksek verimi ve saflığı elde etmek için, indüksiyon, ekspresyon ve saflaştırma koşulları optimize edildi. Üretilen GST-EpEX, glutatyon-agaroz taneleri kullanılarak afinite kromatografisiyle saflaştırıldı. Saflaştırılan rekombinant EpEX moleküllerini boncuklara bağlı kalan GST etiketinden ayırmak için HRV3C proteaz enzimi kullanıldı. Son olarak, izole edilen protein SDS-PAGE yapılarak doğrulandı.

Anahtar Kelimeler: Epitelyal hücre adheyton molekülü (EpCAM), Rekombinant protein üretimi, moleküler klonlama

CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER ONE
INTRODUCTION

1.1 Cell Adhesion Molecules (CAMs)

Cell adhesion molecules are abbreviated as (CAMs) are a group of adhesion molecules existing on the cell membrane included in binding with other cells or with the extracellular matrix (ECM) in the process called cell adhesion. Basically, CAMs assist cells to adhere to each other and to their neighboring. CAMs have a crucially important role in both function and maintaining tissue structure. In animals, they play an intrinsic function in forming movement and force thus confirm that organs are able to perform functions. Additionally, helping as "molecular adhesive molecules", CAMs has a vital function in contact inhibition, moving cellular pathways of growth, and apoptosis. Many times, the abnormal overexpression of CAMs will be pathologically significant in cancers (Fuss, 2012).

1.1.1 The Structure of CAMs

CAMs are individual-pass transmembrane receptors (Membranome, 2019) which composed of three domains:

- An extracellular domain
- An intracellular domain
- A transmembrane domain.

The interactions of CAMs have different ways (Chothia, 1997). The first way by homophilic binding, which CAMs bind with the same CAMs. The second way by heterophilic binding means a CAM of a cell binds with another CAMs on a different cell. The third way via cell-substrate, when an extracellular domain binds two different CAMs.

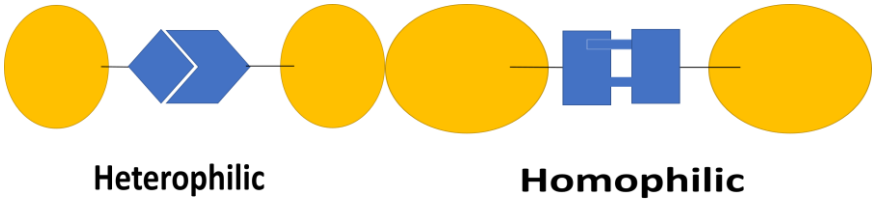

Figure 1.1 Homophilic and heterophilic interactions

1.1.2 CAMs Classification

There are major superfamilies of CAMs: Cadherins, Integrins, and the Superfamily of C-type of lectin-like domains proteins (CTLDs), and the immunoglobulin superfamily of cell adhesion molecules (IgCAMs). Proteoglycans are also one of the CAMs classes. One classification system involves the distinction between calcium-independent CAMs and calcium-dependent CAMs (Brackenbury, 1981). Integrins and the Ig-superfamily CAMs are Ca^{2+} independent while cadherins and selectins are both Ca^{2+} dependent. Only integrins participate in cell-matrix interactions, but other CAM families participate in cell-cell interactions (Lodish, 2000).

1.1.2.1 CAMs That do not Rely on Calcium (Integrin and IgSF)

Integrins, one of the major receptor types within the ECM (Brown,2007), mediates cell–ECM interactions with fibrinogen, collagen, and vitronectin (Humphries, 2006). Integrins create significant links between the extracellular environment and the intracellular signaling pathways, which may have roles in cell behaviors such as differentiation, apoptosis, transcription, and survival, Integrins considerably heterodimeric, and consist of an alpha and beta subunit (García, 2005). Integrin combinations consist of 18 α subunits with 8 β subunits to organize 24 variety of integrin combinations. There are a long extracellular domain, a short cytoplasmic domain, and a transmembrane domain involved within every alpha and beta subunits (Vinatier,1994). With the extracellular domain, some divalent cations are responsible for the binding of the ligands, and it goes this way:

- Mg^{2+} promotes adhesion to cells
- Mn^{2+} increases affinity

- Ca^{2+} decreases cell

The activity of integrins is regulated by changing the conformations.

Immunoglobulin superfamily CAMs (IgSF CAMs) could be considered highly assorted CAMs superfamily type. The extracellular domain of this family gives its definition Ig-like domain. Both types of binding heterophilic or homophilic in addition to bind integrins and different types of IgSF CAMs.

1.1.2.2 CAMs That Rely on Calcium (Cadherins and Selectins)

Cadherins are glycoproteins depend on Ca^{2+}, homophilic binding type. Cadherins exist in the intermediate cell junction with high concentration, and these types of cadherins are E-, P-, and N-. Catenin is a protein which links both cadherins and actin filament network (Buxton, 2010). Embryonic development concludes cadherins; the formation of the three types of the derma (ectoderm, endoderm, and mesoderm). Nervous system development also concludes cadherins, which have an essential function in the process of synaptic stabilization. Each type of cadherins has a special type of tissue pattern with a unique distribution, which is controlled by calcium. Cadherins family involves in the brain, muscle, retinal, placental, and epithelia (Ward, 2011).

Selectins are carbohydrates dependent, heterophilic binding. Selectins family include three types:

- Leukocyte
- Endothelial
- Platelet

Selectins roles are notable in immunity and the function of the immune system through white blood cells.

1.1.3 Biological Functions of CAMs

One of the reasons that make CAMs important as a therapeutic agent is the functionality of in different cancers metastasis, which blocking the cell cancers' capability to travel towards different sites throughout the body. It was obviously explained in the metastasis of melanoma which sharpens to the lungs. From the experiments that carried out in mice's lung endothelium, specific antibodies were used against CAMs, the notable elimination of metastatic numbers was exhibited (Andreoli, 2013).

1.2 Epithelial Cell Adhesion Molecule (EpCAM)

Epithelial tissues cover the internal and external surfaces forming the main barricade against any exchange internally or externally in eukaryotic organisms. Epithelium classified into stratified epithelium and simple epithelium. Stratified epithelial tissue has several cell layers, and can be detected in the cornea, the oral cavity, and the skin. Simple epithelium can be detected in the organs in which filtration, secretion, or absorption take a place in the intestine, the kidney, and pancreas. Simple epithelium comprised of the monolayer that strongly packed and organized, highly exist in cell-cell junction that on the lateral domain, and the attachment of a cell to basal membrane (Figure 1.2). Any dysfunction occurs in cellularly or molecularly levels might be a reason for cancer.

Tight junction creates a non-permeable or semi-permeable, which adhesive between the basolateral domain of the cell and the apical, therefore the epithelial barrier function could be maintained. The polarity of the cell is regulated by cell adhesion, thus reflects cellular function with cell migration, proliferation, and differentiation whenever contact inhibition occurs (Susann, 2011). Contact inhibition is pivotal in embryonic development, and maturation and proliferation of wounds (Fagotto, 1996). Decreasing or losing of contact inhibition might cause uncontrolled proliferation and eventually metastasizes by invading the close tissues (Takai, 2008). There are four subclasses of cell adhesion molecules (Figure 1.2) written below:

- Selectins
- Cadherins
- Integrins
- Immunoglobulin

Additionally, to the previous, some families are classified individually (Trzpis, 2007). Basically, cell adhesion molecules can establish two types of adhesion:

- Cell and cell adhesion
- Cell and matrix adhesion

Epithelial cell adhesion molecule has an individual classification, which does not rightly place in any four cell adhesion molecules classes. It was found out for the first time as cancer diagnostic markers (Köhler, 1975).

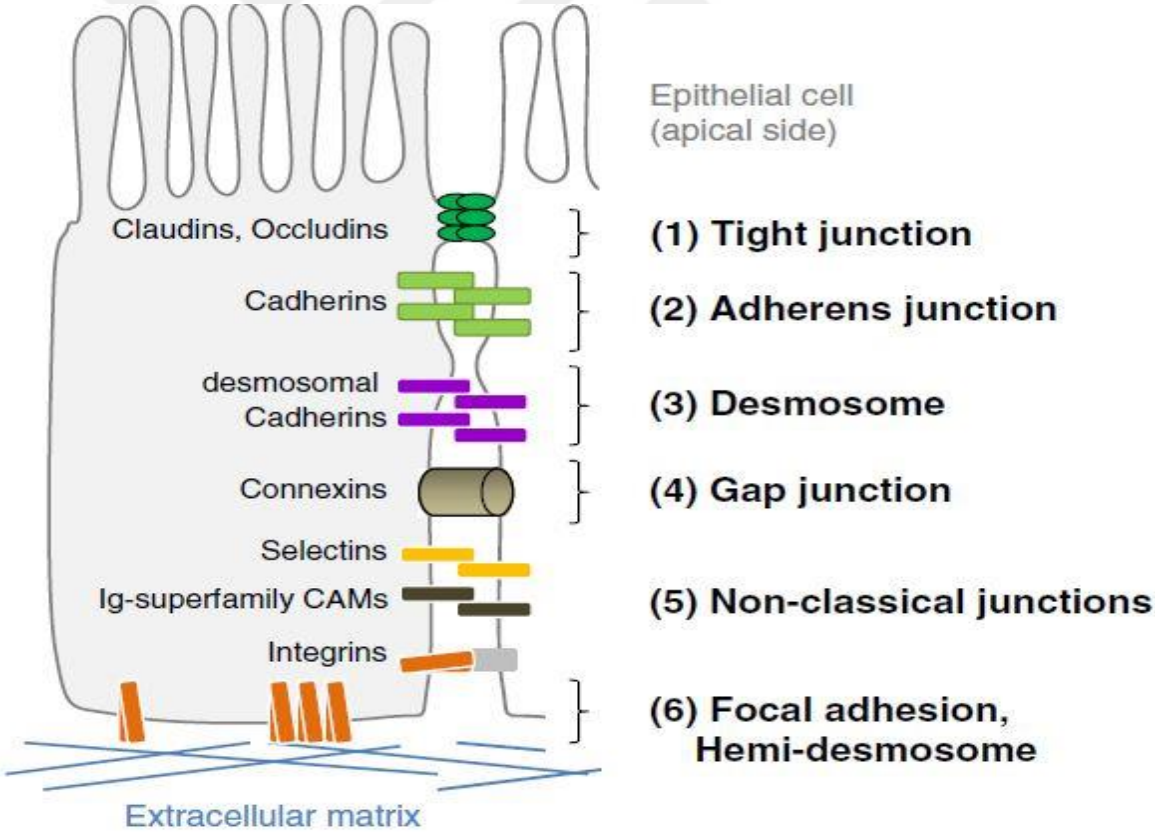

Figure 1.2 CAMs and epithelial adhesion types (Schnell, 2013)

Because of the overexpression of EpCAM in carcinoma, EpCAM's importance in the therapy of different epithelioma and diagnosis takes place (Fagerberg, 2010).

EpCAM works as a regulator morphologically in stem cells and normal epithelia, it may deliver progression in cancer cells. (Spizzo, 2002). Previous studies revealed; any mutations occur in the gene of EpCAM have an impact on intestinal epithelium which causes diarrhea in infants (Sivagnanam, 2008). As a result of EpCAM's function, it is included in different processes such as adhesion, cell signaling (Balzar,1999). The high expression of EpCAM induces proliferation (Münz, 2004).

1.2.1 EpCAM as a Gene, as Protein, and as Structure

The EpCAM of human is defined as a protein which consists of three hundred fourteen amino acids, divided into an extracellular domain of two hundred forty-two amino acids and an intracellular domain of twenty-six amino acids and the middle domain is cytoplasmic consists of twenty-three amino acid (Figure 1.3) (Strnad, 1989). The location of the EpCAM gene is on the second chromosome (2p21) and size of fourteen kb (Balzar, 1999 & Szala, 1990). Genomic studies have proved that EpCAM encoding gene has nine exons; from exon one to six encodes extracellular domain, the seventh exon encodes the transmembrane site, the rest eighth and ninth and the intracellular domain (Linnenbach, 1989). The EpCAM gene most often is preserved within different organisms such as a rat. The protein of EpCAM is extremely preserved between higher vertebrates such that, eighty-one percent of the amino acid sequence among human and mouse seems homology and ninety-nine percent between human and gorilla.

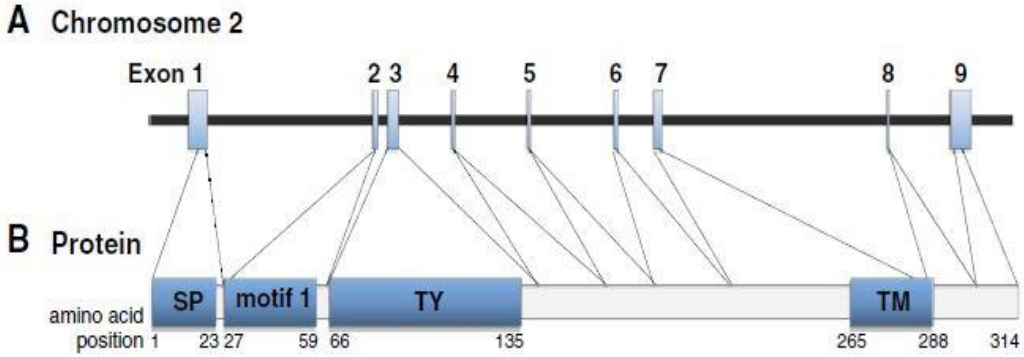

Figure 1.3 EpCAM gene (Schnell, 2013)

1.2.1.1 Human EpCAM Extracellular Domain and Intracellular Domain

The extracellular domain of human EpCAM (EpEX) begins with an individual peptide that is shredded mostly among alanine and glutamine (Figure 1.4) Serval patterns of the tertiary structure of the exodomain of EpCAM were enhanced, introducing three motifs. Some studies have suggested that EpEX contains an epidermal growth factor (EGF) tandemly repeated from the amino acids twenty-seventh to fifty-ninth and from sixty-sixth to one hundred thirty sixth; which nearly match with the 4^{th} and 5^{th} EGF-like domain. With the first motif disulfide bond in EpCAM's ectodomain does not match an EGF-like domain, and the second motif does not display an EGF-like repeat, however, it does match a thyroglobulin type 1 A (TY). TY 1 domains have the ability to bind and inhibit definite cathepsins which are included in cancer progression. A third motif is cysteine-free and irrelevant with any other molecules. EpCAM has N-glycosylation sites. One of the significant importance for the cellular membrane expression of EpCAM as well as the stability of the protein is the glycosylation of Asn 198. The intracellular domain of EpCAM consists of twenty-six amino acids, fourteen of them are charged.

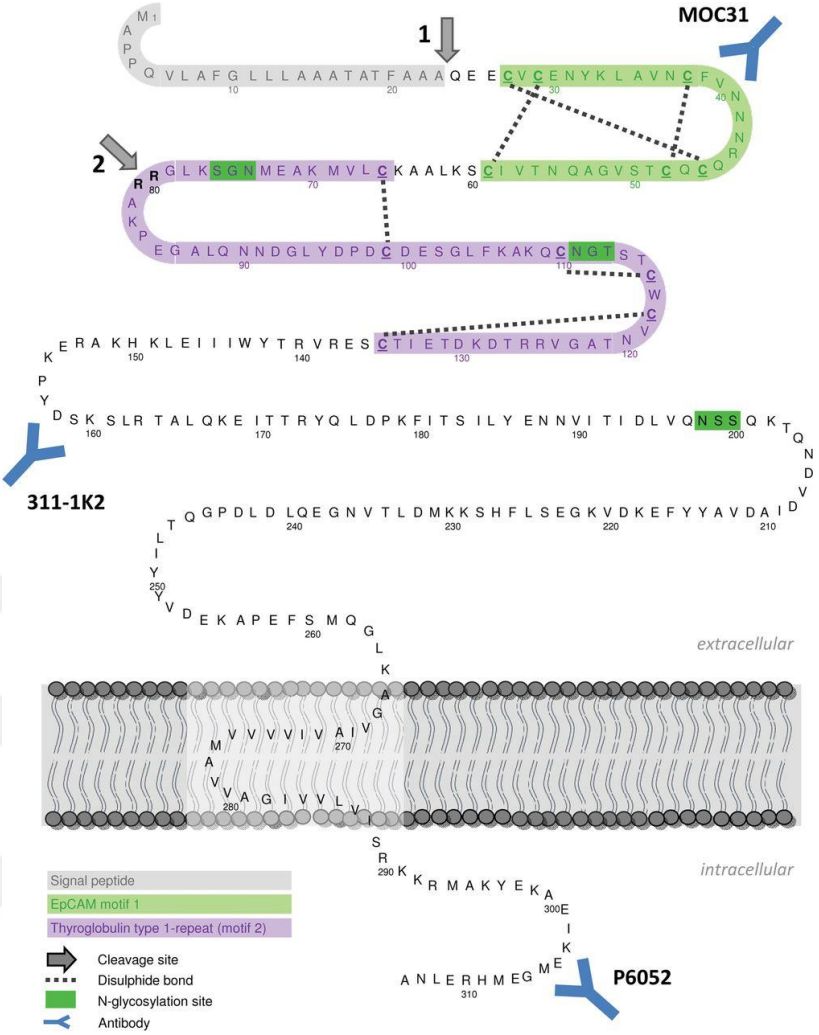

Figure 1.4 The sequence of EpCAM

1.2.1.2 The Cleavage of EpCAM

The cleavage of EpCAM in different sites was imposed on the function. A signal peptide of EpCAM is cleavage by signal peptidase as shown in (Figure 1.4, the first arrow). In addition to cleaving between arg80,81, located in the 2ed motif. The molecular weight of the cleaved part is almost six kDa. there is a link still exist by a disulfide bond between that part six kDa and thirty-two kDa part. The disulfide bond occurs between both sixth and seventh Cys. Remarkable, most of the monoclonal antibodies that target EpCAM bind to the cleaved part six kDa. In some cancers that overexpress EpCAM such as ovarian, colon, and breast cancers, proteases serine,

chymotrypsin, and trypsin can carry out the cleavage of N-terminus. The information about the post-translational modifications of EpCAM still very limited.

