CONSTRUCTION OF VARIOUS FUSION PROTEINS OF RECOMBINANT CITRATE SYNTHASE FROM *THERMOPLASMA VOLCANIUM*

A THESIS SUBMITED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

SEDA ÖZDOĞAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS OF THE DEGREE OF
MASTER OF SCIENCE
IN
THE DEPARTMENT OF BIOLOGY

MARCH 2004

Approval of the Graduate School of Natural and Applied Sciences.

_____

Prof. Dr. Canan Özgen

Director

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

_____

Prof. Dr. Semra Kocabıyık

Head of Department

This is to certify that we had read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

_____

Prof. Dr. Semra Kocabıyık

Supervisor

Examining Committee Members

Prof. Dr. Semra Kocabıyık _____

Prof. Dr. Aykut Kence _____

Assoc. Prof. Dr. Meral Kence _____

Assoc. Prof. Dr. Fatih İzgü _____

Assoc. Prof. Dr. Pınar Çalık _____

ABSTRACT

CONSTRUCTION OF VARIOUS FUSION PROTEINS OF RECOMBINANT CITRATE SYNTHASE FROM *THERMOPLASMA VOLCANIUM*

Özdoğan, Seda

M.S., Department of Biology

Supervisor: Prof. Dr. Semra Kocabıyık

March 2004, 139 pages

In this study, a strategy called gene splicing by overlap extension, "Gene SOEing", was used for the construction of the fusion proteins with the purpose of increasing the thermostability of mesophilic enzymes by incorporation of stability domain from a thermostable enzyme.

Gene SOEing is a PCR-based approach for recombining DNA molecules at precise junctions irrespective of nucleotide sequences at the recombination site and without the use of restriction endonucleases or ligase. In fusion constructs, as the stability determinant *Thermoplasma volcanium* citrate synthase (CS) large domain has been used. This gene has recently been cloned in our laboratory. In two different fusions, as fusion partners, dehalogenase II (*dehC*II) gene of *Pseudomonas* sp. CBS3 and aminoglycoside-3′-phosphotransferase-II (APH(3′)-II) gene of *E. coli* were employed. Following the Gene SOEing, two fusion products, 1722 bp long CS Large Domain–*dehC*II and 1750 bp long CS Large Domain–APH(3′)-II were constructed. Also a 1586 bp long *dehC*II–APH(3′)-II fusion was prepared. Three fusion

constructs were cloned in *E. coli*. Cloning was confirmed in each case, by restriction analysis of the isolated plasmids from recombinant colonies. APH(3′)-II gene associated with CS Large Domain–APH(3′)-II and *dehC*II–APH(3′)-II fusion constructs were successfully expressed in *E. coli* as revealed by enzyme assay and antibiotic agar plate assay. CS Large Domain–APH(3′)-II fusion protein retained 9.4% of the original APH(3′)-II activity after 10 minutes at 60ºC. However, CS Large Domain–*dehC*II and *dehC*II–APH(3′)-II fusions did not display any dehalogenase activity.

Keywords: Fusion proteins, citrate synthase, haloalkanoic acid dehalogenase, aminoglycoside-3′-phosphotransferase-II, Gene SOEing, *Thermoplasma volcanium*

ÖZ

THERMOPLASMA VOLCANİUM'UN REKOMBİNANT SİTRAT SENTAZI İLE ÇEŞİTLİ FÜZYON PROTEİN TASARIMLARI

Özdoğan, Seda
Yüksek lisans, Biyoloji Bölümü
Tez Yöneticisi: Prof. Dr. Semra Kocabıyık

Mart 2004, 139 sayfa

Bu çalışmada, mezofilik enzimlerin thermostabilitesini arttırmak amacı ile termostabil bir enzimin stabilite bölgesini bu enzimlere eklemek suretiyle füzyon protein oluşturmak için "Gen Birleştirme" stratejisi kullanılmıştır.

Gen birleştirme, bağlanma yerlerindeki nükleotid dizilerini dikkate almaksızın ve restriksiyon endonükleazlar ya da ligaz kullanımı olmaksızın DNA moleküllerini belirli bölgelerde birleştiren PCR-temelli bir yaklaşımdır. Bu füzyon yapılarda, *Thermoplasma volcanium*'un sitrat sentaz geninin büyük fragmenti stabilite belirleyicisi olarak kullanılmıştır. Bu gen kısa bir süre önce laboratuarımızda klonlanmıştır. İki farklı füzyon tasarımında, füzyon partnerleri olarak *Pseudomonas* sp. CBS3 dehalogenaz II (*dehC*II) geni ve *E. coli* aminoglikozit-3′-fosfotransferaz-II (APH(3′)-II) geni kullanılmıştır. "Gen Birleştirme" reaksiyonundan sonra 1722 bç büyüklüğünde CS Büyük Fragment–*dehC*II ve 1750 bç büyüklüğünde CS Büyük Fragment-APH(3′)-II elde edilmiştir. Ayrıca bir 1586 bç uzunluğunda *dehC*II–APH(3′)-II füzyon tasarlanmıştır. Üç füzyon tasarım *E. coli*'de klonlanmıştır. Her

defasında, klonlama rekombinant kolonilerden elde edilen plazmitlerin restriksiyon analizi ile doğrulanmıştır. CS Büyük Fragment-APH(3′)-II ve *dehC*II–APH(3′)-II füzyonlara ilişkin APH(3′)-II geni enzim tayini ve antibiyotikli agar yöntemleri ile gösterildiği gibi *E. coli*'de başarılı bir şekilde ekspres edilmiştir. CS Büyük Fragment-APH(3′)-II füzyon proteini 60^{o}C'de 10 dakika inkübasyondan sonra orijinal aktivitesinin %9.4'sini korumuştur. Ancak CS Büyük Fragment–*dehC*II ve *dehC*II–APH(3′)-II füzyonları herhangi bir dehalogenaz aktivitesi göstermemiştir.

Anahtar Kelimeler: Füzyon proteinleri, sitrat sentaz, haloalkanoik asit dehalogenaz, aminoglikozit-3′-fosfotransferaz-II, Gen birleştirme, *Thermoplasma volcanium*

ACKNOWLEDGEMENTS

I would like to express my special thanks and appreciation to my supervisor Prof. Dr. Semra Kocabıyık for her valuable guidance, continuous encouragement and patience during this study.

I also want to send my very special thanks to my dear labmates Dr. F. İnci Özdemir and Hatice Özel for their kind help and moral support whenever I need.

I would like to thank Graduate School of Natural and Applied Sciences METU (Project number: BAP-2003-07-02-00-31) for the financial support they provided.

I express my deepest gratitude to my mom Fügen Özdoğan, my dad M. Erhan Özdoğan and my sister Esra Özdoğan for sharing my thesis life and providing me their endless love and support.

I also send my best wishes and thanks to my best friend İpek Ekmen for her encouragement and endless support.

To My Family

TABLE OF CONTENTS

LIST OF TABLES

TABLE

LIST OF FIGURES

FIGURE

LIST OF ABBREVIATIONS

Tp. acidophilum	*Thermoplasma acidophilum*
Tp. volcanium	*Thermoplasma volcanium*
CS	Citrate synthase
***dehC*II**	2-haloalkanoic acid dehalogenase II
APH(3´)-II	Aminogylcoside 3´-phosphotransferase II
***Ta*CS**	*Thermoplasma acidophilum* citrate synthase
***Pf*CS**	*Pyrococcus furiosus* citrate synthase
***Tv*CS**	*Thermoplasma volcanium* citrate synthase
SPA	Staphylococcal protein A
GST	Glutathione *S*-transferase
MBP	Maltose-binding protein
GFP	Green fluorescent protein
Gene SOEing	Gene splicing by overlap extension
LB	Luria-Bertani medium
X-Gal	5-Bromo-4-chloro-3-indolyl β-D-galactopyranoside
IPTG	Isopropyl β-D-1-thiogalactopyranoside
MCA	Monochloroacetic acid
DTT	Dithiothreitol
PIPES	Piperazine-N,N′-bis-2-ethanesulfonic acid
PEP	Phosphoenolpyruvate

CHAPTER I

INTRODUCTION

Following the discovery in the late 1960s of the tools for manipulating DNA sequences *in vitro*; genetic engineering has provided scientists and technologists with the opportunity to express genes in prokaryotic systems. Among the many systems available for expression of recombinant genes, with its well-characterized genetics and the availability of an increasing large number of cloning vectors and mutant host strains as well as its ability to grow rapidly and at high density on low-cost substrates, the Gram-negative bacterium *Escherichia coli* (*E. coli*) remains one of the most attractive and the most widely used host (Baneyx, 1999). On the other hand, use of this simplest and most inexpensive system as a key technique in the understanding of molecular processes as well as for providing the proteins with desired quality and quantity, brings along some fundamental problems. For a long period of time, inefficient translation initiation of eukaryotic mRNAs on bacterial ribosomes, the formation of inclusion bodies of recombinant proteins in the *E. coli* cytoplasm, instability of the recombinant protein, failure of the complete removal of the amino-terminal initiator methionine residues, poor folding efficiency with insufficient amounts of folding factors upon overexpression and the need to develop custom purification or refolding schemes empirically for each new protein product have been the troubles to deal with for the heterologous gene expression in *E. coli* . A popular strategy, the link of the gene of interest to a second 'carrier' or 'partner' gene which is already known to be expressed in a microbial system, to generate a fusion protein, has been designed to overcome these limitations since the synthesis of a functional protein depends upon these basic features (LaVallie and McCoy, 1995). Today, gene fusion systems have become the most significant part of the recombinant technology.

The gene fusion strategies can be categorized into six types according to their suitable efficacy for the situation (Uhlén and Moks, 1990).

1. The most simple fusion strategy is splicing the recombinant gene of interest directly after a suitable signal sequence (Secretion strategy). It is possible to produce recombinant protein with a native N terminus if the signal peptide is correctly processed during transport.

2. Fusion of a gene product to itself is another simple fusion strategy (Polymerization strategy). Using such an approach, the yield and the half–life of the gene of product can be increased to satisfactory levels since self-polymerization results in formation of much more stable products.

3. C-terminal or N-terminal fusions are two of the most common fusion strategies, where the fusion partner encodes an affinity handle to facilitate purification (C-terminal fusion and N-terminal fusion strategies). In C-terminal fusions, the recombinant product is positioned at the C-terminal side of the fusion partner. The advantage of the C-terminal fusions is that the promoter and the translation initiation signals are all integrated in the 5′ end of the gene and are not changed by different fusions at 3′ end. So expression level is relatively predictable and different promoters for each gene product can easily be tested. For the N-terminal fusions, direct N-terminal sequencing of the gene fusion product can be easily performed and various biological expression assays can be designed, but has the disadvantages that a product-specific transcriptional and translational start must be engineered in the 5′ end of the gene of interest and when chemical methods are used to release the gene product, usually a nonnative protein is obtained.

4. Secretion-affinity fusion strategy combines the advantages of secretion and affinity purification. Secretion to the culture medium provides generous opportunities for various procedures in which the product is continuously

recovered from the medium during fermentation and this would be very useful in large-scale purification.

5. The dual-affinity fusion strategy is based on fusing the gene of interest between two heterologous domains, with specific affinity for two different ligands. The advantage of this approach is that the full-length protein can be selectively recovered by two subsequent affinity purification steps. This approach is suitable for the expression of proteins which are highly susceptible to proteolysis. However, two separate site-specific cleavages are necessary to cleave off the flanking domains.

6. The last fusion strategy combines a secretion signal with a sequence domain that inserts the gene product into the cell wall or one of the cellular membranes. This method can be helpful for vaccine development and other systems to generate immunogenic complexes since the approach can be used to expose receptors or antigens on the outer surface of bacteria or to assemble fusion proteins into virus like particles.

Gene fusions have been used in the first described systems for heterologous bacterial expression of proteins. There are several reasons for this, such as to avoid rapid degradation of foreign proteins by host proteases, to obtain efficient purification schemes which allow rapid recovery of gene products, to localize to different compartments of the cell, to develop more reliable and reproducible method to obtain a native protein i.e., *in vitro* cleavage of a fusion protein.

Now, affinity fusion systems in conjunction with genetic engineering have been used for the production of recombinant proteins on an industrial scale. The addition of specifically designed tags or the modification of the sequences within the target-gene product has enabled the development of novel strategies for downstream processing that can be employed for efficient recovery of both native and modified proteins (Nygren *et al.*, 1994; Stahl and Nygren, 1997).

One recent application of the gene fusion technology has been its use in prolonging the *in vivo* half-life of pharmaceutical proteins by administrating the therapeutic protein as fused to a protein with expanded half-life. This technology also facilitates subunit vaccine production. As a powerful strategy in subunit vaccine development, a dual expression system was devised, combining the two similar affinity-fusion expression systems for the parallel expression of immunogenic peptides and proteins. After affinity purification one fusion protein is used for immunization and the second fusion protein is used to analyze the induced antibody response to the fused peptide (Nygren *et al.*, 1994; Stahl and Nygren, 1997).

The other important application is the display of heterologous proteins or peptides on bacterial surface. This technology is important for the development of live bacterial vaccines that confer immunity against infectious agents, the display and screening of peptide and antibody libraries, for bacteria-based solid-phase immuno-assays and for the production of whole cell adsorbents which can be used in biocatalysis, bioremediation or in drug hunting (Stathopoulos *et al.*, 1996).

An interesting additional application of gene fusion is the display of peptides or proteins in filamentous phage surfaces. Phage particles decorated in such manner can be enriched *in vitro* from a background of irrelevant phages by biopanning techniques, employing an immobilized ligand capable of binding to the displayed protein. Moreover, the genetic information of the displayed protein is carried inside the protein coat of the captured phage, which is later identified by DNA sequencing. Therefore, the phage display system not only gives the ability to select peptides or proteins of desired function but also allows the power of bacterial genetics to be applied to protein structure studies, antibody selection, protein design and the research studies of discovery of the biological function of proteins where the gene sequence is known, but the physiological function of a protein is unknown (Nygren *et al.*, 1994; Stahl and Nygren, 1997). Bacteriophage λ coat proteins (M13 gene III (Scott and Smith, 1990), M13 gene VIII (Greenwood *et al.*, 1991), M13 geneVI (LaVallie and McCoy, 1995)) bacteriophage receptor (*lamB*) (Charbit *et al.*, 1988),

lacI (Cull *et al.*, 1992), a *lpp-ompA* (Francisco *et al.*, 1992) hybrid protein, and a flagellin-thioredoxin hybrid protein (LaVallie and McCoy, 1995) have been developed to use as fusion partners in order to screen diverse protein or peptide libraries for desired binding specificities (LaVallie and McCoy, 1995).

1.1 Gene Fusion Partners

For the expression, purification, secretion and detection of a variety of recombinant proteins, a large number fusion partners that range in size from one amino acid to whole protein have been developed and employed up to date (Table 1.1). Every fusion partner has its advantages and limitations and the choice of the fusion partner depends on these properties and final use of the gene product.

1.1.1 Gene Fusions and Recombinant Protein Expression and Folding

Many heterologous proteins that are overexpressed lead to the formation of insoluble protein aggregation known as inclusion bodies. Eventhough fusion proteins were originally constructed to facilitate protein purification and immobilization and to couple the activity of enzymes acting in a single metabolic pathway, it soon became clear that certain fusion partners could greatly improve the solubility of fused or passenger proteins. The *Staphylococcus* protein A (SPA) (Nilsson and Abrahmsén, 1990), *Schistosoma japonicum* glutathione *S*-transferase (GST) (Smith and Johnson, 1988), *E. coli* maltose-binding protein (MBP) (Guan *et al.*, 1988), and the *E. coli* thioredoxin (LaVallie *et al.*, 1993) are the four suitable and successful gene fusion expression partners for this purpose to produce correctly folded and soluble heterologous proteins that would otherwise accumulate within inclusion bodies in the bacterial cytoplasm. The most probable reason for improved folding of the fused or passenger proteins is that the fusion partner efficiently and rapidly reaches a native conformation as it emerges from the ribosome (or soon after its release), and promotes the acquisition of the correct structure in downstream folding units by favoring on-pathway isomerization reactions. In the case of unfused cytoplasmic

Table 1.1 Fusion partners used for gene fusion (Stevens, 2000)

Partner	Size	Fusion tag location	Tag type	Comments
His-tag	6,8 or 10 aa	N, C, internal	Purification	Most common purification tag used for IMAC one-step purification. Purification is possible even under denaturing conditions. Tag possibly influences crystallization.
FLAG™ peptide (DYKDDDDK)	8 aa	N, C	Purification	Ca^{2+}-dependent monoclonal antibody purification with EDTA elution. Tag cleavable with enterokinase.
Thioredoxin	109 aa (11.7 kDa)	N, C	Purification and Enhanced expression	Affinity purification with phenylarsine oxide-modified (ThioBond) resin. Also purification by heat treatment, osmotic shock or freeze/thaw treatment is possible.
lacZ (β-galactosidase)	116 kDa	N, C	Purification	Purification using r-amino-phenyl-β-D-thiogalactosidase-modified sepharose. Classical tag used for protecting peptides from proteolytic degradation. Fusion proteins with this tag have a high tendency to be insoluble. Active enzyme is a tetramer.
Chloramphenicol acetyltransferase	24 kDa	N	Secretion, purification and detection	Chloramphenicol-sepharose purification. Enzymatic assay possible for quantification.
trpE	27 kDa	N	Purification	Often form insoluble precipitates. Hydrophobic interaction chromatography purification.
Avidin/streptavidin/ *Strep*-tag			Purification and secretion	Biotin affinity purification and streptavidin affinity purification (Strep-tag).
Staphylococcal protein A	14 kDa (or 31 kDa)	N	Purification and secretion	IgG antibody affinity purification is possible (denaturing low pH elution needed). Secretion due to protein A signal sequence.
Streptococcal protein G	28 kDa	N, C	Purification and secretion	Albumin affinity purification, low pH elution needed. Fusion protein secretions due to protein G signal sequence.
Glutathione-*S*-transferase (GST)	26 kDa	N	Purification	Glutathione affinity or GST antibody purification. Enzymatic activity assay possible for quantitative analysis. Fusion proteins form dimers.
Maltose-binding protein (MBP)	40 kDa	N, C	Purification and secretion	Amylose affinity purification with maltose elution.
Galactose-binding protein			Purification	Galactose-sepharose purification.

Table 1.1 (continued)

Partner	Size	Fusion tag location	Tag type	Comments
Calmodulin-binding protein	4 kDa	N, C	Purification and Detection	Calmodulin/Ca2+ affinity purification with EDTA elution. Can potentially assay expression levels with 32P-cAMP kinase.
Green fluorescent protein	220 aa	N, C	Detection	Used as reporter gene fusion for detection purposes. Used at one time for possible refolding tag.
OmpT/OmpA /PelB /DsbA/DsbC	22 aa/21 aa /20 aa 208aa/236 aa	N	Secretion	Periplasmic leader sequences for potential protein export and folding, as well as potential disulfide bond formation and isomeration.
lac repressor			Purification	*lac* operator affinity purification.
Green fluorescent protein	220 aa	N, C	Detection	Used as reporter gene fusion for detection purposes. Used at one time for possible refolding tag.
OmpT/OmpA /PelB /DsbA/DsbC	22 aa/21 aa /20 aa 208aa/236 aa	N	Secretion	Periplasmic leader sequences for potential protein export and folding, as well as potential disulfide bond formation and isomeration.
lac repressor			Purification	*lac* operator affinity purification.
His-patch thioredoxin	109 aa (11.7 kDa)	N, C	Purification and enhanced expression	Use of His-patch modified thioredoxin for immobilized metal-affinity chromatography (IMAC) purification.
Dihydrofolate reductase (DHFR)	25 kDa	N	Purification	Methotrexate-linked agarose used for purification.
Cellulose-binding protein (CBP)	156 aa/ 114 aa/ 107 aa	N N C	Purification and secretion	Cellulose-based resins used for affinity purification with water elution. Different constructs available for cytoplasmic or periplasmic expression. Fusion proteins susceptible to proteolysis between the fusion partners.
Polycystein	4 aa	N	Purification	Thiopropyl-sepharose purification.
Polyarginin	5-15 aa	C	Purification	S-sepharose (cationic resin) purification. Fusion proteins potentially insoluble.
Chitin-binding protein		N, C	Expression	Used in Impact™ system, with intein-based expression constructs.
Ubiquitin	76 aa	N	Possible enhanced solubility	Ubiquitin fusions observed to increase *E.coli* expressed recombinant protein solubility.

MBP fusion, the proper folding requires the molecular chaperons like ATP-dependent DnaK-DnaJ-GrpE and GroEL-GroES which may recruit chaperons in the vicinity of the passenger protein. It has also been proposed that these fusion partners' high stability and solubility, as well as strong folding characteristics let them to act functionally *in vivo* as covalently linked 'chaperons'. As a result, fusion partners may directly interact with fused or passenger proteins that may provide proteins to fold correctly in a soluble form (Baneyx, 1999).

1.1.2 Gene Fusions and Recombinant Protein Purification

Developing fast, efficient, powerful and convenient purification methods of proteins has been the most frequent application of genetic engineering. This is beneficial for different purposes such as industrial production of recombinant proteins or facilitating functional and structural studies of proteins derived from the rapidly growing number of genes coming out of programs such as the Human Genome Project. Today, presence and use of genetically fused affinity fusion partners allow much simpler purification strategies irregardless of the properties of the target protein.

The main idea is based on passing a cell lysate or a culture medium containing the fusion protein through an affinity column containing a ligand that specifically interacts with the affinity handle. Fusion protein binds to the ligand while all other unwanted proteins are washed out using an appropriate buffer. After elution; a chemical or enzymatic method is used to cleave the purified fusion protein at the junction between the two protein moieties. The cleavage mixture is again passed through the column to allow the affinity handle to bind, while the target protein is collected in the flowthrough fraction (Figure 1.1). However, for many applications the affinity tag may be left on the target protein, e.g., 1. to function as immunogens for the generation and purification of antibodies, 2. when the biological activity of the target protein is unaffected by the affinity fusion partner, 3. for the directed immobilization of the target protein (Uhlén and Moks, 1990).

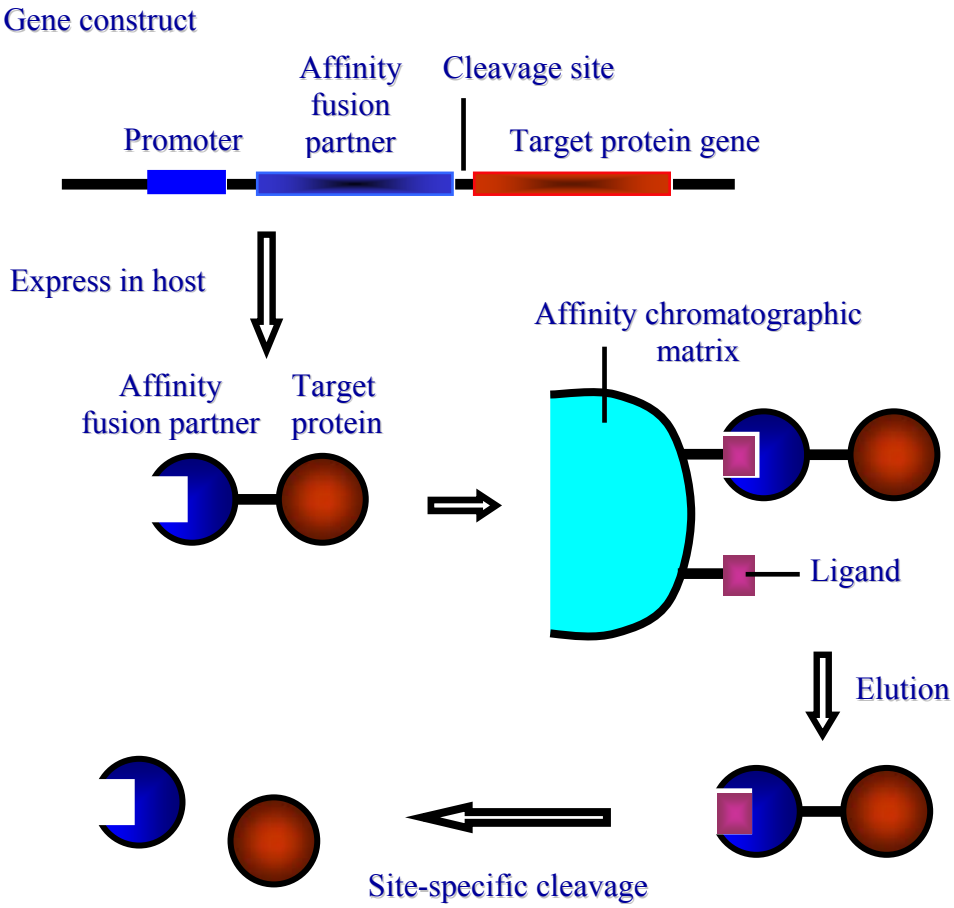


Figure 1.1 The purification of the target protein by using affinity fusion partner. The affinity fusion partner binds to the specific ligand. After elution, the pure target protein can be obtained by performing appropriate site-specific cleavage method.

Different types of interactions including, protein-protein, enzyme-substrate, bacterial receptor-serum protein, polyhistidines-metal ion, and antibody-antigen have been reported for these systems. The conditions for purification differ from system to system and factors such as protein localization, costs for the affinity matrix and buffers, the possibilities of removing the fusion partner by site-specific cleavage and the environment tolerated by the target protein must be considered for deciding which affinity fusion partner to choose. Numerous affinity tags have been developed for affinity purification; these permit several alternative purification protocols. A selection of these systems is, the staphylococcal protein A (Nilsson and Abrahmsén,

1990), glutathione *S*-transferase (Smith and Johnson, 1988), maltose-binding protein (Guan *et al.*, 1988), and thioredoxin (LaVallie *et al.*, 1993), polyhistidine tag (Hochuli *et al.*, 1988), FLAG system (Hopp *et al.*, 1988), Strep-tag system and *in vivo* biotinylated affinity tags (Schatz, 1993).

Production of the recombinant protein with the affinity fusion partner has different application areas. Firstly, the affinity interaction can be used to immobilize enzymes and receptors on a solid support such as biosensors or affinity columns without any prior purification of the protein. Second, immobilization of the enzymes can be used to obtain bioreactors or to facilitate removal or reuse of the biocatalyst and to alter the characteristics of the immobilized enzyme or the intended reaction. Third, the immobilized fusion protein can be eluted and used directly for structural or functional studies if the target protein is unaffected by the affinity fusion partner or be used as immunogens for the generation and purification of antibodies.

1.1.2.1 Site-Specific Removal of Affinity Handles

For certain applications, it is necessary to remove the affinity fusion partner after the purification. A fusion partner can cause unwanted immune responses, alter the properties, or complicate a structural determination of a fused protein. The introduction of a recognition sequence between the fusion partner and the target protein allows for site-specific cleavage of the fusion protein to remove the affinity fusion partner. Table 1.2 summarizes several methods for site-specific cleavage to obtain the native gene product based on chemical and enzymatic treatments together with their cleavage specificities and cleavage conditions.

Chemical cleavage methods are relatively inexpensive but the rather harsh reaction conditions may lead to protein denaturation or amino acid side chain modifications of the target protein. Such methods are, therefore, primarily used in the preparation of peptides and small proteins. Still, use of chemical-cleavage method for certain products offers the attraction of being cost-effective and relatively easy to scale up

Table 1.2 Methods for Site-Specific Cleavage of Affinity Fusion Proteins (Nilsson *et al.*, 1997)

Cleavage agent	Cleavage specificity	Cleavage conditions
Chemical agents		
CNBr	-XM ↓ X-	70%formic acid, RT
Hydroxylamine	-XN ↓ GX-	pH 9, 45°C
Enzymes		
Enterokinase	-X$(D)_4$K ↓ X-	pH 8.0, 37°C
Factor Xa	-XIEGR ↓ X-	pH 7.2, RT
H64A substilisin	-XFAHY ↓ X-	pH 8.6, 37°C
Thrombin	-XLVPR ↓ GSX	pH 8.0, 25°C
IgA protease	-XPAPRPP ↓ TX-	pH 7.4, 37°C
Gst^b-protease 3C	-XLETLFQ ↓ GX-	pH 8.2, 4°C
ABT-protease 3C-His_6	-XLEALFQ ↓ GPX-	pH 8.0, RT

(Nygren *et al.*, 1994). Enzymatic methods are much preferred because they are generally more specific and the cleavage can usually be performed under mild conditions. Unfortunately endopeptidases also suffer from some limitations: (i) the presence of peptide secondary cleavage site activity leading to proteolytically damaged products, (ii) incomplete sample cleavage leading to product homogeneity which hampers crystallization and (iii) inhibition of cleavage by properly folded proteins requiring partial denaturation for successful fusion-tail cleavage. To date, viral proteases have proven to be the most selective and useful for structural biology studies (Stevens, 2000).

There are interesting examples that combine the affinity purification and protease activity together. Pharmacia Biotech presented a kit containing 3C protease of rhinovirus fused to an affinity partner (GST) to be used for site-specific cleavage of GST-affinity fusion proteins (PreScisson system). By applying simple affinity chromatography steps the target protein can be easily released with the affinity-

tagged protease and the tag removal. Moreover, incase the affinity ligand is insensitive to the proteases; the on-column cleavage of an affinity fusion protein immobilized onto an affinity column can be performed. A system, Impact I system, (New England Biolabs) has been described for use of intein from *Saccharomyces cerevisiae*, which was fused between the target protein at its N-terminus and a chitin-binding domain from *Bacillus circulans,* is capable of catalyzing peptide bond cleavage at one of its termini and separates the target protein from the fusion tag without the need for cleavage by a protease (Chong *et al.*, 1996).

1.1.3 Gene Fusions and Protein Secretion and Detection

Sometimes it is important for the gene fusion system to contain a secretion system in order to localize the product to a specific compartment eventhough the highest overall yield is accomplished in the bacterial cytoplasm. Such systems offer advantages for proteins requiring disulfide bond formation that are unstable against degradation due to its inefficient folding *in vivo* and for host systems if the recombinant protein is toxic to the cell. Normally secreted proteins staphylococcal protein A (SPA) (Nilsson and Abrahmsén, 1990), and maltose-binding protein (MBP) (Guan *et al.*, 1988) are the fusion partners that can be used to direct the secretion of recombinant proteins into the oxidizing environment of the *E. coli* periplasm. SPA has also been used to direct complete secretion into the growth medium (LaVallie *et al.*, 1993). In addition an *E. coli* thioredoxin homolog, DsbA can be used as a secretory fusion partner. DsbA, which functions both *in vivo* and *in vitro*, is the most important oxidase of free sulfudryl group in periplasm and promotes disulfide bond formation. Winter and co-workers (2000) reported the successful production of proinsulin in *E. coli* as a C-terminal fusion to DsbA.

Secretion-affinity fusion strategy combines the advantages of secretion and affinity purification. Secretion to the culture medium provides ample opportunities for continuous recovery of the product from the medium during fermentation in case of large-scale purification of the recombinant proteins.

For detection of the recombinant proteins calmodulin and *Aequorea victoria* green fluorescent protein (GFP) (Chalfie *et al.*, 1994) are the popular fusion partners. A highly acidic protein, calmodulin, displays calcium-dependent binding to peptide ligands, has been used as an affinity tag for single-chain Fv antibody fragments in *E. coli*. These calmodulin-tagged fusions can be efficiently visualized using fluorescent peptides, allowing convenient visualization of antibody-antigen interactions for application such as gel-shift assays and fluorescent-activated cell sorting (LaVallie *et al.*, 1993). GFP, a protein of 238 amino acids, is capable of producing a strong green fluorescence when illuminated with UV light when expressed in prokaryotic and eukaryotic cells. Since GFP appears to be non-toxic to the cells, both amino- and carboxyl-terminal fusions retain their fluorescence, fluorescence is stable and requires no additional gene products, it is an excellent portable marker for monitoring gene expression and protein localization in living cells (Chalfie *et al.*, 1994). In addition it has been reported that luciferases, capable of emitting visible light as one of the end-products of their catalysis, can be used for protein localization as well as in the development of sensitive bioluminescent enzyme immunosystems (Karp and Oker-Blom, 1999).

General overview of the most commonly used fusion partners gives us a chance to know them better and analyze their characteristics.

1.1.4 Commonly Used Fusion Proteins

1.1.4.1 Staphylococcal Protein A

Staphylococcal Protein A is a protein present on the cell wall of the pathogenic gram-positive bacterium *Staphylococcus aureus*. The strong and specific interaction between the SPA and the constant part (Fc) of the certain immunoglobins (IgG) has made it useful for the detection and purification of the fusion proteins in a single step by IgG affinity chromatography. SPA is capable of binding to the Fc portion of IgG from a large number of different species including man (Nilsson and Abrahmsén,

1990). Furthermore, several additional properties such as high stability against proteolysis, the absence of disulfides, the solubility of the N- and C-termini of the three-helix bundle structure of an individual IgG-binding domain which facilitates independent folding of the product and SPA, the possibility to introduce different protease recognition sequences for site-specific cleavage in order to release the target protein, SPA fusion secretion to the periplasmic space or in to the medium and high solubility have made SPA particularly suitable as fusion partner for the production of recombinant proteins. Domain B of SPA lacking methionine was designed to accomplish the site-specific cleavage with both CNBr and hydroxylamine (Stahl and Nygren, 1997). In most recently constructed SPA affinity fusion vectors, constructs with different multiplicities of this Z domain have replaced the original SPA fusion partner.