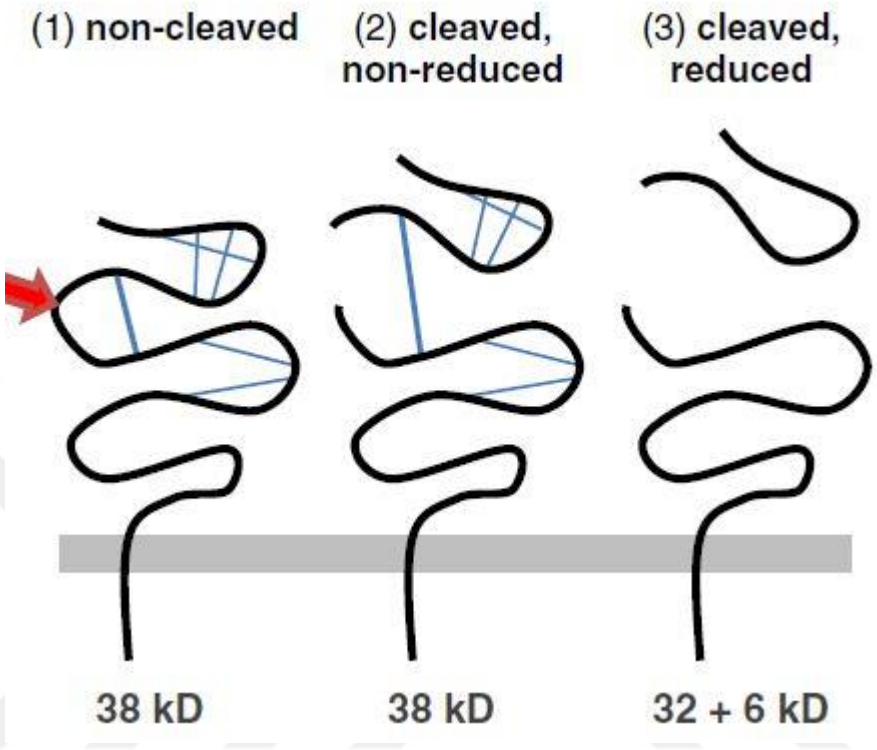

Figure 1.5 The cleavage of EpCAM both parts 6 kDa and 32 kDa

1.2.2 EpCAM Function

At the beginning of discovering EpCAM as a diagnostic marker for tumorigeneses, EpCAM's function still widely limited. Some claims that it could be including in cell-cell or cell-matrix adhesions. It was interpreted by Litvinov and his Colleagues that, EpCAM mediates homophilic cell-cell interaction without depending on calcium. The cells that do not express EpCAM led to cell aggregations and form of cell-cell contact. In a blender of EpCAM expressive cells and EpCAM non-expressive cells, will finally lead to homophilic interaction. In the deletion mutant EpCAMs, Balzar and his colleagues have explained that EpCAM's extracellular domain motifs are the crucially required for cell-cell interaction and adhesion (Balzar, 1999). Furthermore, the intracellular domain in most cases has actin cytoskeleton connection by one of the

actinin types, it is essentially important for the capability of EpCAM to adhere through cell-cell contact (Balzar, 2001).

In addition to dominating over the homophilic adhesion, EpCAM does not belong to any cell adhesion molecules family, neither structure relativity nor serving with any conventional structure via a junction. In epithelial colon cancers and cell lines, the detection of EpCAM seems to be extremely difficult with the adhesive cell-matrix, and tight junction. Whilst there is slightly colocalization with epithelial cadherin within the lateral membrane. Epithelial cadherin is usually stronger cell-cell adhesion than EpCAM (Litvinov, 1994). Cells that express epithelial cadherin are strongly connected because of the adhesive junction, cells that express EpCAM are broadly interlinked (Litvinov, 1994). From the epithelial characteristics, both polarization and contact inhibition, are stronger in epithelial cadherin than epithelial cell adhesion molecules (Litvinov, 1997). All of this information ensures that EpCAM behaves as a negative adhesive regulator. Epithelial cadherins are vital in maintaining epithelial tissues (Gaiser, 2012). Lack of the expression of epithelial cadherin can subdue the contact inhibition and polarity of the cells, uncontrolled growth, and metastasis (Buono, 1999). Throughout the development and progression, there is a co-expression between epithelial cadherin and epithelial cell adhesion molecule. In whatever way, an overexpression of EpCAM mostly serves with the travel of the tumor from tissue to another tissue, in contrast, epithelial cadherin reveals inhibition and suppression for both tumors and function of the growth. In malignancy cases, cell-cell adhesion of cadherins gets weaker, controversy, EpCAM brings more cell elasticity enhancement, which eventually leads to overrating the proliferation and cell death. Furthermore, it was shown that EpCAM acts as a regulator of integrity by its effect on the function of the junction through the interaction with claudins family. As depicted in (Figure 1.6) the cleavage of the total EpCAM through ADAM17 unleashing the extracellular domain (EpEX). The behavior of EpEX seems to be homophilic. The rest is the cytoplasmic domain and an intracellular domain (EpICD), both of them slightly short domains. EpICD is cleaved by gamma-secretase.

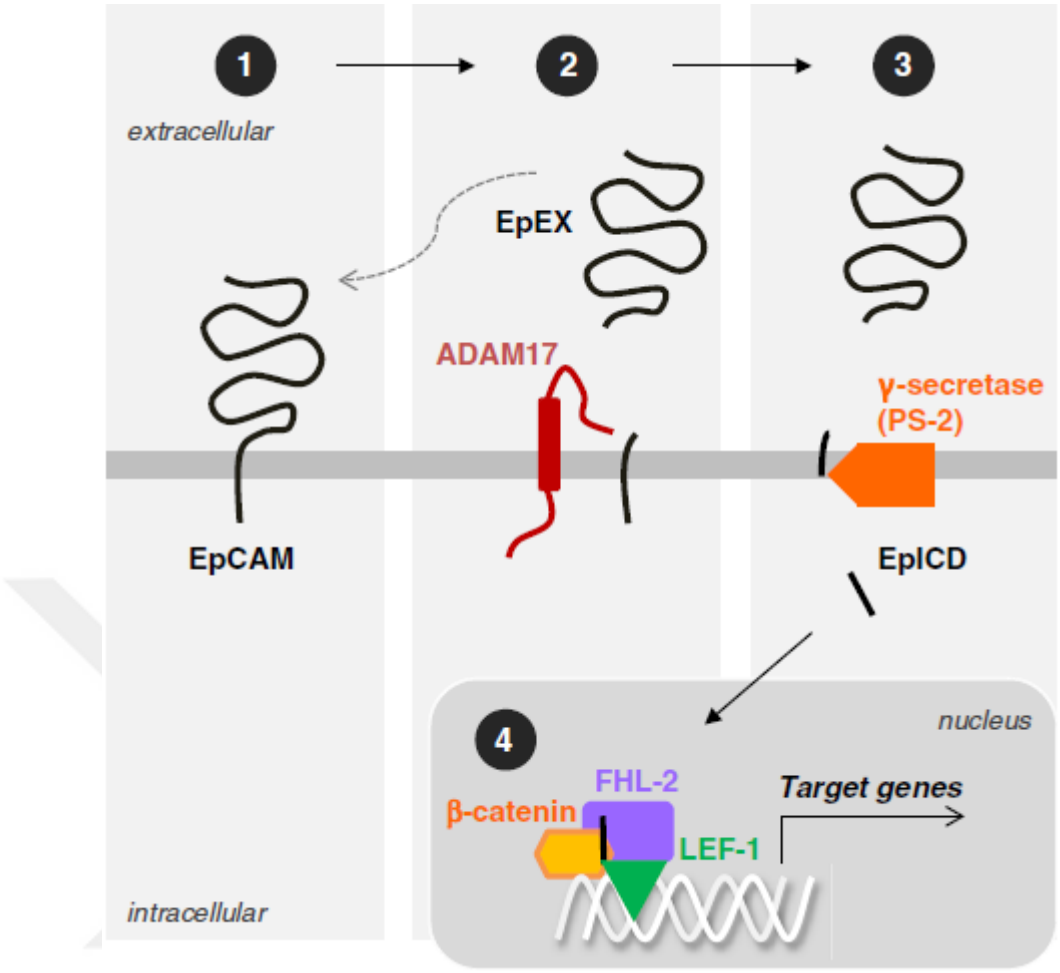

Figure 1.6 The cleavage of EpCAM

1.2.3 EpCAM in Cancer

According to the previous background about the expression of EpCAM in normal and malignant cells and the rate of expression between both them. The overrate expression assists with prognosis in some tumors. The epithelial cadherin was found out that involving with claudin-7 (Rahel, 2014), which supply cell-cell adhesive homotypic, throughout the wnt pathway the boost of different hallmarks occurs (Yu, 2017). EpCAM has a duplicated role as a marker, anti-tumor target with a number of clinical trials (Xiaoa, 2017).

Most of EpCAM studies show that the activation of EpCAM through regulated intramembrane proteolysis (RIP). The initiation of cleaving EpCAM is facilitated via

one of the protease types called ADAM17 (Baeuerle, 2007). The mechanism of ADAM17 protease which associate with EpCAM is to cleave the extracellular domain of EpCAM (EpEX). Another protease that associate with the cleavage of EpCAM is gamma-secretase, its cleavage site within the intracellular domain (EpICD) (Denzel, 2009). The cleavage of extracellular domain occurs between the amino acids Alanine 23 and Glutamic 24 by a signal peptide. The tertiary structure of the extracellular domain has some recent development, also the motifs were identified in this order (Giepmans, 2013):

- Epidermal growth factor (EGF)-like domain
- Thyroglobulin (TY)-like domain
- Cysteine free domain

There are 3 different glycosylation regions within the extracellular domain at Asparagine 74, 111 and 198, the glycosylation of these region included in tumors, leads to the occurrence of 3-fold of the EpCAM average life from seven hours to twenty-one hours. Not only EpCAM is glycosylated in tumorigenesis, but also some receptors such as (CD44, Notch, and integrins), which reflects the importance of EpCAM might be a comparable factor in healthy and non-healthy tissues. In breast cancer, the adhesive capability was decreased because of mutation in glycosylation factor (Liu, 2017). Later after the discovery of epithelial cell adhesion molecules, and the identification of the cleavage between Arginine 80/81, the extracellular domain bounding by a disulfide bond in the TY-like domain, which can be broken down by a reduction condition (Giepmans, 2013).

The cornerstone of EpCAM's role in cancer is the correlation between epithelial cell adhesion molecule from a side and the epithelial cadherins from the cross side. Repeatedly, when EpCAM mediates cell and cell adhesion, this prohibits metastases, which means EpCAM a suppression factor. What occurs if the cleavage is activated by epithelial cadherins? the promotion of metastases will facilitate (Gastl, 2000). The extreme expression of EpCAM in epithelial tumors comparable with the downregulation in normal epithelia is used as a diagnostic tumor marker. Previous

studies demonstrated that EpCAM can serve as an attractive anti-tumor target. It could be recognized by chimeric antigen receptor T cell and RNA aptamer (Figure 1.7).

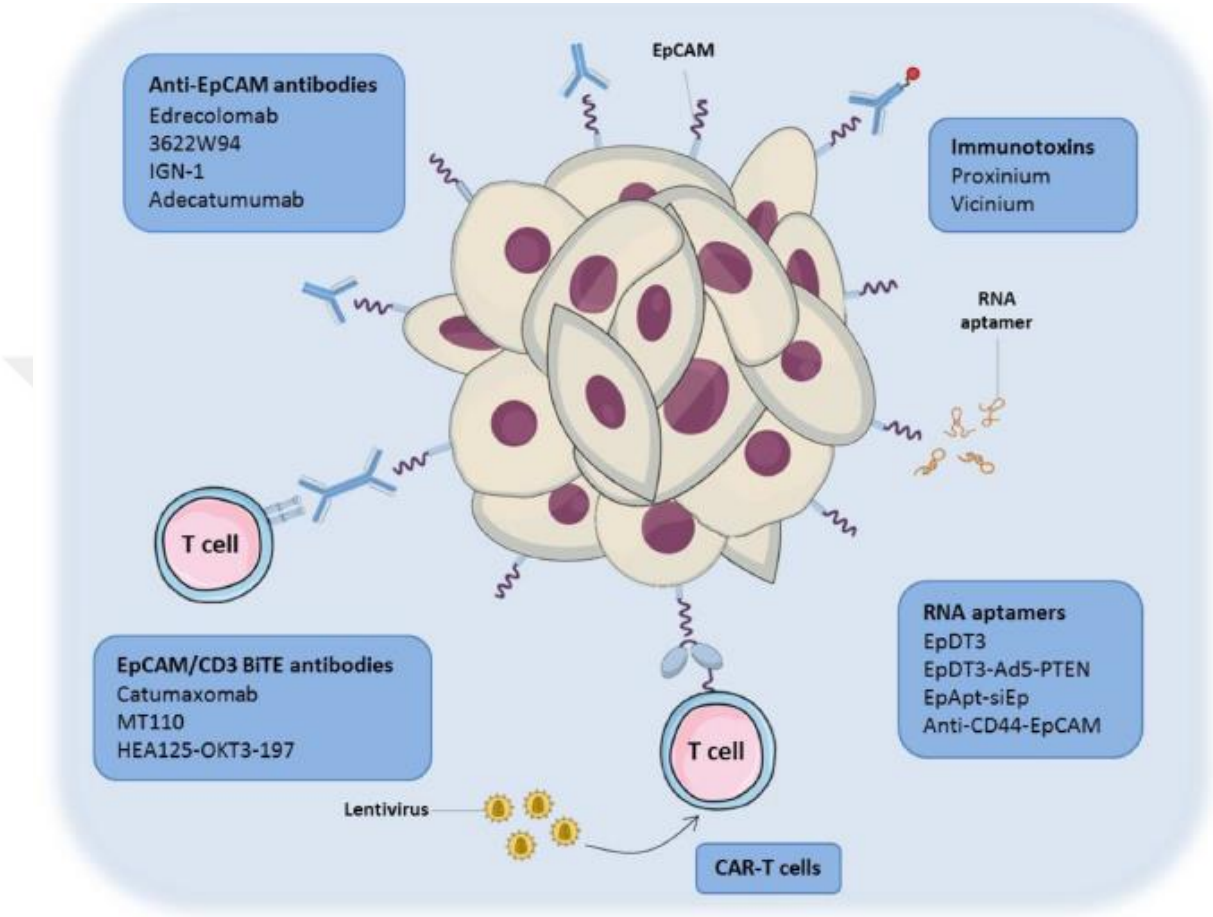

Figure 1.7 Therapeutics strategies that can be associated with epithelioma based on the overexpression of EpCAM (Herreros-Pomares, 2018)

1.3 Recombinant Protein Production and Purification

1.3.1 Production of Proteins from Cloned Genes

In biotechnology, recombinant protein production is being used. The definition of biotechnology in the modern era is the utilization of all biological tools industrially and technologically. In the archaeological archive, the history of biotechnology has been for 4000 years ago, in ancient Egypt, yeast was used to the fermentation bake bread and beer. Sir Alexander Fleming and penicillin discovery as an antibiotic as one

of his attributions that opened the door of the modern biotechnology in the production of antibiotics in both large and low scales (Stanbury, 2017). initially, the growth of microorganisms was carried out in big vessels, the cells have been expelled and purification steps performed to obtain the antibiotic (Figure 1.8) Which called batch culture. That batch culture has been replaced by an ongoing culture called continuous culture, allowing the utilize of fermenter with non-stopping production. Molecular gene cloning is a reason for paying the attention towards biotechnology in the last decades. The uniqueness of productivity by microorganisms such as biopharmaceuticals.

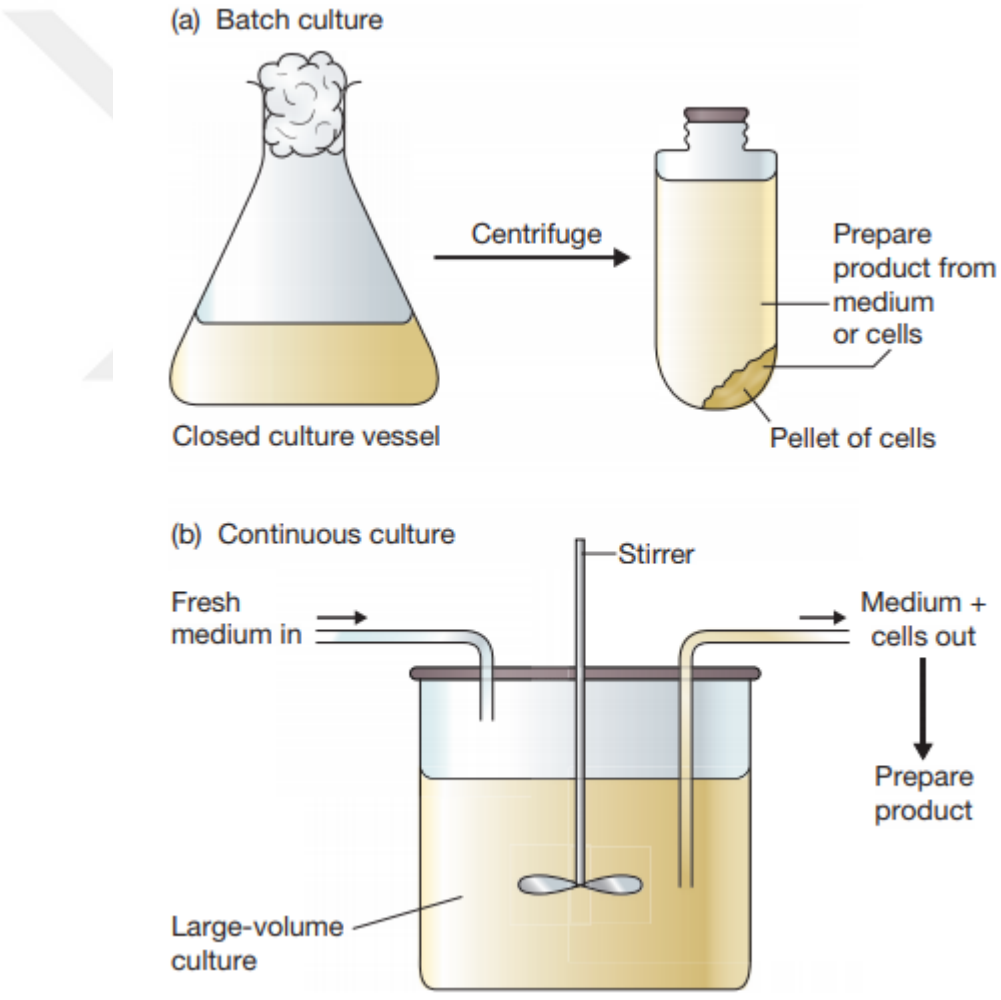

Figure 1.8 Both continuous and batch culture systems (Brown, 2016)

Currently, we have the ability to clone a gene, to produce a protein of a plant or animal, for example, the specific gene can be obtained, inserted into a plasmid vector, and introduced into yeast or bacteria according to the purpose (Figure 1.9). The proper manipulation is crucially required for synthesizing the protein by the bacterial host. The possibility of getting a large amount. The recombinant is not that easy as it seems. There are so many questions related to the required yields of production, the plasmid vector, and the host (Brown, 2016)

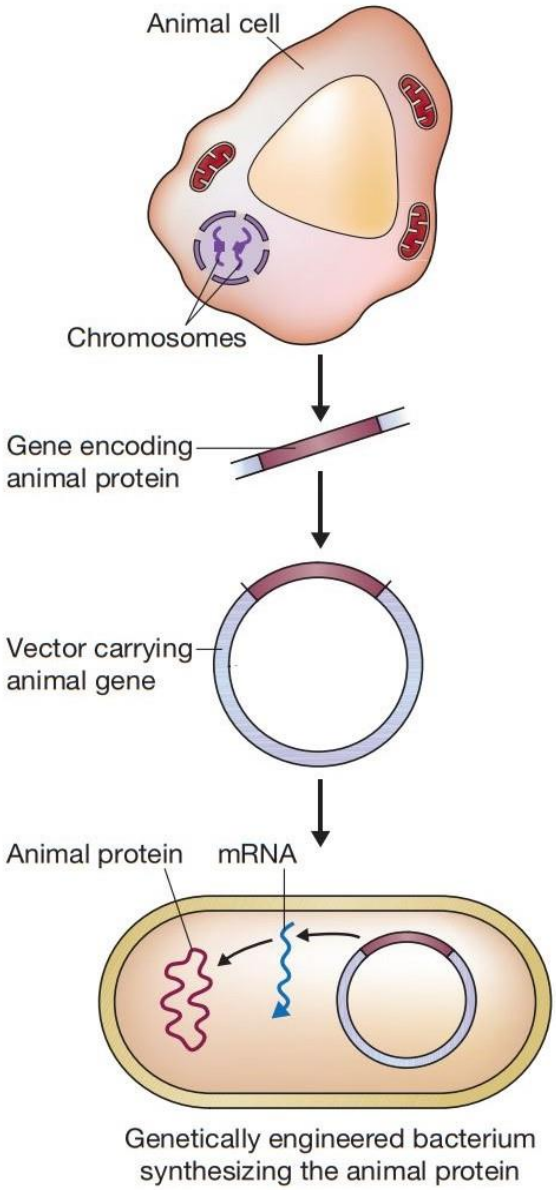

Figure 1.9 An example of Eukaryotic protein produced by bacteria (Brown, 2016)

1.3.2 Plasmid Vectors and E. coli

If there is a eukaryotic gene that inserted with a plasmid and introduced to E. coli, most probably the synthesized amount will be low. The reason is reliable because of the differences between the expression machine in eukaryotic and prokaryotic systems, in another meaning bacteria do not have a post-translational modification.

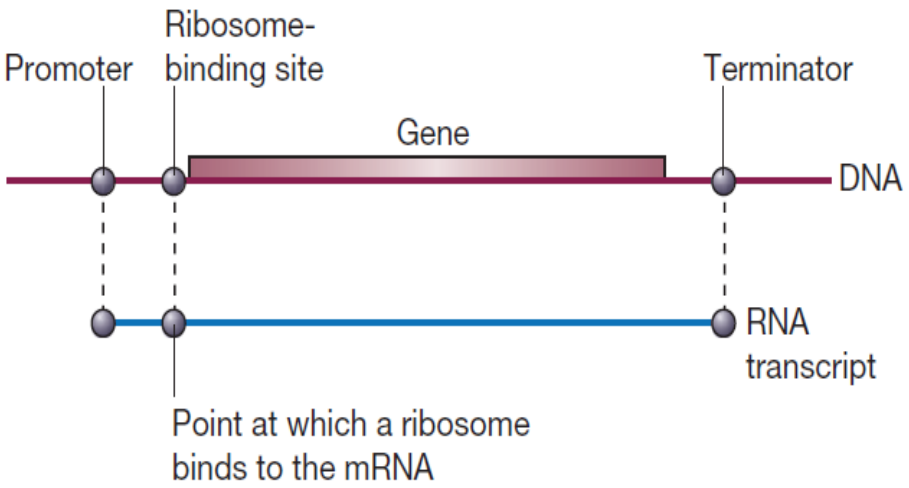

Figure 1.10 The significant regions of E. coli gene

E. coli gene has three significant regions (Figure 1.10):

- The first region is the promoter, it is the starter of the gene, is regulated by the sigma subunit with the control of RNA polymerase.
- The second region the terminator, the transcriptional stop points.
- The ribosome-binding site, the area that mRNA and ribosome bind together. In animals there are also expression signals that surround the genes, but totally different from E. coli ones. We can explain this when comparing the promoters in prokaryotic cell (E. coli) and eukaryotic cell (human) (Figure 1.11). the promotor can be similar in both in some points, but it is far that the RNA polymerase of E. coli can match with the promoter of an animal.

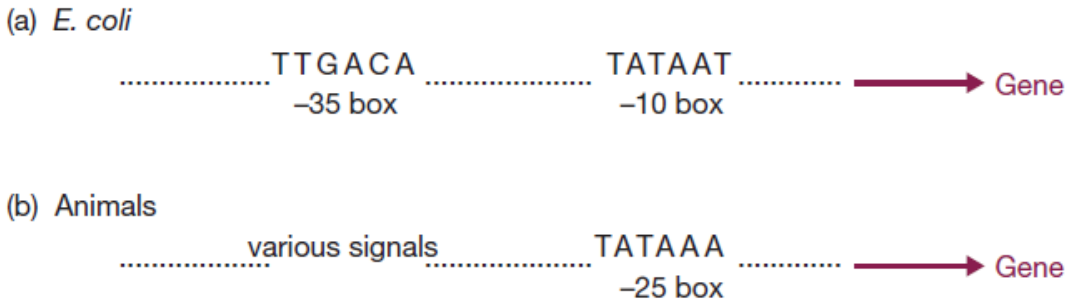

Figure 1.11 Comparison between the promotor in E. coli and animals' genes

This trouble could be solved if we manage to ligate the animal gene into the plasmid in a pattern that will be controlled by *E. coli* expression system. It occurs the gene will have a place in the transcription and translation (Figure 1.12). the vectors that have these characteristics can be used interpreted as expression cloning vectors (Brown, 2016).

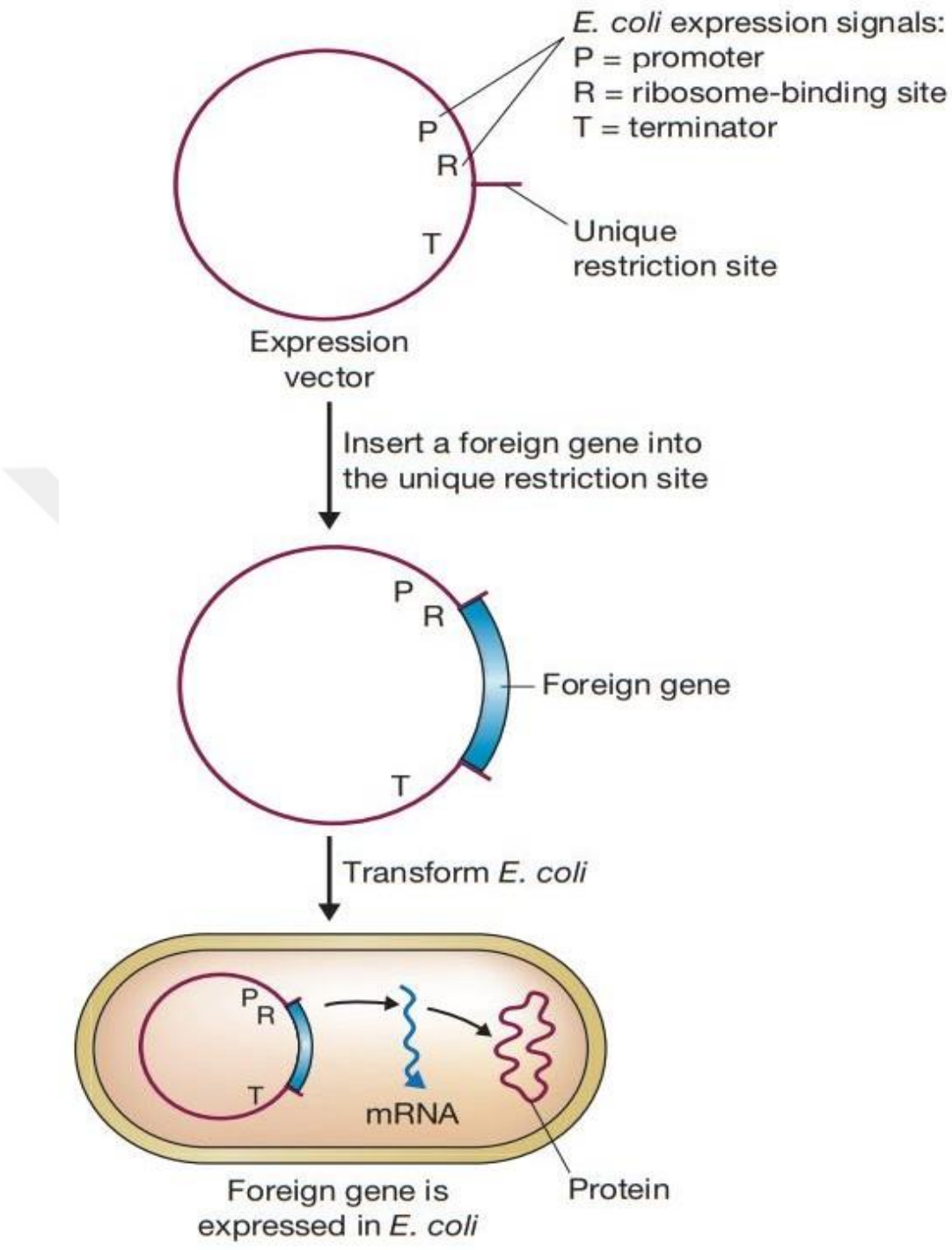

Figure 1.12 The whole recombinant process using *E. coli* with a eukaryotic insert

1.3.3 The Significant Importance of The Promotor as a Part of Bacterial Plasmid

The promoter is considerably the first thing to be looking for while choosing the plasmid expression vector. Due to its controls in the initial stages while RNA polymerase attaches to the DNA, indicating the mRNA synthesized rate. The type of the promoter determines the final amount the protein production, thus, the selection of the promoter should be very careful. There are two types of promotors; strong promoters, bearable the overrate of transcription, in addition to controlling the required gene for high production. On the other hand, weak promoters, lower efficiency, works in a low amount of expression (Figure 1.13). obviously, the plasmid vector must have a strong promoter, which will be resulted in extreme rate with the transcription of the clone.

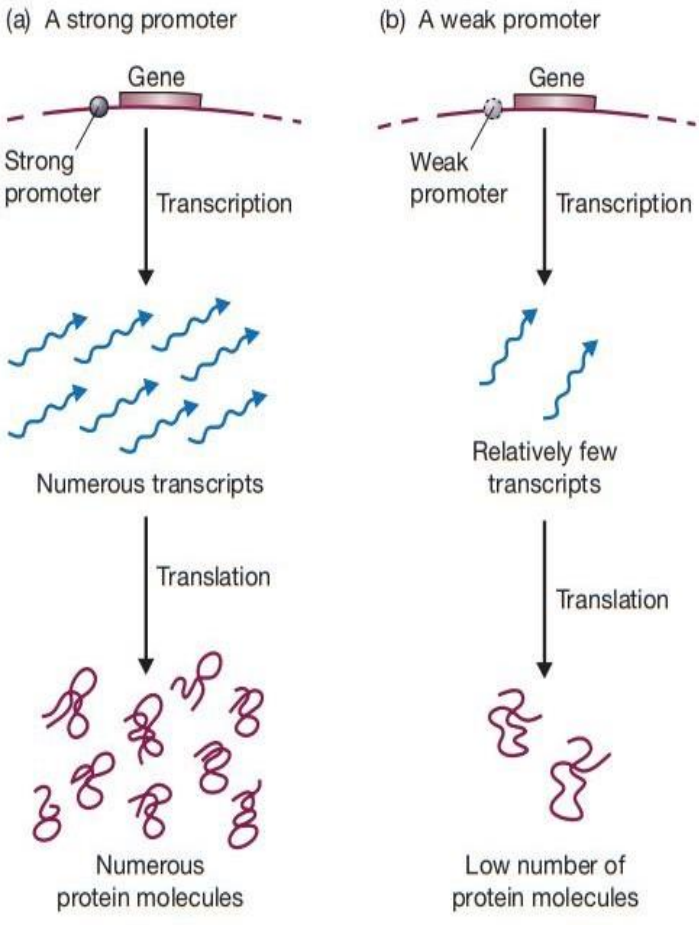

Figure 1.13 Strong and weak promoters (Brown, 2016*)*

Gene regulation in *E. coli* is taken over by switching on (induction) or switching off (repression). therefore, it is essential to add some compound to the bacterial medium. This compound is a substrate for a specific enzyme in the inducible gene (Figure 1.14). The induction and repression sequences are located on the map of the vector, In most cases close to the promoter (Rosano, 2014).

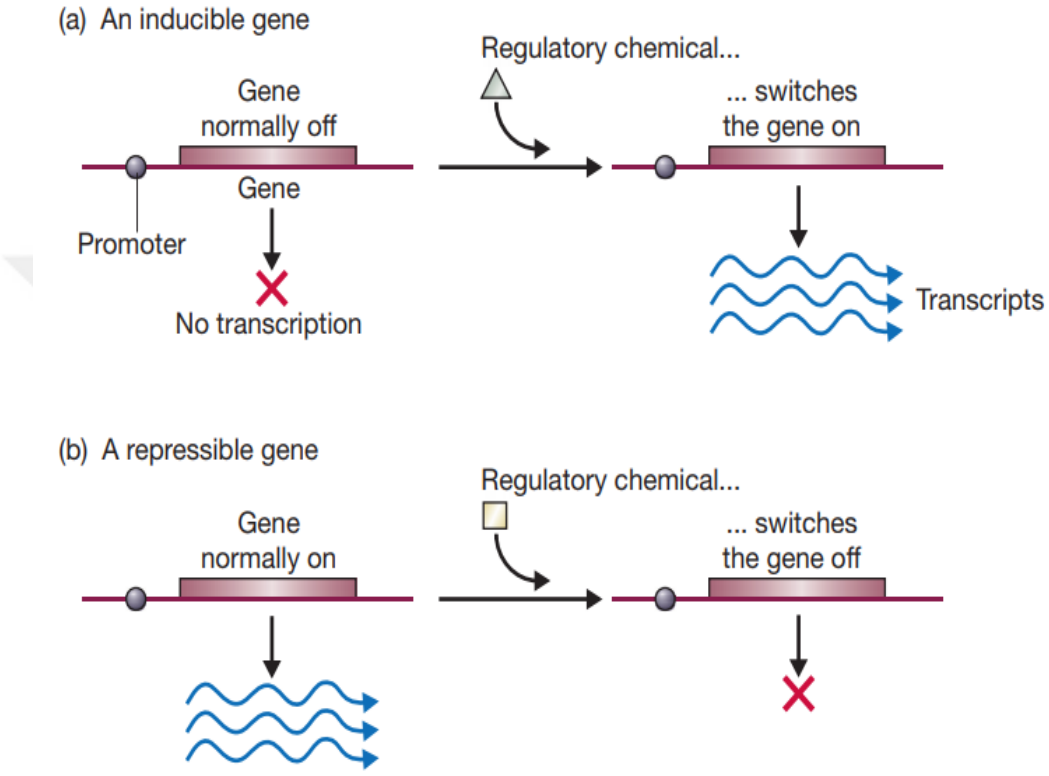

Figure 1.14 A scheme of the induction and repression process (Brown, 2016)

Promoters in *E. coli* sometimes have a combination of all needed characteristics for regulation. In the following lines we will summarize some of the promoters in the plasmid vectors:

- As mentioned before the promoter is the key for the transcription process when RNA polymerase binds, there is a type called Lac promoter, where the LacZ is controlled. And LacZ gene codes for beta-galactosidase. If we need to induce the Lac promoter, isopropyl thiogalactoside (IPTG) should be added to the medium to switch on the induction.
- The second sort of promoters called trp promoter, trp refers to the amino acid tryptophan and this promoter includes in the genes code for some

enzymes. The induction of trp promoter by 3-Beta-indoleacrylic acid, and repression by tryptophan.