1.1.4.2 Glutathione S-Transferase

The glutathione *S*-transferases (GST) are a family of enzymes that can transfer sulfur from glutathione to substances such as nitro and halogenated compounds, leading to their detoxification (Nilsson *et al.*, 1997). GST can be purified by affinity chromatography on immobilized glutathione followed by competitive-elution with excess reduced glutathione and purification of many mammalian GST isozymes has been reported by this method. Based on this specific interaction Smith and co-workers (1988) described a gene fusion system for *E. coli* using a 26-kDa GST, encoded by the parasitic helminth *Schistosoma japonicum*, as a fusion partner. They demonstrated that a variety of eukaryotic polypeptides, fused to C-terminus of the GST, can be purified from *E. coli* lysate by glutathione-affinity chromatography. For the removal of the GST carrier, protease cleavage sites for thrombin and blood coagulation factor Xa have been incorporated into the C-terminus. A series of vectors for the production of the target protein by fusing at both N- and C-terminal of GST have been described (Sharrocks, 1994). Moreover, a dual affinity fusion system having GST at the N-terminal in combination with a His_6 tag at the C-terminus of a tripartite fusion protein has been constructed (Panagiotidis and Silverstein, 1995).

1.1.4.3 Maltose-Binding Protein

Maltose-binding protein, the product of the *mal*E gene of the *E. coli,* is exported to periplasmic space when it binds specifically with high to maltose or maltodextrins to transport these sugars across the cytoplasmic membrane. (Duplay *et al.*, 1984). Purification is accomplished in a single affinity chromatography step on resins containing cross-linked amylose followed by competitive elution with maltose. MBP as a fusion partner promises for development of expression systems, due to simple and feasible purification conditions with the low cost for amylose resins, and the fact that MBP does not contain any cysteine residues, which can interfere with disulfide bond formation within the target protein, supports this idea. The further advantages include the production in the intracellular form and secreted form by presence of a signal peptide (Guan *et al.*, 1988). Expression vectors allowing translational fusions to MBP in all reading frames or containing a cleavage recognition sequence for the factor Xa protease at the C-terminus of MBP were reported (Nilsson *et al.*, 1997). MBP has also been used as an N-terminal constituent in a dual affinity approach, in combination with a His_6 tag at the C-terminus with the choice of different alternatives including transcription promoters, selection markers and protease cleavage sites (Pryor and Leiting, 1997).

1.1.4.4 Thioredoxin

E. coli thioredoxin bears a number of characteristics which make it a suitable choice as a fusion partner. First of all, when overexpressed, thioredoxin can accumulate the target protein up to 40% of the total cellular protein which is still soluble and stable as well as biologically active. In contrast to other systems, since it is a small monomer (11,675 kDa) it compromises a relatively modest portion of any fusion. Both the N- and C-termini of thioredoxin are good positions for potential fusions to other proteins. One of the remarkable characteristics of *E. coli* thioredoxin is its ability to withstand prolonged incubation at elevated temperatures without undergoing irreversible thermal denaturation which enables heat treatments to be

used as an effective purification tool. The other characteristic is related to thioredoxin's cellular localization where it resides at sites around the inner periphery of the cytoplasmic membrane in *E. coli* known as adhesion zones. From this location thioredoxin can be released to the exterior of the cell by simple osmotic shock or freeze/thaw treatment in the presence of EDTA. To separate the thioredoxin from the target protein, a linker peptide which contains a recognition sequence for the intestinal protease enteropeptidase has been introduced between thioredoxin and fused C-terminal domains. In addition to these characteristics thioredoxin contains a permissive active-site loop which can be used for the high-level expression of soluble short peptides that are either rapidly degraded by host peptidases or accumulate as an inclusion body (LaVallie *et al.*, 1993).

1.1.4.5 Polyhistidine Tags

In 1975 Porath and co-workers introduced a new concept for affinity purification of proteins which is today known as immobilized metal ion affinity chromatography (IMAC). IMAC is based on the interaction between the immobilized metal ions (Cu^{2+}, Zn^{2+}, Ni^{2+}) and electron donor groups in the side chains of certain amino acids, particularly histidines, on a protein surface. The metal ions are immobilized on a hydrophilic support by the use of a chelating agent capable of presenting the metals for binding to the protein. Since then, several gene fusion systems employing IMAC by using the strong interaction between the histidine and Ni^{2+} have been described for purification of recombinant proteins. Polyhistidine tags (two to six histidine residues) can be fused to the gene of interest either N- or C-terminally. Adsorption of the poly-His-tagged proteins to a metal-chelate absorbent was performed at neutral or slightly alkaline pH, at which the imidazole group of the histidines is not protonated. Purification was done in a single step using a Ni^{2+}-nitrilotriaceticacid (NTA) and followed by elution with low pH or by competition using imidazole (Hochuli *et al.*, 1988). Alternative chelator like iminodiacetic acid (IDA) (Smith *et al.*, 1988; Ljungquist *et al.*, 1989; Van Dyke *et al.*, 1992) as well as other His-containing tag sequences such as variants also containing Trp residues (Smith *et al.*, 1988; Kasher *et*

al., 1993) or multiple copies of the peptide Ala-His-Gly-His-Arg-Pro (Ljungquist *et al.*, 1989) was also described. An interesting extension of the IMAC technology, termed as His patch thioredoxins, has been described. A metal ion binding site was incorporated onto the solubilizing *E. coli* thioredoxin protein where it combines two beneficial features: high solubility and the possibility of using the affinity chromatography for purification (Lu *et al.*, 1996).

1.1.4.6 FLAG Peptide

"Flag" is a short polypeptide marker, useful for recombinant protein identification and purification. The Flag sequence is a small hydrophilic peptide of eight amino acids, Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys, and it does not interfere with the native folding of the proteins, to which it is attached, can readily interact with the affinity purification substrate and can easily be removed from the target protein. N-terminal fusion of the Flag peptide to a variety of recombinant lymphokines has been reported. The purification is based on immunoaffinity chromatography on immobilized calcium dependent monoclonal antibody, which reacts with the marker sequence, and followed by EDTA elution (Hopp *et al.*, 1988). Alternatively, a different monoclonal antibody, capable of binding to the Flag peptide, is available for the fusion to either C-terminus of the target protein or the N-terminus if it is preceded by a methionine. For elution low pH or competition with an excess of synthetically produced Flag peptide is used since the antibody interacts with the Flag peptide in a non-calcium-dependent manner (Nilsson *et al.*, 1997). Finally, Flag peptide contains an internal enterokinase cleavage recognition sequence, which is specific for the five C terminal amino acids (Asp-Asp-Asp-Asp-Lys) of the marker sequence in order to obtain the target protein (Hopp *et al.*, 1988).

1.1.4.7 *In Vivo* Biotinylated Affinity Tags

The binding affinity between biotin and streptavidin or avidin is among the strongest noncovalent bond known to exits ($K_d = 10^{-15}$ M). Once formed, it is highly resistant

to denaturing reagents, extremes in pH and temperature, protein digestion; therefore biotinylation is frequently used for immobilization, purification, labeling, and detection purposes. For use in such applications, biotinylation is generally accomplished using chemical reagents which lack site specificity. This might lead to the inactivation of proteins. Cronan (1990) demonstrated an elegant strategy for site-specific biotinylation of recombinant proteins during their production, utilizing the *in vivo* biotinylation machinery of *E. coli*. *E. coli* biotin carboxyl carrier protein (BCCP) of acetyl-CoA carboxylase is one of the few proteins that are naturally modified by biotin incorporation. By the action of biotin ligase (BirA), the biotin is covalently attached to a specific lysine residue. The C-terminal 101 residues of the *E. coli* BCCP protein have been utilized successfully as an affinity handle for the production and purification of a Fab antibody fragment in *E. coli*. A 75 amino acid domain of the 1.3S subunit of *Propioniacterium shermanii* transcarboxylase was the previously known as the smallest substrate for biotinylation (Cronan, 1990). Recently, short tags replace the larger *in vivo* biotinylated fusion partners. A 13-residue peptide, selected from a random *lacI* library and capable of mimicking the normal BirA substrate, has been reported for its ability to be biotinylated *in vivo* (Schatz, 1993). Alternatively, a 10 amino acid C-terminal peptide (Strep-tag, nonbiotinylated), selected from a random peptide library, has been described for its ability to bind to streptavidin. This peptide mimics the biotin and used as an affinity fusion partner with the streptavidin-affinity chromatography purification (Nilsson *et al.*, 1997).

1.2 Citrate Synthase: As A Fusion Partner

Extremophiles are a source of enzymes (extremozymes) with their extreme stability and their successful activity under adverse conditions for life. The understanding of the structural basis of both the stability and the activity of these enzymes not only allow the manipulations for further extremes but also give the opportunity to engineer such structural features into mesophilic enzymes of choice (Hough and Danson, 1999).

Citrate synthase is among the enzymes under consideration to investigate the structural basis of thermostability. Citrate synthase, the first enzyme of the citric acid cycle, catalyzes the reversible condensation of the oxaloacetate (OAA) and acetyl coenzyme A (acetyl-CoA) to form citrate and coenzyme A (CoA) (Figure 1.2). This enzyme is one of few enzymes that can directly form a carbon-carbon bond with no requirement of unusual cofactors and metal ions for activity. In addition, it was the first enzyme shown to contain almost entirely of helices (Remington, 1992).

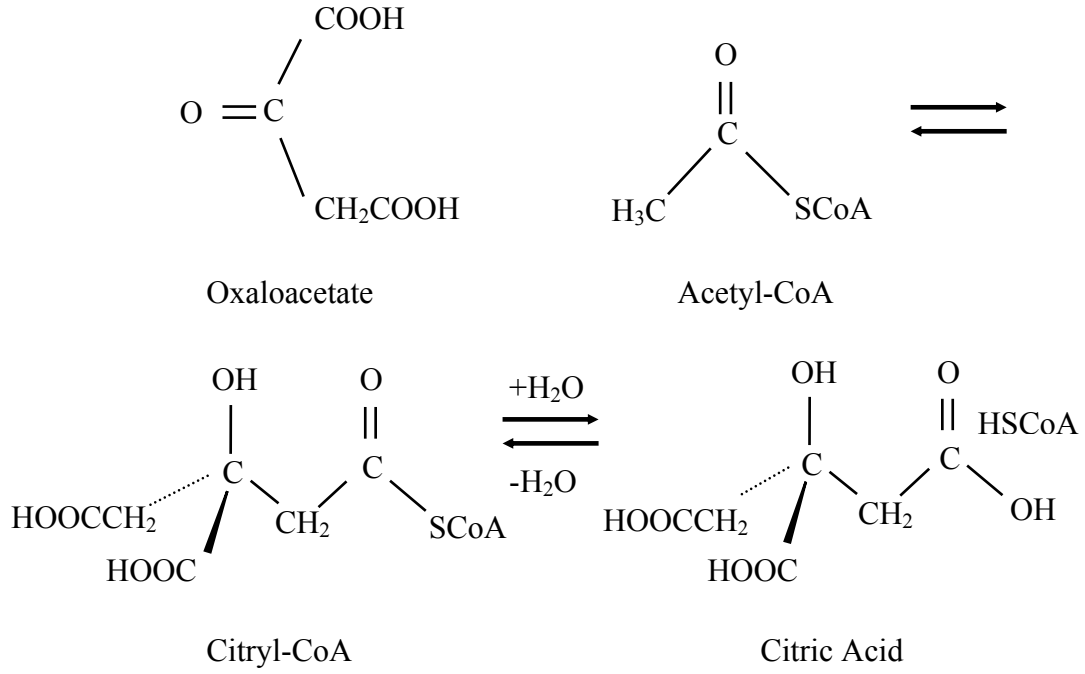


Figure 1.2 Schematic drawing the reaction that is catalyzed by citrate synthase, and thus affects the entry of carbon into the citric acid cycle (Remington, 1992).

Citrate synthase has been studied from organisms that represent all three domains, which possesses a diversity of structure, activity and regulation. The Gram-negative eubacteria carry a homo-hexameric form of citrate synthase, allosterically inhibited by NADH, whereas a dimeric form, isosterically inhibited by ATP and also by 2-oxoglutarate in the case of facultative anaerobes, is found in Gram-positive

eubacteria, archaea and eukaryotes (Smith *et al.*, 1987). Despite this oligomeric variation all citrate synthases have polypeptide subunits of M_r 42 000-49 000, and their sequence is homologous (Gerike *et al.*, 1998).

To better understand the molecular diversity of citrate synthases, and also the correlation of oligomeric structure and function, studies have focused on the enzyme from the archaebacteria. *Thermoplasma acidophilum* (*Tp. acidophilum*) is one of the most studied organisms about citrate synthase. *Tp. acidophilum* is a thermophilic archaebacterium growing optimally at pH 1 to 3 and 55°C to 60°C and possesses a dimeric form of citrate synthase with a subunit M_r 43 000. The gene encodes a protein of 384 amino acids (Sutherland *et al.*, 1990).

The overall folded conformation *Tp. acidophilum* citrate synthase comprises 16 α-helices per monomer, which is equivalent to 57% of residues. Each subunit of dimer consists of two domains, a large domain and a small domain. The larger domain contains 11 helices (residues 1-224 and 326-384; helices C-G, I-M and S), and a small domain contains 5 helices (residues 225-325; helices N-R). A small section of surface-accessible antiparallel β-sheet, present as 3 small strands is formed by residues 21-36 (Figure 1.3a). The active site of citrate synthase situated in the cleft between the large and small domains of each monomer, and each contributes functional groups to the active site of the other (Russell *et al.*, 1994). The thermostability studies indicated that the large domain, which is responsible for forming the majority of the inter-subunit contacts, is being the main determinant of the enzymes' thermostability where the function of the small subunit is primarily catalytic (Arnott *et al.*, 2000).

Upon dimerization, 20% of the accessible surface area of each monomer is calculated to be buried. The subunit interface consists of an 8 α–helix sandwich, with four pairs of antiparallel helices (FF', GG', LL' and MM') whose axes lie approximately perpendicular to the two-fold axis relating the monomers. A pair of antiparallel helices (I and S) is wrapped around the sandwich and both of them bend smoothly (Figure 1.3b) (Russell *et al.*, 1994).

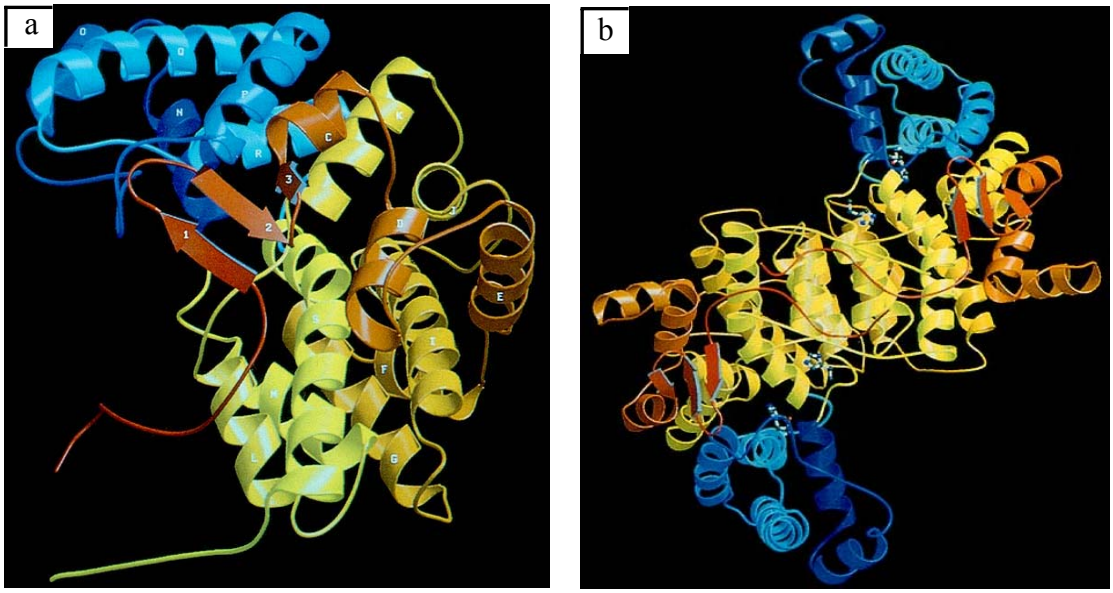


Figure 1.3 (a) Schematic representation of the *Tp. acidophilum* citrate synthase monomer **(b)** Schematic representation of the *Tp. acidophilum* citrate synthase dimer (Russell *et al.*, 1994).

Detailed comparison of the 3D-crystal structures of the CSs from hyperthermophilic *Pyrococcus furiosus* (Russell *et al.*, 1997), thermophilic *Thermoplasma acidophilum* (Russell *et al.*, 1994) and *Sulfolobus solfataricus* (Bell *et al.*, 2002), and mesophilic pig heart (Remington *et al.*, 1982) and psychrophilic *Arthrobacter* strain DS2-3R (Russell *et al.*, 1998) has allowed a structural analysis to investigate the structural mechanisms underlying protein thermostability.

Despite their close degree of similarity, the studies revealed that there is a concomitant reduction in the size of the loop regions, an increase in packing density and a reduction in both the size and the number of internal cavities, and also greater surface complimentarity at the dimer interface. All these measures serve to reduce the conformational flexibility of the protein and therefore increase its thermostability (Russell and Taylor, 1995). Apart from these differences, a marked increase in aromatic interactions within the small domain of *Tp. acidophilum* citrate synthase

and a complete absence of cysteine is observed since thermolabile amino acid residues have a tendency to undergo oxidation (Cys, Met) or deamination (Asn, Gln) at high temperatures (Russell *et al.*, 1994).

"The more rigid assembly of the monomer subunits without losing activity" strategy was employed by thermostable CSs as a thermostability and has been investigated the most intensively (Russell *et al.*, 1997; Russell *et al.*, 1994; Bell *et al.*, 2002; Russell *et al.*, 1998). *Tp. acidophilum*'s dimer interface as compared to that of pig CS is more hydrophobic as a consequence of serine/threonine to alanine exchanges in helices G and M (from three to eight) of the 8-helical sandwich (Figure 1.4) (Russell *et al.*, 1994). Alanine is an excellent helix-forming amino acid and higher

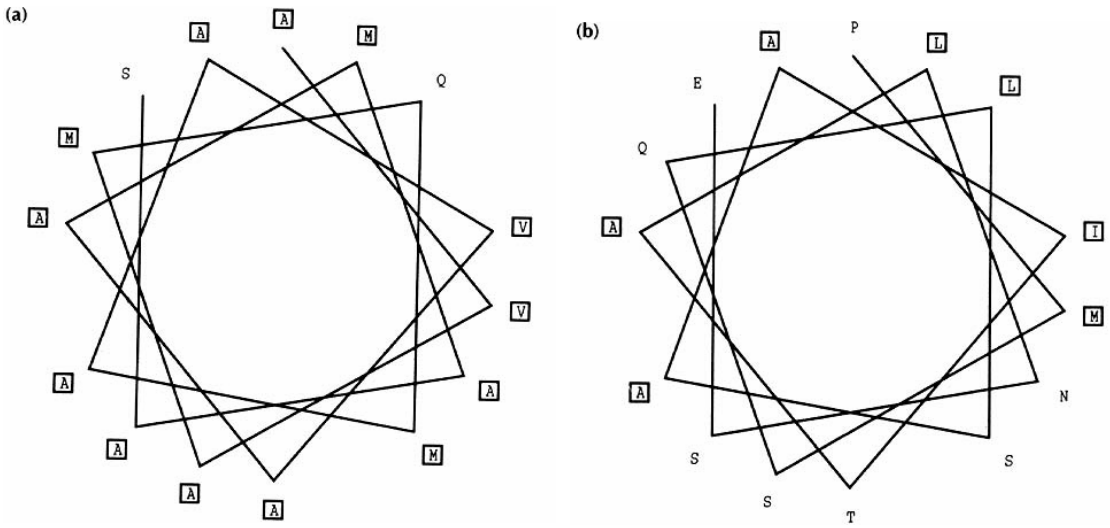


Figure 1.4 Helical wheel of helix G **(a)** *Tp. acidophilum* (residues 95-109) and **(b)** pig heart (residues 137-151) citrate synthases. Boxed residues indicate hydrophobic residues (Russell *et al.*, 1994).

content of alanine residues contribute to the thermostability not only by significantly increasing the hydrophobicity of the subunit interaction, but also through the stabilization of the individual helices. A site specific mutagenesis study showed that

the exchange of Gly209, which is located in an interface-helix (helix M) of *Tp. acidophilum* citrate synthase, by alanine, increased the hydrophobicity in the subunit interface, thus stabilize the α-helix and therefore improve the thermostability of the citrate synthase by 23-fold (Erduran and Kocabıyık, 1998). Another site-directed mutagenesis study revealed that the replacement of Gly196 by Val on the helix L of the subunit interface of citrate synthase from *Tp. acidophilum* not only decreased the thermal and chemical stabilities, but also reduced the catalytic activity since valine is invariably a poor helix former, and therefore leads to helix destabilization and diminishes the hydrophobic effect (Kocabıyık and Erduran, 2000).

On the other hand, there is a significant increase in the intersubunit ion pairs no increase in hydrophobicity in the extremely thermophilic *Pyrococcus furiosus* CS (Figure 1.5) (Russell *et al.*, 1997). Many of the additional ion-pairs in *Pf*CS compared to pig CS are located primarily in helix G and helix M. In *Sulfolobus* which has growth optimum between *Tp. acidophilum* and *Pyrococcus furiosus* shows increased hydrophobic and ionic interactions as compared to *Ta*CS in the same helices of the 8-helix sandwich (Bell *et al.*, 2002).

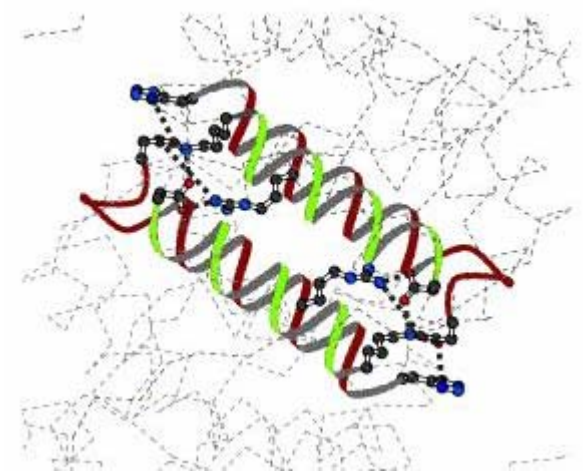


Figure 1.5 View down the twofold axis of the *Pyrococcus furiosus* dimer, showing the five-membered ionic network at each end of helices G (brown) and M (green), which constitute two of the four antiparallel of helices at the subunit interface (Danson and Hough, 1998).

1.3 Haloalkanoic Acid Dehalogenases: As A Passenger Protein

Halogenated organic compounds play a major part in the industrialized world. Although a portion of these chemicals are generated by naturally occurring biotic and abiotic processes in the oceans and atmosphere, the wide spread use of halogen-based chemistry in industrial-scale chemical processing over the past 100 years has introduced many additional man-made halocarbons into the environment. They are widely used not only because of their chemical and thermal stability but also they find applications as solvents (e.g. chloroform, trichloroethylene and chlorobenzene), degreasers, flame-retardants, insulating liquids (polychlorinated biphenyls), wood preservatives (pentachlorophenol), herbicides, fungicides, pesticides, starting materials for the synthesis of various pharmaceuticals and chemical intermediates used in the manufacture of high-volume chemicals. They are produced in large quantities and represent an important class of environmental pollutants that are ubiquitous in the environment and affect the health of humans and wild life (Swanson, 1999). Halogenated hydrocarbons are extremely persistent in nature. The biological recalcitrance is related to the number, species, and position of halogen substituents on the molecule (Hardman, 1991).

Biocatalysis of the carbon-halogen bond is achieved by a variety of microbial enzyme systems that have been found to catalyze the cleavage of carbon-halogen bonds by specific enzymes called “dehalogenases” and provide the means for these compounds to be utilized as carbon source or alternative electron acceptors. As a general rule, the cleavage of the carbon-halogen bond has been recognized as the most critical step in the microbial dehalogenation of these compounds; once the halogen substituent has been removed, the product can easily be assimilated by the organisms’ central metabolism (Hardman, 1991). Halogen removal reduces both recalcitrance to biodegradation and the risk of forming toxic intermediates during subsequent metabolic steps (Janssen *et al.,* 2001). Since nature itself produces a variety of halogenated compounds in large amounts, microorganisms have evolved a diverse potential to transform and degrade halogenated compounds.

Studies on dehalogenases have revealed much about the catabolic systems involved in the cleavage of the carbon-halogen bonds as well as the molecular mechanisms. The 2-haloacid dehalogenases, for being one of the best studied among the dehalogenases in terms of microbiology, biochemistry, reaction mechanism and genetics, are focus of our research. The economic and environmental importance of these enzymes has driven researches for industrial synthetic processes and into environmental protection technology.

The dehalogenases are classified into three groups according to their mode of action: (i) Reductive dehalogenations are co-metabolic processes, or they are specific reactions involved in substrate utilization (carbon metabolism) or coupled to energy conservations (RCH_2-X $\Rightarrow$ RCH_3) (ii) In dehydrohalogenase-catalyzed dehalogenation, halide elimination leads to the formation of a double bond ($RCHXCH_3$ $\Rightarrow$ $RCHCH_2$) (iii) Substitutive dehalogenation in most cases is a hydrolytic process, catalyzed by halidohydrolases (RCH_2-X $\Rightarrow$ RCH_2OH) (Fetzner, 1998).

The largest of the dehalogenase group enzymes are hydrolytic dehalogenases. In the reactions catalyzed by these enzymes, halogen substituent is replaced in a nucleophilic substitution reaction by a hydroxyl group, which is derived from water. Hydrolytic dehalogenation of heterocyclic, aromatic, and aliphatic compounds has been reported. Hydrolytic dehalogenases that have recently been studied in more detail are the haloacid dehalogenases, which can be divided into at least two groups: group I and II (Hill *et al.,* 1999). Group II enzymes use covalent catalysis, and usually convert L-2-chloropropionic acid, but not the D-2-chloropropionic acid, with an inversion of configuration at the chiral center. The group II dehalogenases have a structure that is very similar to phosphatases, the catalytic subunits of P-type ATPases, and several other enzymes of the haloacid dehalogenase superfamily. They have a conserved nucleophilic aspartate residue that is located close to the N-terminus and is involved in formation of the covalent intermediate.

Group I haloacid dehalogenases, which include a remarkable enzyme selective for D-2-chloropropionic acid and an enzyme that converts trichloroacetate to carbon monoxide, do not use a covalent mechanism for catalysis. Nardi-Dei *et al.* (1999) proposed that the dehalogenase directly activates a water molecule to attack the α-carbon of the 2-haloalkanoic acid, thereby displacing the halogen atom.

Microorganisms including *Pseudomonas* sp. (Liu *et al.*, 1994; Murdiyatmo *et al.*, 1992; Smith *et al.*, 1990; Motosugi *et al.*, 1982; Weightman *et al.*, 1982), *Xanthobacter autotrophicus* GJ10 (Van Der Ploeg *et al.*, 1991) and also *Moraxella* sp. strain B (Kawasaki *et al.,* 1992) are reported as the producers of these 2-haloacid dehalogenases. More than 10 L-2-haloacid dehalogenases have been cloned and sequenced. Although DNA sequence homology among these dehalogenase genes is low, the amino acid sequence homology for the enzymes is high, ranging from 37 to 68% (Pang and Tsang, 2001).

A large group of substrate for hydrolytic dehalogenases is halogenated aliphatic hydrocarbons, which are widely used as herbicides and solvents. Some of these compounds are carcinogenic (e.g., vinyl chloride), whereas others (e.g., 1,2-chloroethane) can be biologically or chemically converted into mutagenic products and two major groups of halogenated aliphatic compounds are of current importance: the haloalkanes and the haloalkanoic acids (chloroacetic acid, dichloroacetic acid, bromoacetic acid, fluoroacetic acid 2-choloropropionic acid, and 2,2-dichloropropionic acid) (Müller and Lingens, 1986).

Haloalkanoic acids are widely used as herbicides, fungicides, and insecticides and they are the most susceptible group of compounds to enzymatic degradation. The classification of haloalkanoic acid dehalogenases based on substrate activity divides into two groups: haloacetates and halopropionates.

Hydrolytic dehalogenases have been attracting a great deal of attention because of their possible application to find chemical synthesis and bioremediation of halo compound-polluted environment.

Optically active 2-hydroxylalkanoic acids are useful as starting materials for the synthesis of various pharmaceuticals and agrochemicals. Bacterial 2-haloacid dehalogenases catalyze the hydrolytic dehalogenation of 2-haloalkanoic acids, and are applicable to the production of optically active 2-hydroxylalkanoic acids (Motosugi *et al.*, 1984). Dehalogenases are also expected to play a role in bioremediation of halo compound-polluted environments because they detoxify various organohalogen compounds (Hardman, 1991).

Pseudomonas sp. strain CBS3 possesses four different dehalogenating enzymes. One of these enzymes called 4-chlorobenzoate dehalogenase, whereas the other one is a two-component enzyme system termed 4-chlorophenylacetic acid 3,4- dioxygenase. The last two of them are the two 2-haloalkanoic acid dehalogenases (*dehC*I and *dehC*II) that utilizes monochloroacetate and 2-monochloropropionate as a carbon and energy source. Both enzymes dehalogenate only L-2-monochloropropionate to D-lactate, but neither of them showed activity with D-2-monochloropropionate. The molecular weight of these two enzymes is around 25 kDa and *dehC*II has been used as a fusion partner in our construction experiments (Schneider *et al.*, 1991).

1.4 Aminoglycoside-3′-Phosphotransferase-II: As A Second Passenger Protein

Antibiotics, which are typically antibacterial drugs, have proven to be of great value in medicine and have revolutionized the treatment of infectious diseases. The largest class of clinically relevant antibiotics, apart from the β-lactams, function by inhibiting protein translation. The RNA and protein machinery of the prokaryotic ribosomes is sufficiently different from the analogous eukaryotic machinery to make protein synthesis a viable target for the design of antibiotics, and there are several families of compounds that function at this site, including the macrolides of the erythromycin class, the tetracyclins and the aminoglycosides (Walsh, 2000).

The aminogylcosides are a class of broad-spectrum antimicrobial compounds, originally isolated from soil bacteria including various species of *Streptomyces* and

Micromonospora (Table 1.3). Streptomycin, the parent member of the family, was the first chemotherapeutic agent to be effective against *Mycobacterium tuberculosis*. Gentamycin, neomycin, amikacin and kanamycin are the other members of the family and they generally consist of a central aminocyclitol ring (typically known as the B ring) with two to three substituted aminoglycan rings (A, C and D) attached at different positions (Figure 1.6) (Nurizzo *et al.*, 2003).

Table 1.3 Aminoglycoside antibiotics and their source organisms (Davies and Wright, 1997)

Aminoglycoside	**Source organism**
Kanamycin	*Streptomyces kanamyceticus*
Streptomycin	*Streptomyces griseus*
Gentamicin	*Micromonospora purpurea*
Spectinomycin	*Streptomyces spectabilis*
Butirosin	*Bacillus circulans*
Tobramycin	*Streptomyces tenebrarius*
Neomycin	*Streptomyces fradiae*
Amikacin	Semisynthetic derivative of kanamycin
Netilmicin	Semisynthetic derivative of sisomicin
Isepamicin	Semisynthetic derivative of gentamicin B

Aminoglycoside antibiotics are used extensively for the treatment of many bacterial infections and combination therapy with penicillin, ampicillin and vancomycin is often recommended to treat serious enterococcal infections, endocarditis, urinary tract, and intraabdominal infections. Apart from their use for the treatment of a variety of diseases, aminoglycosides have numerous applications such as an anthelmintic agent and as glycohydrolase inhibitors (Davies and Wright, 1997).