- A third sort is tac promoter, it is a combination between the last two promoters both lac and trp. From its characteristics is the strength is higher, but the induction by isopropyl thiogalactoside (IPTG).
- The fourth sort of promoters is λPL promoter very specific for one type of DNA called λ DNA. Highly strong and can be recognized by RNA polymerase.
- The fifth sort of promoter is the most renowned one T7 bacteriophage which called T7 promoter. The induction can be carried out by IPTG to the medium.

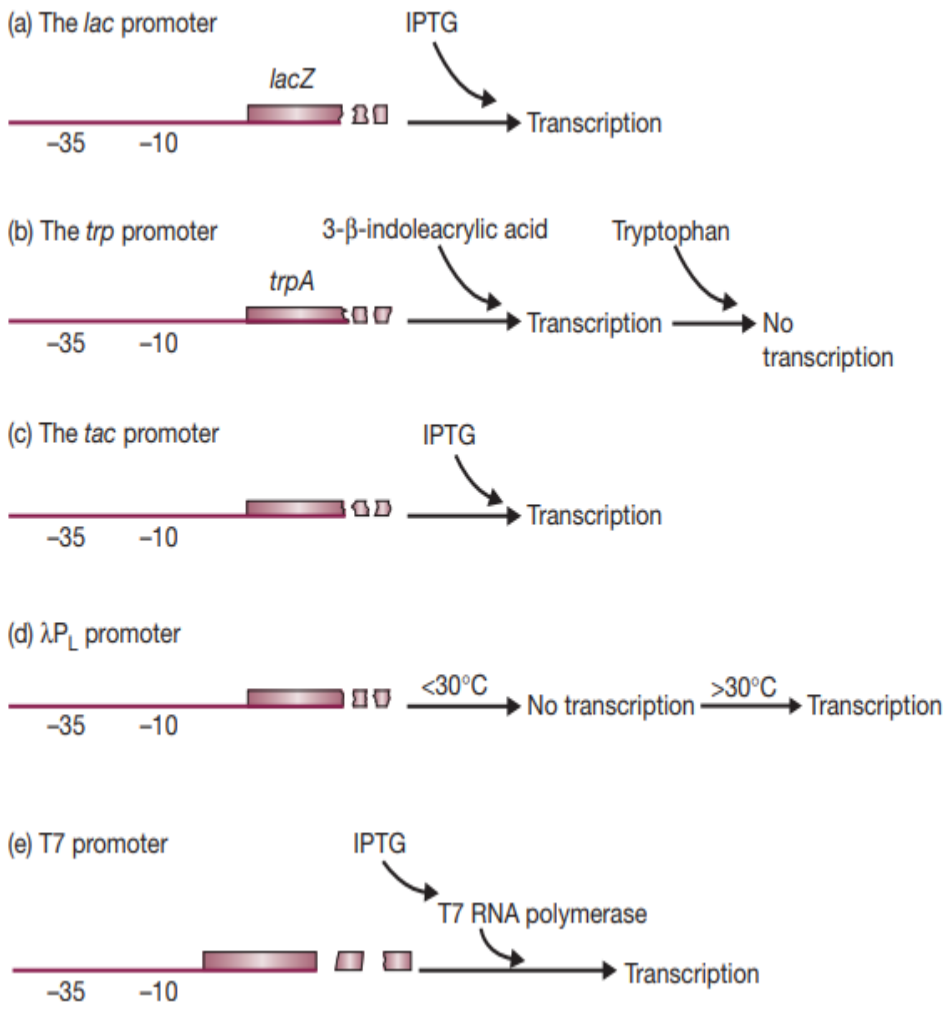

Figure 1.15 Promoters' types with the way of induction

1.3.4 Common Troubles in Recombinant Protein Production

Regardless of the uniqueness of the plasmid vectors, we still face difficulties served with the protein production using E. coli. these difficulties can be classified into two troubles: (1) troubles that are because of insertion and its sequence; and (2) troubles because of the E coli machinery itself. In the next paragraph, we will reveal some of these troubles with how to avoid or to solve them.

(1) Troubles that are because of insertion and its sequence, there are several points that do not give an efficiency within the expression, will be briefly summarized:

- If the insertion concludes any introns, it will be the main trouble, and *E. coli* genes do not have such introns and thus the bacteria will not be able to splice those introns (Figure 1.16)
- In some cases, the insertion gene may have a sequence that its behavior like a terminator maker in the *E. coli* (Figure 1.16). these sequences are absolutely nontoxic in the host, but toxic in *E. coli*, this will lead at the end to wastage of the expression.
- Codon bias emerges when there is a difference between the same or equivalent codons in the insert and the host as shown (Figure 1.16), also the bias specifically occurs with each organism against favorite codons.

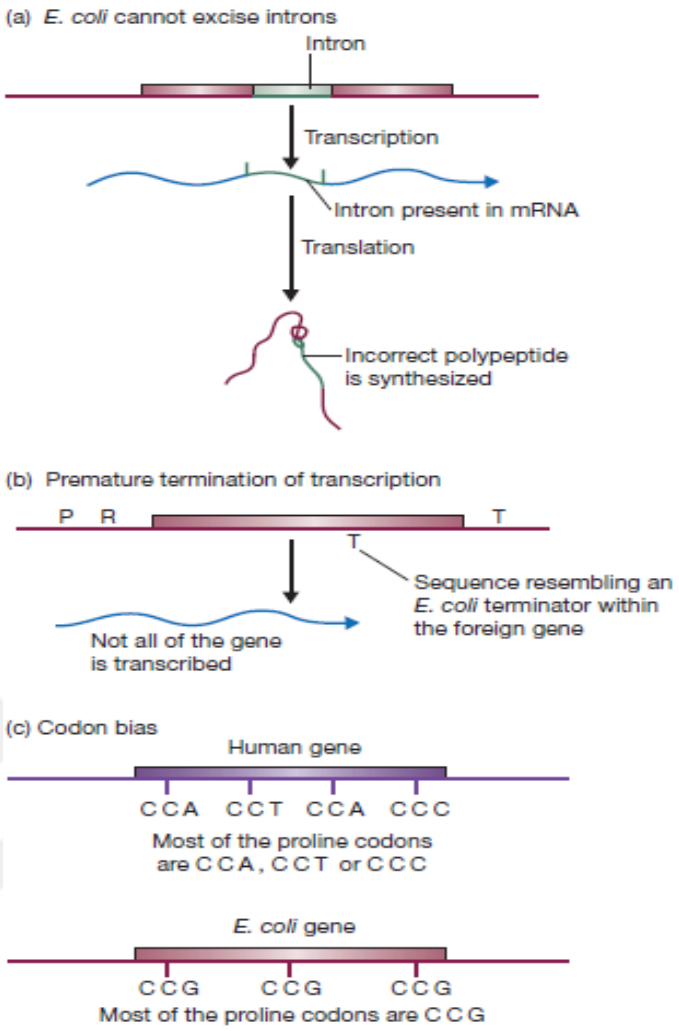

Figure 1.16 Three of the problems that could be encountered when an insertion expresses within E. coli (Brown, 2016)

(2) troubles because of the *E. coli* machinery itself. *E. coli* may make mistakes while recombinant proteins. In eukaryotic protein is subsequent of the translation process after a significant modification that happens to the amino acids, usually called post-translational modifications to provide each protein with specific function cellularly or genetically, some examples of the post-translational modifications (PTM) such as, methylation, lipidation, glycosylation, and phosphorylation. glycosylated proteins that acquire a carbohydrate group to be attached post-translation. In bacteria glycosylated proteins in most cases have a glycosylation error. Folding of the protein may not also be correct within E. coli, and if the proteins do not reach to tertiary structure folding, it goes to be insoluble protein creating inclusion bodies inside the

bacteria (Figure 1.17). it is not a big deal to recover the protein from inclusion bodies, but the issue is having the correct folding and the activity of the protein as desired.

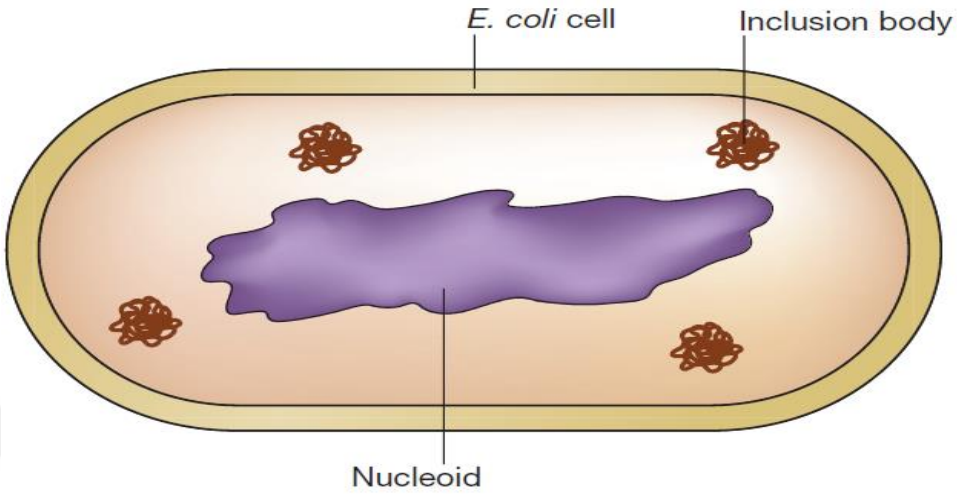

Figure 1.17 Inclusion bodies (Brown, 2016)

These drawbacks are not difficult to solve than the sequence problems mentioned in the last part but can be alleviated to a certain extent by using special *E. coli* strains. If the mutant strain of *E. coli* the degradation of proteins can be eliminated. Selecting the host strain is crucial. The main drawback, however, is nonattendance of glycosylation. Attempts have been made to solve this problem with *E. coli* strains that contain cloned genes for enzymes that carry out glycosylation in other organisms. These include *Campylobacter jejuni*, one of the few bacteria that have any glycosylation activity. So far, however, this approach has had limited success, and *E. coli* is generally looked on as only convenient for the production of those eukaryotic proteins that do not necessary to be glycosylated (Brown, 2016).

CHAPTER TWO
MATERIAL AND METHODS

2.1 Cell Lines and Culturing

The selected cell lines were ovarian (OVCAR and A2780cis) were grown in 90% RPMI medium which supplemented with 10% fetal bovine serum (FBS) and 1% penicillin at 37° C and 5% CO2 in a humidified incubator. Hemocytometer was used as a counting method for both cell lines OVCAR ($10x10^6$ cells) and A2780cis ($5x10^6$).

2.2 RNA Isolation

TRIzoL™ reagent was purchased from ThermoFisher Scientific for isolating a high-quality total RNA from cells. TRIzoL™ reagent is a monophasic solution of phenol, guanidine isothiocyanate which facilitates the isolation of a variety of RNA species of large or small molecular size.

2.2.1 Required Materials for RNA Isolation

Cells were pelleted by centrifugation, then the pellet was homogenized with 0.75 mL TRIzoL™ reagent, 0.2 mL chloroform was added to separate the homogenate into three phases the upper aqueous layer is RNA, the rest contains DNA and proteins (Figure 2.1). 0.5 mL of isopropanol was added to precipitate the RNA and centrifuged 10 minutes at 12,000 x g at 4° C, the total RNA precipitated as a white gel-like pellet at the bottom of the tube. Pellet was washed with 75% ethanol.

2.2.2 Determination of Isolated RNA Yield

The RNA concentration was measured using NanoDrop™ Spectrophotometer at the absorbance of 260 nm for total nucleic acid content and 280 nm to determine the sample purity. Note. The sample was diluted in RNase-free water before the measurement.

2.3 cDNA Synthesis

Applied Biosystems^{TM} High-Capacity cDNA Reverse Transcription kit was used including random primers for the starting of cDNA synthesis.

2.3.1 Preparation of 2X RT Master Mix

The condition was optimized as shown in the (Table 2.1).

Table 2.1 The 2X Master mix

Component	**Volume**
10X RT Buffer	4.0 µL
25X dNTP mix	2.0 µL
10X RT Random Primers	4.0 µL
Reverse Transcriptase	2.0 µL
RNase Inhibitor	2.0 µL
Nuclease-free H2O	7.0 µL
Total per reaction	20.0 µL

2.3.2 Preparation of Reverse Transcription Reaction

10 µL of 2X RT master mix was pipetted and added into individual PCR tubes, then 10 µL of OVCAR isolated RNA and 10 µL of A2780cis isolated RNA were added.

2.3.3 Thermal Cycling Condition

The condition was optimized as shown in the (Table 2.2).

Table 2.2 Thermal cycling condition

settings	Step 1	Step 2	Step 3	Step4
Temp.	25° C	37° C	85° C	4° C
Time	10 minutes	120 minutes	5 minutes	∞

2.4 Polymerase Chain Reaction (PCR)

2.4.1 EpCAM Primer Designing and Preparation

CLC Main Workbench bioinformatic tool was utilized to design the primers (Figure 2.1 and 2.2). The concentration of the primer was prepared by dissolving the powder primer in suitable volume DNase free water (Table 2.3). primers were ordered from Sentegen and their size ranges approximately between 17-35 nucleotides.

Table 2.3 The EpCAM primers with the final concentration of nucleic acids

Name	**5'**	**Sequence**	**3'**	**bp**
FWD.1		CGGCGACGGCGACTTTT		17
REV.1		CAGCAACAACTGCTATCACCAC		22
FWD.2		CGGCGACTTTTGCCGC		16
REV.2		GTTCCCTATGCATCTCACCCAT		22
FWD.3		GATCCTCGAGCAGGAAGAATGTGTCTG		27
REV.3		GATCCTCGAGCTATTTTAGACCCTGCATTGA		31

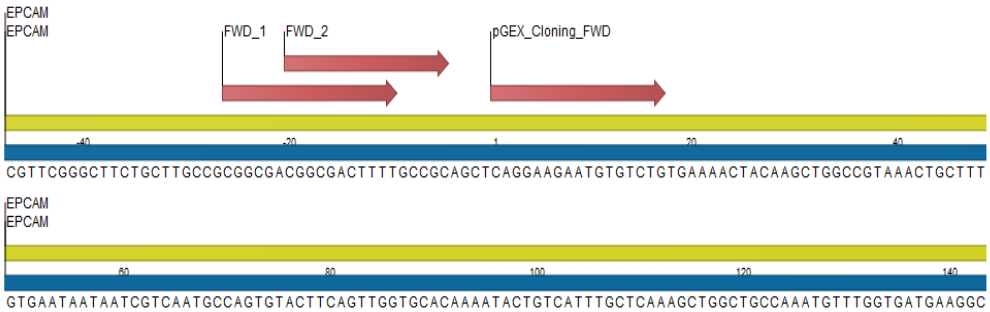

Figure 2.1 EpEX forward primers

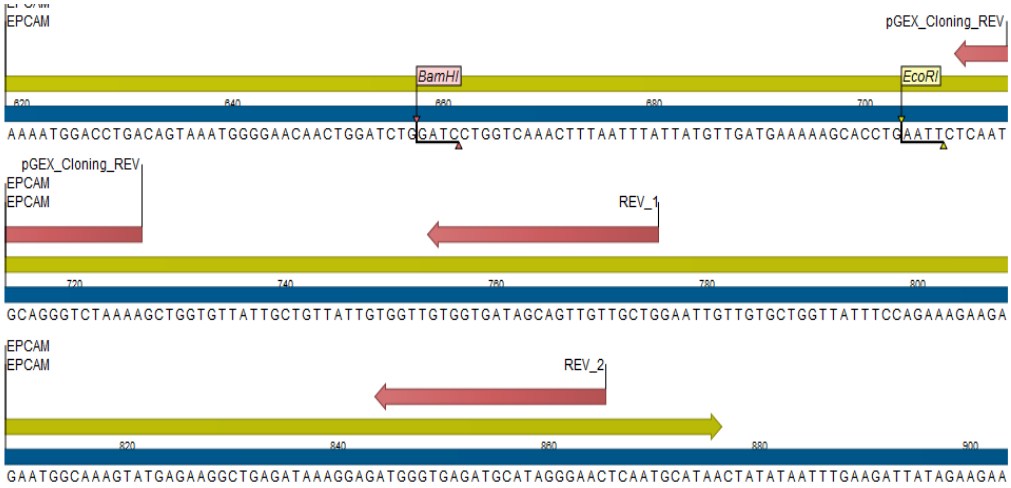

Figure 2.2 EpEX reverse primers

2.4.2 PCR Reaction Setup and Thermocycling Conditions

For molecular cloning applications which require high-fidelity PCR, Phusion high-fidelity DNA polymerase was used. All the reaction components were assembled on ice and quickly transferred to thermocycler preheated to the denaturation temperature (98°C). All components were mixed and centrifuged prior to use. Phusion DNA Polymerase was added the last in order to avoid any primer degradation caused by 3´→ 5´ exonuclease activity.

Two PCR reactions were carried out, the first reaction (FWD.1, FWD.2, REV.1, and REV.2 primers) were used in order with the template which obtained from cDNA synthesis reaction in the previous experiment (Table 2.5). The second reaction (FWD.3, and REV.3 primers) were used with amplification of the first reaction after purifying using (Gene jet purification kit).