Aminoglycosides exert their antimicrobial effects by binding to the 30S subunit of the bacterial ribosome and obstructing protein biosynthesis. High resolution crystallographic studies have recently determined the binding sites for paromomycin,

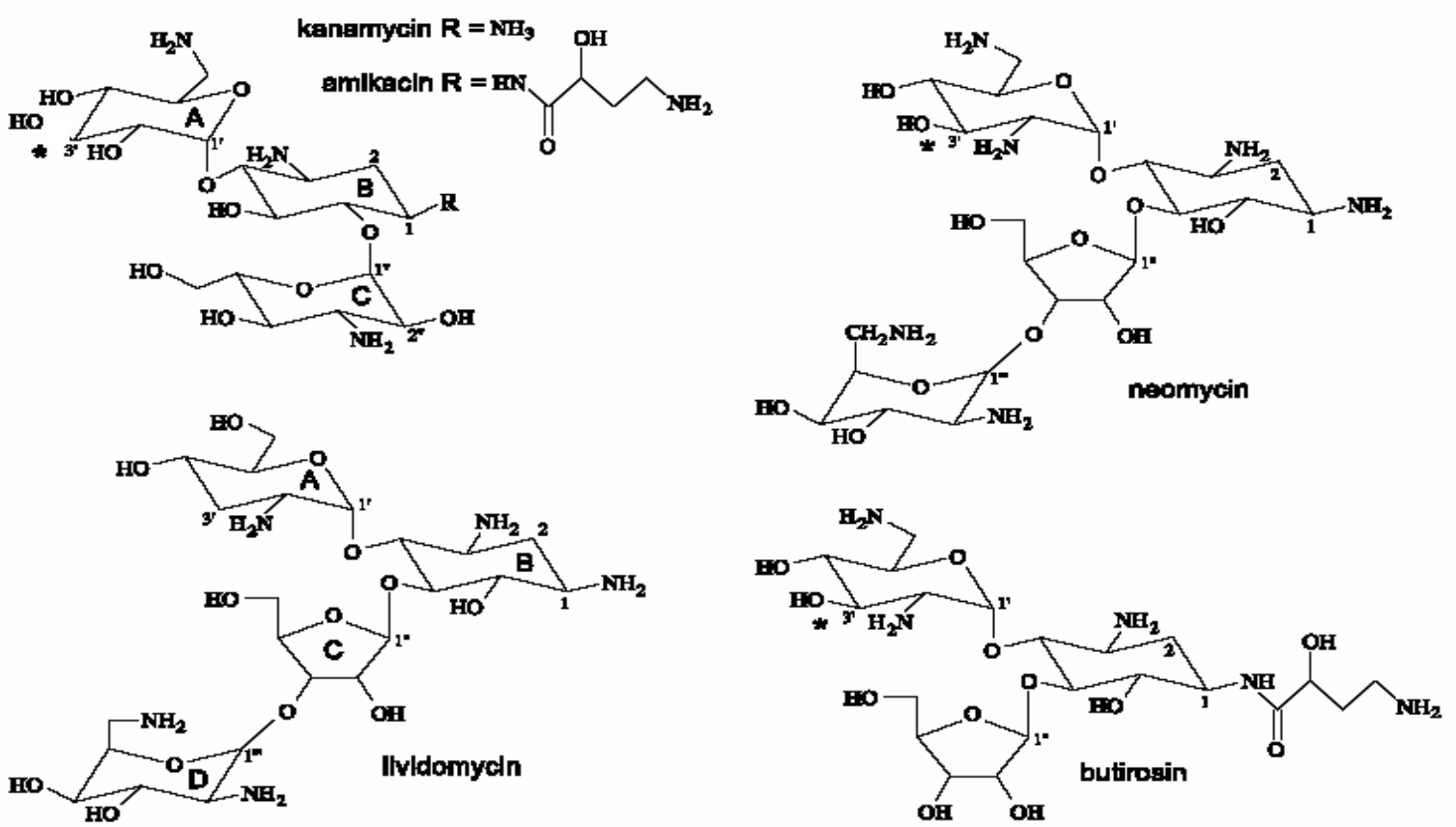


Figure 1.6 Structures of some aminoglycoside antibiotics. The ring nomenclature used here (A, B, C and D) is shown for kanamycin and lividomycin (Nurizzo *et al.*, 2003).

streptomycin, and spectinomycin (Carter *et al.*, 2000). Although the exact mechanism still remains unclear it is known that drug binding decreases the dissociation rate of aminoacyl-tRNA, promoting miscoding. It has also been suggested that aminoglycoside binding leads to the formation of non-functioning initiation complexes possibly due to a tighter binding of the 30S and 50S ribosomal subunits.

The huge consumption of antibiotics, both in man and animals, have resulted the scientific and popular press to report the emergence and spread of a vast amount of antibiotic resistance determinants among bacterial populations over the last years. The mechanisms of antibiotic resistance developed by bacteria are highly diverse. Resistance to aminoglycosides can occur by four different mechanisms: enzymatic modification of the antibiotic, alterations in the influx system, mutation in the ribosomal targets (16S rRNA, proteins) or enzymatic modification of 16S rRNA (Baquero and Blasquéz, 1997). The most prevalent source of clinically relevant resistance is conferred by aminoglycoside-modifying enzymes because these enzymes provide high-level resistance to the drugs that they catalyze cofactor-dependent modification of hydroxyl or amino groups on the aminocyclitol residues and they are encoded on extrachromosomal elements such as transposons and bacterial plasmids They thought to have originated as self-defense mechanisms used by the microorganisms that produce drugs (Davies and Wright, 1997).

Three classes of aminoglycoside-modifying enzymes have been identified for aminoglycoside resistance: ATP-dependent *O*-phosphotransferases (APH), ATP-dependent *O*-adenyltransferases (ANT), acetyl CoA-dependent *N*-acetyltransferases (AAC) (Table 1.4). The aminoglycoside-modifying enzymes are encoded by a large group of genes and over 50 different enzymes have been identified as aminoglycoside modifiers. At the nucleotide sequence level, genes are unrelated and derived from a variety of different microbial organisms. The enzymes are named according to their reaction that they catalyze and the position on the aminoglycoside at which they act (Davies and Wright, 1997).

Table 1.4 Aminoglycoside-modifying enzymes (Davies and Wright, 1997)

Enzyme	Typical substrate
O-Phosphotransferases (APH)	
2′′-b	**Kanamycin, gentamicin, tobramycin**
3′ (-5′′)-	**Kanamycin, neomycin**
3′′-	**Streptomycin**
4-	**Hygromycin**
6-	**Streptomycin**
N-Acetyltransferases (AAC)	
1-	**Apramycin, paromomycin**
2′-	**Gentamicin, tobramycin**
3-	**Kanamycin, gentamicin, tobramycin**
6′-b	**Kanamycin, gentamicin, tobramycin**
O-Adenyltransferases (ANH)	
2′′-	**Gentamicin, tobramycin**
3′′	**Streptomycin**
4′-	**Kanamycin, neomycin**
6-	**Streptomycin**
9-	**Spectinomycin**

Over 20 distinct aminoglycoside phosphotransferases enzymes are known and as a whole they show significant sequence similarity. Beck and co-workers (1982) reported a gene from *E. coli* transposon Tn5 which codes for an aminogylcoside 3′-phosphotransferase II (APH(3′)-IIa) and confers resistance to the aminoglycosides kanamycin and neomycin. The gene codes for a protein of 264 residues with a calculated M_r value of 26,053. The crystal structure study of APH(3′)-IIa revealed that the APH(3′)-IIa molecule is folded into two structural domains (Figure 1.7). The N-terminal domain (residues 10-96) starts with a short α-helix (α1) followed by three antiparallel β-strands (β1- β3). A second α-helix (α2) is inserted between

strands β3 and β4, and the domain is completed by additional strand (β5). A short linker peptide (residues 94-101) joins the N-terminal domain to the C-terminal lobe. The polypeptide chain then forms apart of the C-terminal domain (residues 101-132) comprising a short loop and a long helix (α3), followed by two helixes (α4 and α5), which form part of a separate helical sub-domain (residues 132-182). The remainder of the central core of the C-domain (183-250) comprises a four-stranded β-finger (β6-β9), which lies across the top of the helix α3 and protrudes between α3 and the interdomain linker, followed by a second long helix (α6) which lies antiparallel to

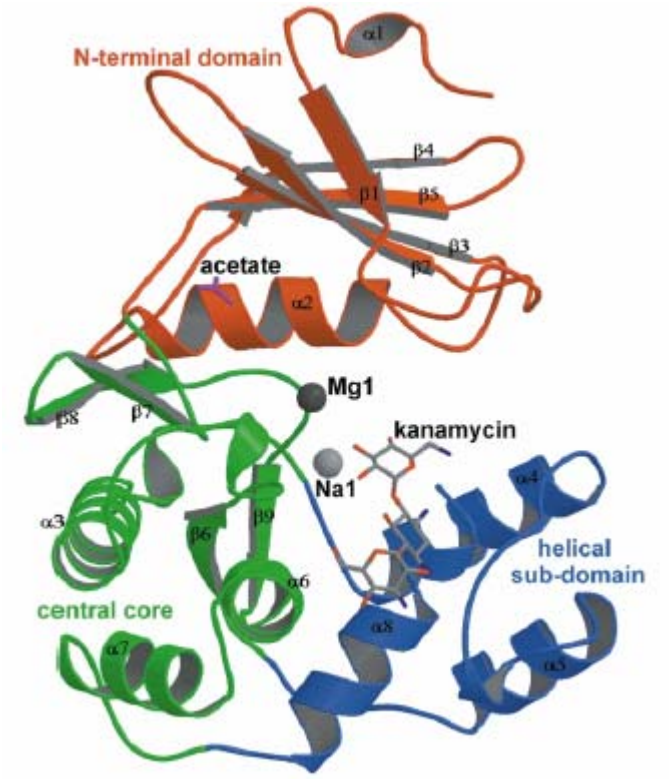


Figure 1.7 Ribbon representation of APH(3′)-IIa, showing the N-terminal domain (red), the central core of the C-terminal domain (green) and the helical sub-domain (blue). The kanamycin substrate is shown sandwiched between the central core and the helical sub-domain, along with the sodium ion (light gray sphere) and the Mg^{2+} (black sphere). The acetate ion bound in the adenine-binding pocket is shown as purple sticks (Nurizzo *et al.*, 2003).

α3, and a short helix (α7). The final C-terminal (α8) completes the small helical sub-domain. The active site lies between the two domains and compromises both the ATP binding site and the aminoglycoside binding site (Nurizzo *et al.*, 2003).

1.5 Scope and Aim of the Study

The extremophilic nature of many Archaea has stimulated intense efforts to understand their ability to remain catalytically active under extremes of temperature, salinity, pH and solvent conditions. Understanding the structural basis of both the stability and the activity of these enzymes under extreme conditions allow manipulating these enzymes to further extremes as well as permitting the attempts to engineer such features into mesophilic enzymes of choice. Thermostability is one of the popular topics under investigation.

The aim of this study is the construction of various fusion proteins by using two strategies (direct ligation and Gene SOEing) with a major purpose of enhancing the thermostability of mesophilic enzymes. As fusion partners, stability domain from *Thermoplasma volcanium* citrate synthase, dehalogenase II from *Pseudomonas* sp. and aminogylcoside 3′-phosphotransferase II of *E. coli* were employed. This approach not only presents a novel application of the gene fusion strategy but might also be used as a tool for the analysis of thermostability.

CHAPTER II

MATERIALS AND METHODS

2.1 Materials

2.1.1 Chemicals, Enzymes and Kits

Ethidium bromide, agarose low gelling temperature, dipotassium hydrogen phosphate ($K_2HPO_4.3H_2O$), potassium acetate (CH_3COOK), phosphoenolpyruvate (PEP), NADH, ATP, ampicillin, neomycin, 5-Bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-Gal), isopropyl β-D-1-thiogalactopyranoside (IPTG), dithiothreitol (DTT), piperazine-N,N′-bis-2-ethanesulfonic acid (PIPES), lactate dehydrogenase and pyruvate kinase were purchased from Sigma Chemical Co., St. Louis, MO, USA. Bromophenol blue ($C_{19}H_{10}Br_4O_5S$), ethylenediaminetetraacetic acid (EDTA), α-D-glucose, sucrose, Tris-base, magnesium chloride hexahydrate ($MgCl_2.6H_2O$), magnesium acetate ($(CH_3COO)_2Mg$), potassium dihydrogen phosphate (KH_2PO_4), potassium hydrogen phthalate, tryptone, yeast extract, sodium chloride (NaCl), sodium hydroxide (NaOH), ferric ammonium sulfate, mercuric thiocyanate, monochloroacetic acid (MCA) and sulfuric acid (H_2SO_4) were from Merck, Darmstadt, Germany. Agarose was from Boehringer Mannheim, Germany. Agar was from Difco, Detroit, USA. Ethanol was from Reidel-de Häen, Seelze, Germany. N,N dimethyl formamide was from Fluka AG, Bucks SG, Switzerland.

Taq DNA polymerase was purchased from Sigma Chemical Co., St. Louis, MO, USA. *Pfu* DNA polymerase was from Promega Co., Madison, WI, USA. T4 DNA Ligase, restriction endonucleases *Eco*RI, *Hin*dIII, *Sph*I, *Sac*I, *Kpn*I, *Pst*I, *Bam*HI,

*Sal*I and their buffers were from MBI Fermentas AB, Vilnius, Lithuania. Restriction endonucleases *Ava*II, *Pvu*II and their buffers were from Roche, Basel, Switzerland.

Wizard® Plus SV Miniprep DNA Purification System Kit and pGEM®-T Vector Systems II Cloning Kit were purchased from Promega Co, WI, USA. Concert Gel Extraction Systems Kit was from Life Technologies Gibco BRL Grand Island, NY, USA. DNA Extraction Kit was from MBI Fermentas AB, Vilnius, Lithuania. QIAGEN MinElute PCR Purification Kit and QIAGEN PCR Cloning Kit were from Qiagen Inc., Valencia, USA. All of the primers were synthesized by Applichem, Darmstadt, Germany.

2.1.2 Buffers and Solutions

The preparation and the composition of the buffers and solutions used in this study are listed in Appendix A.

2.1.3 Plasmid Vector, Molecular Size Markers and Genomic DNA Sequence Data

The map of cloning vector is illustrated in the Appendix B. DNA molecular size markers are given in Appendix C. Nucleotide sequence of citrate synthase from *Thermoplasma volcanium*, 2-haloalkanoic acid dehalogenase II from *Pseudomonas* sp. Strain CBS3, aminogylcoside 3′-phosphotransferase II from *E. coli* transposon Tn5 with marked primer sequences are given in Appendix D.

2.1.4 Strain and Medium

2.1.4.1 Bacterial Strains, Growth and Maintenance

E. coli pBSK(2CS) (expressing recombinant *Tp. acidophilum* CS), *E. coli* pDrive-CS3-1 (expressing recombinant *Tp. volcanium* CS), *E. coli* pBS-aph (expressing recombinant APH(3′)-II) and *E. coli* pUKS 107 (expressing recombinant *dehC*II

from *Pseudomonas* sp. Strain CBS3) cultures were grown at 37°C on Luria-Bertani (LB) agar plates supplemented with ampicillin (final concentration 50 μg/ml) and renewed monthly.

E. coli strain TG1 (*supE hsd*Δ5 *thi*Δ(*lac-proAB*) F[*traD*36 *proAB*$^{+}$ *lacl*^{q}*lac* ZΔM15]) was used as the recipient strain in the transformation studies. *E. coli* cells were grown at 37°C on Luria-Bertani (LB) agar plates supplemented with ampicillin (final concentration 50 μg/ml) and cultures renewed monthly.

2.2 Methods

2.2.1 Isolation of Recombinant Plasmids

Recombinant plasmids, pBS(+) vector containing the complete *Tp. acidophilum* citrate synthase (CS) gene (Figure 2.1, 2.2 and Table 2.1), pUC18 vector containing the complete *Pseudomonas* sp. Strain CBS3 2-haloalkanoic acid dehalogenase II (*dehC*II) gene (Figure 2.3, 2.4 and Table 2.2), and pBS(+) phagemid vector containing the complete *E. coli* transposon Tn5 aminogylcoside 3´-phosphotransferase II (APH(3´)-II) gene (Figure 2.5, 2.6 and Table 2.3) were isolated using Wizard® Plus SV Miniprep DNA Purification System Kit according to the manufacturers' instructions (Promega Co., Madison, WI, USA). *E. coli* cells which are harboring the recombinant plasmids were inoculated into a 10 ml LB broth supplemented with 50 μg/ml ampicillin at a final concentration, and incubated overnight at 37°C with vigorous shaking at 160 rpm (Heidolph UNIMAX1010, Heidolph Instruments GmbH, Kelheim, Germany). The cells were harvested by centrifugation at 4 000 rpm for 15 min (Herause Sepatech, GmbH, Osterode, Germany) and the pellet was thoroughly resuspended with 250 μl Cell Resuspension Solution. The cells were lysed by adding 250 μl Cell Lysis Solution and the proteins were digested by adding 10 μl Alkaline Protease Solution with the incubation at room temperature for 5 minutes. After the addition of 350 μl Neutralization Solution and incubation for 5 minutes on ice, the cell debris was centrifuged at 13 000 rpm for 10 minutes (Herause Sepatech, GmbH, Osterode, Germany). The supernatant was

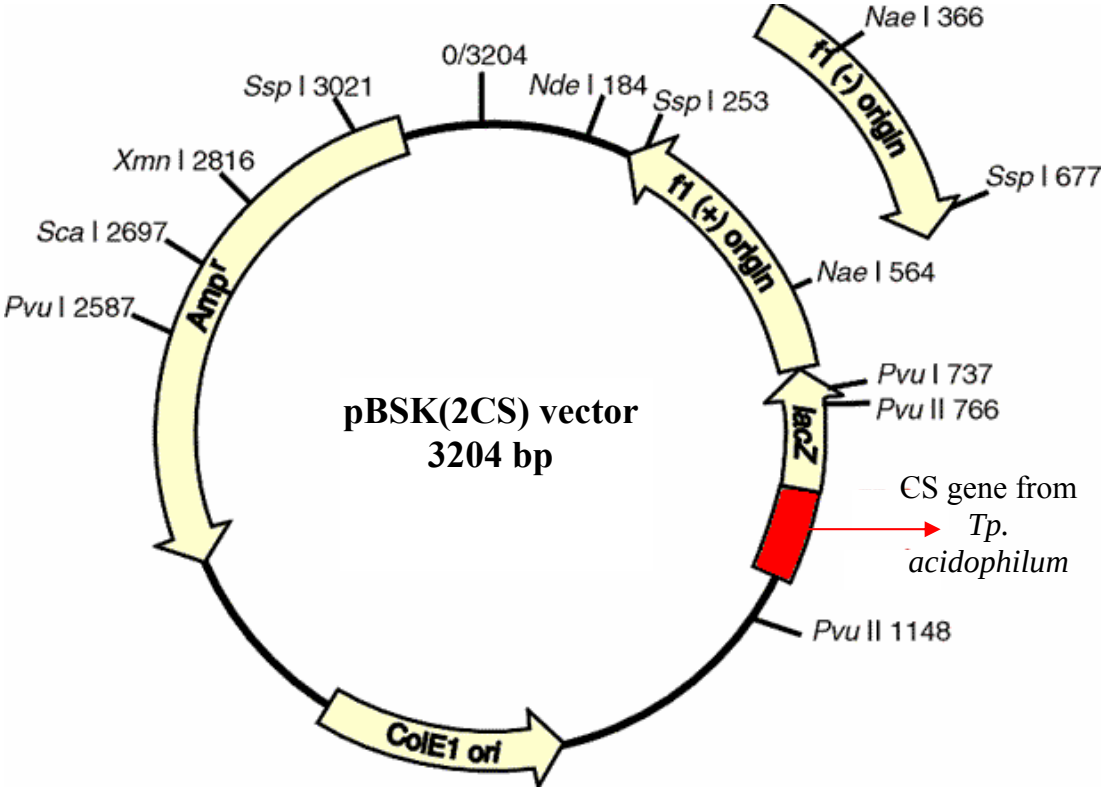


Figure 2.1 The schematic representation of pBSK(2CS) vector harboring *Tp. acidophilum* CS gene.

Table 2.1 Restriction sites and their cut positions in *Tp. acidophilum* CS gene which were determined by RestrictionMapper program.

Noncutter Restriction Enzymes	Cutter Restriction Enzymes	Cut Positions
AvaI	SspI	185
EcoRI	EcoRV	246
KpnI	SacI	466
PvuI	HindIII	1164
SalI	PvuII	536, 948
SmaI	PstI	908, 1391
SphI	AvaII	927, 1383
	BamHI	983, 1515

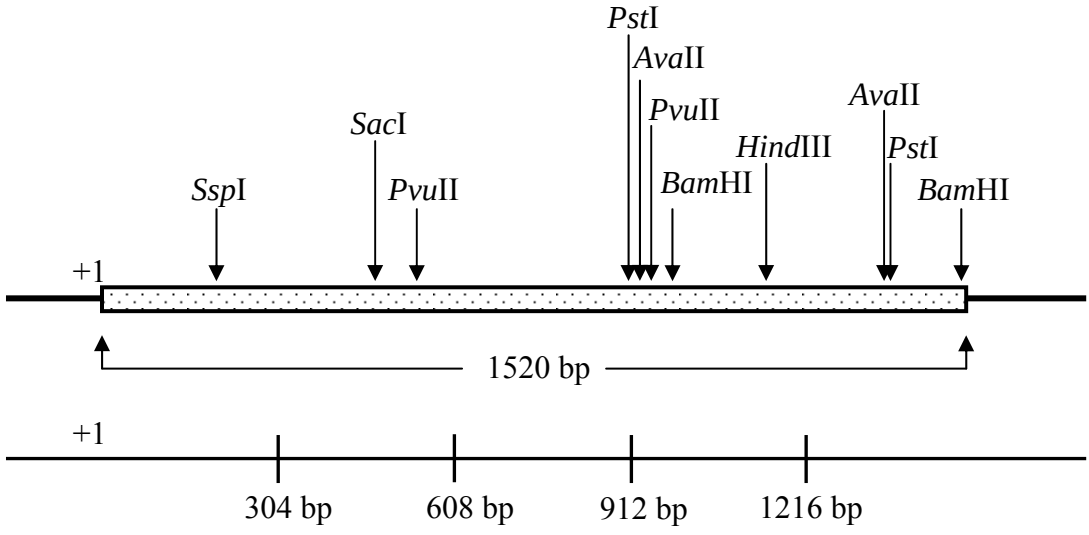


Figure 2.2 The restriction map of *Tp. acidophilum* CS gene.

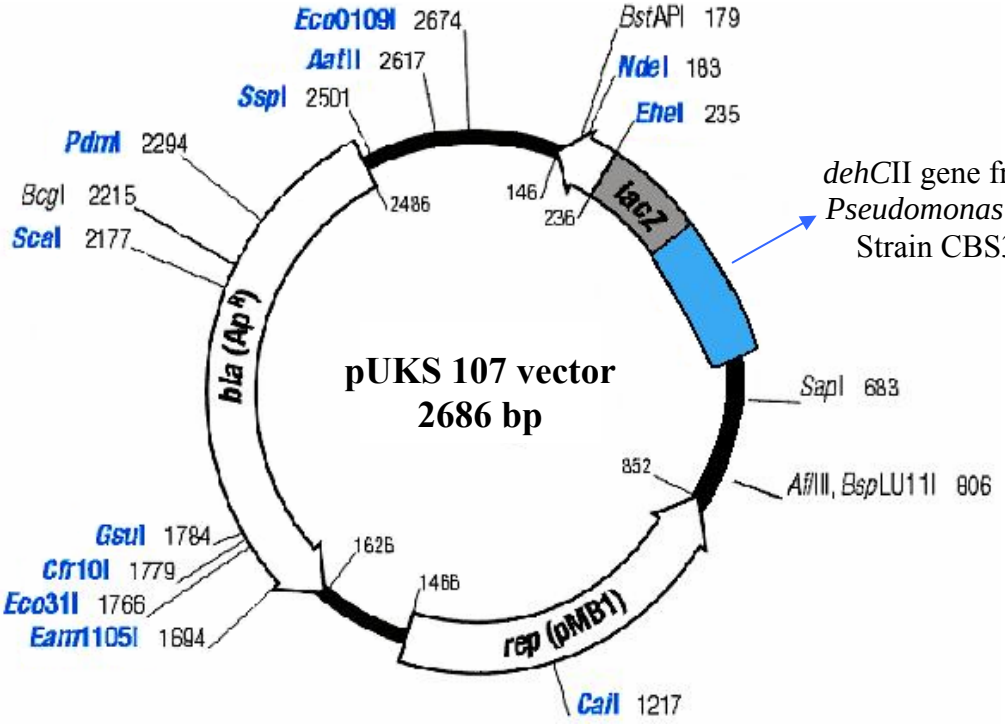


Figure 2.3 The schematic representation of pUKS 107 vector harboring *Pseudomonas* sp. Strain CBS3 *dehC*II gene.

Table 2.2 Restriction sites and their cut positions in *Pseudomonas* sp. Strain CBS3 *dehC*II gene which were determined by RestrictionMapper program.

Noncutter Restriction Enzymes	Cutter Restriction Enzymes	Cut Positions
EcoRI	SphI	74
EcoRV	SspI	153
HindIII	PvuII	214
KpnI	AvaII	781
PstI	BamHI	1, 1033
PvuI	AvaI	381, 681
SacI		
SalI		

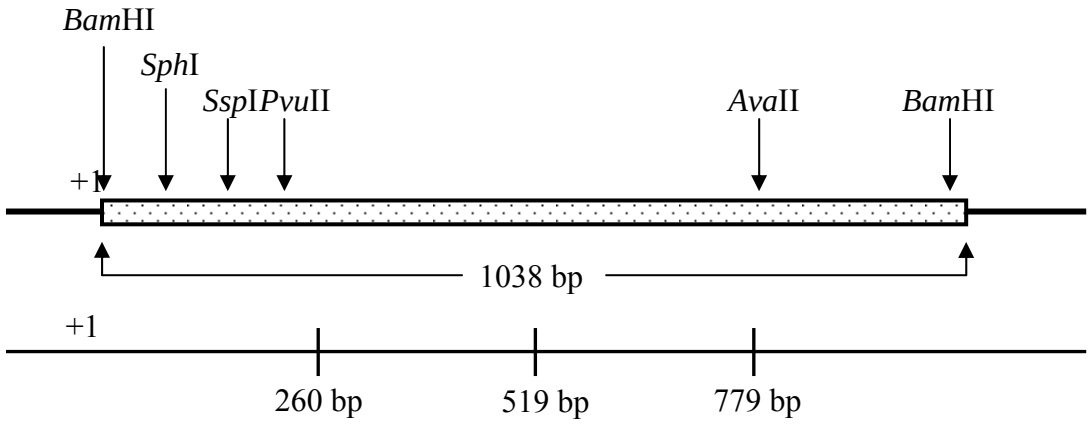


Figure 2.4 The restriction map of *Pseudomonas* sp. Strain CBS3 *dehC*II gene.

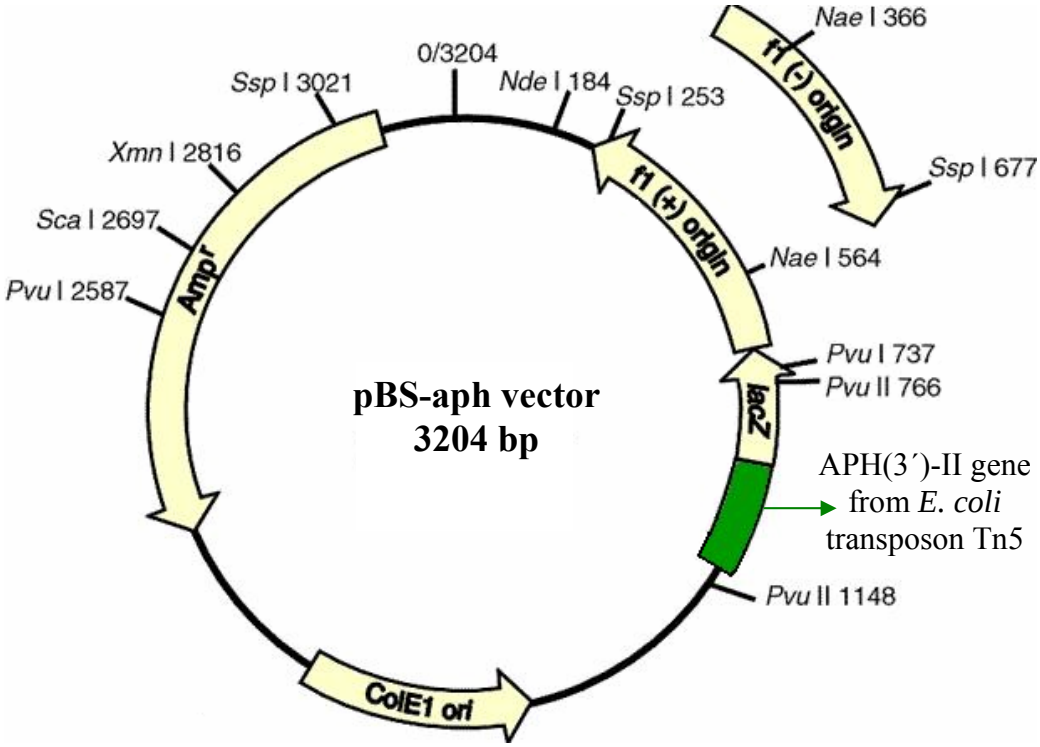


Figure 2.5 The schematic representation of pBS-aph vector harboring *E. coli* transposon Tn5 APH(3´)-II gene.

Table 2.3 Restriction sites and their cut positions in *E. coli* Tn5 APH(3´)-II gene which were determined by RestrictionMapper program.

Noncutter Restriction Enzymes	Cutter Restriction Enzymes	Cut Positions
BamHI	SphI	684
EcoRI	AvaII	795
EcoRV	AvaI	1116
HindIII	SmaI	1118
KpnI	SalI	1284
PvuI	PstI	332, 1255
SacI	PvuII	25, 385, 1147
SspI		

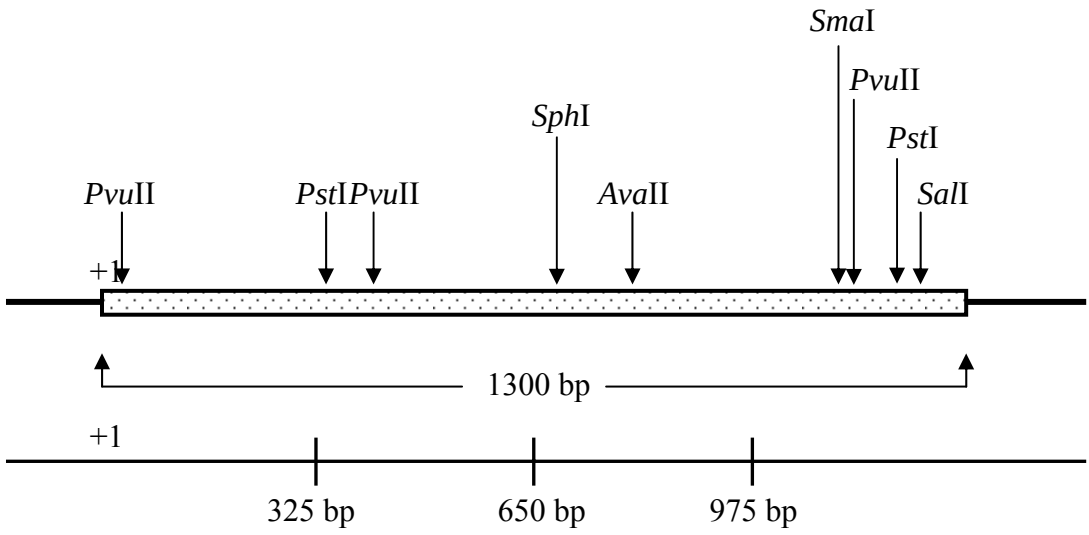


Figure 2.6 The restriction map of *E. coli* Tn5 associated APH(3´)-II gene.

transferred to Spin Column, provided with the Kit, and centrifuged at 13 000 rpm for 1 minute in order to bind the DNA to the column. The bound DNA was washed twice, first with 750 µl and then with 250 µl Wash Solution, and eluted by adding 100 µl of Nuclease-Free Water. Out of 100 µl of plasmid sample approximately 15 µl was run to the agarose gel to check the efficacy of purification. The remaining sample was stored at -20ºC for further use.