Table 2.4 A typical PCR Reaction mixture

Component	**50 µL Reaction**	**5X50 Reaction**
Nuclease-free Water	Up to 50 µL (32.5 µL)	Up to 250 µL (162.5 µL)
5X Phusion HF or GC buffer	10 µL	50 µL
10 mM dNTPs	1 µL	5 µL
10 µM FWD Primer	2.5 µL	12.5 µL
10 µM REV Primer	2.5 µL	12.5 µL
Template DNA	1 µL	5 µL
Phusion DNA Polymerase	0.5 µL	2.5 µL

Table 2.5 Templates and the order of the primers for the first reaction

Templates	tubes	Primers order
A2780cis	1	FWD.1/REV.1
A2780cis	2	FWD.2/REV.2
OVCAR	3	FWD.1/REV.1
OVCAR	4	FWD.2/REV.2

2.4.3 Thermocycling Condition for Both the First Reaction and the Second Reaction

Thermocycle condition for both first and second reactions explained in the (Table 2.6 and 2.7).

Table 2.6 A typical PCR cycle for the first reaction

Step	Temp	Time
Initial Denaturation	98°C	30 sec
35 X	98°C 65°C 72°C	10 sec 30 sec 1 min
Final Extension	72°C	3 min
Hold	10°C	∞

Table 2.7 A typical PCR cycle for the second reaction. First five cycles (annealing temperature 51°C, while the last thirty cycles 72°C

Step	Temp	Time
Initial Denaturation	98°C	30 sec
5 X	98°C 51°C 72°C	10 sec 30 sec 1 min
35 X	98°C 72°C	10 sec 1 min
Final Extension	72°C	3 min
Hold	10°C	∞

2.5 Agarose Gel Electrophoresis (1 %) Preparations

For the preparation of agarose gel: (60 mL 1x TAE buffer, 0.6 mg agarose, 3.6 µL safe view). To prepare 1x TAE buffer from 50x TAE stock (Table 2.8), 20 mL of the stock was diluted into 980 mL of ddH2O.

Table 2.8 One-liter 50x TAE buffer

Tris-base	242 g
Acetate (100% acetic acid)	57.1 mL
EDTA	100 mL
ddH2O	Up to one liter

2.6 PCR Product Purification

Thermo Scientific GeneJET PCR Purification Kit was used to purify after each PCR reaction. PCR EpEX amplified product should be purified to be used for downstream applications restriction digestion and molecular cloning. The used protocol was performed this way; Binding buffer was added 1:1 volume to PCR products (45µL from each), centrifuge step and the flow was discarded, then Wash buffer 700 µL was added, centrifuge step 2 times for 1 min, then the flow was discarded, ultimately, Elution by ddH2O shown (Figure 2.3).

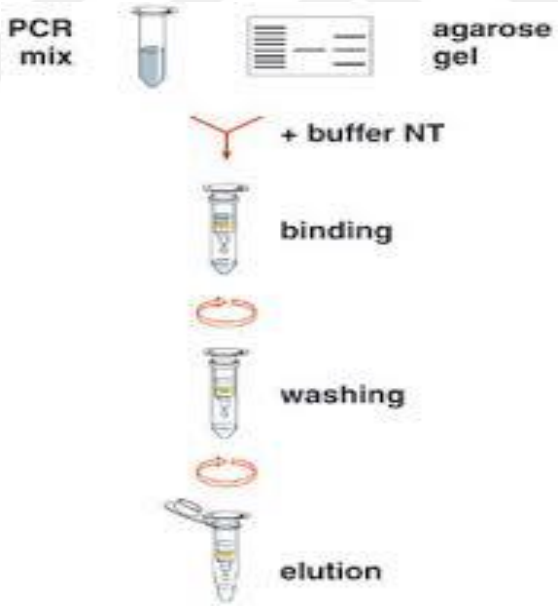

Figure 2.3 PCR products purification steps

2.7 Restriction Digestion of EpEx PCR

For restricting digestion, a master mixture in (Table 2.9) was prepared and incubated at 37 °C for 2 hours. XhoI cut mechanism.

5' C↓ T C G A G 3'
3' G A G C T↑ C 5'

Table 2.9 Restriction digest of EpEX PCR product

EpEx PCR Product	25 µL
10 X FastDigest Buffer	3 µL
XhoI	1 µL
ddH2O	1 µL
Total volume	30 µL

2.8 Ligation

2.8.1 pGEX-6P-1

The pGEX-6P-1 plasmid vector (Figure 2.4) was restriction digested with XhoI overnight prior to a ligation step. XhoI Restriction site cuts at 969→970, and complement 973→974.

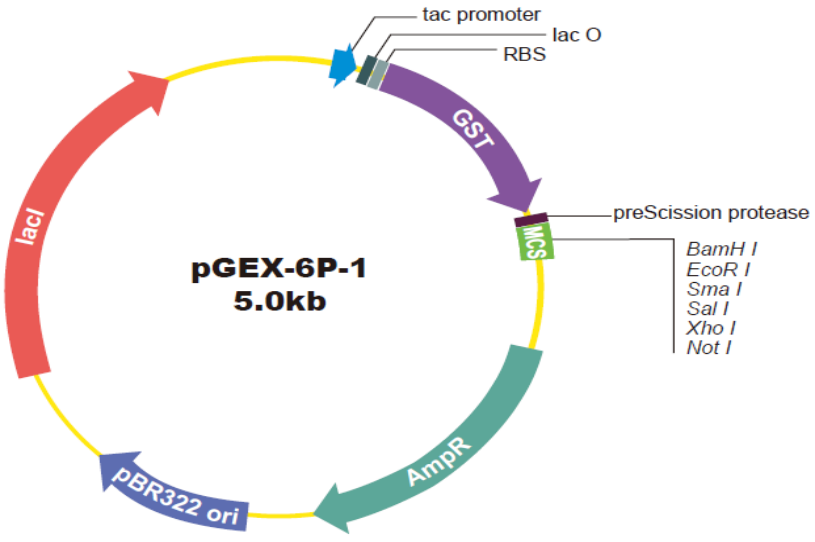

Figure 2.4 pGEX-6P-1 plasmid map

2.8.2 Ligation Reaction

Before ligation, both plasmid and the inserted DNA (EpEX) were digested. To avoid self-ligation, plasmid ends needed dephosphorylation by utilizing 5 units of

alkaline phosphatase at 37°C for 15 mins. To achieve the ligation reaction, 60 ng plasmid and 1X equimolar ratio of the insertion (EpEX) were utilized for Ligation of the DNA at 22°C for 1 hour (Table 2.10).

Table 2.10 A typical ligation reaction

Vector DNA	1.5 µL (60 ng)
Insert DNA (EpEX)	2.5 µL (20 ng)
Ligase Buffer 10X NEB	2 µL
T4 DNA Ligase	0.5 µL (2.5 units)
ddH2O	13.5 µL
Total volume	20 µL

2.9 Transformation of The Competent *E. coli*

2.9.1 Bacterial Strains and Media

Table 2.11 Bacterial strains

BL21(DE3) *E.Coli* cells Invitrogen
Top10 *E.Coli* cells Invitrogen

Table 2.12 Media

LB medium, 1liter	10g tryptone, 5g yeast extract, 5gNaCl, 1ml 1N NaOH (autoclaved)
LB plates, 1 liter	10g tryptone, 5g yeast extract, 5g NaCl, 15g agar (autoclave in a flask, cool down to 40- 45°C, add antibiotics and pour into Petri dishes.

2.9.2 Transformation

Chemical Competent Bacterial Stocks the following steps were carried out:

1. For the preparation of chemically competent TOP10 E. coli cells, 50ml of LB medium is inoculated with a bacterial colony grown overnight on an LB plate.
2. Cells are grown at 37°C by shaking 180rpm until OD600 value reaches 0.4. (3-6 hours)
3. Cells are spun down in a 50ml conical tube 15 min 3000rpm at 4˚C.
4. Pellet is resuspended in 10ml sterile ice-cold 0.2µm filtered 0.1M CaCl2 solution and cells are kept on ice for 15 min to 4 h.

5. Cells are spun down as before and the pellet is resuspended in 2ml CaCl2. Sterile glycerol is added to an end concentration of 10%.
6. Cells are shock frozen by plunging 50- 100µl aliquots in liquid nitrogen.
7. Aliquots are stored at –80°C until use.
8. The last aliquot is used to prepare new stocks.
9. Competent cells are thawed on ice and mixed with 3µL ligation reaction and that step Before the heat shock of the transformation.
10. Cells are incubated on ice for 30 min.
11. The heat shock is performed by incubating the cells at 42°C water bath for 60 seconds.
12. LB plates with ampicillin- containing were prepared and cells were spread on.

2.9.3 Colony PCR

To screen the colonies for the presence of inserted gene colony PCR was required to accomplish that. Primers were designed forward from the vector and reverse from EpEX gene to detect the presence of the inserted EpEX (Figures 2.5 and 2.6), PCR reaction as shown in (Table 2.13), and Thermocycling condition of colony PCR (Table 2.14), ultimately, first eight pools products were used to run on agarose gel 1 %.

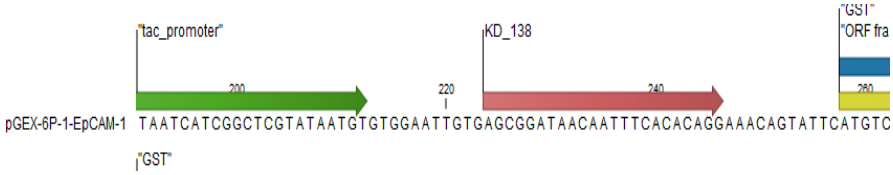

Figure 2.5 Forward primer (KD_138)

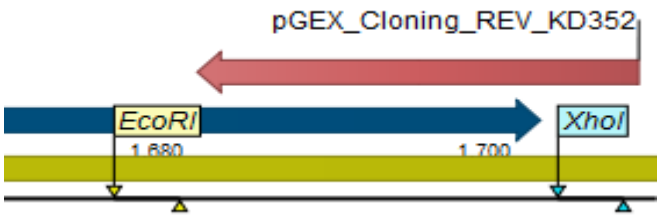

Figure 2.6 Reverse primer (KD_352)

Table 2.13 Colony PCR reaction

Component	20 µL Reaction	10X20 Reaction
Nuclease-free Water	Up to 20 µL (16.65 µL)	Up to 200 µL (166.5 µL)
Template	0.5 µL	5 µL
10 mM dNTPs	1 µL	5 µL
FWD Primer (KD_138)	0.2 µL	2 µL
REV Primer (352)	0.2 µL	2 µL
10X AMP Buffer	2 µL	20 µL
AMP Taq	0.05 µL	0.5 µL

Table 2.14 Thermocycling condition of colony PCR

Step	Temp	Time
Initial Denaturation	95°C	30 sec
35 X	95°C 50°C 72°C	15 sec 30 sec 1.5 min
Final Extension	72°C	3 min
Hold	10°C	∞

2.10 Expression and Purification of Recombinant EpEX

For bacterial expression of recombinant EpEX, genes cloned into pGEX-6P-1 vector was used. This vector introduces a GST tag to the cloned DNA (Figure 2.7), one by one. EpEx protein was expressed in BL21 strain *E. coli* cells.

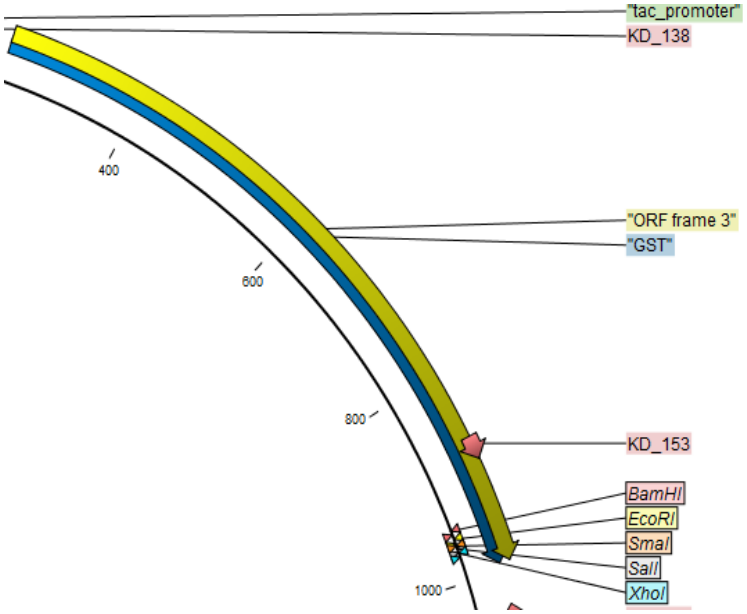

Figure 2.7 "GST" region on pGEX-6P-1 vector yellow color

2.10.1 Attempts EpEX Affinity Purification

10 mL LB-Ampicillin starter culture was inoculated with the glycerol stock and grown overnight. starter culture was used to inoculate 250 mL LB-Ampicillin media and grown at 200 rpm shaking at 73°C for 1-2 hours. OD 600 value was monitored until it reached 0.6-0.8 Log phase of growth.

Protein expression was induced by adding 200 µM final concentration of IPTG and cells was grown further 4 hours at 30°C with 160 rpm shaking. Cells were pelleted by centrifugation at max speed in Eppendorf cooled centrifuge. Cells were resuspended in 4 mL PBS, 1 mM EDTA by pipetting or shaking on ice in 5 mL tubes. Lysozyme was added to 100 µg/mL final concentration (4µL lysozyme), mixed and incubated on ice 20 minutes. Sonication of the cells was carried out 3 minutes total time at 50 amplitude, with 15 seconds active cycle and 30 seconds off intervals. 10 units of piercing universal nuclease, 2mM final concentration of $MgCl_2$ and 1 % TritonX-100 were added and then incubated on rotating wheel for 30 minutes. The lysate was spun down with maximum speed to clarify 25000g for 30 minutes, the 100µL lysate was taken as aliquot. 50µL bed Glutathione beads and 2 mM EDTA was added into the cleared lysate then incubated in cold room for 2 hours on a rotating wheel, and spun down 5 minutes at 1000 rpm. 3 times washing was performed 10 minutes with PBS, 1mM EDTA, 1% TritonX-100. One-time wash like the previous step but without TritonX-100. Total of 100 µL was resuspended and 10 µL aliquot was taken. Finally, the bound proteins were eluted using 100 µL elution buffer (Table 2.15) at the room temperature for 1 hour on the rotating wheel, then a post elution aliquot of 10 µL was taken.

In the second attempt of EpEx affinity purification before adding the lysozyme cells were resuspended in 4 mL binding buffer (50mM Tris-Cl pH 8.0, 100 mM NaCl, 1 mM PMSF, 1 mM DTT) by vortexing.

Table 2.15 Elution buffer was used in the first attempt

10X PBS	30 µL
100 mM Glutathione	75 µL
100 mM DTT	3 µL
ddH2O	152 µL
Total	300 µL

The steps of the first trial:

1. Inoculate a 2ml LB-Ampicillin starter cultures with the glycerol stock and grow overnight.
2. In the morning, use 2ml of the starter culture for inoculating 100 ml LB-Amp media and grow at 200 rpm shaking at 37˚C for 1-2 hours. Monitor the OD600 value until it reaches the 0.6-0.8 Log phase growth.
3. Induce protein expression by adding 200 µM final concentration of IPTG and grow cells further 4 hours at 30˚C with 160rpm shaking.
4. Cells were pelleted using a centrifuge at max speed in Eppendorf cooled centrifuge.
5. Resuspend cells in 4 ml PBS, 1mM EDTA by pipetting or shaking on ice in 5 ml tubes.
6. Add lysozyme to 100ug/ml final concentration. Mix and incubate on ice at least 20 minutes.
7. Sonicate using tip MS 1.5, at 50% amplitude for a total of 2 minutes, such as 15 seconds finger controller sonication and chill on ice 30 seconds. Insert the tip of the probe about 2cm into the bacterial mix. Keep the tube in ice during sonication.
8. Add 10 Units of Pierce Universal Nuclease, 2mM final concentration of MgCl2 and 1% TritonX-100. Lyse in cold room by incubating on rotating wheel for 30minutes.
9. Spin down the lysate maximum speed to clarify. 25000g for 30 minutes. Take 100 ul of the lysate as aliquot.
10. Add 50 ul bed volume Glutathione beads into the cleared lysate. Add to additional 2mM EDTA. Incubate in cold room/fridge for 2hours on a rotating wheel.
11. Spin down 5 minutes at 1000rpm.

12. Wash 3 times 10 minutes with PBS, 1mM EDTA, 1% TritonX-100.
13. Wash once with PBS, 1mM EDTA (no triton). Resuspend in a total of 100ul and take 10ul aliquot.
14. Elute the bound proteins using 100 ul elution buffer [PBS, 25mM reduced glutathione, 1mM DTT, pH 7.5-8.0]. Elute at room temperature for 1hour on a rotating wheel. Take a post elution aliquot of 10 ul.

The steps of the second purification trial:

1. Inoculate a 10ml LB-Ampicillin starter cultures with BL21 cell colonies on LB plates and grow overnight.
2. In the morning, use 10ml of the starter culture to inoculate 250 ml LB-Amp media and grow at 200 rpm shaking at 37˚C for 1-2 hours. Monitor the OD600 value until it reaches the 0.6-0.8 Log phase growth.
3. Induce protein expression by adding 100 µM final concentration of IPTG and grow cells further 4 hours at 30˚C with 160rpm shaking.
4. Pellet the cells by centrifugation at max speed in Eppendorf cooled centrifuge. Freeze pellet at -20. Thawed washed pellet with 10 ml PBS. Refrozen.
5. Resuspend cells in4 ml of binding buffer (50 mM TrisCl pH 8.0, 100mM NaCl, 1mM PMSF, 1 mM DTT) by vertexing. Transfer into a 5 ml microfuge tube.
6. Add lysozyme to 100ug/ml final concentration. Mix and incubate on ice at least 20 minutes.
7. Sonicate using tip MS 1.5, at 50% amplitude for a total of 2 minutes, such as 15 seconds finger controller sonication and chill on ice 30 seconds. Insert the tip of the probe about 2cm into the bacterial mix. Keep the tube in ice during sonication.
8. Add 10 Units of Pierce Universal Nuclease, 1mM final concentration of MgCl2 and 1% TritonX-100. Lyse in cold room by incubating on rotating wheel for 30minutes.
9. Spin down the lysate maximum speed to clarify. 25000g for 30 minutes. Take 100 ul of the lysate as aliquot.