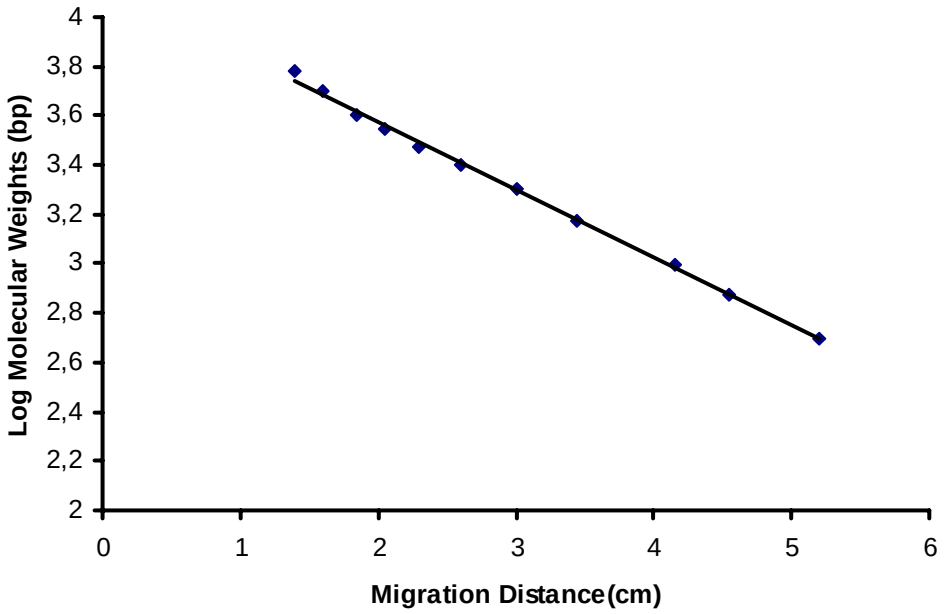
2.2.2 Agarose Gel Electrophoresis

Plasmid samples, restriction endonuclease digested DNA fragments, and PCR amplicons were analyzed by using Submarine agarose gel apparatus (Mini Sub™ DNA Cell, Bio Rad, Richmond, CA, U.S.A) with 1% (w/v) agarose gel (Boehringer Mannheim, Germany) and 1x TAE as the running buffer. DNA samples of 10-20 µl were mixed with 1/10 volume of tracking dye and loaded on the gel. Electrophoresis was carried out at 60 volts. The bands were visualized with a UV transilluminator (Vilber Lourmat, Marne La Vallée Cedex, France) and gel photographs were taken by using a gel imaging and documentation system (Vilber Lourmat Gel Imaging and Analysis System, Marne La Vallée Cedex, France). The molecular sizes of DNA fragments were determined by referring to calibration curves, which were obtained by plotting the log molecular weights of known marker fragments against migration distance on the gel. Figure 2.7 shows calibration curves for Gene Ruler 1 kb DNA Ladder (MBI Fermentas AB, Vilnius, Lithuania) and Lambda DNA/*Eco*RI+*Hin*dIII, molecular size marker (MBI Fermentas AB, Vilnius, Lithuania).

2.2.3 Restriction Endonuclease Treatment of the Isolated Recombinant Plasmids

Isolated pBSK(2CS), pUKS 107 and pBS-aph vectors were subjected to restriction enzyme digestion for characterization of the genes which they possess. Single or double digestions with the restriction enzymes, *Bam*HI, *Eco*RI, *Pst*I, *Hin*dIII and *Sal*I, were performed following the instructions of the manufacturers'. To map the sites for restriction endonucleases, RestrictionMapper program was used. These informations were used in the fusion experiments.

A)



B)

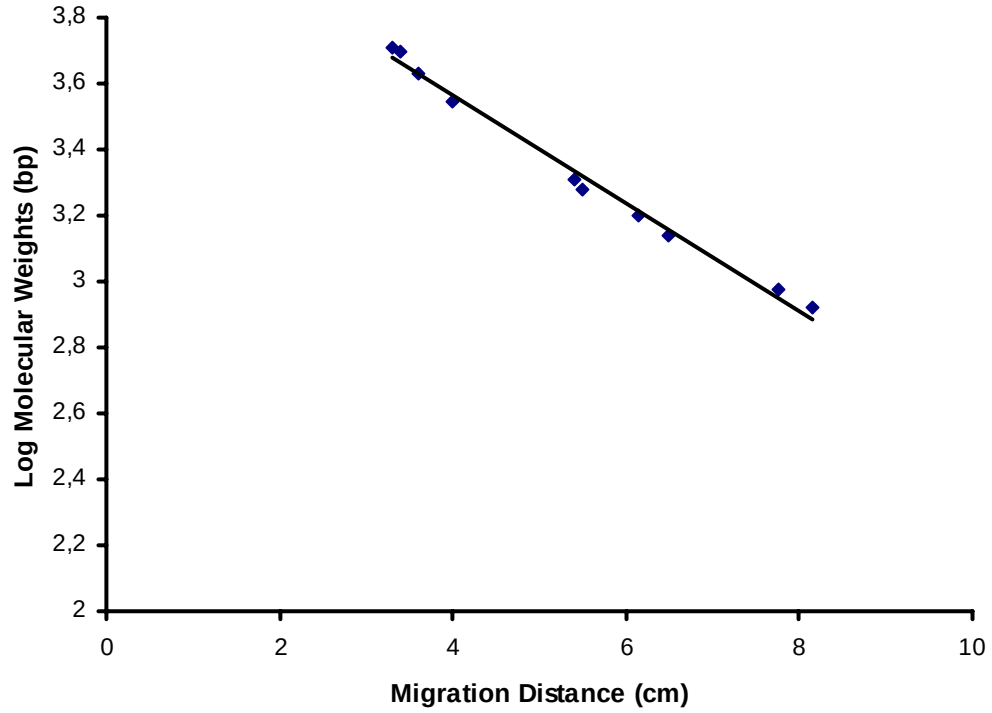


Figure 2.7 Calibration curves for **A)** Gene Ruler 1 kb DNA Ladder (MBI Fermentas AB, Vilnius, Lithuania) and **B)** Lambda DNA/*Eco*RI+*Hin*dIII, molecular size marker (MBI Fermentas AB, Vilnius, Lithuania).

2.2.4 Strategies for the Construction of Citrate Synthase Fusion Proteins

Two strategies were used for the construction of fusion proteins:

a) Direct ligation of recombinant *dehC*II gene to *Tp. acidophilum* CS Large Domain,
b) Gene SOEing

2.2.4.1 Strategy 1: Direct Ligation of *Pseudomonas dehC*II gene to *Tp. acidophilum* CS Large Domain

Direct ligation strategy experiment is based on fusing the *Pseudomonas dehC*II gene near to the *Tp. acidophilum* CS Large Domain which is being the main determinant of the enzymes' thermostability (Arnott *et al.*, 2000) (Figure 2.8).

*Bam*HI site was used for this purpose because still the remaining pBSK(2CS) vector has the CS Large Domain after digestion. *Tp. acidophilum* CS gene and *Pseudomonas dehC*II was digested with *Bam*HI restriction endonuclease at 37ºC for 1.5 hours. Two volumes of ice cold absolute ethanol and 1/10 volume of Buffer B were added to the tube of CS fragment, followed by incubation at -20ºC for overnight. The precipitated DNA was collected by centrifugation at 13 000 rpm for 25 minutes (Herause Sepatech, GmbH, Osterode, Germany). The supernatant was discarded and the pellet was washed with 70% (v/v) ethanol. After additional incubation at -80ºC for 20 minutes and centrifugation at 13 000 rpm for 20 minutes (Herause Sepatech, GmbH, Osterode, Germany), the supernatant was poured and the remaining ethanol was removed by air drying. The DNA pellet was dissolved in 15 µl of sterile ddH_2O and stored at -20ºC until use. The *Bam*HI-digested *dehC*II sample was run on the 1% Low Gelling Temperature Agarose (Sigma Chemical Co., St. Louis, MO, USA) for the isolation of the fragment from gel.

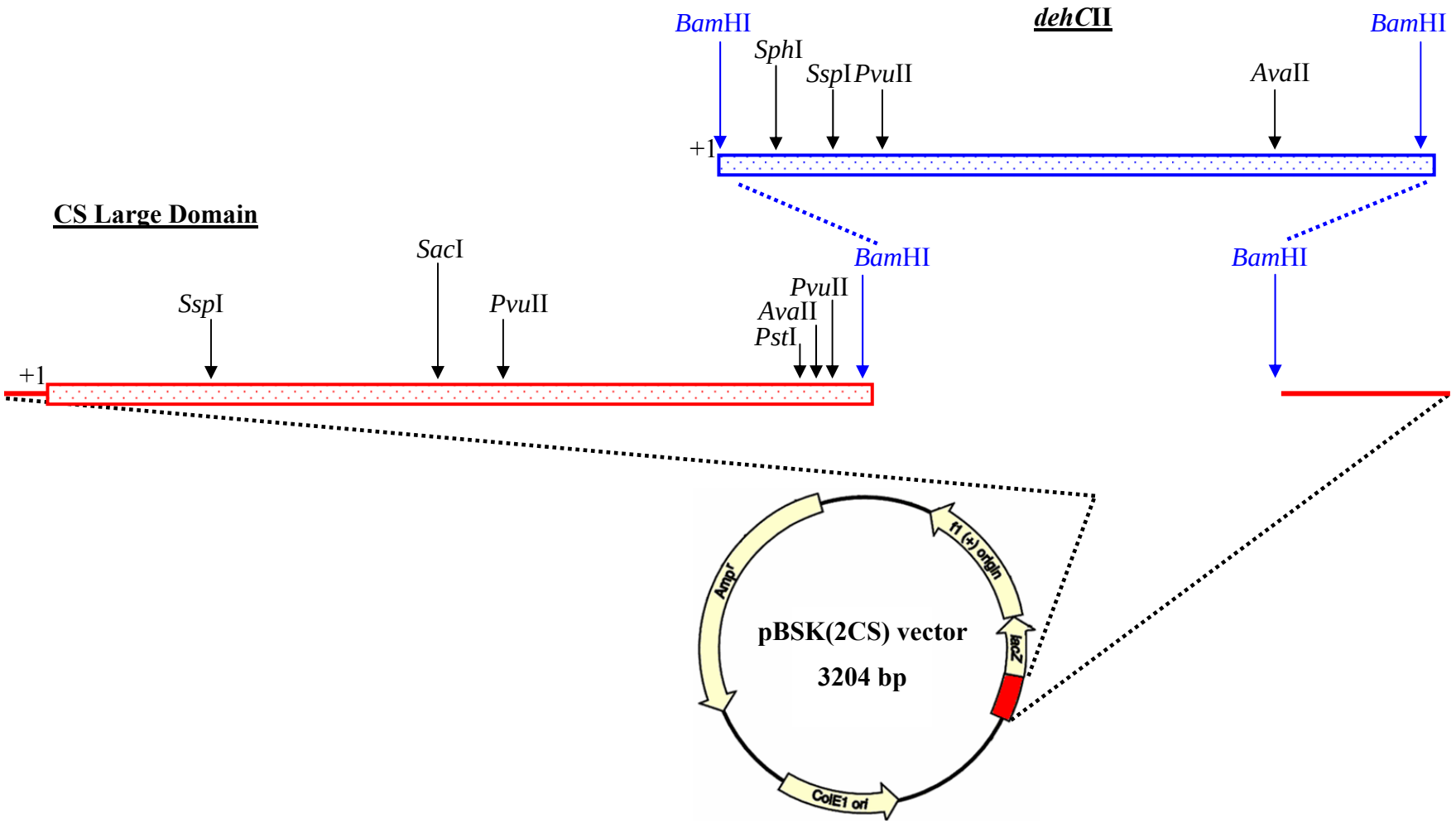


Figure 2.8 The schematic representation of direct ligation strategy.

2.2.4.1.1 Isolation of DNA Fragments from Agarose Gel

Concert Gel Extraction Systems Kit (Life Technologies Gibco BRL Grand Island, NY, USA) and MBI Fermentas DNA Extraction Kit (MBI Fermentas AB, Vilnius, Lithuania) were used according to the manufacturers' instructions for the isolation of the DNA fragments. The gel slice containing the DNA band was excised, placed into an eppendorf tube, weighed and dissolved in Binding Solution (3 volumes of buffer/1 volume of gel) with the incubation at 55ºC for 5 minutes. Then, 2 µl of Silica Powder Suspension was added per 1µg of DNA and incubated at 55ºC for additional 10 minutes. To keep the silica powder in suspension, the mixture was mixed gently for every 2 minutes. Silica powder/DNA complex was pelleted by centrifugation at 13 000 rpm (Herause Sepatech, GmbH, Osterode, Germany) for 1 minute and the supernatant was removed. The silica resin was washed three times with 500 µl ice cold Wash Buffer as the pellet was resuspended completely and supernatant was removed during each washing. Then, the silica was air dried for 10-15 minutes. For the elution of DNA, the resin was suspended into water and incubated at 55ºC for 5 minutes followed by centrifugation at 13 000 rpm (Herause Sepatech, GmbH, Osterode, Germany) for 1 minute.

2.2.4.1.2 Fusion of *dehC*II Fragment with CS Large Domain

2.2.4.1.2.1 Preparation of Competent *E. coli* Cells

The overnight culture of *E. coli* TG1 cells (1/100 volume) was transferred into 20 ml LB broth. The cells were grown at 37°C with vigorous shaking at 171 rpm (Heidolph UNIMAX1010, Heidolph Instruments GmbH, Kelheim, Germany). The growth was monitored by measurement of the optical density of cell culture at 600 nm using the Shimadzu UV-160A double beam spectrophotometer (Shimadzu Analytical Co., Kyoto, Japan). When the logarithmic phase was reached, the cells were collected by centrifugation at 4°C, 4 000 rpm for 10 minute (Sigma 3K30 Centrifuge, Sigma Chemical Co., St. Louis, MO, USA). The supernatant was discarded and the pellet

was dissolved in 1/10 volume TSS solution. The solution was distributed as 100 µl into eppendorf tubes and stored at -80°C until use.

2.2.4.1.2.2 Ligation, Transformation and Screening

Ligation mixture containing 1 volume of plasmid sample/5 volume of fragment sample, 1 µl T4 DNA Ligase and 2 µl T4 DNA Ligase Buffer was prepared and incubated overnight at 4°C.

Competent *E. coli* cells were removed from -80°C deep freezer and thawed on ice to be used in transformation. An aliquot (10 µl) of overnight ligation mixture was added and mixed gently with the following incubation on ice for 30 minutes. Cells were transferred to glass tubes containing 0.6 ml LB broth supplemented with glucose and incubated at 37°C with vigorous shaking at 225 rpm for 1 hour (Heidolph UNIMAX1010, Heidolph Instruments GmbH, Kelheim, Germany). Appropriate dilutions of competent cells were spread onto selective LB agar plates containing ampicillin, IPTG and X-Gal and grown at 37°C for overnight.

Cells from appropriate colonies were collected and plasmid isolation was carried out using Wizard® Plus SV Miniprep DNA Purification System Kit (Promega Co., Madison, WI, USA) to find out the recombinant ones.

2.2.4.2 Strategy 2: Gene SOEing

Thermoplasma volcanium (*Tp. volcanium*) citrate synthase gene has recently been cloned in our laboratory and it was used as a partner in fusion construction by Gene SOEing (Figure 2.9, 2.10 and Table 2.4). Sequence alignment of *Tp. volcanium* with *Tp. acidophilum* citrate synthases was done by using the ClustalW program (http://www.ebi.ac.uk/clustalw) for determination of the homology and conserved regions, as well as for the large domain and the small domain.

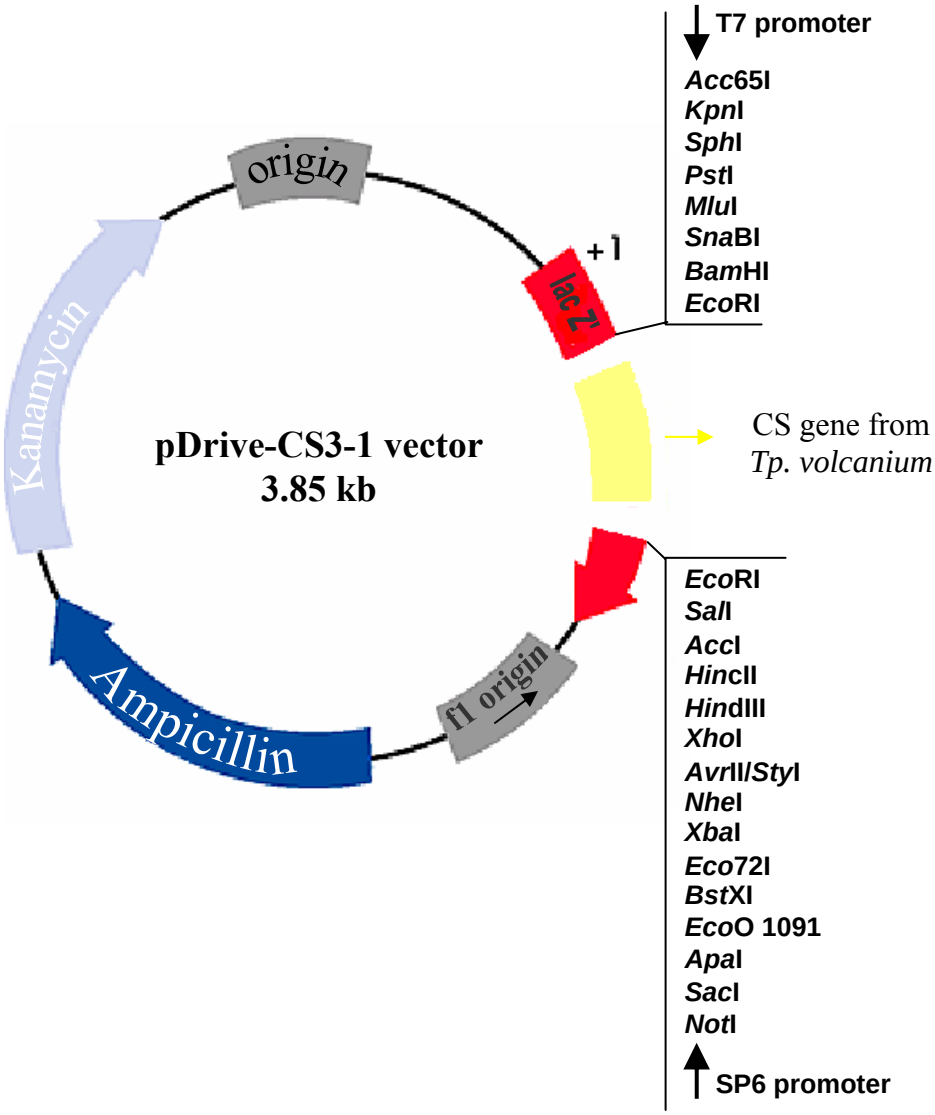


Figure 2.9 The schematic representation of pDrive-CS3-1 harboring *Tp. volcanium* CS gene.

Table 2.4 Restriction sites and their cut positions in *Tp. volcanium* CS gene which were determined by RestrictionMapper program.

Noncutter Restriction Enzymes	Cutter Restriction Enzymes	Cut Positions
AvaI	SspI	135
BamHI	PvuII	692
EcoRI	HindIII	727
EcoRV	AvaII	1153
KpnI	PstI	559, 838, 931
PvuI		
SacI		
SalI		
SmaI		
SphI		

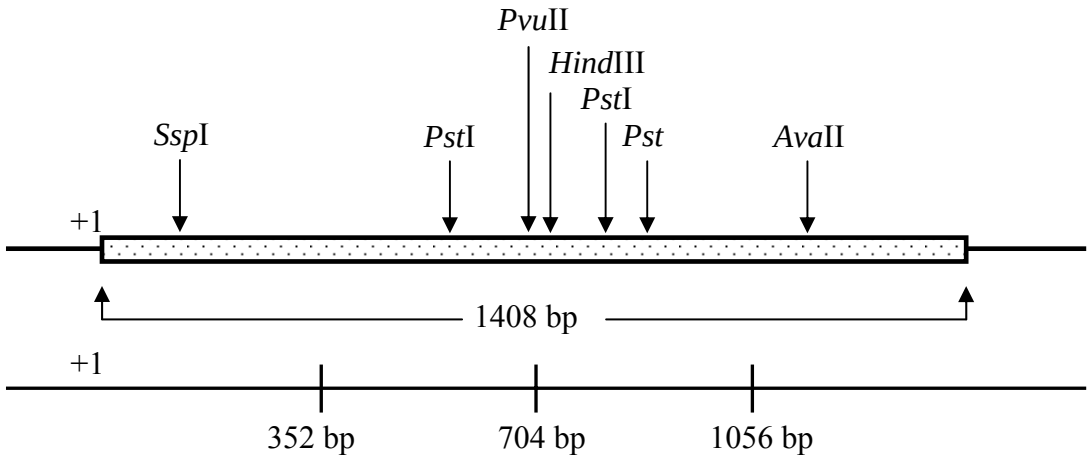


Figure 2.10 The restriction map of *Tp. volcanium* CS gene.

Gene SOEing or gene splicing by overlap extension strategy, a PCR-based approach, is based on recombining DNA molecules at precise junctions irrespective of nucleotide sequences at the recombination site and without depending on restriction sites and ligase. The strategy was carried out according to the method of Horton and co-workers (1989).

The general mechanism of the strategy is illustrated in Figure 2.11. The main idea is joining the DNA sequences which originated from different genes by incorporating complementary sequence into the 5′-end of one primer in order to overlap PCR products and recombine. The first step is the PCR reaction. The segments to be joined (Fragment AB from Gene 1 and Fragment CD from Gene 2) are amplified in separate PCRs (‘1’ and ‘2’). The complementary sequences added to the 5′-end of the primer (‘b’) become incorporated into the end of the product. This provides the PCR products of these first reactions overlap since the PCR amplified segment shares homologous sequence with other segment. The strategy follows with the SOE reaction. These products are mixed and then subjected to repeated rounds of denaturation, reannealing, and primer extension, as in PCR. One strand from each fragment contains the overlap sequence at the 3′-end, and these strands act as primers for DNA polymerase. Extension of this overlap by DNA polymerase yields the recombinant product. The presence of the appropriate primers (‘a’ and ‘d’) in the reaction allows the recombinant product to be amplified by PCR as soon as it is formed.

2.2.4.2.1 Design of PCR Amplification Primers

The primers were designed by consulting to the amino acid sequences of *Tv*CS (available in NCBI-National Center for Biotechnology Information-Nucleic Acid Sequence Database), *dehC*II (Schneider *et al.*, 1990), and APH(3´)-II (Beck *et al.*, 1982) (See Appendix D). These primers were used for adding appropriate sequences at the 5´-ends of the PCR products in order to make the fusion, as well as for the amplification of the created recombinant product. The list of PCR primers is given Table 2.5, 2.6 and 2.7. The added overlapping sequences are indicated in blue color.

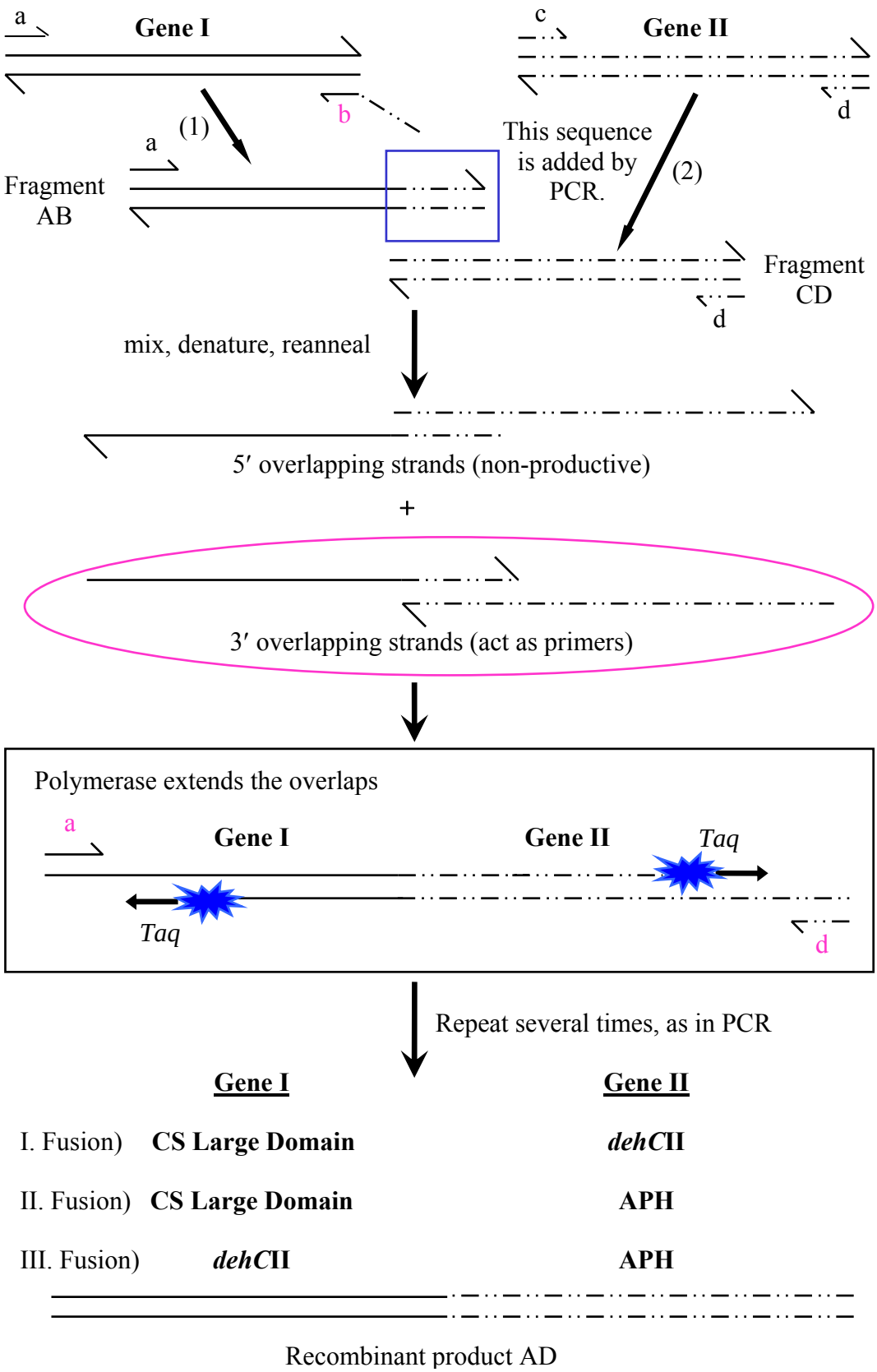


Figure 2.11 Mechanism of gene splicing by overlap extension.

Table 2.5 Primers for generation and amplification of CS Large Domain-*dehC*II fusion

Primer	Sequence
CS-Fr-1	5′-GGTTGACAATTACATAGACGTTTC-3′
CS-Rv-1	5′-CTCCCCGGATCTCCTGCATATGGAGCGGGCCCTTGAG-3′
*dehC*II-Fr-1	5′-ATGCAGGAGATCCGGGGAGTTG-3′
*dehC*II-Rv-1	5′-CCCAATATGGCAATCTTCATGG-3′

Table 2.6 Primers for generation and amplification of CS Large Domain-APH(3′)-II fusion

Primer	Sequence
CS-Fr-2	5′-GGTTGACAATTACATAGACGTTTC-3′
CS-Rv-2	5′-CAATCCATCTTGTTCAATCATATGGAGCGGGCCCTTGAG-3
APH-Fr-1	5′-ATGATTGAACAAGATGGATTGCAC-3′
APH-Rv-1	5′-CTTGGTCGGTCATTTCGAACC-3′

Table 2.7 Primers for generation and amplification of *dehC*II-APH(3′)-II fusion

Primer	Sequence
*dehC*II-Fr-2	5′-GGATCCATTTCTGCCTGGCTTTTGG-3′
*dehC*II-Rv-2	5′-CAATCCATCTTGTTCAATCATTCTATCCGCGGATGCGAAG-3
APH-Fr-2	5′-ATGATTGAACAAGATGGATTGCAC-3′
APH-Rv-2	5′-CTTGGTCGGTCATTTCGAACC-3′

2.2.4.2.2 Construction of Fusion Proteins by Using Gene SOEing Strategy

Three sets of fusion proteins including CS Large Domain-*dehC*II, CS Large Domain-APH(3′)-II, and *dehC*II-APH(3′)-II, were constructed. The method was two-staged: the PCR reaction and the SOE reaction. To eliminate the interference of the excess

primers and template DNA, PCR products were purified after the PCR reaction.

For the PCR reaction, the isolated, highly purified recombinant plasmids of pDrive, pUC18, and pBS(+) containing the complete gene of CS, *dehC*II, and APH(3′)-II from *Tp. volcanium* , *Pseudomonas* sp. Strain CBS3, and *E coli* transposon Tn5 were used as the template for performing the PCR. PCR reaction mixtures consist of 300 ng of the template DNA, 10 µM of 1x PCR buffer, 1.5 mM $MgCl_2$, 200 µM each of deoxyribonuclease triphosphate (dNTP), 3 µM of each primers and 2.5 units of *Pfu* polymerase (Promega Co., Madison, WI, USA) in a final volume of 100 µl. PCR conditions were as follows: 94ºC for 3 minutes, 25 cycles of 94ºC for 1 minute, 50ºC for 2 minutes, and 72ºC for 3 minutes, and 72ºC for 10 minutes and Techgene Thermal Cycler (Techgene, Techne Inc. NJ. USA) was used for this purpose.

The purification of PCR products was done by using Qiagen MinElute PCR Purification Kit (Qiagen Inc., Valencia, USA).

The SOE reactions were carried out using Techgene Thermal Cycler (Techgene, Techne Inc. NJ. USA) for 25 cycles, each consisting of 1 minute denaturation at 94ºC, 2 minutes annealing at 50-67ºC, and 3 minutes polymerization at 72ºC, and followed by 10 minute incubation at 72ºC. Denaturation at 94ºC for 5 minutes was performed before PCR was started. The PCR reaction mixture in a total volume of 100 µl contained 300-1000 ng of the template, 10 µM of 1X PCR buffer, 1.5 mM $MgCl_2$, 200 µM each of deoxyribonuclease triphosphate (dNTP), 3 µM of each primers and 2.5 units of *Taq* polymerase (Sigma Chemical Co., St. Louise, MO, USA).

2.2.5 Cloning of the Fusion Proteins

All the three fusion constructs, CS Large Domain-*dehC*II, CS Large Domain-APH(3′)-II, and *dehC*II-APH(3′)-II, were cloned using the QIAGEN PCR Cloning Kit (Qiagen Inc., Valencia, USA). They were ligated to the pDrive Cloning Vector, provided with the Kit, following the manufacturers' instructions.

The pDrive Cloning Vector contains several unique restriction endonuclease recognition sites, a T7 and SP6 promoter in *lacZ* which allows easy restriction analysis of recombinant plasmids and *in vitro* transcription of cloned PCR products as well as sequence analysis using standard sequencing primers. This vector also allows ampicillin and kanamycin selection, as well as blue/white colony screening.

Ligation mixtures contained 1 µl pDrive Cloning Vector, 2 µl PCR products, 5 µl 2x Ligation Master Mix (provided by the Kit) and 3 µl distilled water (provided by the Kit) in a total volume of 10 µl. The incubation was carried out at 11ºC for 2 hours. Competent *E. coli* strain TG1 was used for transformation. The transformation assays were carried out following the same method used in Strategy 1. Cells from the white colonies were picked up and their plasmids were isolated using Wizard® Plus SV Miniprep DNA Purification System Kit (Promega Co., Madison, WI, USA).

The CS Large Domain-*dehC*II fusion construct was also cloned using the pGEM®-T Vector Systems II Cloning Kit (Promega Co., Madison, WI, USA). Ligation to the pGEM®-T Vector was done according to the manufacturers' instructions. The 10 µl ligation mixture contained 5 µl 2X Rapid Ligation Buffer, 1 µl pGEM®-T Vector (50 ng), 3µl PCR product, 1 µl T4 DNA Ligase (3 weiss units/µl). Positive control reaction mixture was prepared as described above except that Control Insert DNA was used instead of PCR product. The reaction mixtures were incubated at 4ºC for overnight. Both JM109 High Efficiency Competent Cells, provided by pGEM®-T Vector Systems II Cloning Kit (Promega Co., Madison, WI, USA), and *E. coli* strain TG1 were used as recipients during the transformation assays. For transformation using the JM109 High Efficiency Competent Cells, 2 µl of ligation mixture was added to sterile 1.5 ml microcentrifuge tube on ice. Frozen JM109 High Efficiency Competent Cells were removed from -80ºC storage and thawed on ice. 50 µl of cells were carefully transferred into the tube containing the ligation mixture and incubated on ice for 20 minutes. Then the cells were heat-shocked for 45-50 seconds in water-bath at exactly 42 ºC. The tubes were immediately returned to ice for 2 minutes. 950 µl room temperature SOC medium was added to the tubes and they were incubated at

37°C with vigorous shaking at 150 rpm for 1.5 hours (Heidolph UNIMAX1010, Heidolph Instruments GmbH, Kelheim, Germany). 100 µl of transformation culture was spread onto selective LB agar plates containing ampicillin, IPTG and X-Gal and grown at 37°C for overnight. The transformation assay using the competent *E. coli* strain TG1 was carried out following the same method used in Strategy 1. Plasmid isolation from the cells of white colonies was carried out again by using Wizard® Plus SV Miniprep DNA Purification System Kit (Promega Co., Madison, WI, USA).