10. Add 100 ul bed volume Glutathione beads into the cleared lysate. Add to additional 1mM EDTA. Incubate in cold room/fridge for 2hours on a rotating wheel.
11. Spin down 5 minutes at 1000rpm.
12. Wash 3 times 10 minutes with 5ml binding buffer + 1% TritonX-100.
13. Wash once with binding buffer (no triton). Resuspend in a total of 200ul and take 10ul aliquot.
14. No elution was performed
15. Transfer the beads into a 1.5 ml microfuge tube. Wash once with binding buffer. Wash once with 500 ul ddH2O. Wash once with 1X HRV3C Rxn Buffer. Remove all buffer by the help of a syringe to avoid aspiration of beads.
16. Add 200 ul of HRV3C Rxn Buffer. Take 10 ul as pre-cleavage aliquot. It would contain 3.3 ul beads.
17. Add 2 units of HRV3C protease. Mix well. Incubate at 4-5 C over the weekend. (i.e. 60 hours)
18. Recover supernatant that contains cleaved proteins (1^{st}). Add an additional 200 ul Rxn Buffer. Vortex and recover again (2^{nd}). Wash the beads with 500 ul ddH2O and add 100ul 1X SB.

The steps of the third trial purification:

1. Inoculate a 10ml LB-Ampicillin starter cultures with BL21 cell colonies on LB plates and grow overnight.
2. In the morning, use 10ml of the starter culture to inoculate 250 ml LB-Amp media and grow at 200 rpm shaking at 37˚C for 1-2 hours. Monitor the OD600 value until it reaches the 0.6-0.8 Log phase growth.
3. Cool down the culture to RT by submerging flasks in a water bath.
4. Induce protein expression by adding 100 µM final concentration of IPTG and grow cells further 16 hours at 18˚C with 160rpm shaking. (Note: Culture temperature could not be reduced below 27C during the first hour after induction)
5. Pellet the cells by centrifugation.

6. Resuspend cells in4 ml of binding buffer (50 mM TrisCl pH 8.0, 100mM NaCl, 1mM PMSF, 1 mM DTT, 5% Glycerol) by vertexing. Transfer into a 5 ml microfuge tube.
7. Add lysozyme to 100ug/ml final concentration. Mix and incubate on ice at least 20 minutes.
8. Sonicate using tip MS 1.5, at 50% amplitude for a total of 2 minutes, such as 15 seconds finger controller sonication and chill on ice 30 seconds. Insert the tip of the probe about 2cm into the bacterial mix. Keep the tube in ice during sonication.
9. Add 10 Units of Pierce Universal Nuclease, 1mM final concentration of MgCl2 and 1% TritonX-100. Lyse in cold room by incubating on rotating wheel for 30minutes.
10. Spin down the lysate maximum speed to clarify. 25000g for 30 minutes. Take 100 ul of the lysate as aliquot.
11. Add 100 ul bed volume Glutathione beads into the cleared lysate. Add to additional 1mM EDTA. Incubate in cold room/fridge for 2hours on a rotating wheel.
12. Spin down 5 minutes at 1000rpm.
13. Wash 3 times 10 minutes with 5ml binding buffer + 1% TritonX-100.
14. Wash once with binding buffer (no triton). Resuspend in a total of 200ul and take 10ul aliquot.
15. Transfer the beads into a 1.5 ml microfuge tube. Wash once with binding buffer. Wash once with 500 ul ddH2O. Wash once with 1X HRV3C Rxn Buffer. Remove all buffer by the help of a syringe to avoid aspiration of beads.
16. Add 200 ul of HRV3C Rxn Buffer. Take 10 ul as pre-cleavage aliquot. It would contain 3.3 ul beads.
17. Add 2 units of HRV3C protease. Mix well. Incubate at 4-5 C overnight. (i.e. 16 hours)

    Recover supernatant that contains cleaved proteins (1^{st}). Add an additional 200 ul Rxn Buffer. Vortex and recover again (2^{nd}). Wash the beads with 500 ul ddH2O and add 100ul 1X SB.

2.10.2 Analysis of Induction by SDS-PAGE

SDS-PAGE (Polyacrylamide gel electrophoresis) is a widely used biochemical technique for separation and analysis of protein samples. A mini protean II gel casting apparatus (Biorad) has been used for the preparation of 15% gels according to the (Table 2.16) below.

Table 2.16 Preparations of SDS gel

Stacking gel 4.50%		**Separating Gel 15%**	
Acrylamide/bis-acrylamide 30%	1.5 mL	Acrylamide/bis-acrylamide 30%	5.0 mL
1 M TrisCl pH 6.8	1.25 mL	1 M TrisCl pH 6.8	3.75 mL
10 % SDS	100 µL	10 % SDS	100 µL
10 % (w/v) APS	50 µL	10 % (w/v) APS	50 µL
TEMED	15 µL	TEMED	5 µL
ddH2O	7.1 mL	ddH2O	1.1 mL
Total	10 mL	Total	10 mL

2.11 Gene Jet Plasmid Miniprep Kit

Thermo Scientific GeneJET plasmid miniprep kit was used to purify the pGEX-6P-1 for restriction analysis and sequencing. 5 mL of *E. coli* culture was used for purification a high copy number of the plasmid. 250 µL resuspension solution was added to the pelleted cells and the cell suspension was transferred to a microcentrifuge tube. 250 µL lysis solution was added the tube and mixing 4-6 times by inverting the tube till the solution became viscous. 350 µL of neutralization solution was added and mixing by inverting the tube 4-6 times. Spinning down 5 minutes by centrifuge, then transferring the supernatant to the supplied GeneJET spin column. Spinning down 1 minute and the flow-through was discarded. 500 µL wash solution was added to the GeneJET spin column and centrifugation for 1 minute and flow-through was discard, then this step was repeated to avoid residual ethanol in the plasmid. Ultimately the plasmid was eluted by elution buffer and stored at -20°C.

2.12 Restriction Analysis

For recovering the plasmid from the bacterial culture, 20 µL typical restriction digestion reaction was prepared (Table 2.17). after incubation for 1 hour at 37°C was loaded on the agarose gel 1%.

Table 2.17 Restriction analysis reaction master mix

Plasmid	8 µL (~1000ng plasmid)
Cut smart buffer	2 µL
BamH I RE	0.5 µL
ddH2O	9.5 µL

CHAPTER THREE
RESULTS AND DISCUSSION

3.1 Isolated RNA Concentrations

The absorbance of isolated RNA has been taken at 260/280 nm for both DNA and RNA purities (Table 3.1).

Table 3.1 Isolated RNA concentrations

Sample	Nucleic acid concentration	unit	A260	A280	A260/280	260/230	Sample type	factor
OVCAR $10x10^6$ cells	756.1	ng/µL	18.902	9.227	2.05	2.10	RNA	40.0
A2780cis $5x10^6$ cells	143.2	ng/µL	3.581	1.748	2.05	1.67	RNA	40.0

3.2 Nucleic Acids Concentrations After cDNA Synthesis Reaction

After using random primers in RT-PCR, the concentrations of nucleic acids have been raised to be ready for further application. The nucleic acid concentrations were measured (Table 3.2).

Table 3.2 Concentrations of cDNA reaction

Sample	Nucleic acid concentration	unit	A260	A280	A260/280	260/230	Sample type	factor
OVCAR	1554.7	ng/µL	31.093	16.904	1.84	2.27	DNA	50.0
OVCAR diluted	141.4	ng/µL	2.827	1.576	1.79	1.74	DNA	50.0
A2780cis	1523.5	ng/µL	30.470	16.531	1.84	2.21	DNA	50.0
A2780cis diluted	142.6	ng/µL	2.852	1.591	1.79	1.70	DNA	50.0

3.3 PCR

3.3.1 Primer Concentrations

All the primers were resuspended in the proper amount of ddH2O for the amplification reaction after resuspending concentrations were measured using Nano drop.

Table 3.3 Primers concentrations

Name	5'	Sequence	3'	bp	scale	purification	Nucleic Acid Conc. ng/µl (NanoDrop)
FWD.1		CGGCGACGGCGACTTTT		17	50	Standard	420.1
REV.1		CAGCAACAACTGCTATCACCAC		22	50	Standard	577.5
FWD.2		CGGCGACTTTTGCCGC		16	50	Standard	373.1
REV.2		GTTCCCTATGCATCTCACCCAT		22	50	Standard	563.6
FWD.3		GATCCTCGAGCAGGAAGAATGTGTCTG		27	50	Standard	695.7
REV.3		GATCCTCGAGCTATTTTAGACCCTGCATTGA		31	50	Standard	845.3

3.3.2 PCR Reactions Using Agarose Gel Electrophoresis 1%

As shown in (Figure 3.1) We have obtained two positive bands for A2780cis, and negative bands with OVCAR; which means the specific primers have matched very well with the first and did not match with the second. As expected, a band with (A2780cis FWD.1/REV.1) was in the size of 801bp, (A2780cis FWD.2/REV.2) was in the size 884 bp. In the second PCR reaction, the third primers were used to match and amplify more (Figure 3.2), As expected, a band with (A2780cis FWD.3/REV.3) was is the size of 786 bp, which means well matched.

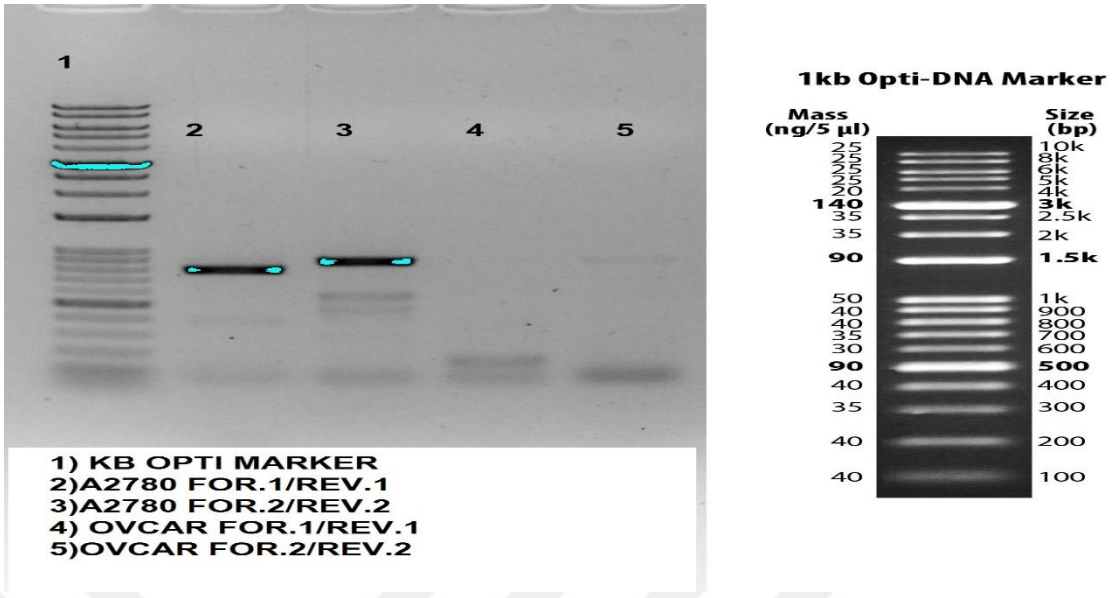

Figure 3.1 EpEX amplification first PCR reaction; agarose gel electrophoresis 1 %

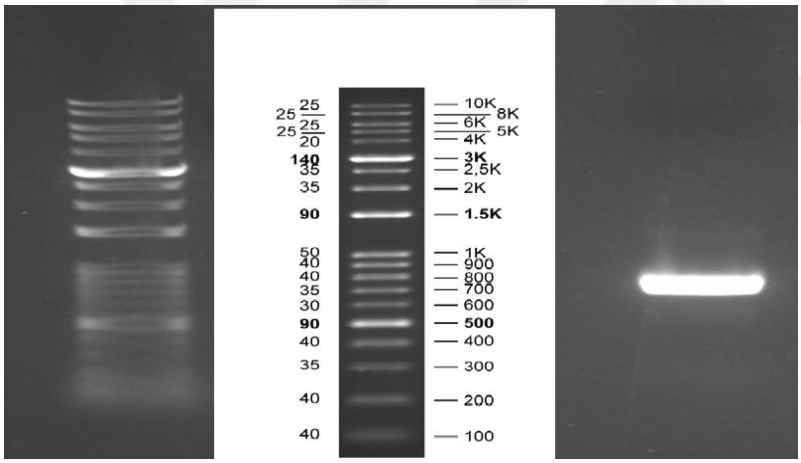

Figure 3.2 EpEX amplification second PCR reaction: agarose gel electrophoresis 1% with a positive band at the size of 786 bp

3.4 Digestion Analysis

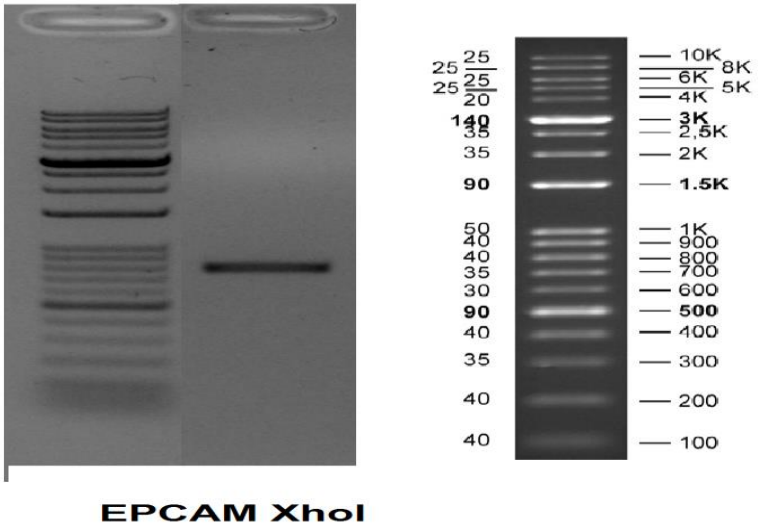

Figure 3.3 Agarose gel electrophoresis of recombinant pGEX-6P-1-EpEX plasmid, using restriction enzyme Xhol

3.5 Expression Screening

Colony PCR was performed to select the positive colonies for sequencing and recombinant expression steps as depicted (Figure 3.4) only two positive colonies were obtained the second colony and the sixth colony.

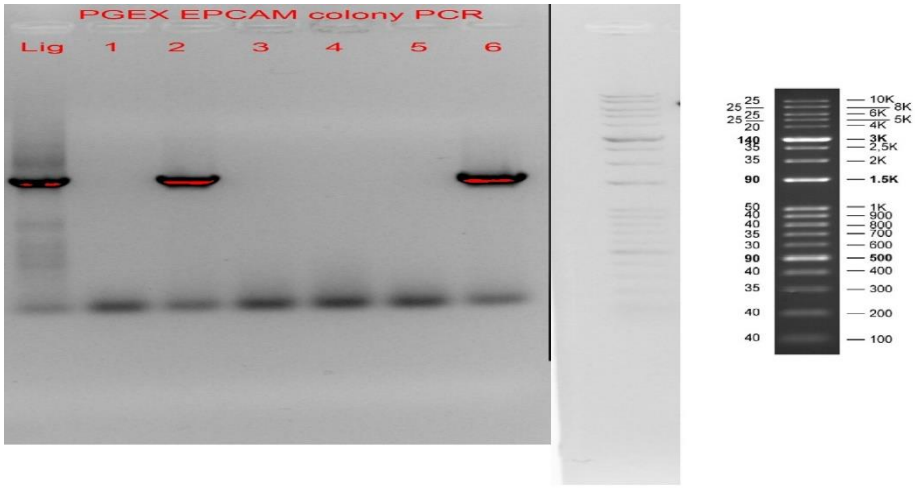

Figure 3.4 Colonies 2 and 6 are positive to be screened for further application

Table 3.4 Concentrations of the plasmids

Sample	**Nucleic acid concentration**	**unit**	**A260**	**A280**	**A260/280**	**260/230**	**Sample type**	**factor**
pGEX-6P-1 EpEX	142.0	ng/µL	2.839	1.644	1.73	0.96	DNA	50.0

3.6 Plasmid Restriction Analysis

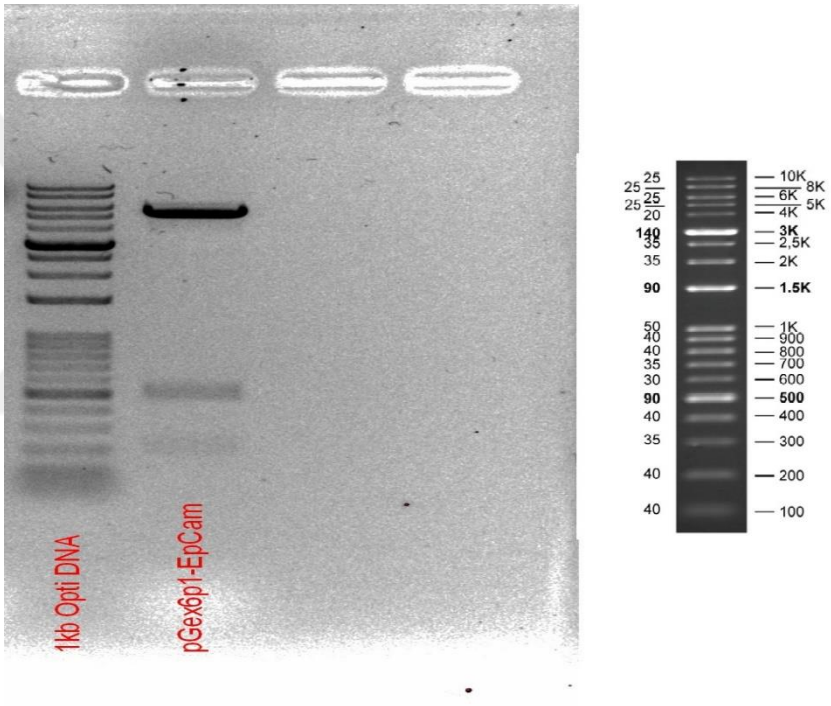

Figure 3.5 Restriction analysis of pGEX-6P-1-EpEX

3.7 Densitogram Alignment of Cloned EpEX

In the verification of the pGEX-6P-1-EpEX sequence by Sanger sequencing (Figure 3.6), the second clone only was sequenced. It seems there are two missense mutations that result in amino acid substitutions. The first mutation is probably bearable by the mRNA clones, but the second mutation might be a potential problem as it causes Arginine (positively charged) to Tryptophan (strongly hydrophobic) amino acid substitution. It was expected that those mutations might be in the Genome of the cell

line may arise during Reverse transcription, or less likely they may arise during PCR amplification.