2.2.6 Preparation of Crude Extracts of Fusions Containing Dehalogenase Gene

E. coli cells harboring the pDrive-CS Large Domain-*dehC*II and pDrive-*dehC*II-APH(3′)-II fusions were cultured in 10 ml modified LB broth supplemented with ampicillin (50 µg/ml), in which NaCl had been replaced by 0.3% MCA as an inducer and incubated overnight at 37ºC. Overnight culture (2 ml) was transferred to 300 ml of fresh LB broth with MCA, supplemented with ampicillin (50 µg/ml), and incubated overnight at 37ºC. The cells were harvested by centrifugation at 10 000 rpm for 15 minutes at 4ºC (Sigma 3K30 Centrifuge, Sigma Chemical Co., St. Louis, MO, USA) and washed twice first with 200 ml then with 100 ml of 200 mM Tris/SO_4 (pH 7.9). After being suspended in 10 ml 200 mM Tris/SO_4 (pH 7.9), the cells were disrupted by ultrasonication (SonicatorVC100, Sonics and Materials, CT, USA). The sonicate was centrifuged at 13 000 rpm for 2 hours at 4ºC (Sigma 3K30 Centrifuge, Sigma Chemical Co., St. Louis, MO, USA). The supernatant was supplemented with 20 µl of 0.5 M DTT and stored at -20ºC until use for the enzyme assays.

2.2.7 Determination of Dehalogenase Activity

The enzymes were assayed by quantitation of the halide ions released from the substrate. The assay mixture in a total volume of 2.5 ml contained 50 µl of MCA, 1.200 µl Tris/SO_4 (pH 7.9), and 50 µl enzyme. After incubation at 30ºC at known time intervals, the reaction was terminated by addition of 500 µl of 1.5 M H_2SO_4.

The chloride ions were quantitated spectrophotomerically at 460 nm with mercuric thiocynate and ferric ammonium sulfate as described by Iwasaki *et al*. (1956) using Shimadzu UV-1601A double beam spectrophotometer (Shimadzu Analytical Co., Scientific Instruments Division, Kyoto, Japan).

2.2.8 Determination of Neomycin Resistance by Agar-Plate Assay

To determine the antibiotic resistance of the CS Large Domain-APH(3′)-II and *dehC*II-APH(3′)-II fusions, the following LB-agar was supplemented with different concentrations of neomycin (final concentration 20 μg/ml, 50 μg/ml, 100 μg/ml and 200 μg/ml) and of ampicillin (final concentration 50 μg/ml) and poured into sterile glass plates. *E. coli* cells harboring the pDrive-CS Large Domain-APH(3′)-II and pDrive-*dehC*II-APH(3′)-II fusions were cultured in 10 ml LB broth, supplemented with ampicillin (50 μg/ml), and incubated overnight at 37ºC. Also for positive and negative controls as well as for fusion constructs, LB-agar was supplemented with only neomycin (100 μg/ml) or with only ampicillin (50 μg/ml). The overnight cultures (100 μl) were spread onto the glass plates and the plates were incubated at 37ºC for 16 hours. Antibiotic resistance was revealed by the formation of colonies on the antibiotic supplemented agar-plates.

2.2.9 Preparation of Crude Extracts of Fusions Containing APH(3′)-II Gene

E. coli cells harboring the pDrive-CS Large Domain-APH(3′)-II and pDrive-*dehC*II-APH(3′)-II fusions were cultured in 10 ml LB broth supplemented with ampicillin (50 μg/ml) and incubated overnight at 37ºC. The overnight culture (1.5 ml) was transferred to 150 ml of fresh LB broth, supplemented with ampicillin (50 μg/ml), and incubated at 37ºC overnight. The cells were harvested by centrifugation at 10 000 rpm for 15 minutes at 4ºC (Sigma 3K30 Centrifuge, Sigma Chemical Co., St. Louis, MO, USA) and washed with 30 ml of TMND buffer containing 20 mM Tris.HCl (pH 7.4), 8.6 mM $MgCl_2$ and 1.1 mM DTT. After being suspended in 10 ml TMND buffer, the cells were disrupted by ultrasonication (Sonicator VC100, Sonics

and Materials, CT, USA). The sonicate was centrifuged at 13 000 rpm for 2 hours at 4ºC (Sigma 3K30 Centrifuge, Sigma Chemical Co., St. Louis, MO, USA) and the supernatant was stored at -20ºC until use for the enzyme assays.

2.2.10 APH(3′)-II Assay

The APH(3′)-II activity was measured by using a pyruvate kinase/lactate dehydrogenase coupled assay system, which links the release of ADP to the oxidation of NADH to NAD^+, and the resulting decrease in absorbance at 340 nm was monitored spectrophotometrically using a Shimadzu UV-1601A double beam spectrophotometer (Shimadzu Analytical Co., Scientific Instruments Division, Kyoto, Japan). Reactions contained a 932 µl of assay buffer (400 mM potassium acetate, 200 mM magnesium acetate, 0.1 M PIPES, pH 7.0, 1 mM NADH, 10 mM ATP, 20 mM PEP, 50 U pyruvate kinase, 100 U lactate dehydrogenase) and 63 µl of 5 mM neomycin. The mixtures were preincubated at 33°C for 10 min and then the reactions were initiated by the addition of 5 µl of cell-free extract (10.6 mg/ml protein). One unit of enzyme was defined as nmoles ADP release per minute.

2.2.11 Thermostability

CS Large Domain-APH(3′)-II fusion was analyzed for its thermostability property in terms of its resistance to irreversible thermal inactivation. The samples were incubated at 60°C for 10 minutes and denatured heat-labile proteins were removed by centrifugation at 12 000 rpm for 30 minutes at 4°C. The remaining enzymatic activity was assayed at 33°C as described in APH(3′)-II assay.

CHAPTER III

RESULTS

3.1 Isolation of Recombinant Plasmids

Recombinant plasmids, pBSK(2CS) vector expressing the complete *Tp. acidophilum* citrate synthase (CS) gene, pUKS 107 vector expressing the complete *Pseudomonas* sp. Strain CBS3 2-haloalkanoic acid dehalogenase II (*dehC*II) gene, and pBS-aph vector expressing the complete *E. coli* transposon Tn5 aminogylcoside 3´-phosphotransferase II (APH(3´)-II) gene were isolated as described in the Materials and Methods and visualized by agarose gel electrophoresis. Figure 3.1 shows the isolated plasmids.

The restriction maps of pBSK(2CS), pUKS 107 and pBS-aph recombinant vectors are shown in the Figures 2.2, 2.4, 2.6 and Tables 2.1, 2.2, 2.3; respectively.

Single and double digestions with the restriction enzymes, *Bam*HI, *Eco*RI, *Pst*I, *Hin*dIII and *Sal*I, were performed to check the eligibility of the isolated plasmids for use in fusion experiments during Strategy 1.

The recombinant vector pBSK(2CS) was linearized (4724 bp: 3204 bp pBSK(2CS) + 1520 bp *Ta*CS gene) when digested by *Eco*RI, at MCS, since *Ta*CS gene does not have any *Eco*RI site. There is a *Hin*dIII site within the gene at position 1164 bp. Therefore, *Eco*RI/*Hin*dIII double digestion products were 3560 bp vector DNA and 1164 bp *Ta*CS gene. Besides the vectors' MCS, there are two more *Bam*HI sites in the *Ta*CS gene at positions 983 bp and 1515 bp. *Bam*HI digestion just released a 532 bp fragment from the recombinant vector (Figure 3.2).

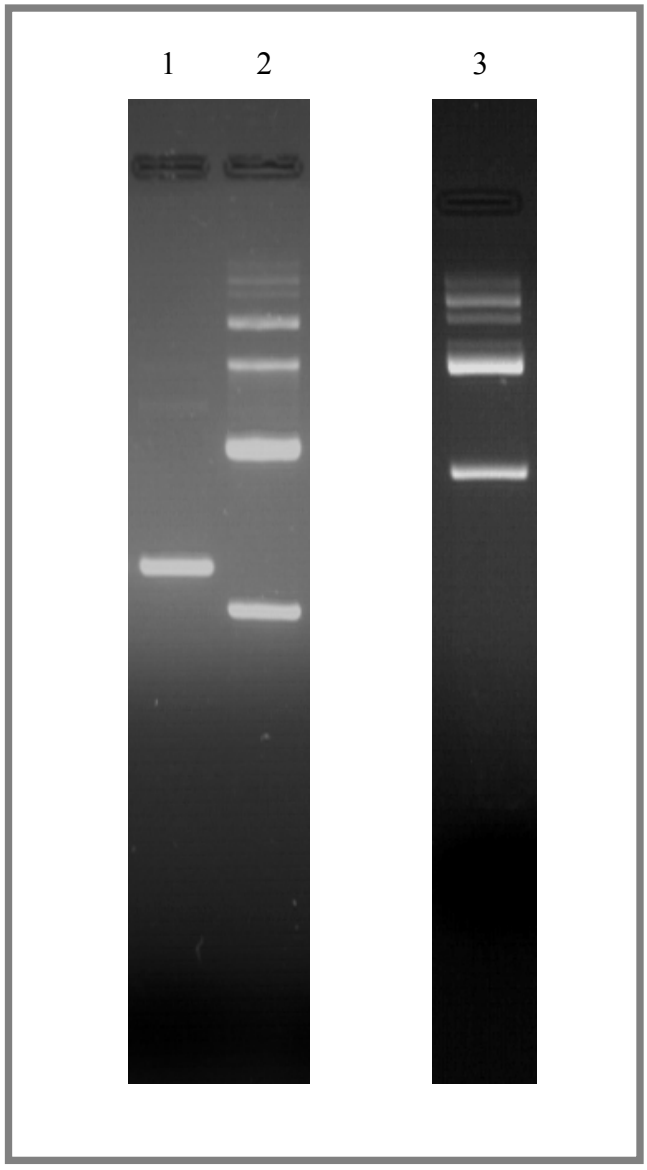


Figure 3.1 Isolation of recombinant plasmid vectors, lane 1: pBSK(2CS) (harboring *Tp. acidophilum* CS gene); lane 2: pUKS 107 (harboring *Pseudomonas* sp. Strain CBS3 *dehC*II gene); and lane3: pBS-aph (harboring *E. coli* transposon Tn5 APH(3´)-II gene).

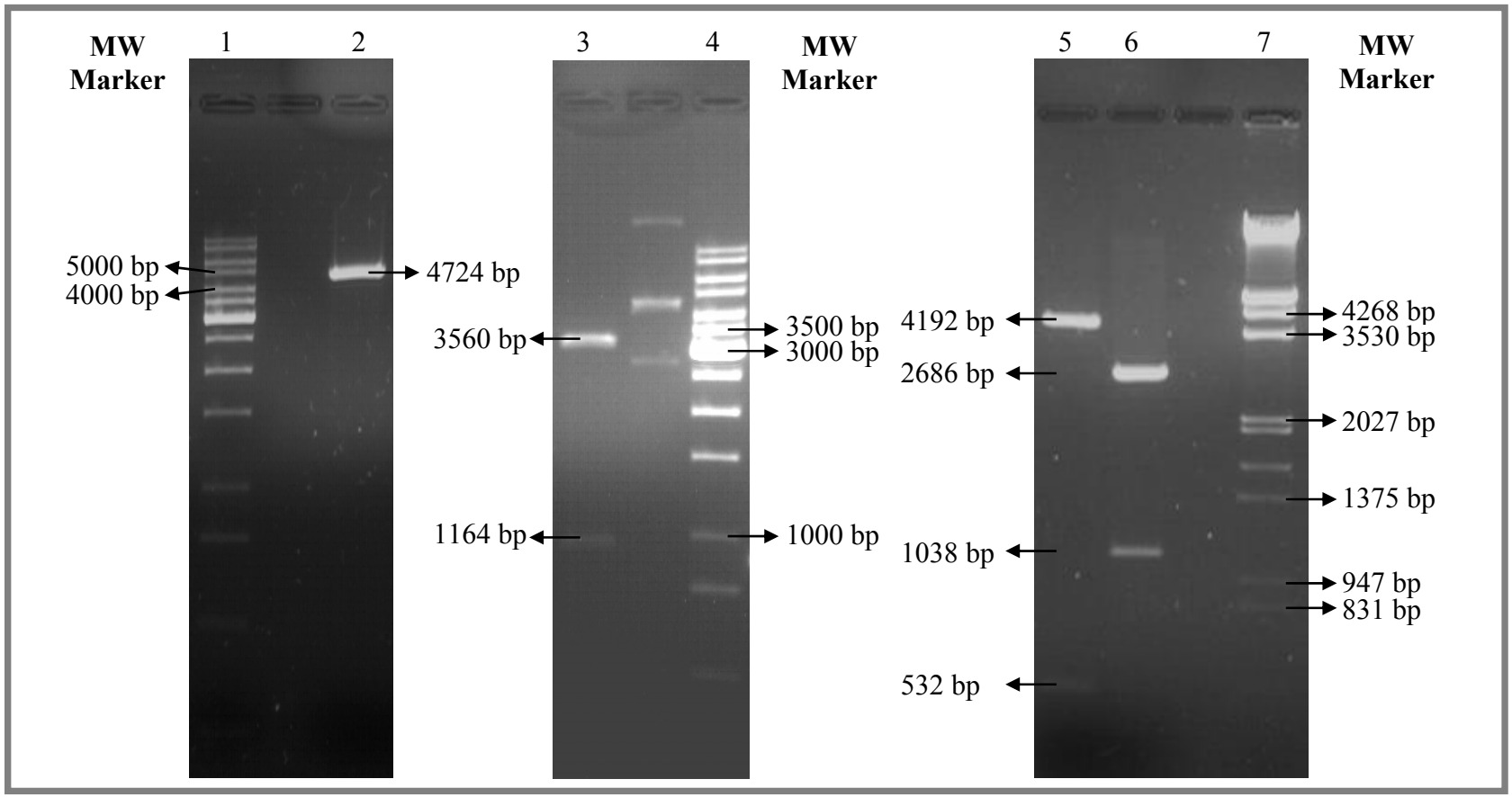


Figure 3.2 Restriction analyses of the pBSK(2CS) vector and pUKS 107 vector. pBSK(2CS) vector is subjected to *Eco*RI single digestion (lane2), *Eco*RI/*Hin*dIII double digestion (lane 3) and *Bam*HI single digestion (lane 5). pUKS 107 vector is digested with *Bam*HI (lane6). Lane 1 and lane 4, Gene Ruler 1 kb DNA Ladder; lane 7, Lambda DNA/*Eco*RI+*Hin*dIII, molecular size marker.

Pseudomonas sp. Strain CBS3 *dehC*II gene was excised from pUKS 107 vector, by cutting at flanking *Bam*HI sites, as a 1038 bp fragment (Figure 3.2).

Digestion of pBS-aph vector with *Hin*dIII, which is located at MCS, linearized the recombinant vector: 4504 bp: 3204 bp phagemid DNA + 1300 bp APH(3′)-II gene. Similarly, single digestion with *Sal*I that is found at 1284 bp yielded linear 4504 bp recombinant phagemid DNA. Cleavage with *Pst*I, positioned at 332 bp and 1255 bp, splited the pBS-aph vector into two fragments: 3581 bp vector DNA and 923 bp APH(3′)-II gene (Figure 3.3).

3.2 Strategy 1: Direct ligation of *Pseudomonas dehC*II gene to *Tp. acidophilum* CS Large Domain

For ligation of the complete *dehC*II gene to the pBSK(2CS) vector, which is containing the CS Large Domain after digestion, the only appropriate restriction site to be used was *Bam*HI. *Bam*HI-digested pBSK(2CS) vector yielded a 4192 bp fragment (3204 bp pBSK(2CS) vector + 988 bp *Ta*CS Large Domain) while *Bam*HI-digested pUKS 107 vector resulted in the excision of *dehC*II gene which was 1038 bp in length (Figure 3.2).

4192 bp fragment of pBSK(2CS) vector was concentrated and 1038 bp *dehC*II gene fragment was isolated from agarose gel (Figure 3.4A and Figure 3.4B). The sticky-end ligation was carried out as described in the Materials and Methods to generate *Ta*CS Large Domain-*dehC*II fusion. Competent *E. coli* strain TG1 was used as recipients in the transformation studies. High numbers of colonies were picked up and plasmids isolated from 45 transformant colonies were screened to find out fused recombinants. However, restriction analysis revealed that none of the plasmids proved to contain the fusion construct (Figure 3.5A and Figure 3.5B).

Based on these informations, it was concluded that direct ligation strategy could not achieve its goal because direct selection of recombinants by blue/white screening was not possible and self-ligation was much more favored. Due to these drawbacks

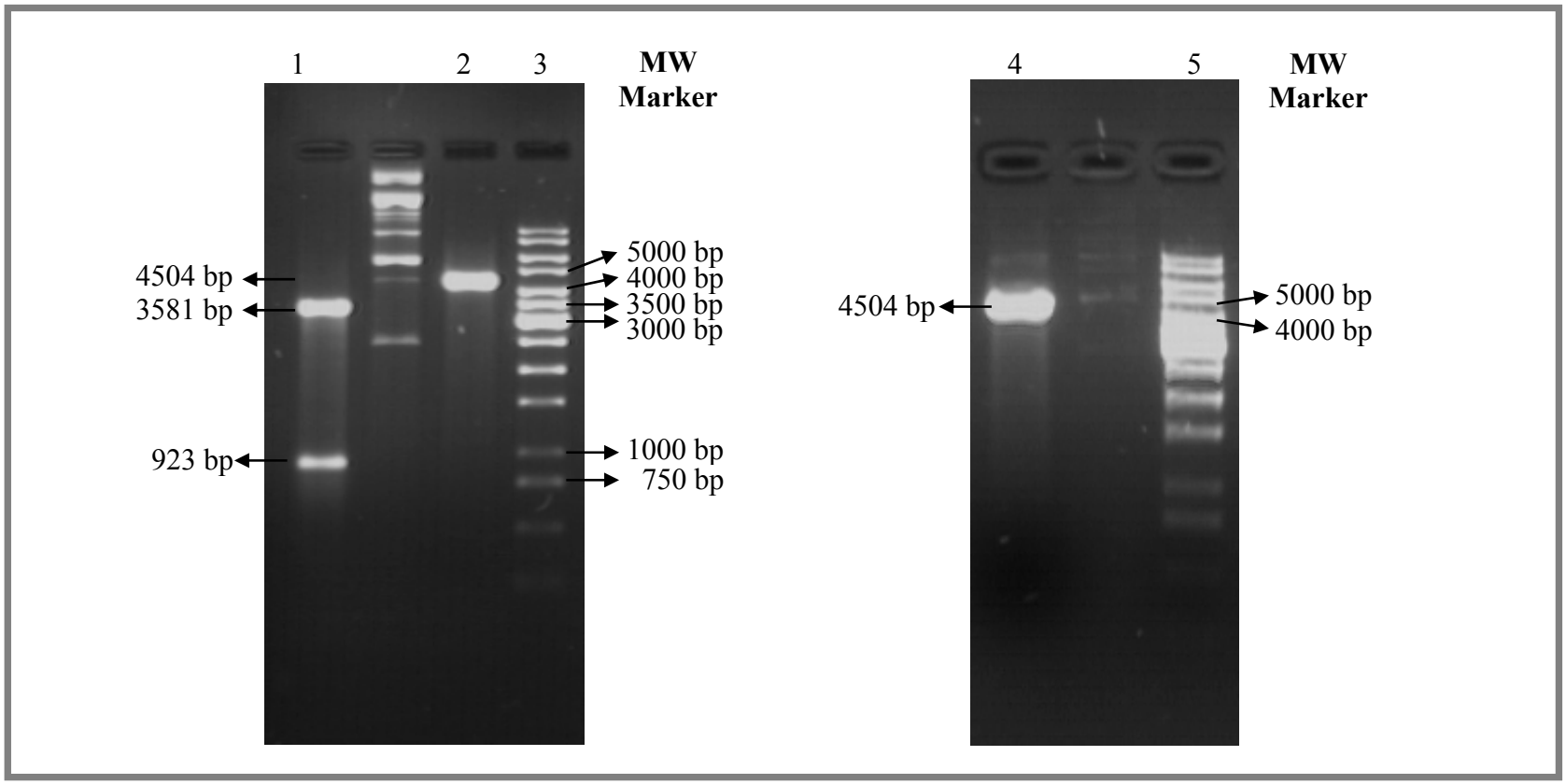


Figure 3.3 Restriction analysis of the pBS-aph vector: Lane 1, *Pst*I-digested pBS-aph vector; lane 2, *Hin*dIII-digested pBS-aph vector; lane 4, *Sal*I-digested pBS-aph vector. Lane 3 and lane 5, Gene Ruler 1 kb DNA Ladder.

direct ligation strategy was time-consuming and tedious. Therefore, it was replaced by an alternative approach, "Gene SOEing".

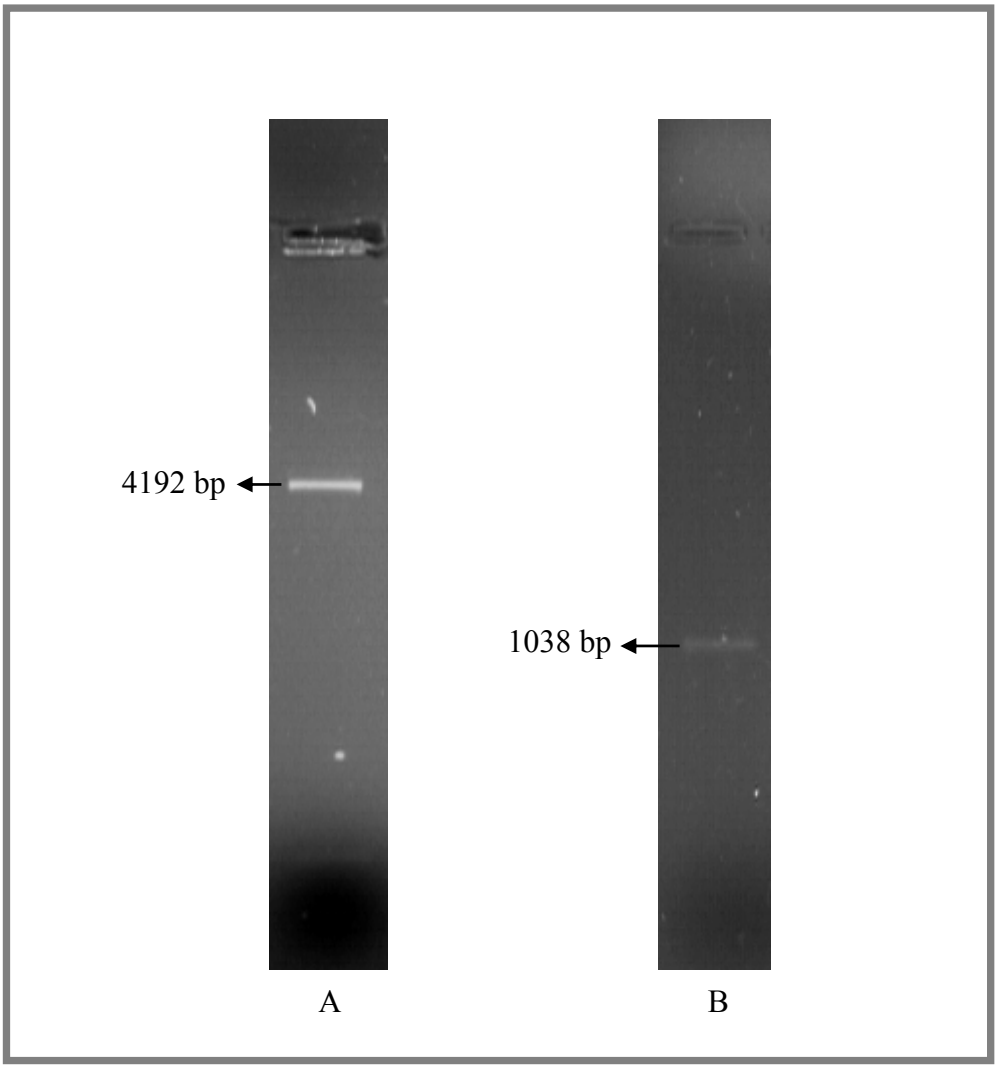


Figure 3.4 A) Concentrated *Bam*HI-digested pBSK(2CS) vector possessing CS Large Domain. B) Gel isolated *dehC*II gene released from pUKS 107 through *Bam*HI digestion.

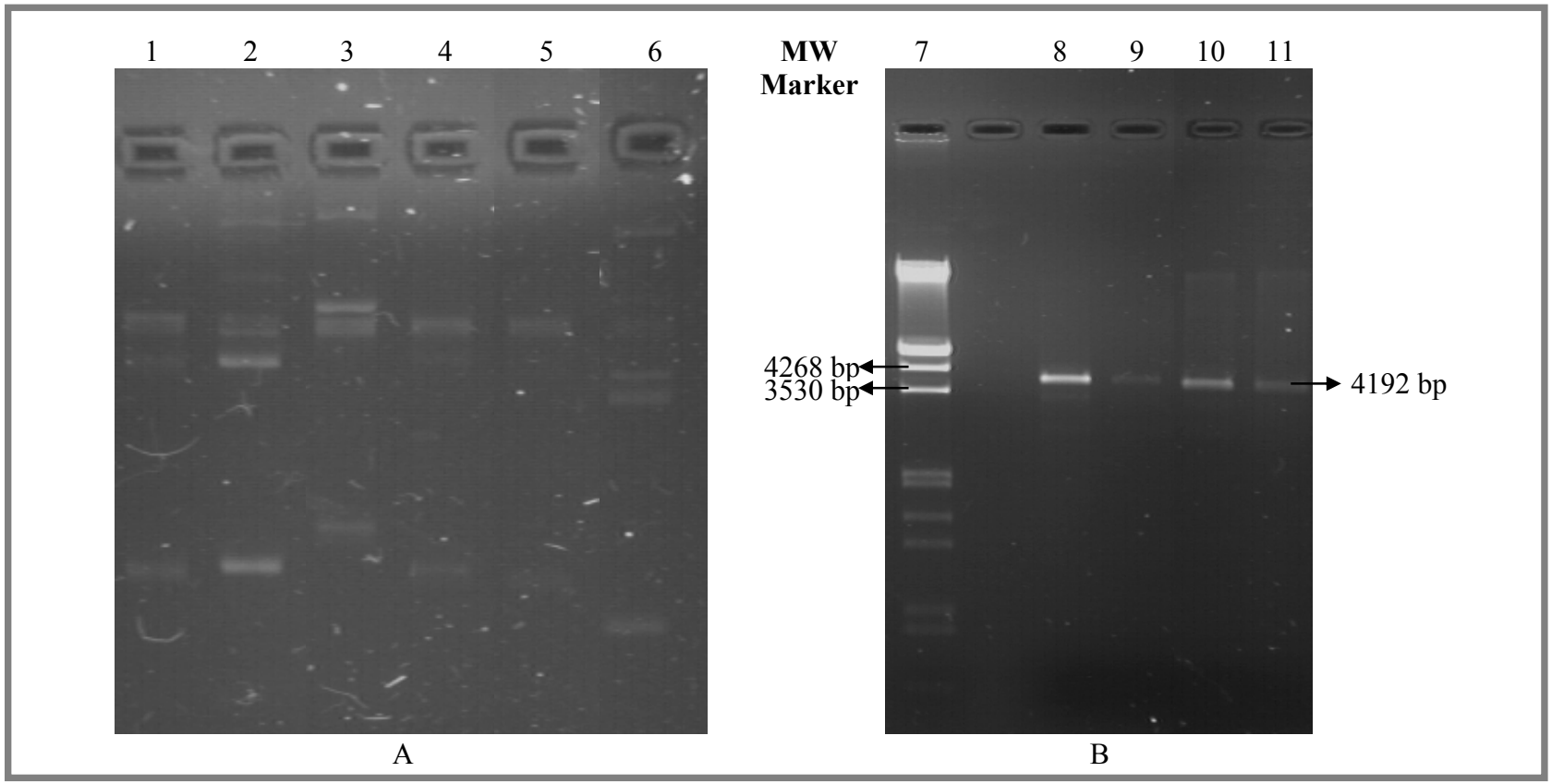


Figure 3.5 A) Plasmid isolation from putative recombinant colonies containing *Ta*CS Large Domain-*dehC*II gene fusions: Lane 1, S-19; lane 2, S-20; lane 3, S-9; lane 4, S-10; lane 5, S-18; lane 6, S-6. B) Restriction digestions of the putative recombinant plasmids with *Bam*HI: Lane 8, S-19; lane 9, S-20; lane 10, S-9; lane 11, S-10. Lane 7, Lambda DNA/*Eco*RI+*Hin*dIII, molecular size marker.

3.3 Strategy 2: Gene SOEing

3.3.1 Amino Acid Sequence Alignment of *Tp. volcanium* CS with *Tp. acidophilum* CS

Thermoplasma volcanium citrate synthase is a protein comprised of 386 amino acids (M_r~43 000). Amino acid sequence alignment of *Tp. volcanium* CS with *Tp. acidophilum* CS showed that there is 87% homology between two enzymes (Figure 3.6). The alignment data also clearly illustrated that there is a high conservation of functional amino acids (e.g. active site residues) and domains of citrate synthase enzymes from two *Thermoplasma* species. *Tp. acidophilum* CS larger domain contains helices C-G, I-M and S (residues 1-224 and 326-384), and small domain contains helices N-R (residues 225-325). Therefore we have decided to use *Tp. volcanium* CS, which was recently cloned in our laboratory as a fusion partner in fusion construction experiments by Gene SOEing.

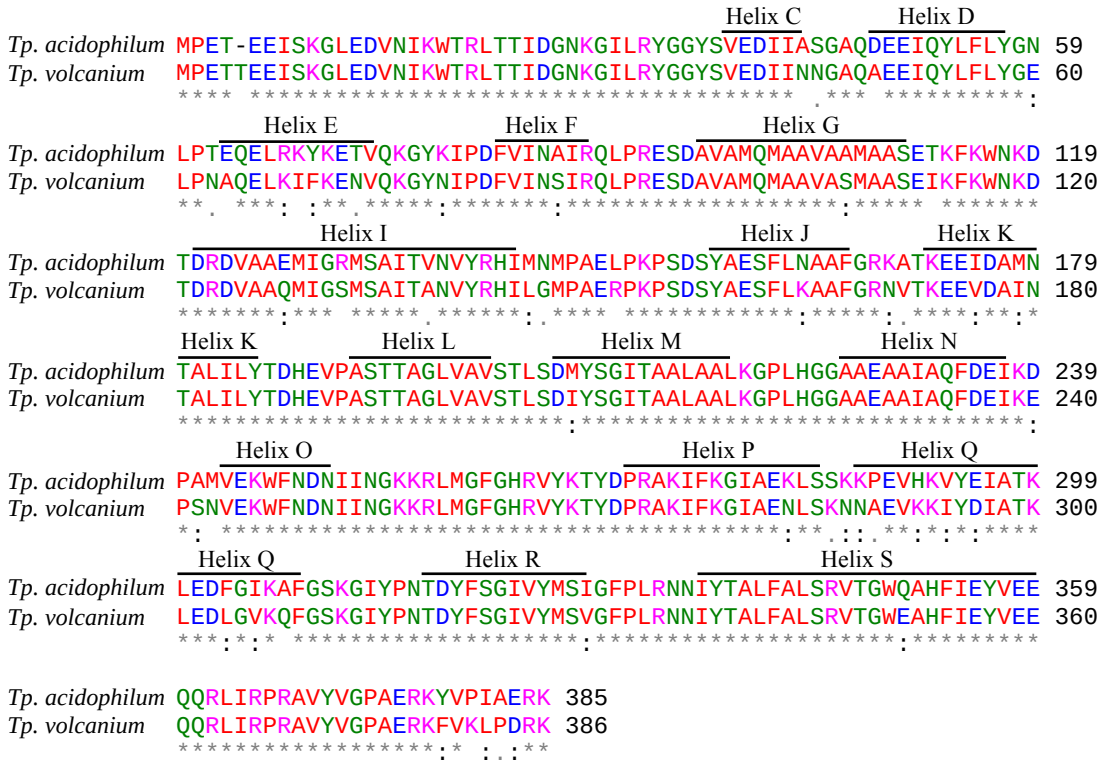


Figure 3.6 Sequence alignment of *Tp. volcanium* CS with *Tp. acidophilum* CS. Asterisks indicate identical residues, colons point out conservative and dots show semi-conservative substitutions. Helices are marked with horizontal lines.

3.3.2 PCR Amplifications and Cloning of PCR Amplified Fusion Constructs

3.3.2.1 PCR Reactions and SOE Reactions of Fusion Constructs

The general mechanism of this strategy is illustrated in the Materials and Methods (Figure 2.10).

To construct the CS Large Domain-*dehC*II fusion in first place, *Tv*CS Large Domain including the upstream promoter region and *dehC*II structural gene sequence were to be amplified. Therefore, the CS-Fr-1 and CS-Rv-1 primers were designed for amplification of CS Large Domain from pDrive-CS3-1 vector to cover a genome region of 922 bp extending from -250 position upstream to +672 position while the *dehC*II-Fr-1 and *dehC*II-Rv-1 primers for amplification of *dehC*II gene from pUKS 107 plasmid covered a region of 800 bp extending from +1 position to +800 position of the gene. In addition to this, the CS-Rv-1 primer was designed to contain a 5′ extension which was overlapping with 3′-end of the *dehC*II segment. The PCR reactions, yielded fragments with expected sizes of 922 bp and 800 bp, for CS Large Domain and *dehC*II gene, respectively (Figure 3.7, panel A).

To construct the CS Large Domain-APH(3′)-II fusion, as the first step, CS Large Domain region and APH(3′)-II should have been amplified. To this end, the CS-Fr-2 and CS-Rv-2 primers were used for CS Large Domain amplification. The APH-Fr-1 and APH-Rv-1 primers for APH(3′)-II amplification from pBS-aph vector covered a region of 828 bp extending from +1 position to +828 position downstream. Again the CS-Rv-2 primer was designed to contain extra sequences at 5′-end which is complementary to the 3′-end of the resulting PCR amplified APH(3′)-II segment. The PCR reactions yielded fragments with expected sizes of 922 bp and 828 bp which are that of CS Large Domain and APH(3′)-II gene, respectively (Figure 3.7, panel B).

To construct the *dehC*II-APH(3′)-II fusion, primers were designed to amplify *dehC*II structural gene and promoter region from pUKS 107 vector up to translation stop

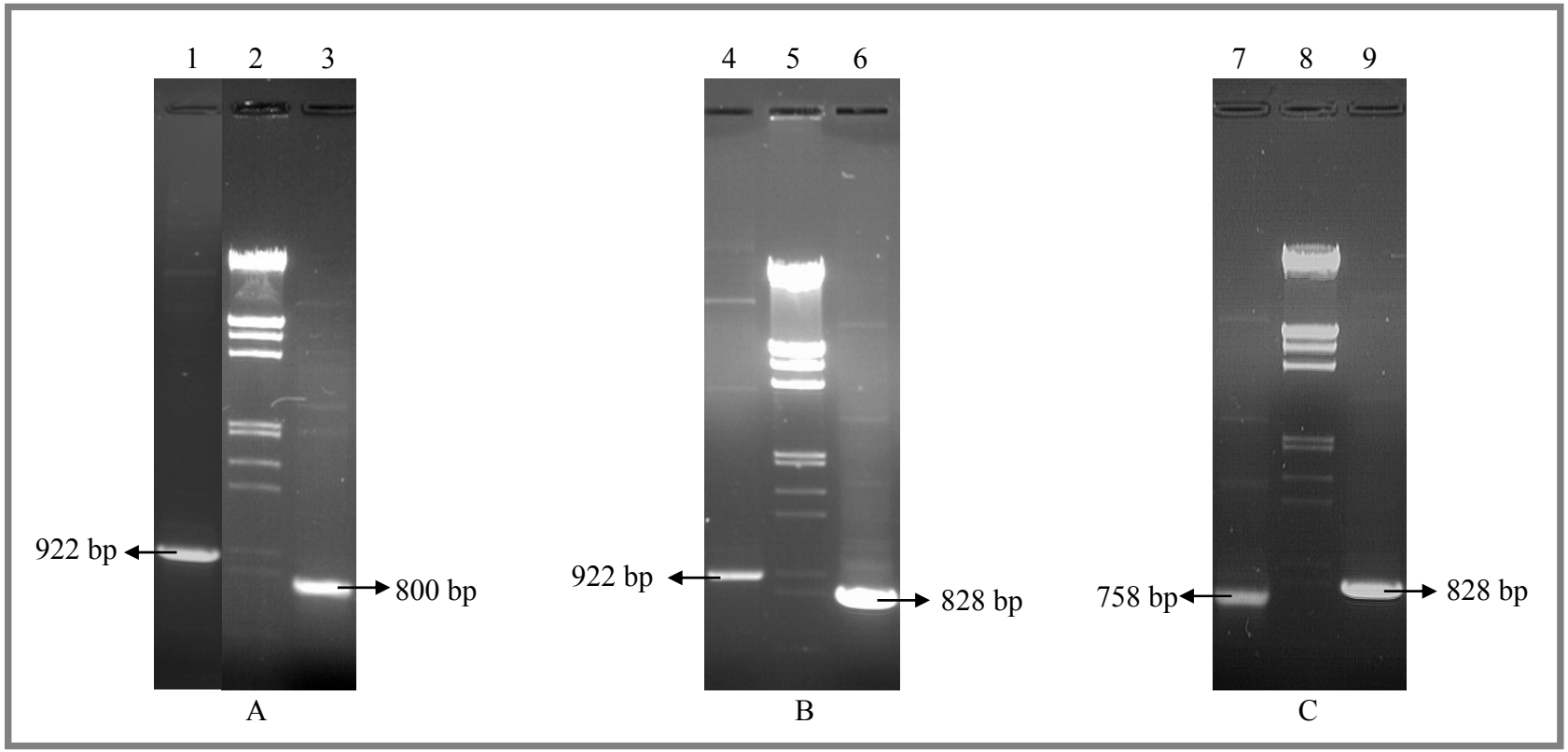


Figure 3.7 PCR amplifications for construction of CS Large Domain-*dehC*II (panel A), CS Large Domain-APH(3′)-II (panel B) and *dehC*II-APH(3′)-II (panel C) fusions at annealing temperature of 50ºC. Lane 1, CS Large Domain fragment; lane 3, *dehC*II fragment; lane 4, CS Large Domain fragment; lane 6, APH(3′)-II fragment; lane 7, *dehC*II fragment; lane 9, APH(3′)-II fragment. Lanes 2, 5 and 8, Lambda DNA/*Eco*RI+*Hin*dIII, molecular size marker.

codon and APH(3′)-II structural gene. The *dehC*II-Fr-2 and *dehC*II-Rv-2 primers for *dehC*II amplification covered a region of 758 bp extending from -72 position upstream to +686 position while the APH-Fr-2 and APH-Rv-2 primers for APH(3′)-II amplification covered a genome region of 828 bp extending from +1 position to +828 position downstream. For SOEing, *dehC*II-Rv-2 primer was designed to have an 5′ extension which was overlapping with the 3′-end of APH(3′)-II segment. The PCR reactions for amplifications of *dehC*II and APH(3′)-II sequences yielded fragments with the expected sizes of 758 bp and 828 bp, respectively (Figure 3.7, panel C).

CS Large Domain (922 bp)-*dehC*II (800 bp) fusion through SOEing was achieved in a PCR reaction using CS-Fr-1 and *dehC*II-Rv-1. For CS Large Domain (922 bp)-APH(3′)-II fragment (828 bp) fusion CS-Fr-2 and APH-Rv-1 were used in the SOE reactions. *dehC*II (758 bp) and APH(3′)-II (828 bp) fragments were fused by using *dehC*II-Fr-2 and APH-Rv-2 primers in the SOE reactions. The SOE reaction products were 1722 bp long CS Large Domain-*dehC*II fusion, 1750 bp long CS Large Domain-APH(3′)-II fusion and 1586 bp long *dehC*II-APH(3′)-II fusion (Figure 3.8, 3.9 and 3.10).

In order to obtain specific amplification products and get rid of the non-specific hybridizations in SOE reactions, the PCR reaction products that would be used as templates were purified by using Qiagen MinElute PCR Purification Kit as described in the Materials and Methods and different annealing temperatures were tried. Although minor PCR products were observed after SOE reactions; bands of the expected sizes were the major products in most of the reactions. The annealing temperature was increased gradually from 50ºC to 67ºC during the construction of the CS Large Domain-*dehC*II fusion and optimal annealing temperature to amplify this construct was found to be 67ºC. However, the product yield of the CS Large Domain-APH(3′)-II construct was quite low at annealing temperature of 67ºC. Optimal annealing temperature to amplify the CS Large Domain-APH(3′)-II was found to be 65ºC while that for *dehC*II-APH(3′)-II fusion was 60ºC.

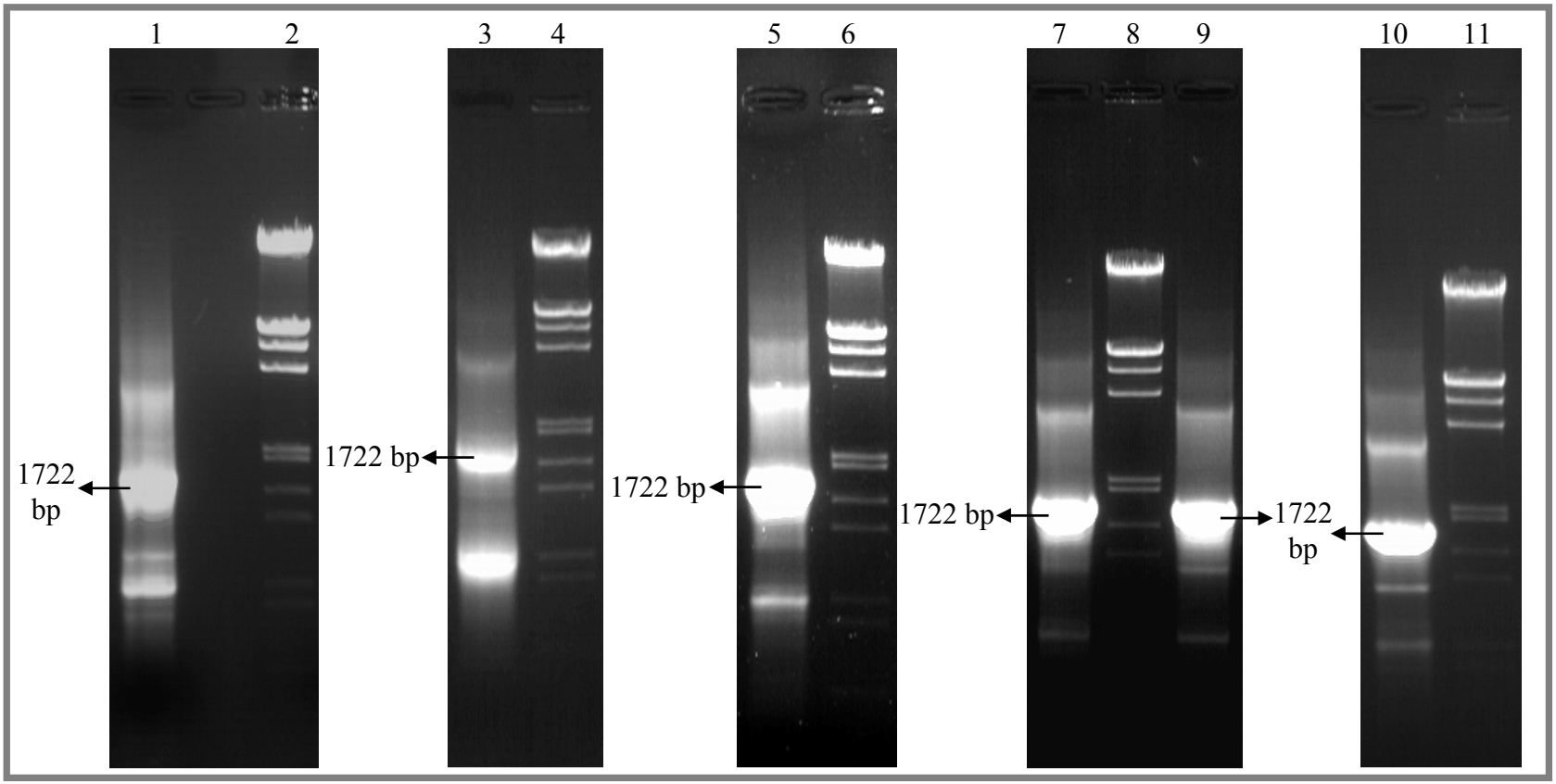


Figure 3.8 SOE reactions for constructing CS Large Domain-*dehC*II fusion at different annealing temperatures: Lane 1, 50ºC; lane 3, 55ºC; lane 5, 60ºC; lane 7, 65ºC; lane 9, 63ºC; lane 10, 67ºC. Lanes 2, 4, 6, 8 and 11, Lambda DNA/*Eco*RI+*Hin*dIII, molecular size marker. The PCR amplification product obtained at 65ºC annealing temperature was used in the cloning experiments.

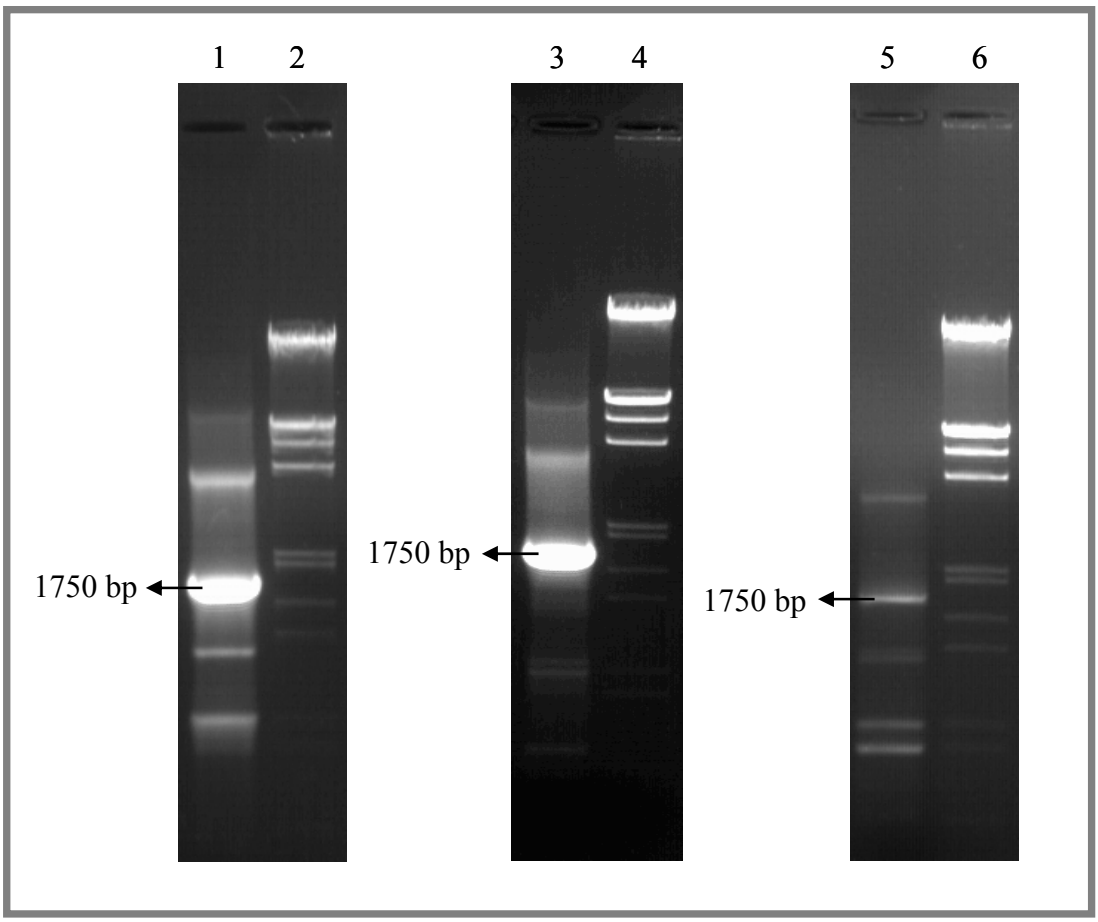


Figure 3.9 SOE reactions for constructing CS Large Domain-APH(3′)-II fusion at annealing temperature of 65ºC (lane 1 and lane 3) and 67ºC (lane 5). Lanes 2, 4 and 6, Lambda DNA/*Eco*RI+*Hin*dIII, molecular size marker. The PCR amplification product obtained at annealing temperature of 65ºC (in lane 3) was used in the cloning experiments.

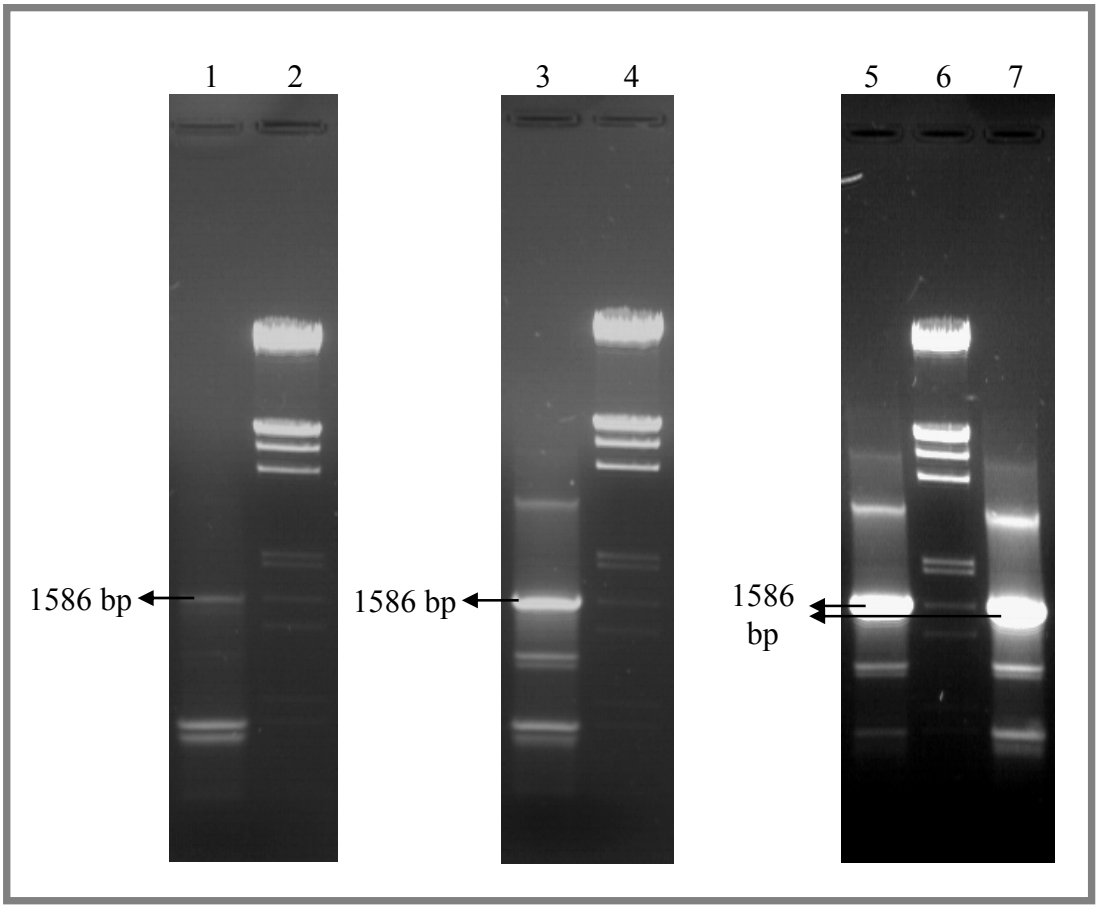


Figure 3.10 SOE reactions for construction of *dehC*II-APH(3′)-II fusion at annealing temperature of 65ºC (lane 1) and 60ºC (lane 3, 5, and 7). Lanes 2, 4 and 6, Lambda DNA/*Eco*RI+*Hin*dIII, molecular size marker. The fusion construct showed in lane 5 was used in the cloning experiments.

3.4 Cloning of PCR Amplified Fusion Constructs

PCR amplified fusion constructs, 1722 bp long CS Large Domain-*dehC*II fusion, 1750 bp long CS Large Domain-APH(3′)-II fusion and 1586 bp long *dehC*II-APH(3′)-II fusion, were cloned in *E. coli* by using Qiagen PCR Cloning Kit (Qiagen Inc., Valencia, USA). Qiagen PCR Cloning Kit uses the advantage of highly efficient cloning of PCR products, which are generated using *Taq* and other non-proofreading DNA polymerases, through UA-based ligation (Figure 3.11, 3.12 and 3.13). Available in the kit, is the pDrive Cloning Vector in a linear form with a U overhang at each end and hybridizes with the single A overhang at each end of PCR products which is added by *Taq* and other non-proofreading DNA polymerases.

The recombinant colonies were selected through blue/white screening and total 8 white colonies of CS Large Domain-*dehC*II fusion, 70 white colonies of CS Large Domain-APH(3′)-II fusion, 104 white colonies of *dehC*II-APH(3′)-II fusion were obtained. Plasmids of 8 CS Large Domain-*dehC*II fusion transformants, 30 CS Large Domain-APH(3′)-II fusion transformants and 28 *dehC*II-APH(3′)-II fusion transformants were isolated as described in the Materials and Methods in order to check the presence of fusion constructs (Figure 3.14).

In addition to this, another cloning experiment was performed for CS Large Domain-*dehC*II fusion using pGEM®-T Vector Systems II Cloning Kit (Promega Co., Madison, WI, USA). High numbers of white colonies were picked up and plasmids of the 12 putative white colonies were isolated as described in the Materials and Methods.

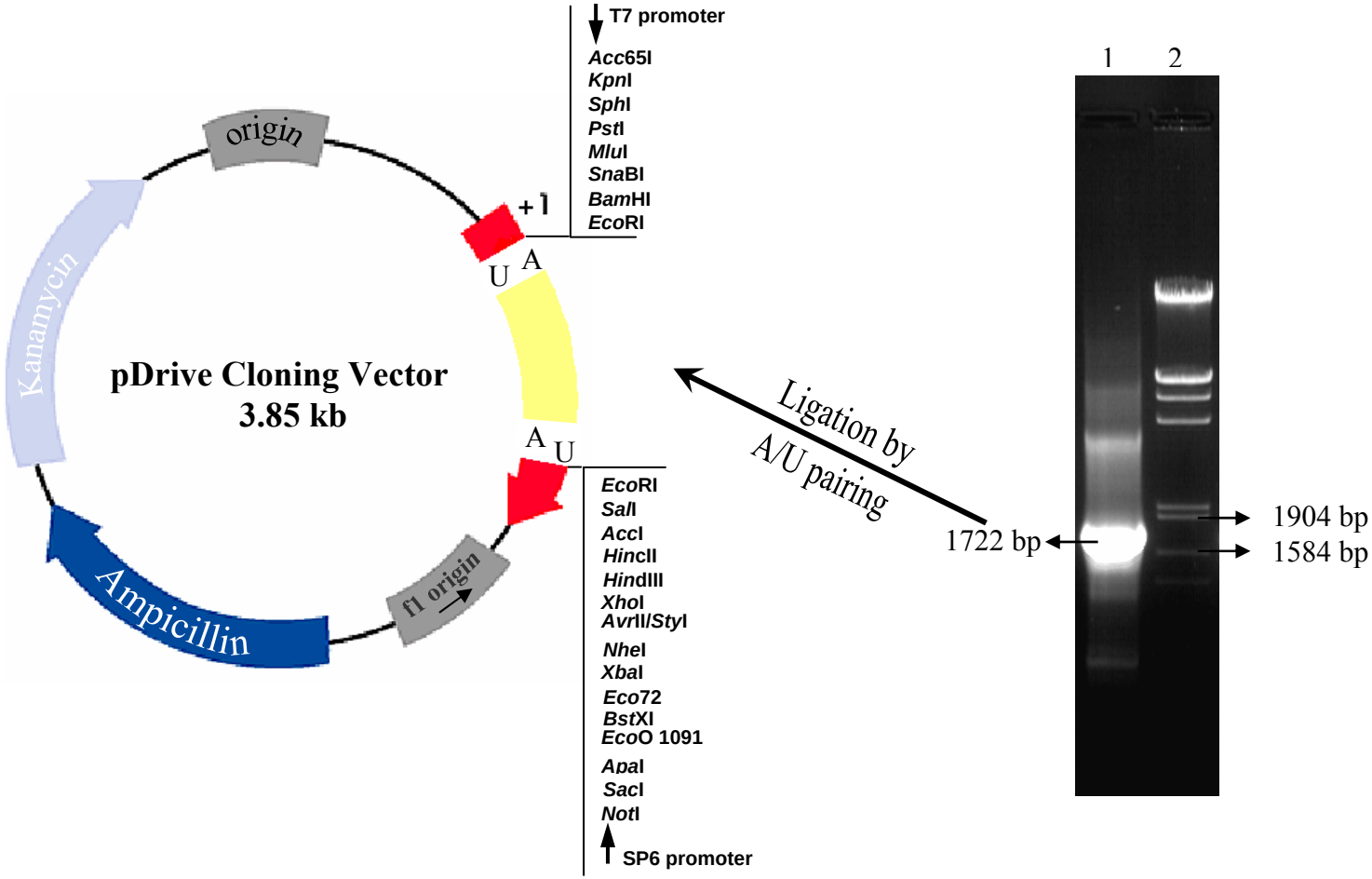


Figure 3.11 The schematic representation of cloning of CS Large Domain-*dehC*II fusion construct. Lane 1, 1722 bp PCR amplified CS Large Domain-*dehC*II fusion construct and lane 2, Lambda DNA/*Eco*RI+*Hin*dIII, molecular size marker.

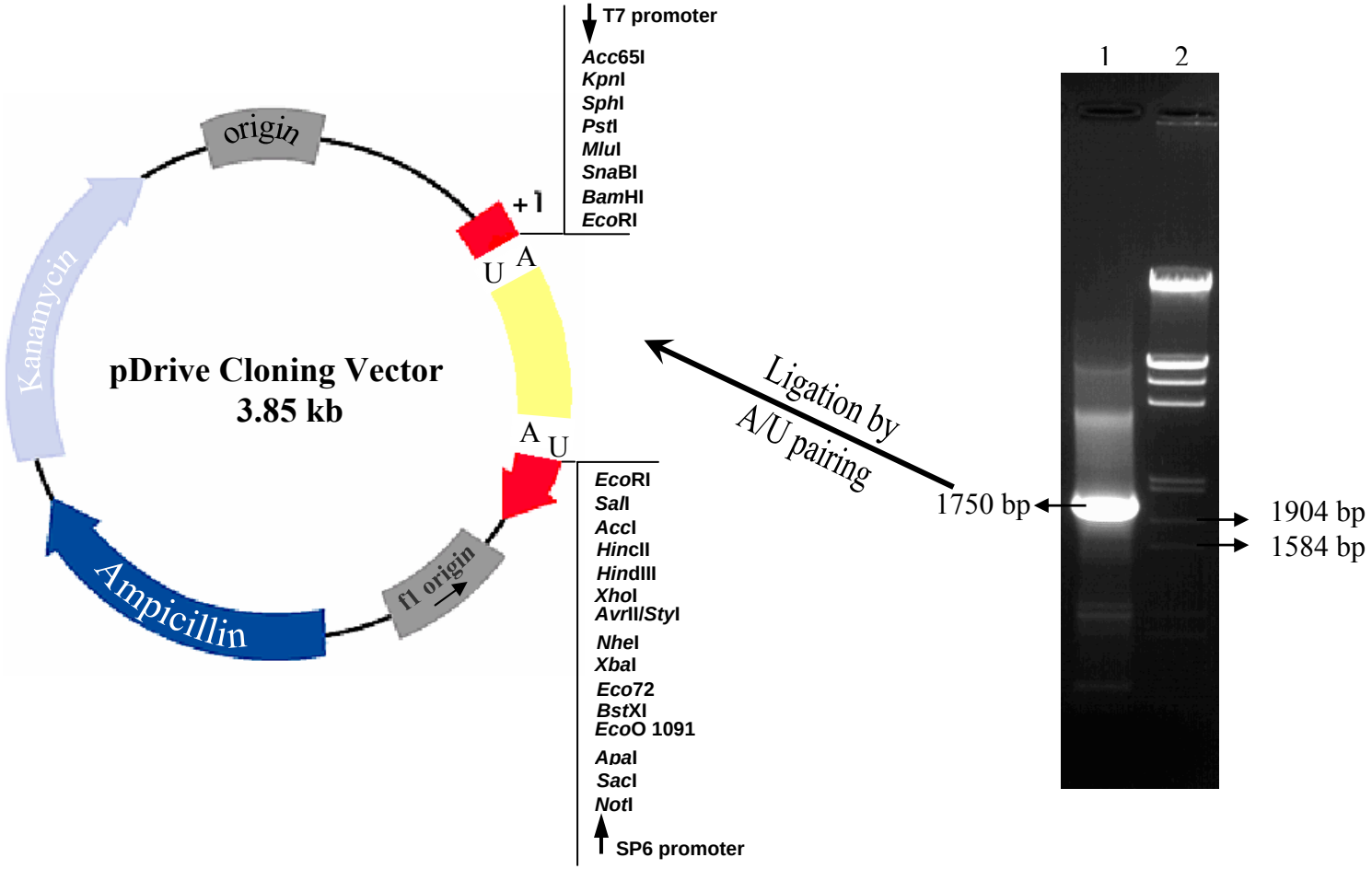


Figure 3.12 The schematic representation of cloning of CS Large Domain-APH(3′)-II fusion construct. Lane 1, 1750 bp PCR amplified CS Large Domain-APH(3′)-II fusion construct and lane 2, Lambda DNA/*Eco*RI+*Hin*dIII, molecular size marker.

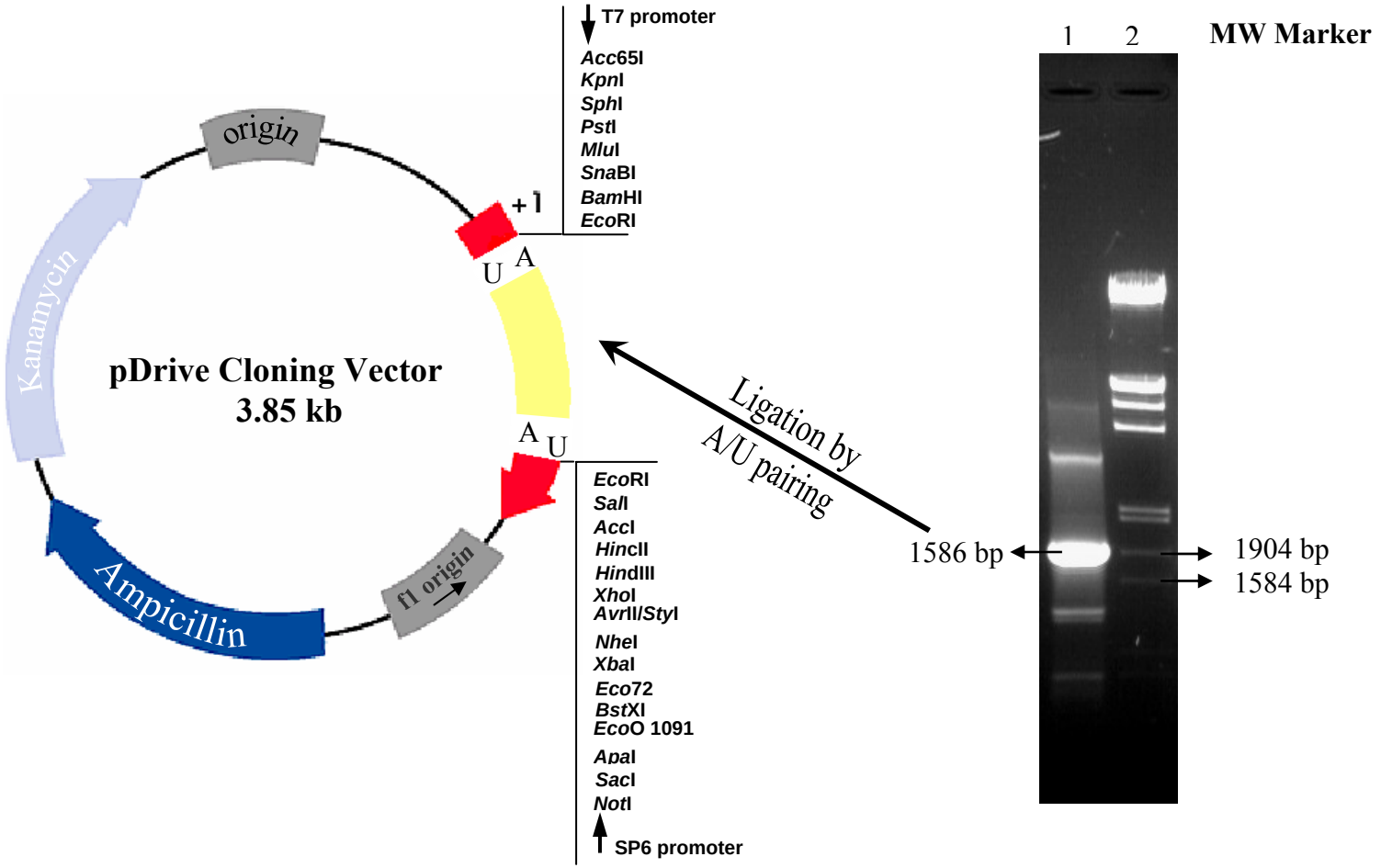


Figure 3.13 The schematic representation of cloning of *dehC*II-APH(3′)-II fusion construct. Lane 1, 1586 bp PCR amplified *dehC*II-APH(3′)-II fusion construct and lane 2, Lambda DNA/*Eco*RI+*Hin*dIII, molecular size marker.