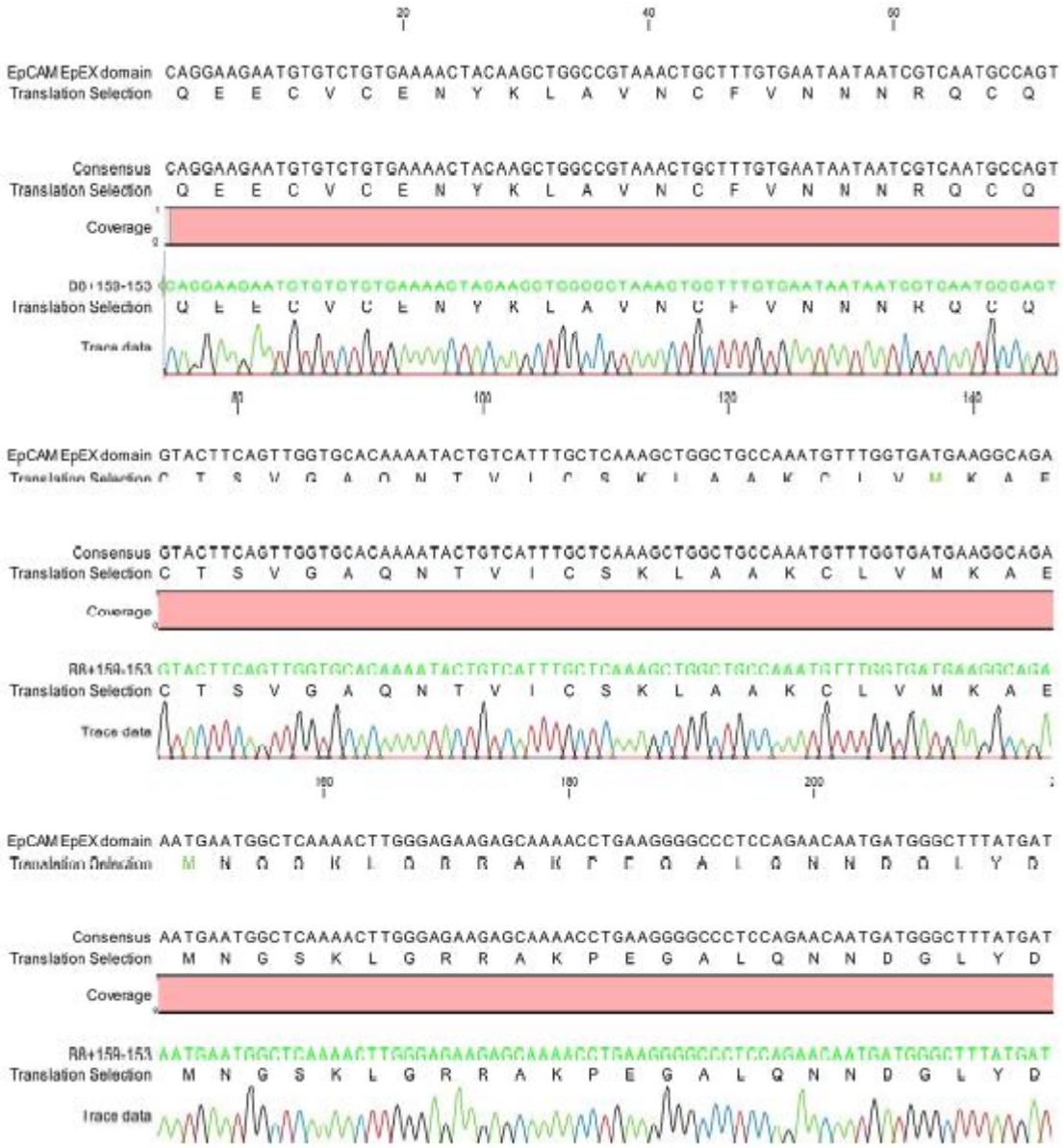

Figure 3.6 Conformation of a pGEX-6P-1-EpEX sequence of the second clone using Sanger sequencing method

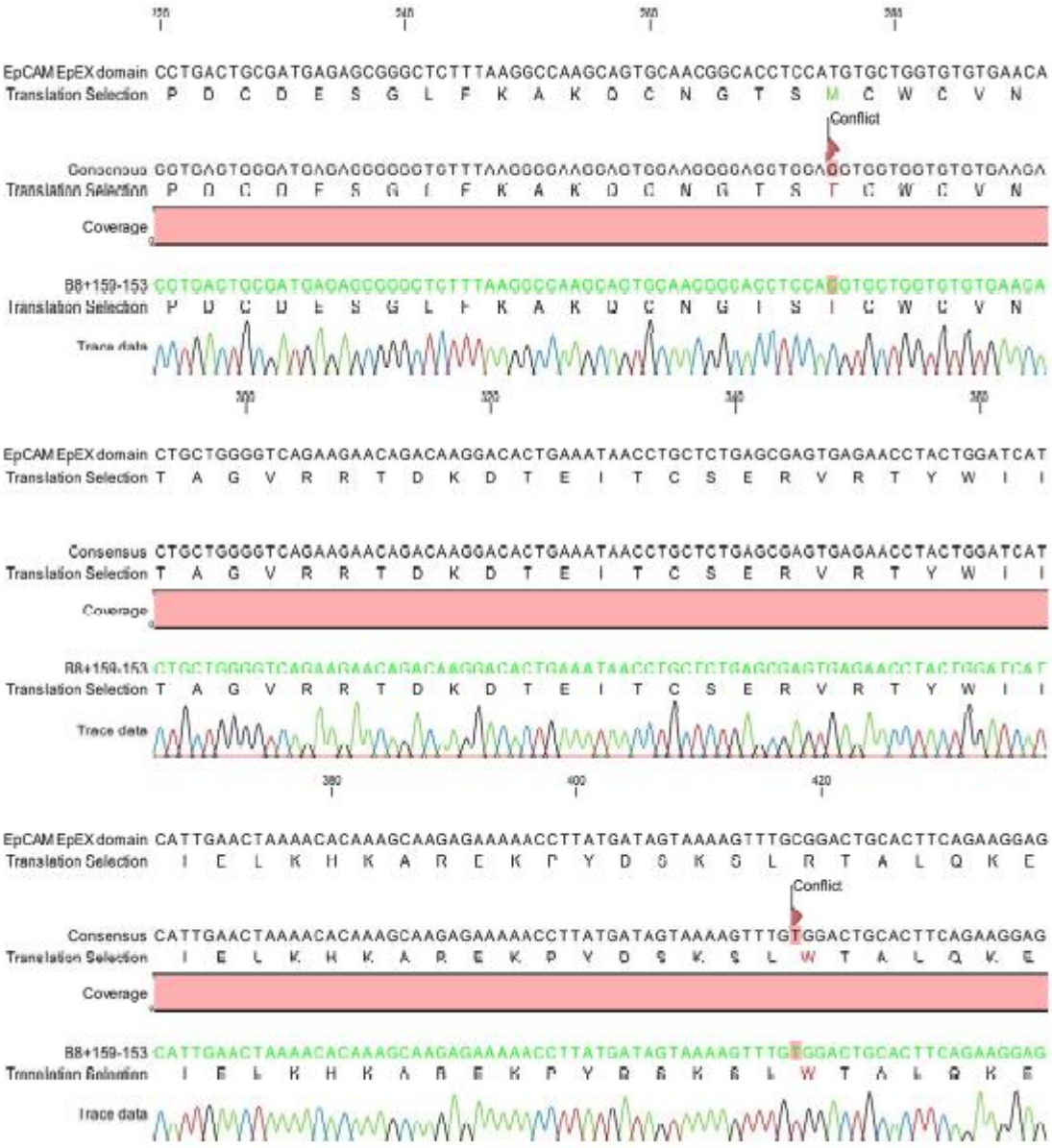

Figure 3.6 continues

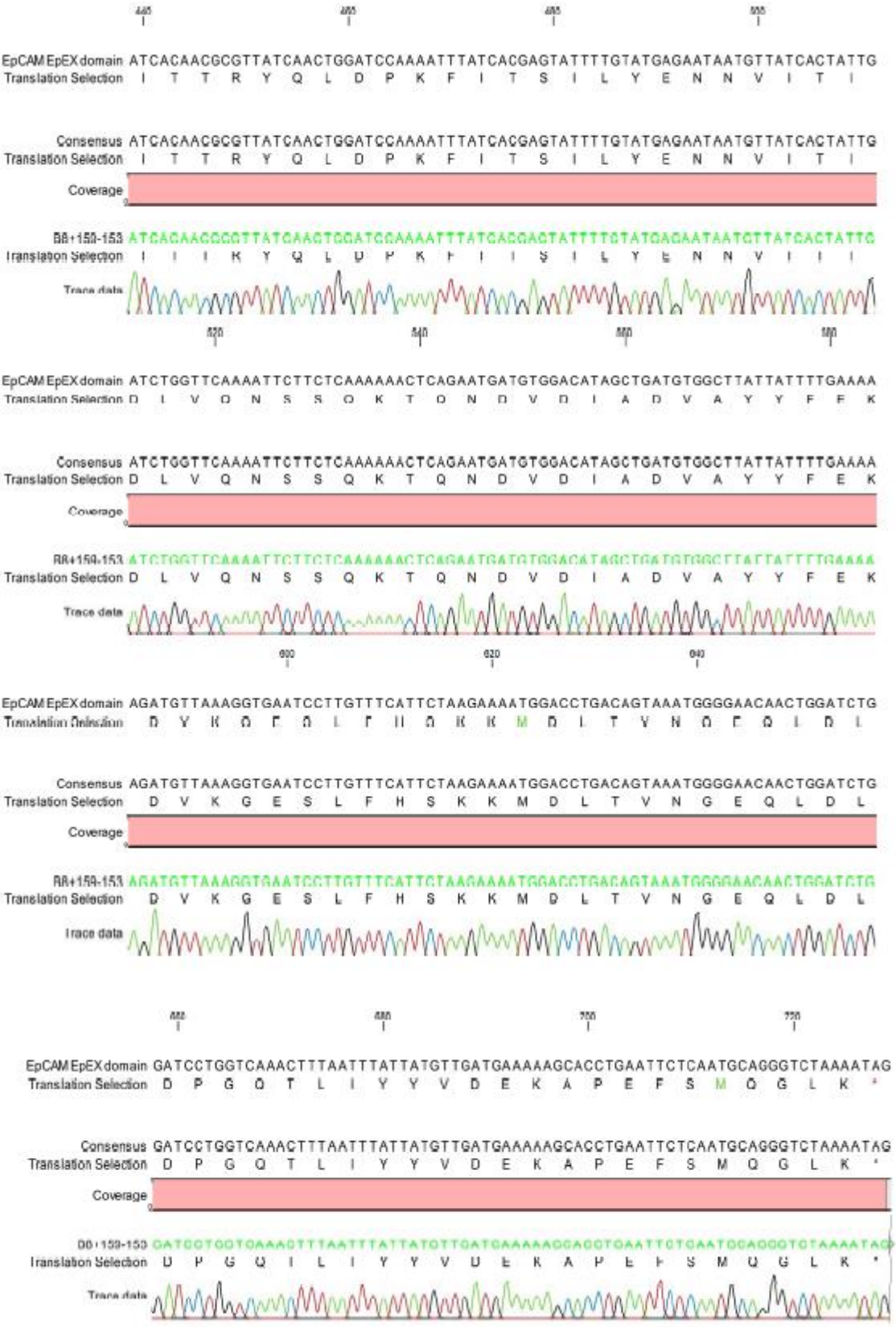

Figure 3.6 continues

3.8 GST-EpEX Purification

First, we tried to analyze the induction of recombinant GST-EpEX fusion protein. As a positive control, another protein whose induction was characterized previously (GST-ODF), was used. Upon addition of IPTG, GST-ODF was produced in high quantities as expected (asterisk lanes 3 and 5). However, induction of GST-EpEx was relatively weak (asterisk in lane 7).

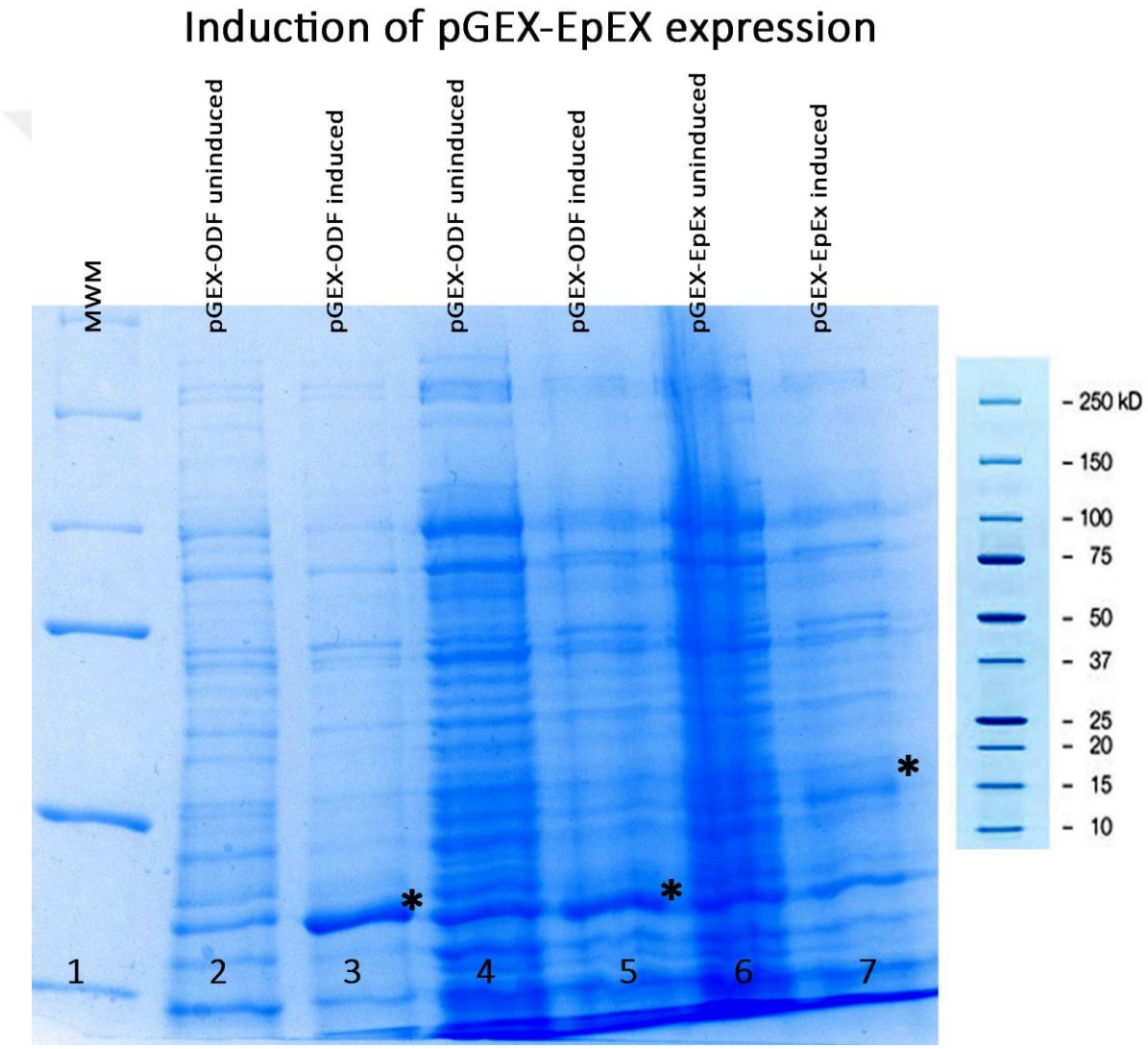

Figure 3.7 SDS-PAGE Control

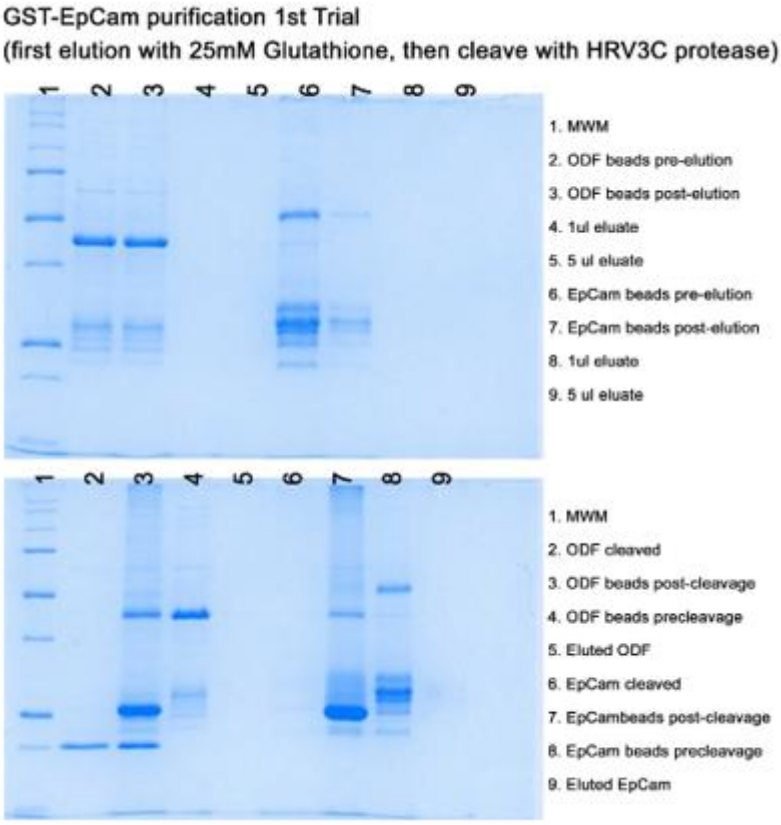

Figure 3.8 SDS-PAGE First Trial

It was found out the desired protein EpEX expression level significantly low with the pGEX-6P-1 plasmid. As depicted in (Figure 3.7) SDS-PAGE electrophoresis indicated the molecular weight of EpEx. A trial with the second positive colony was performed, and another trial with the sixth positive colony was performed. Induction of the expression of GST-EpEx cannot be readily observed when comparing proteins extracts from uninduced and induced bacterial cultures (lanes 7 and 8 respectively). Although affinity purification yields a protein of correct molecular weight (lane 4, asterisk), there is a significant amount of degradation product below the full-length protein. After cleavage by HRV3C protease, an intense GST band of 25 kDa size is visible in lane 5, suggesting that, although not readily visible in cell extracts, induction takes place. Proteolytic cleavage yields a small amount of recombinant EpEx protein (lane 2, asterisk) of the expected molecular weight of 28 kDa. Results suggest that recombinant EpEx is degraded while being produced in bacterial cells.