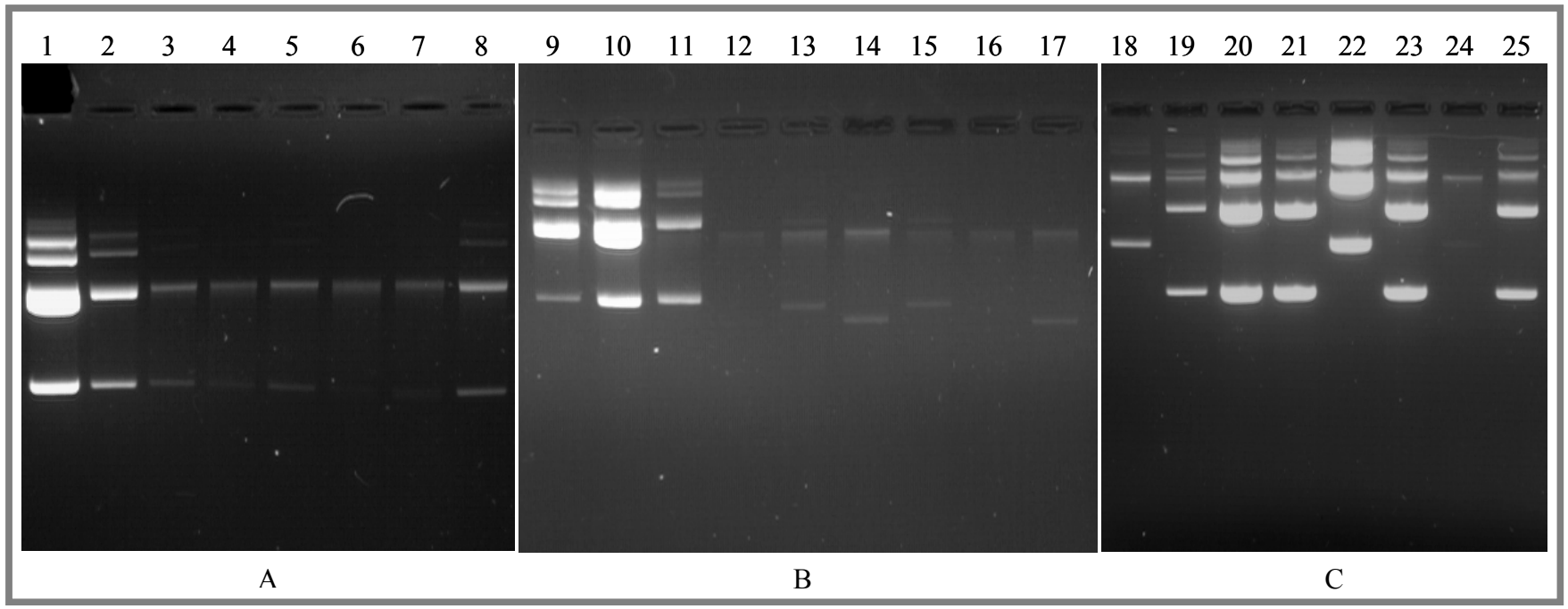


Figure 3.14 Plasmids isolated from putative recombinant colonies containing fusion constructs. Panel A: Plasmids harboring CS Large Domain-*dehC*II fusion construct (Lane 1, M6-1; lane 2, M6-2; lane 3, M6-3; lane 4, M6-4; lane 5, M6-5; lane 6, M6-6; lane 7, M6-7; lane 8, M6-8), Panel B: Plasmids harboring CS Large Domain-APH(3′)-II construct (Lane 9, S5-13; lane 10, S5-39; lane 11, S5-42; lane 12, S5-48; lane 13, S5-50; lane 14, S5-51; lane 15, S5-56; lane 16, S5-57; lane 17, S5-59), Panel C: Plasmids harboring *dehC*II-APH(3′)-II fusion construct (Lane 18, E5-15; lane 19, E5-4; lane 20, E5-16; lane 21, E5-23; lane 22, E5-25; lane 23, E5-30; lane 24, E5-36; lane 25, E5-53). Qiagen PCR Cloning Kit was used for cloning.

3.5 Characterization of Cloned Fusion Constructs

Restriction enzyme digestions were performed to confirm the presence of CS Large Domain-*dehC*II, CS Large Domain-APH(3′)-II fusion, and *dehC*II-APH(3′)-II fusion inserts in the putative recombinant plasmids which were isolated from the white colonies. Based on the restriction maps, which were constructed using RestrictionMapper program as described in the Materials and Methods (Figure 3.16, 3.20, 3.23 and Table 3.1, 3.2, 3.3), initial screening of putative recombinant plasmids containing all three fusion constructs was according to the *Eco*RI digestion pattern (Figure 3.15). There are two *Eco*RI sites in the MCS of the pDrive Vector flanking the insertion site so these digestions resulted in the excision of the fusion inserts which were 1722 bp for CS Large Domain-*dehC*II fusion construct, 1750 bp for CS Large Domain-APH(3′)-II fusion construct, and 1586 bp for *dehC*II-APH(3′)-II fusion construct each together with 3850 bp linear pDrive Cloning Vector. According to these results, three plasmids (M6-1, M6-2, M6-8) containing CS Large Domain-*dehC*II fusion construct, three plasmids (S5-13, S5-39, S5-42) containing CS Large Domain-APH(3′)-II fusion construct, and four plasmids (E5-15, E5-25, E5-37, E5-58) containing *dehC*II-APH(3′)-II fusion construct were identified as true recombinants.

For further characterization, recombinant plasmids M6-1 and M6-2 for CS Large Domain-*dehC*II fusion construct, recombinant plasmids S5-39 for CS Large Domain-APH(3′)-II fusion construct, recombinant plasmids E5-15 and E5-25 for *dehC*II-APH(3′)-II fusion construct were subjected to a number of single and double digestions with several enzymes including: *Eco*RI, *Hin*dIII, *Pvu*II, *Kpn*I, *Bam*HI, *Ava*II, *Sph*I, *Sal*I and *Pst*I.

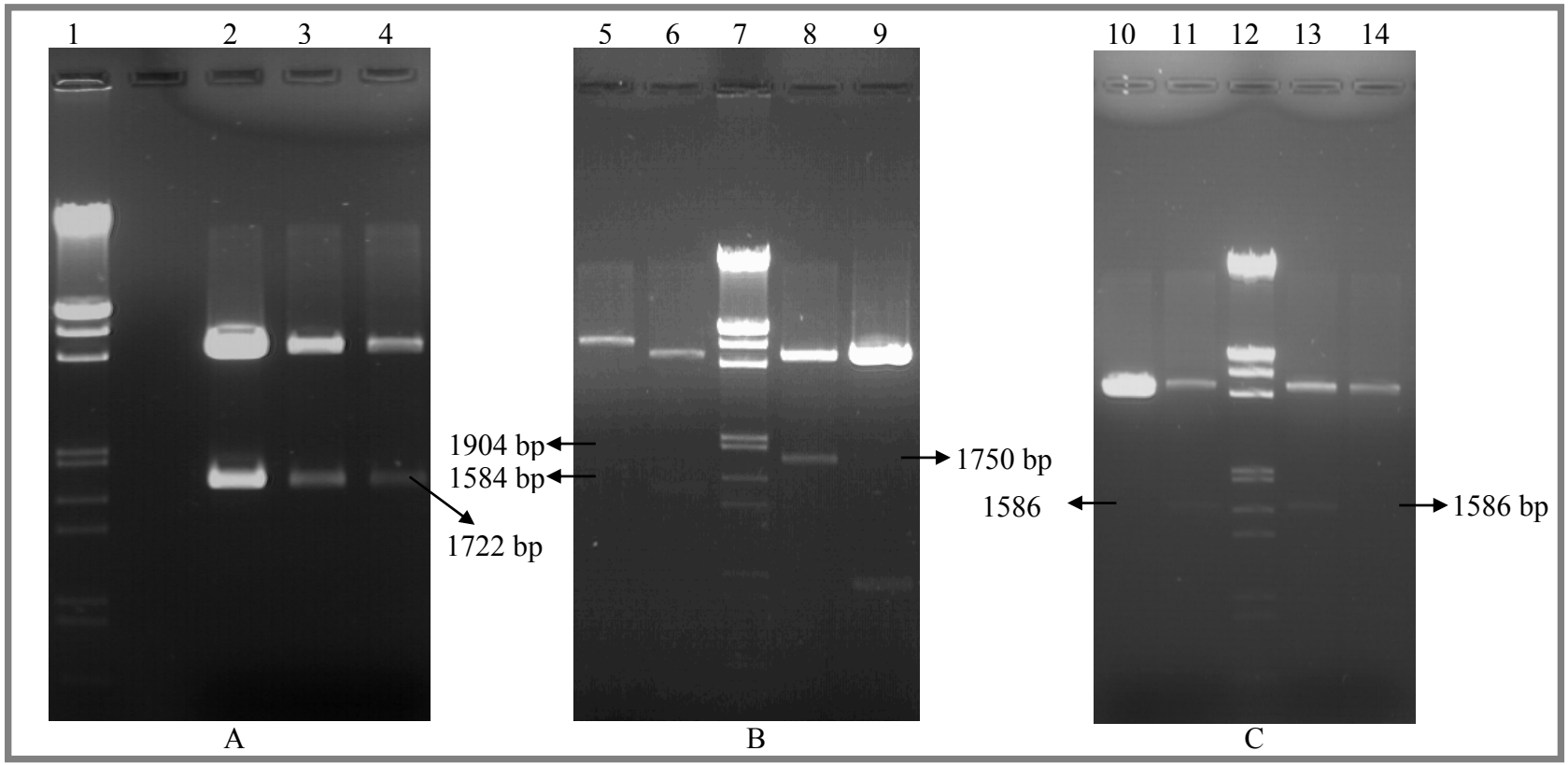


Figure 3.15 *Eco*RI digestions of putative recombinant plasmids isolated from white colonies. Panel A: Recombinant plasmids harboring CS Large Domain-*dehC*II fusion construct (Lane 2, M6-1; lane 3, M6-2; lane 4, M6-8), Panel B) Recombinant plasmids harboring CS Large Domain-APH(3′)-II fusion construct (Lane 5, S5-4; lane 6, S5-8; lane 8, S5-13; lane 9, S5-17), Panel C: Recombinant plasmids harboring *dehC*II-APH(3′)-II fusion construct (Lane 10, E5-17; lane 11, E5-15; lane 13, E5-58; lane 14, E5-62). Lanes 1, 7 and 12, Lambda DNA/*Eco*RI+*Hin*dIII, molecular size marker.

Recombinant plasmids containing CS Large Domain-*dehC*II fusion construct were digested by *Eco*RI, *Hin*dIII, *Ava*II, and *Hin*dIII/*Kpn*I and *Hin*dIII/*Bam*HI combinations. As indicated before there are two *Eco*RI sites in the vector flanking the insert, therefore *Eco*RI digestion yielded two bands: a 1722 bp long PCR amplified CS Large Domain-*dehC*II fusion construct and linear pDrive Cloning Vector of 3850 bp in length. There is one *Hin*dIII site at position 730 bp within the CS Large Domain-*dehC*II fusion construct and another one on the vectors' MCS, located in the upstream of the insert; so *Hin*dIII digestion gave rise to two bands in the sizes of 730 bp fusion fragment and 4815 bp consisting of the remaining fusion fragment and the vector. There is a single site for *Ava*II within the fusion construct at position 1632 bp and two more sites on the vectors' MCS. As a result, digestion with *Ava*II excised three fragments in the sizes of 2909 bp, 2426 bp and 222 bp. Besides the single sites both for *Hin*dIII and *Kpn*I on the vectors' MCS; there is one more *Hin*dIII site at position 730 bp and one more *Kpn*I site at position 6 bp within the CS Large Domain-*dehC*II fusion construct. Therefore *Hin*dIII/*Kpn*I double digestion resulted in three fragments of 3787 bp, 1028 bp, and 724 bp in length. There is one *Hin*dIII site at position 730 bp within the CS Large Domain-*dehC*II fusion construct and another one on the vectors' MCS. There is also a site for *Bam*HI on the vectors' MCS so *Hin*dIII/*Bam*HI double digestion cut out three fragments in the sizes of 3816 bp, 999 bp, and 730 bp. Restriction digestion profile for the CS Large Domain-*dehC*II fusion construct is shown in Figure 3.17. The digestion pattern is also shown schematically in Figure 3.18.

According to this restriction analysis, it was concluded that the 1722 bp CS Large Domain-*dehC*II fusion construct was cloned successfully and was inserted into the vector in the opposite orientation relative to the *lac* promoter.

The resulting restriction map was in good agreement with the restriction map of the CS Large Domain-*dehC*II fusion construct which was derived from the known sequences of the *Tp. volcanium* CS gene and *Pseudomonas* sp. Strain CBS3 *dehC*II gene using RestrictionMapper program.

Table 3.1 Restriction sites and their cut positions in CS Large Domain-*dehC*II fusion construct which were determined by RestrictionMapper program.

Noncutter Restriction Enzymes	Cutter Restriction Enzymes	Cut Positions
BamHI	KpnI	6
EcoRI	HindIII	730
EcoRV	AvaII	1632
PvuI	SspI	138,1004
SacI	PvuII	695, 1065
SalI	PstI	562, 841
SmaI	AvaI	1253, 1532
SphI		

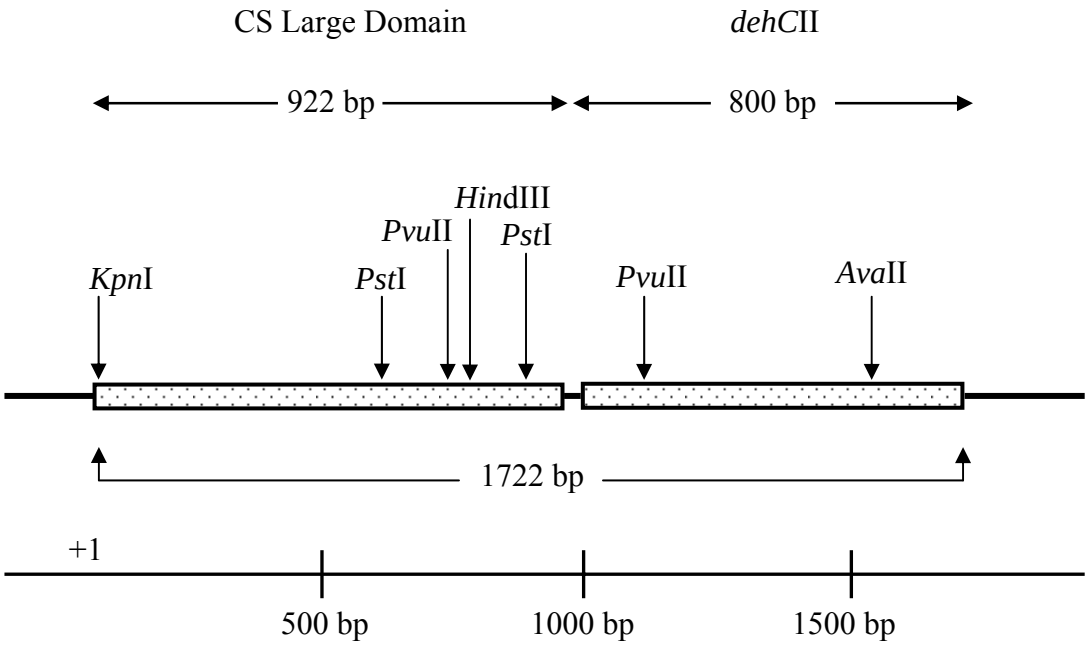


Figure 3.16 Restriction map of the CS Large Domain-*dehC*II fusion construct by software analysis.

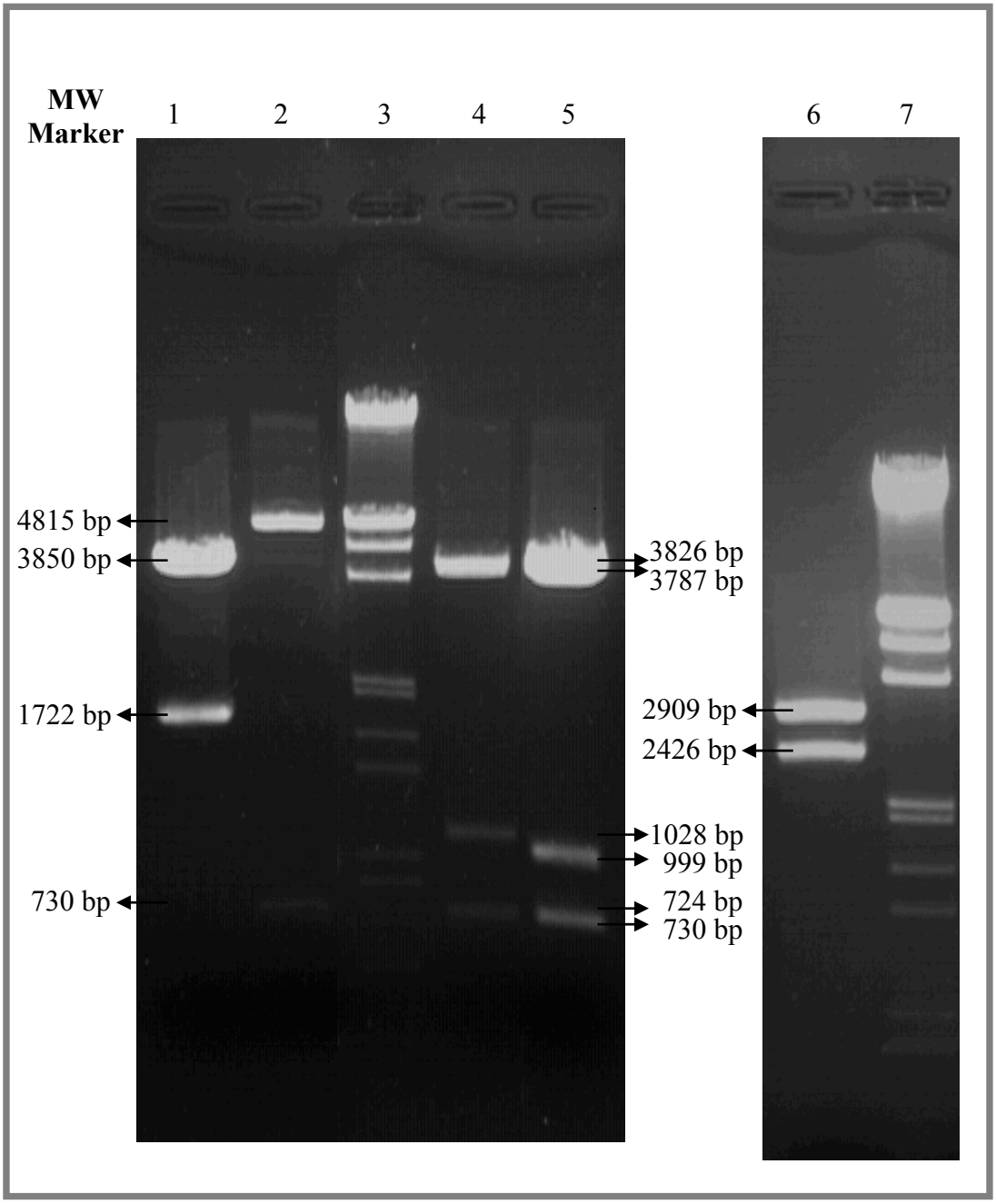
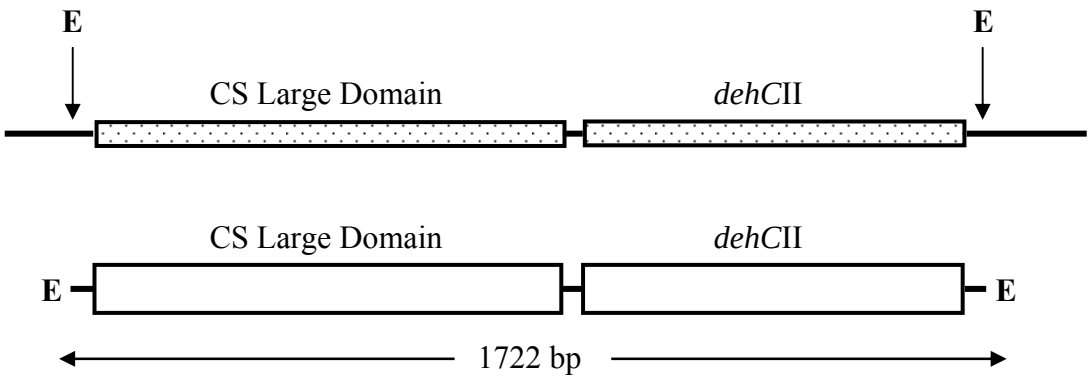
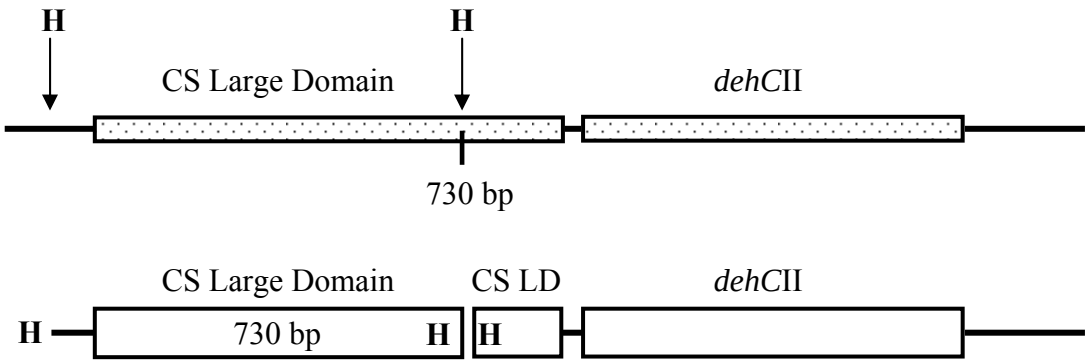


Figure 3.17 Restriction analyses of recombinant plasmids containing CS Large Domain-*dehC*II fusion construct. *Eco*RI digestion excised the complete 1722 bp CS Large Domain-*dehC*II fusion construct (lane1), digestion with *Hin*dIII excised a 730 bp fragment (lane 2), *Hin*dIII/*Kpn*I double digestion excised 1028 bp and 724 bp fragments (lane 4), *Hin*dIII/*Bam*HI double digestion excised 999 bp and 730 bp fragments (lane 5), and digestion with AvaII excised 2909 bp and 2426 bp fragments (lane 6). Lane 3 and lane 7, Lambda DNA/*Eco*RI+*Hin*dIII, molecular size marker.

*Eco*RI Digestion



*Hin*dIII Digestion



*Hin*dIII/*Bam*HI Digestion

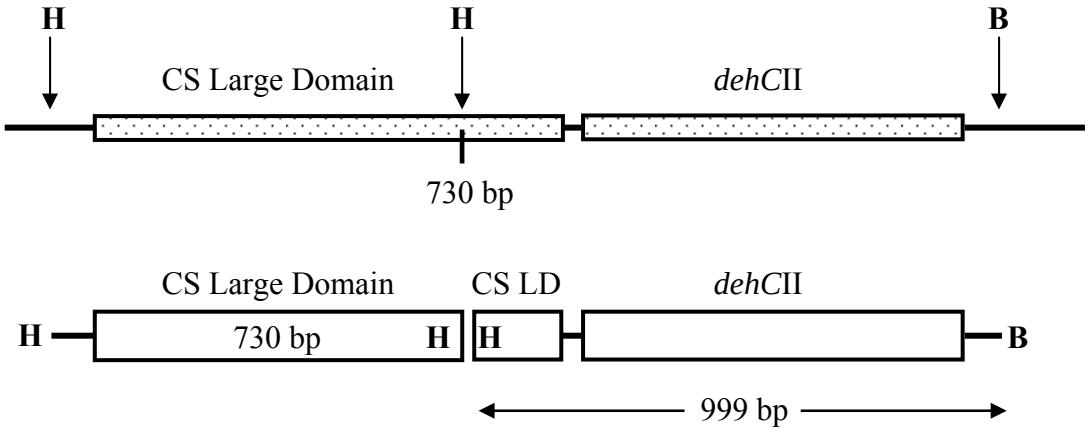
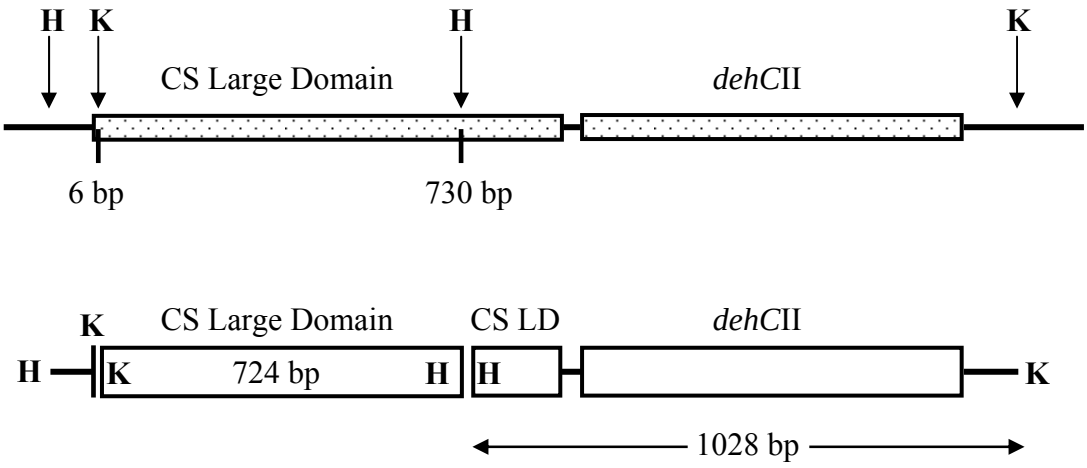
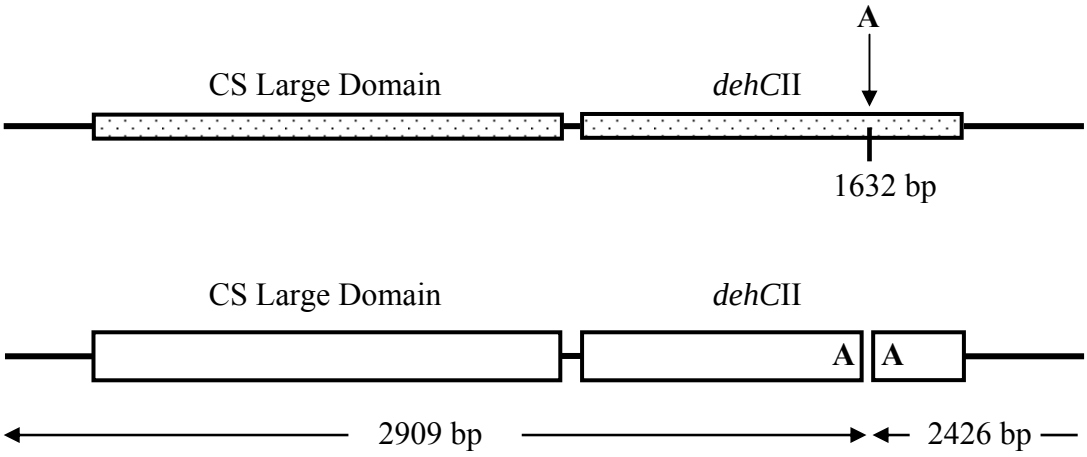


Figure 3.18 Schematic representations of Restriction Enzyme Digestion Sites for CS Large Domain-*dehC*II fusion on pDrive Cloning Vector. **E:** *Eco*RI, **H:** *Hin*dIII, **B:** *Bam*HI, **K:** *Kpn*I, **A:** *Ava*II.

*Hin*dIII/*Kpn*I Digestion



*Ava*II Digestion



Overall

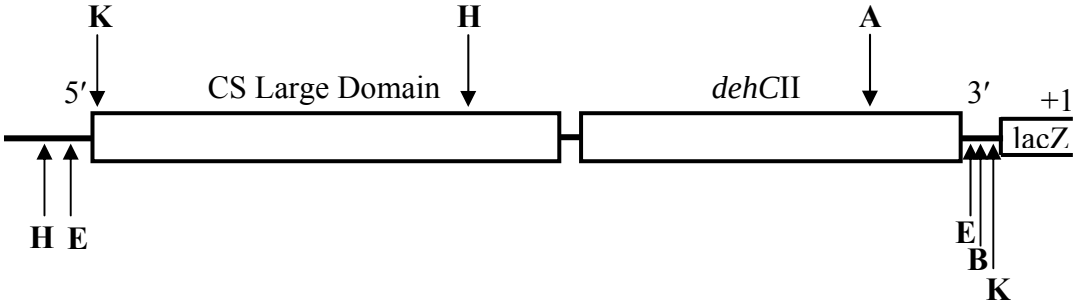


Figure 3.18 continued

The cloning experiments done for CS Large Domain-*dehC*II fusion construct using the pGEM®-T Vector Systems II were unsuccessful. Eventhough high numbers of white colonies were obtained, the restriction digestion results revealed that none of the putative recombinant colonies isolated from these white colonies were found to be true recombinants (Figure 3.19).

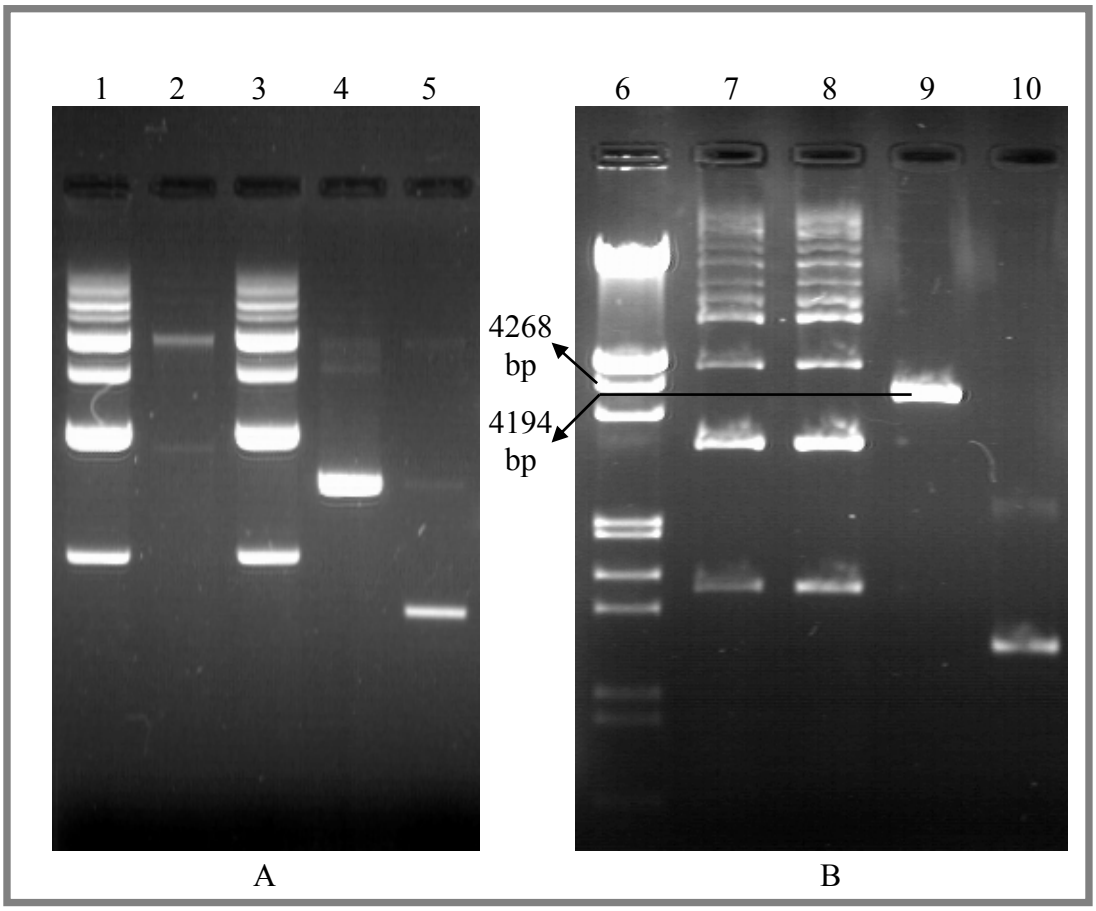


Figure 3.19 A) Plasmids isolated from putative recombinant colonies containing 1722 bp CS Large Domain-*dehC*II fusion construct: Lane 1, M-3; lane 2, M-5; lane 3, M-14; lane 4, M-21; lane 5, M-30. Promega the pGEM®-T Vector (3000 bp) Systems II Cloning Kit was used in the cloning and transformation experiments. B) Restriction analysis of the putative recombinant plasmids. Plasmids were single digested with *Sal*I. Lane 7, M-3; lane 8, M-14; lane 9, M-21; lane 10, M-30. Lane 6, Lambda DNA/*Eco*RI+*Hin*dIII, molecular size marker.