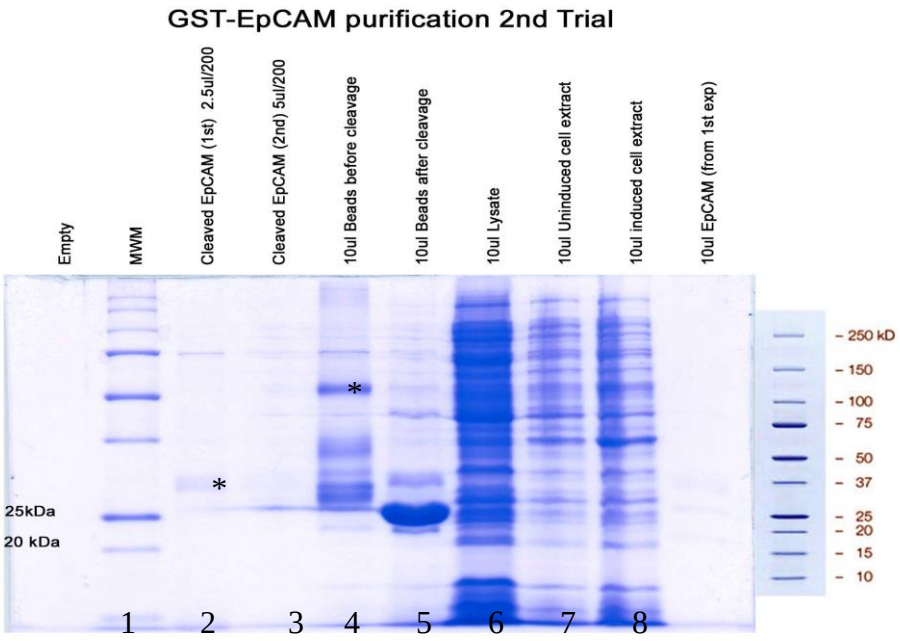

Figure 3.9 SDS-PAGE Second Trial

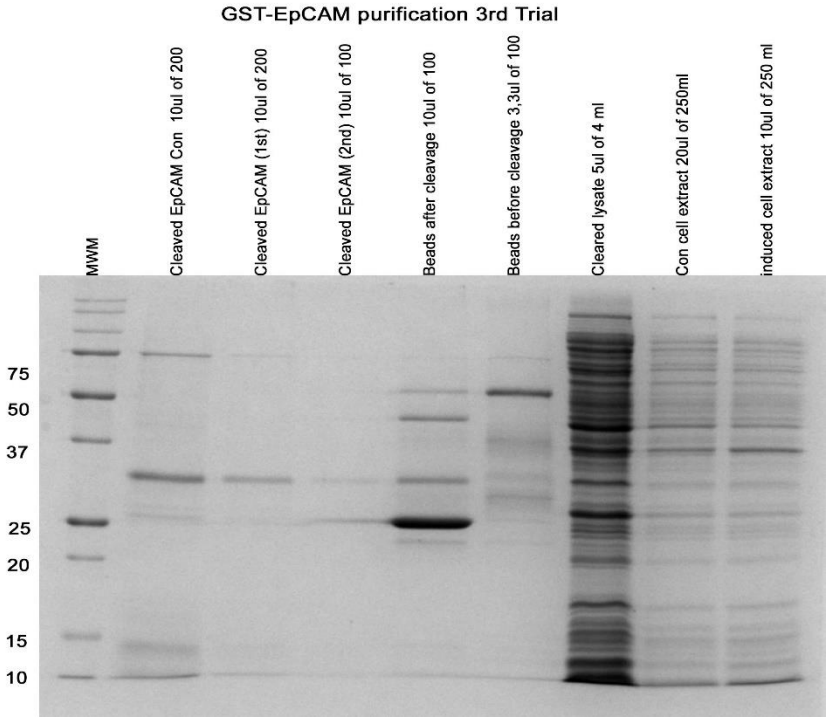

Figure 3.10 SDS-PAGE Third Trial

CHAPTER FOUR
CONCLUSION

The focus was on the molecular cloning and recombinant production of the extracellular domain of epithelial cell adhesion protein. Initially, we started by identifying cell adhesion molecules and the classification of their families; some of them rely on calcium to mediate cell and cell interaction or cell and matrix interaction, some of them do not rely on calcium, some of them bind homophilic, and some of them heterophilic. Epithelial cell adhesion molecule individually classified, means does not belong to any of cell adhesion molecule family. Does not depend on calcium to mediate homotypic cell and cell adhesion.

The genetic location of EpCAM's gene in the second chromosome 2P21 site and the gene contains nine exons which ultimately encode for the three different domains. Glycosylation which gives the tertiary structure functionality to the EpCAM. Subsequent cleavage, each cleaved part share in a cellular pathway or as an ectodomain.

All the stages of recombinant protein production were carried out using E. coli strain bacterial production. pGEX-6P-1 plasmid vector which contains GST-tag was utilized. It was found out the desired protein EpEX expression level significantly low with the pGEX-6P-1 plasmid. Three different trials of purification was performed to improve the expression level. Conclusively, recombinant EpEx is degraded while being produced in bacterial cells. This study was paving the way for more future studies on the production of different cell adhesion molecules using bacterial hosts for obtaining high scale that might be exploited as a diagnostic marker or for discovering the biochemical structure.

REFERENCES

Andreoli, T. B. (2013). *Molecular Biology of Membrane Transport Disorders Plenum Press.* New York: Springer Science & Business Media.

Balzar, H. B. (2001). Cytoplasmic tail regulates the intercellular adhesion function of EpCAMs. *Molecular Cell Biologyl*,18.

Balzar, Bakker HA, Briaire-de-Bruijn IH, Fleuren GJ, Warnaar SO, Litvinov SV. (1999). The biology of the 17-1A antigengen EpCAM. *Molecular Cell Biologyl*, *77*, 699–712.

Baeuerle, P.A.O. G. (2007). EpCAM (CD326) finding its role in cancer. *British Journal of Cancer*, 96, 417–423.

Brackenbury, U Rutishauser, and Edelman (1981). Distinct calcium-independent and calcium-dependent adhesion systems of chicken embryo cells. *Proceedings of the National Academy of Sciences* USA *78*, 387-391.

Brown, T. (2016). *Gene cloning and DNA analysis an introducation*. West Sussex, UK: John Wiley & Sons, Ltd.

Brown, M.S. J. Y. (2000). Regulated intramembrane proteolysis: a controlmechanismconserved from bacteria to humans. *Cell 100* , 391–398.

Buono, Del, M. P. (1999). The role of the E-cadherin complex in gastrointestinal cell differentiation. *Cell Proliferation 32*, 79–84.

Buxton, R.S. A.I.Magee. (2010). Structure and interactions of desmosomal and other cadherins. *Seminars in Cell Biology*. London: Elsevier, 157-167.

Chaves-Pérez, A. B. (2012). EpCAM regulates cell cycle progression via control of cyclin D1 expression. *Oncogene 32*, 1-10.

Chothia, C. E. Y. (1997). The Molecular Structure of Cell Adhesion Molecules. *Annual Review of Biochemistry*, 823-862 .

Denzel, S. D. M. (2009). Initial activation of EpCAM cleavage via cell-to-cell contact,. *BMC Cancer 9* , 402.

Fagerberg, G. R. (2010). Side by Side analysis of five clinically tested anti-EpCAM monoclonal antibodies. *Cancer Cell International.10*, 44.

Fagotto, B. G. (1996). Cell contact-dependent signaling. *Developmetnal Biology.180*, 445–454.

Fong, M. S. (2008). Ep-CAM expression in pancreatic and ampullary carcinomas: frequency and prognostic relevance. *Journal of Clinical Pathology.61*, 31–35.

Fuss, R. H. (2012). *Cell Adhesion Molecules Basic Neurochemistry. Principles of Molecular, Cellular, and Medical Neurobiology*, *Academic Press [8. ed]* (165-179). Chicago,USA.

García, J. A. (2005). Get a grip: integrins in cell–biomaterial interactions. *Biomaterials*, 7525-7529.

Gastl, G. S. (2000). Ep-CAM overexpression in breast cancer. *Lancet, 356*, 1981–1982.

Gaiser, M.R.T. L. (2012). Cancer-associated epithelial cell adhesion molecule (EpCAM; CD326) enables epidermal Langerhans cell motility and migration in. *National Academey of Science. U.S.A. 109*, 889–897.

Giepmans, U. S.-A. (2013). Absence of cell-surface EpCAM in congenital tufting enteropathy. *Human Molecular Genetics,22,* 2566–2571.

Herreros-Pomares, A. A. (2018). EpCAM duality becomes this molecule in a new. *Oncology / Hematology*, *126*, 52-63.

Herlyn, M. Z. S. (1979). Colorectal carcinoma-specific antigen: detection by means of monoclonal antibodies. *Proceedings of the National Academy of Sciences USA.76*, 1438–1442.

Humphries, Jonathan D.A. B. (2006). Integrin ligands at a glance. Journal of Cell Science. *Journal of Cell Science.119*, 3901-3903

Karabulut, S. T. (2014). The diagnostic, predictive, and prognostic role of serum epithelial cell adhesion molecule (EpCAM) and vascular cell adhesion molecule-1 (VCAM-1) levels in breast cancer. *Tumour Biology.35*, 8849–8860.

Karen, E.Brown, K. M. (2007). The role of integrins during vertebrate development. Seminars in Developmental Biology. From the Laboratory of Developmental Biology, National Institute of Dental Research, *National Institutes of Health, Bethesda, MD 20892,* USA: Elsevier. 69-77

Khosla, R. R. (2017). EpCAM+ liver cancer stem-like cells exhibiting autocrine Wnt signaling. *Stem Cells Translation. 3*, 807–818.

Köhler, C. M. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature 256,* 495-497.

Linnenbach, J. S. (1989). Sequence investigation of themajor gastrointestinal tumor-associated antigen gene family, GA733. *Proceedings of the National Academy of Sciences USA. 86,* 27–31.

Litvinov, S. B. (1997). Epithelial Cell adhesion molecule (Ep-CAM) modulates cell–cell interactions mediated by classic cadherins. *Journal of Cell Biology*. *139*, 1337–1348.

Litvinov, S.V. M. V. (1994). Ep-CAM: a human epithelial antigen is a homophilic cell–cell adhesion molecule. *Journal of Cell Biology. 125*, 437–446.

Liu, X. G. (2017). Mutation of N-linked glycosylation in EpCAM affected cell adhesion in breast cancer cells. *Biological Chemistry. 398*, 1119–1126.

Lodish, H, B. A. (2000). *Molecular Cell Biology. 4th edition*. New York: H. Freeman and Company.

Litvinov, Sergey,M. B.-d. (1997). Epithelial Cell Adhesion Molecule (Ep-CAM) Modulates Cell–Cell Interactions Mediated by Classic Cadherins. *Journal of Cell Biology.139,* 1337.

Maetzel, S. D. (2009). Nuclear signalling by tumour-associated antigen EpCAM. *Cell Biology. 11,* 162–171.

Marleen, J E M Gosens, L. C. (2007). Loss of membranous Ep-CAM in budding colorectal carcinoma cells. *Modern Pathology. 20*, 221–232.

Membranome (2019). Retrieved February 13, 2019 from membranome, College of Pharmacy, University of Michigan., https://membranome.org/protein_classes/3

Münz, M.C. K. (2004). The carcinoma-associated antigen EpCAM upregulates c-myc and induces cell proliferation. *Oncogene, 23,* 5748–5758.

Park, S. B. (2016). Nuclear EpICD expression and its role in hepatocellular carcinoma. *Oncology. 36,* 197-204.

Rahel Philip, S. H. (2014). Claudin-7 promotes the epithelial – mesenchymal transition in human colorectal cancer. *Oncotarget,* 2046-2063.

Ralhan, R. H.-H.-C. (2010). Nuclear and cytoplasmic accumulation of Ep-ICD is frequently detected in human EpCAM. *Plos One 5*.

Rosano, G. L. (2014). Recombinant protein expression in E.coli. *Frontiers in microbiology.172*

Rupp, A.-K. R. (2011). Loss of EpCAM expression in breast cancer derived serum exosomes: role of proteolytic cleavage. *Gynecology & Oncology.122,* 437–446.

Ryan PJ, O. J. (1997). Distinction between pleural mesothelioma and pulmonary adenocarcinoma using MOC31 in an asbestos sprayer. *Respir Med* , 57–60.

Schnell, V. C. (2013). EpCAM: Structure and function in health and disease. *Biochimica et Biophysica Acta,* 1989–2001.

Shield, C. J. (1994). Markers for metastatic adenocarcinoma in serous effusion specimens. *Diagnosis Cytopathol, 237–*45.

Sivagnanam, M. J. M. (2008). Identification of EpCAM as the gene for congenital tufting enteropathy. *Gastroenterology 135*, 429–437.

Spizzo, P. B. (2002). Prognostic significance of Ep-CAM and Her-2/ neu overexpression in invasive breast cancer. *International Journal of Cancer 98,* 883–888.

Susann, K. a. (2011). Cancer Cells Cut Homophilic Cell Adhesion Molecules and Run. *Cancer Reseach,* 303–309.

Strnad, J. A. H. (1989). Molecular cloning and characterization of a human adenocarcinoma/epithelial cell surface antigen complementary DNA. *Cancer Reseach. 49*, 314–317.

Stanbury, Peter F. A. W. (2017). *Principles of Fermentation Technology.* Oxford OX5 1GB, United Kingdom: Elsevier Ltd.

Szala, S. Y. K. (1990). Molecular cloning of cDNA for the human tumor-associated antigen CO-029 and identification of related transmembrane antigens. *Proceedings of the National Academy of Sciences USA. 87*, 6833–6837.

Takai, J. M. (2008). Nectins and nectin-like molecules: roles in contact inhibition of cell movement and proliferation. *Nature Reviews Molecular Cell Biology 9,* 603–615.

Trzpis, M. P. (2007). Epithelial cell adhesion molecule: more than a carcinoma marker and adhesion molecule. *The American Journal of Pathology*, 386–395.

Vinatier, D. (1994). Integrins and reproduction. *European Journal of Obstetrics & Gynecology*, 71-81.

Ward, F. S. (2011). The Function of E-Cadherin in Stem Cell Pluripotency and Self-Renewal. *Genes*, 229-259.

Went, P., Dirnhofer, S., Salvisberg, T., Amin, M., Lim, S., Diener, P.-A., & Moch, H. (2005). Expression of Epithelial Cell Adhesion Molecule (EpCAM) in Renal Epithelial Tumors. *The American Journal of Surgical Pathology, 29,* 83-88.

Winter, M. C.-d. (2007). Cadherins are regulated by Ep-CAM via phosphaditylinositol-3 kinase. *Molecular Cellular Biochemistry*. 302, 17-26.

Xiaoa, Z. L. (2017). Aptamer-mediated gene therapy enhanced antitumor activity against human hepatocellular carcinoma in vitro and in vivo. *controlled Release, 258*, 130-145.

Xu, C. S. (2009). Tetraspanin12 regulates ADAM10-dependent cleavage of amyloid precursor protein. *FASEB, 23*, 3674–3681.

Yu, T. M. (2017). EpCAM intracellular domain promotes porcine cell reprogramming by upregulation of pluripotent Gene expression via beta-catenin signaling. *Scientific Reports, 7*, 46315.

Yu, Y. M. (2017). EpCAM Intracellular Domain Promotes Porcine Cell Reprogramming by Upregulation of Pluripotent Gene Expression via Beta-catenin Signaling. *Scientific Reports, 7,* 46315.

Yáñez-Mó, M. M. G.-L. (2011). Functional interplay between tetraspanins and proteases. *Cellular & Molecular Life Sciences, 68,* 3323–3335.