For restriction mapping, the CS Large Domain-APH(3′)-II fusion construct was digested with *Eco*RI, *Hin*dIII, *Ava*II and *Sph*I/*Sac*I combination. *Eco*RI digestion yielded two bands corresponding to; 1750 bp long PCR amplified CS Large Domain-APH(3′)-II fusion construct accompanied by linear pDrive Cloning Vector of 3850 bp in length since there are two *Eco*RI sites in the vector flanking the insert. There is one *Hin*dIII site at position 730 bp within the CS Large Domain-APH(3′)-II fusion construct and another one on the vectors' MCS, located in the upstream of the insert; so *Hin*dIII digestion gave rise to two bands in the sizes of 730 bp fusion fragment and 4843 bp consisting of the remaining fusion fragment and the vector. There is one *Sph*I site at position 1456 bp within the CS Large Domain-APH(3′)-II fusion construct and another one on the vectors' MCS. There is a site for *Sac*I site on the vectors' MCS as well. Therefore *Sph*I/*Sac*I double digestion excised three fragments of 3744 bp, 1519 bp, and 322 bp in length. Besides the single site for *Ava*II within the CS Large Domain-APH(3′)-II fusion construct at position 1567; there are two more sites on the vectors' MCS. As a result, digestion with *Ava*II cut out three fragments in the sizes of 2844 bp, 2519 bp and 222bp. Restriction digestion profile for the CS Large Domain-APH(3′)-II fusion construct is shown in Figure 3.21. The digestion pattern is also shown schematically in Figure 3.22.

This restriction analysis demonstrated that the orientation of the 1750 bp CS Large Domain-APH(3′)-II fusion construct was opposite to the *lac* promoter of the pDrive Cloning Vector.

The resulting restriction map was in good correlation with the restriction map of the CS Large Domain-APH(3′)-II fusion construct which was derived from the known sequences of the *Tp. volcanium* CS gene and *E. coli* transposon Tn5 APH(3´)-II gene using RestrictionMapper program.

These results showed that CS Large Domain-APH(3′)-II fusion construct was successfully cloned in pDrive Cloning Vector.

Table 3.2 Restriction sites and their cut positions in CS Large Domain-APH(3′)-II fusion construct which were determined by RestrictionMapper program.

Noncutter Restriction Enzymes	Cutter Restriction Enzymes	Cut Positions
AvaI	KpnI	6
BamHI	SspI	138
EcoRI	HindIII	730
EcoRV	SphI	1456
PvuI	AvaII	1567
SacI	PvuII	695, 1157
SalI	PstI	562, 841, 1104
SmaI		

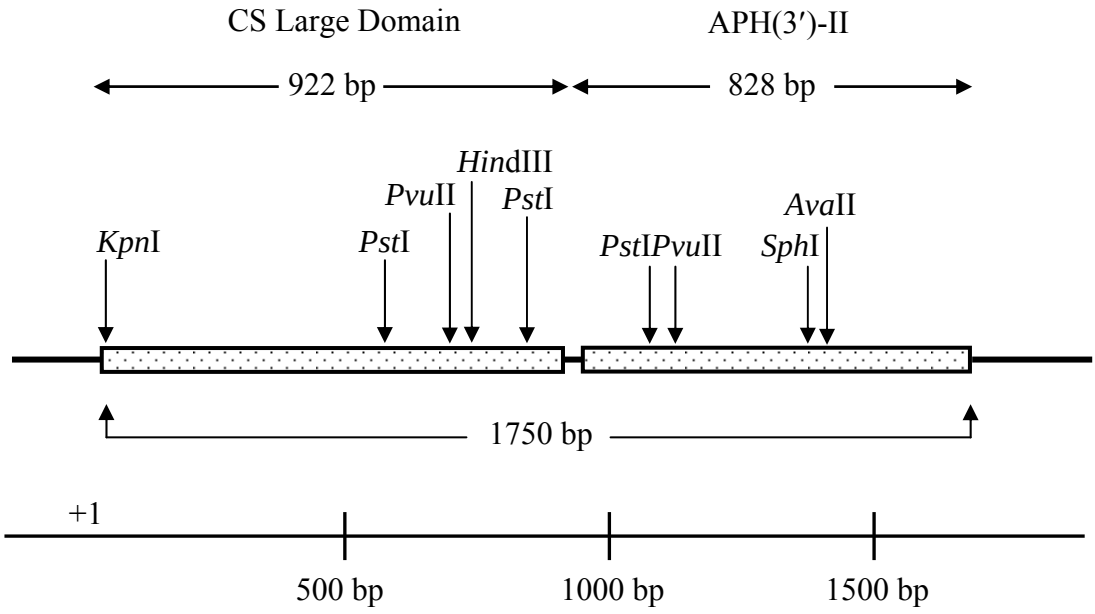


Figure 3.20 Restriction map of the CS Large Domain-APH(3′)-II fusion construct by software analysis.

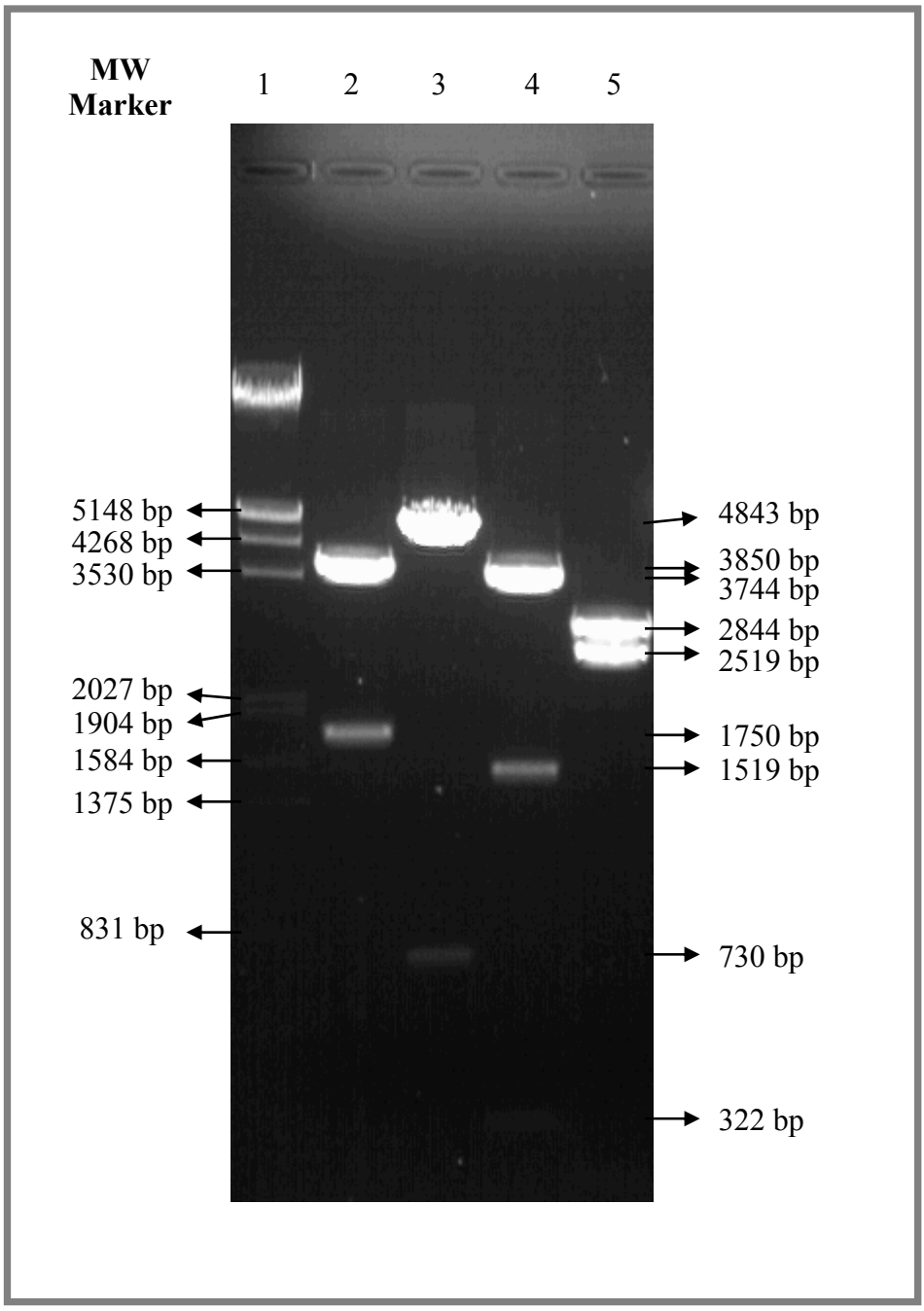
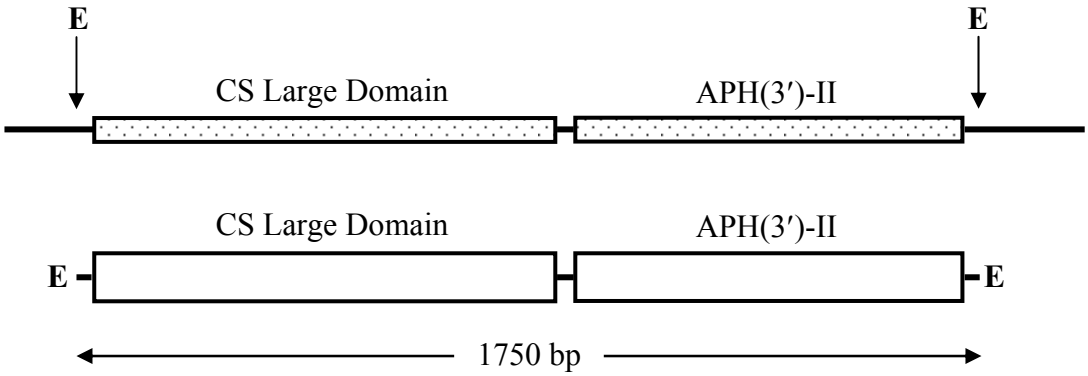
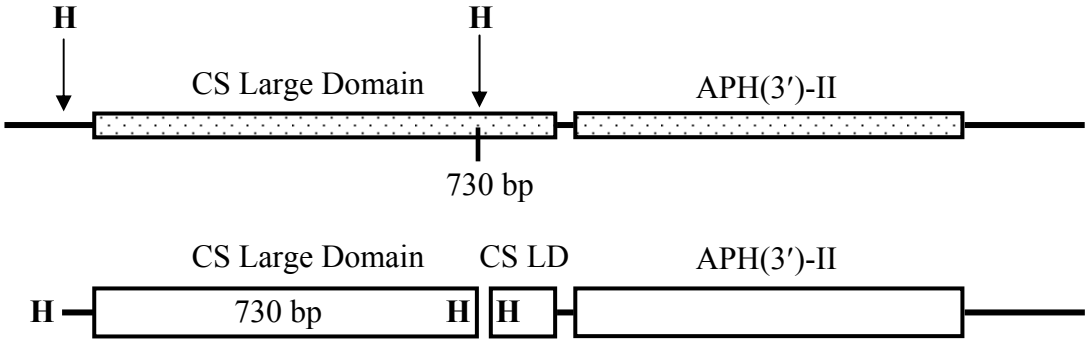


Figure 3.21 Restriction analyses of recombinant plasmids containing CS Large Domain-APH(3′)-II fusion construct. Digestion with *Eco*RI excised the complete 1750 bp CS Large Domain-APH(3′)-II fusion construct (lane2), *Hin*dIII digestion excised 730 bp fragment (lane 3), double digestion with *Sph*I/*Sac*I excised 1519 bp and 322 bp fragments (lane 4), and *Ava*II digestion excised 2844 bp and 2519 bp fragments (lane 5). Lane 1, Lambda DNA/*Eco*RI+*Hin*dIII, molecular size marker.

*Eco*RI Digestion



*Hin*dIII Digestion



*Ava*II Digestion

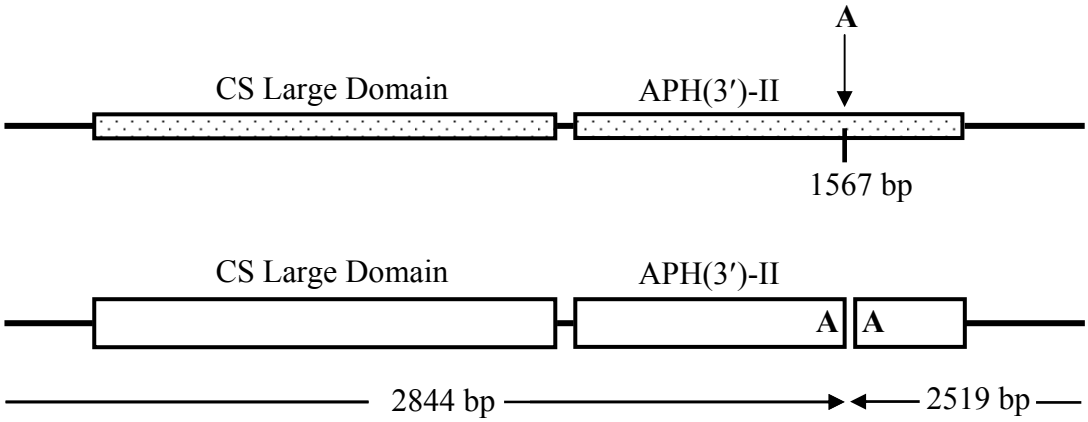
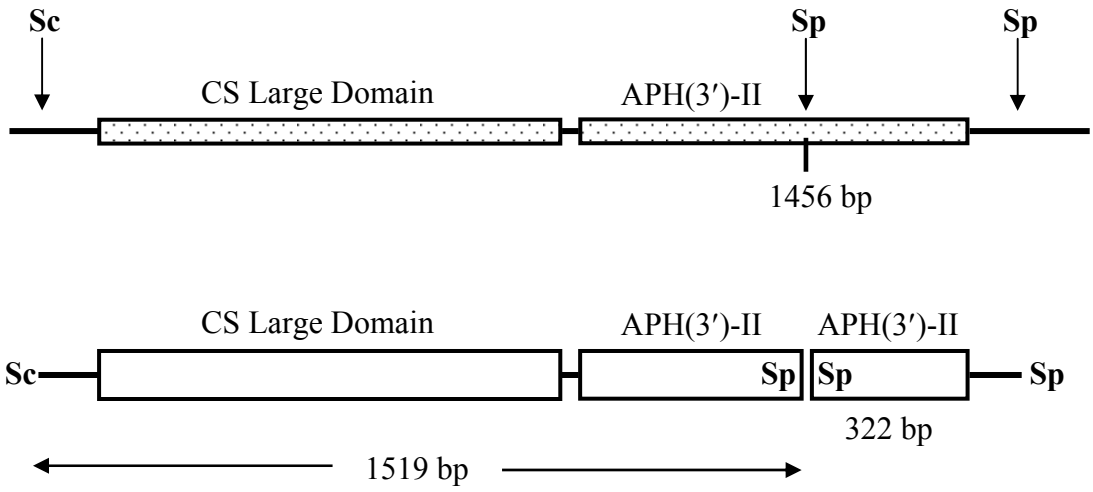


Figure 3.22 Schematic representations of Restriction Enzyme Digestion Sites for CS Large Domain-APH(3′)-II fusion on pDrive Cloning Vector. **E:** *Eco*RI, **H:** *Hin*dIII, **A:** *Ava*II, **Sp:** *Sph*I, **Sc:** *Sac*I.

*Sph*I/*Sac*I Digestion



Overall

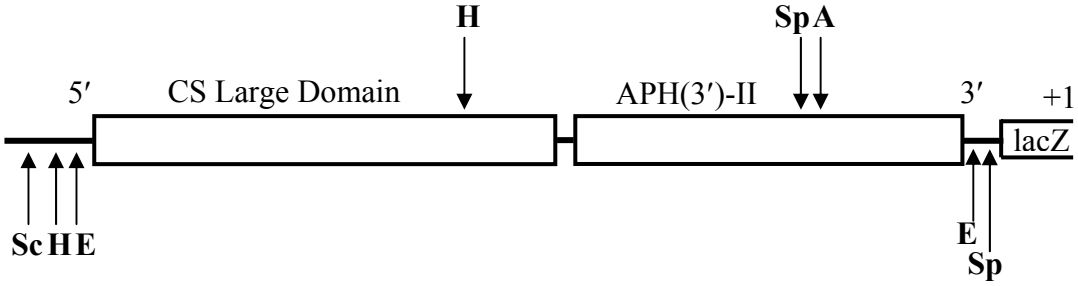


Figure 3.22 continued

For restriction mapping of the *dehC*II-APH(3′)-II fusion construct, recombinant plasmid was digested using *Eco*RI, *Bam*HI, *Pst*I, *Sph*I, *Pvu*II, and *Ava*II. *Eco*RI digestion yielded two bands corresponding to; 1586 bp long PCR amplified *dehC*II-APH(3′)-II fusion construct together with linear pDrive Cloning Vector of 3850 bp in length as there are two *Eco*RI sites in the vector flanking the insert. There is one *Bam*HI site at position 1 bp within the *dehC*II-APH(3′)-II fusion construct and another one on the vectors' MCS, located in the upstream of the insert; so *Bam*HI single digestion excised the *dehC*II-APH(3′)-II fusion construct which was in the size of 1586 bp together with the 3850 bp linear pDrive Cloning Vector. Digestion with *Pst*I resulted in two bands; one of which was the fusion fragment at the size of 669 bp and the other one was the vector plus the remaining fusion fragment at the size of 4752 bp since there is a single *Pst*I site within the fusion construct at position 940 bp and another one on the vectors' MCS. Besides the sites for *Sph*I at position 74 bp and at position 1292 bp within the *dehC*II-APH(3′)-II fusion construct; there is one more *Sph*I site on the vectors' MCS. Therefore *Sph*I single digestion gave rise to three fragments which were 3881 bp, 1218 bp, and 322 bp in length. In the case of *Pvu*II, the single digestion cut out four fragments which were in the sizes of 3387 bp, 845 bp, 779 bp and 410 bp as two of sites are positioned at 214 bp and 993 bp within the *dehC*II-APH(3′)-II fusion construct where the other two of them are located on the vectors' MCS. There are three sites for *Ava*II: two of them positioned on the vectors' MCS and the third one is within the *dehC*II-APH(3′)-II fusion construct positioned at 1403 bp. As a result, *Ava*II single digestion yielded three fragments of 2680 bp, 2519 bp, and 222 bp in length. Restriction digestion profile for the *dehC*II-APH(3′)-II fusion construct is shown in Figure 3.24. The digestion pattern is also shown schematically in Figure 3.25.

Judging along with this restriction analysis, it is quite possible to suggest that the 1586 bp *dehC*II-APH(3′)-II fusion construct was inserted in the opposite orientation with respect to the *lac* promoter of the pDrive Cloning Vector.

The resulting restriction map was in good correlation with the restriction map of the

Table 3.3 Restriction sites and their cut positions in *dehC*II-APH(3′)-II fusion construct which were determined by RestrictionMapper program.

Noncutter Restriction Enzymes	Cutter Restriction Enzymes	Cut Positions
EcoRI	BamHI	1
EcoRV	SspI	153
HindIII	PstI	940
KpnI	AvaII	1403
SacI	PvuII	214, 993
SalI	AvaI	381, 681
SmaI	SphI	74, 1292

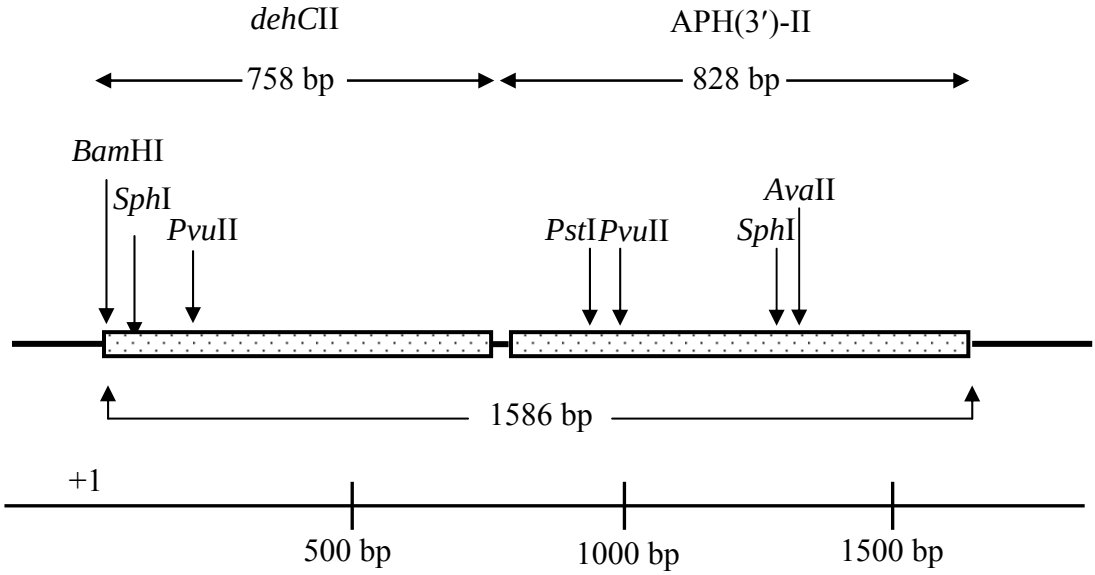


Figure 3.23 Restriction map of the *dehC*II-APH(3′)-II fusion construct by software analysis.

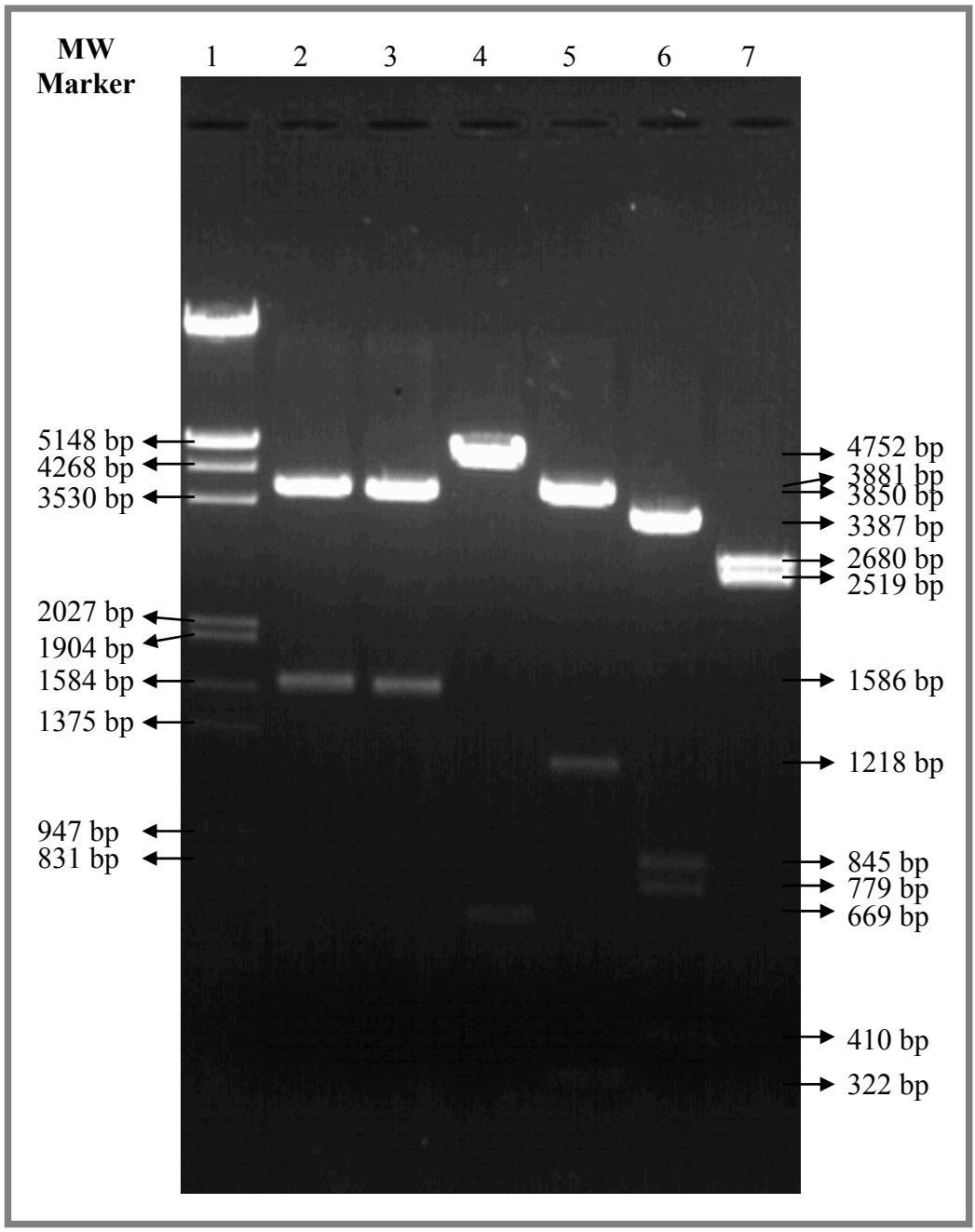
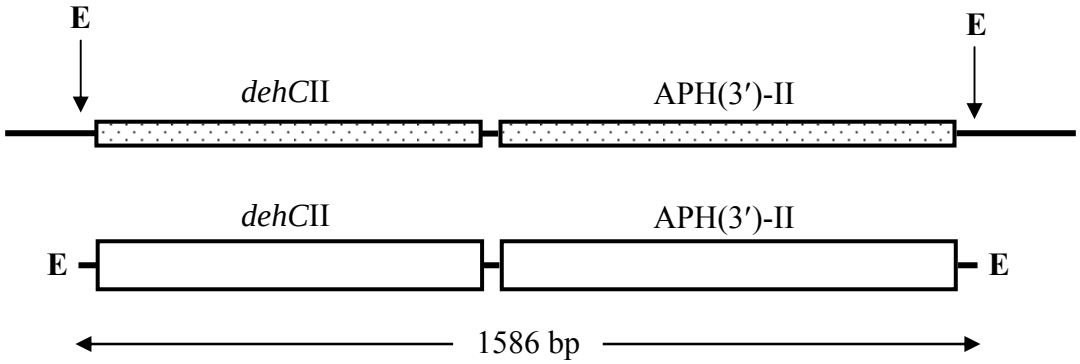
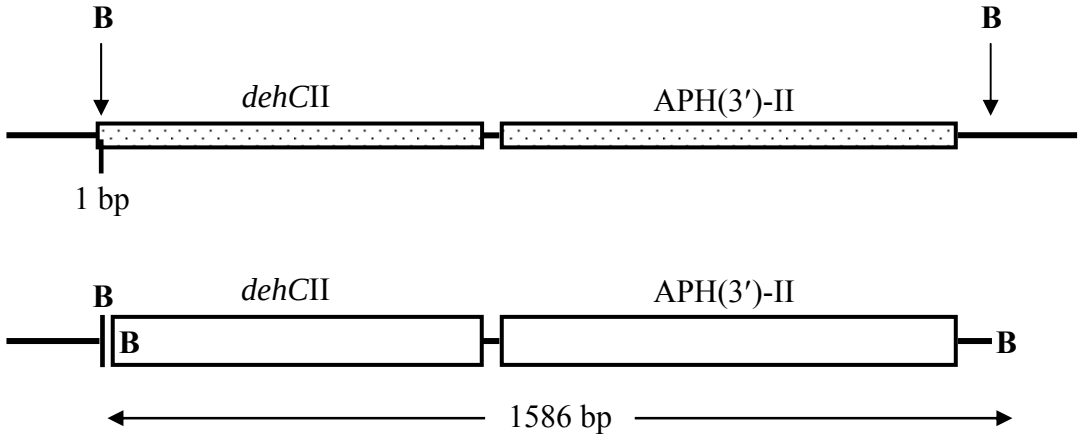


Figure 3.24 Restriction analyses of recombinant plasmids containing *dehC*II-APH(3′)-II fusion. Digestion with both *Eco*RI and *Bam*HI excised the complete 1586 bp *dehC*II-APH(3′)-II fusion construct (lane2 and lane 3), *Pst*I digestion excised 669 bp fragment (lane 4), digestion with *Sph*I excised 1218 bp and 322 bp fragments (lane 5), *Pvu*II digestion excised three fragments in lengths of 845 bp, 779 bp, and 410 bp (lane 6), and digestion with *Ava*II excised 2680 bp and 2519 bp fragments (lane 7). Lane 1, Lambda DNA/*Eco*RI+*Hin*dIII, molecular size marker.

*Eco*RI Digestion



*Bam*HI Digestion



*Pst*I Digestion

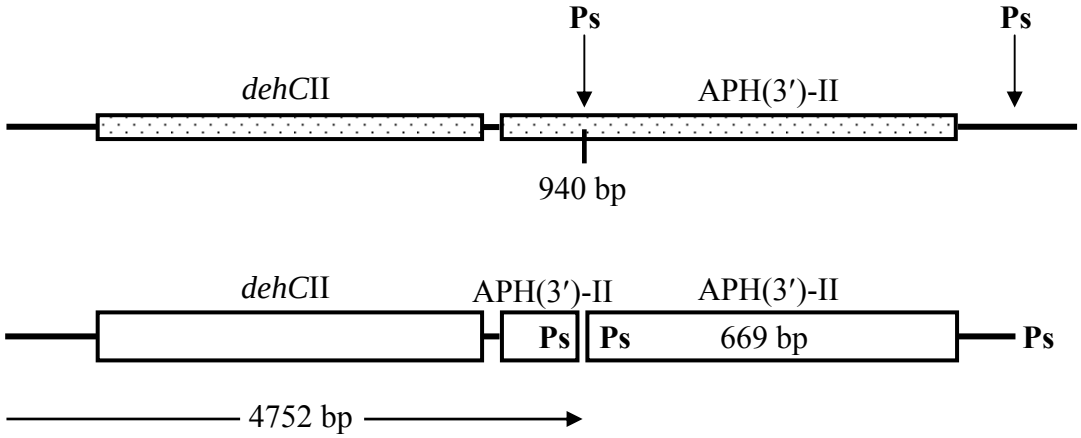
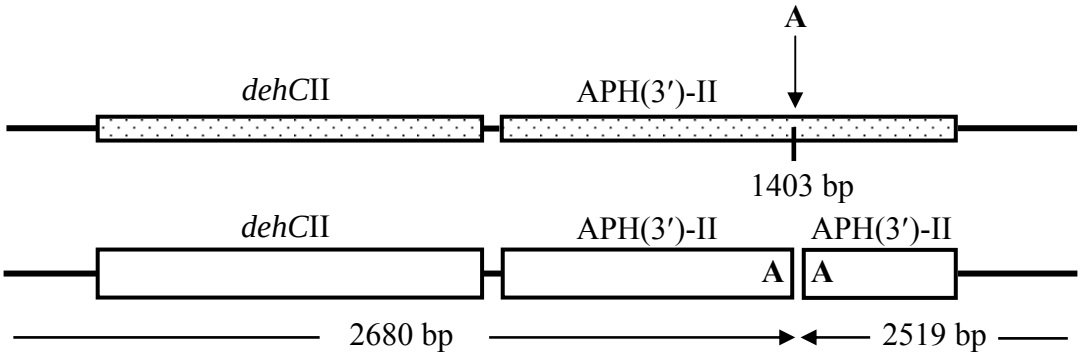
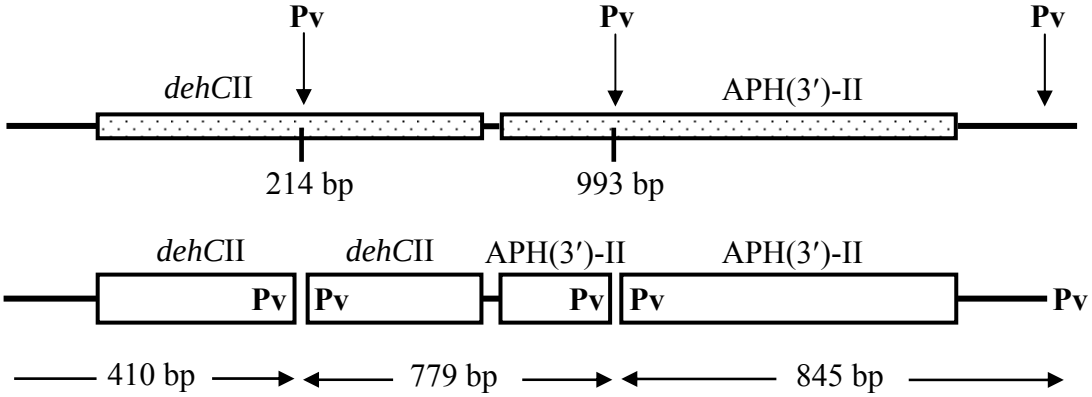


Figure 3.25 Schematic representations of Restriction Enzyme Digestion Sites for *dehC*II-APH(3′)-II fusion on pDrive Cloning Vector. **E:** *Eco*RI, **B:** *Bam*HI**, Ps:** *Pst*I, **A:** *Ava*II, **Pv:** *Pvu*II, **S:** *Sph*I.

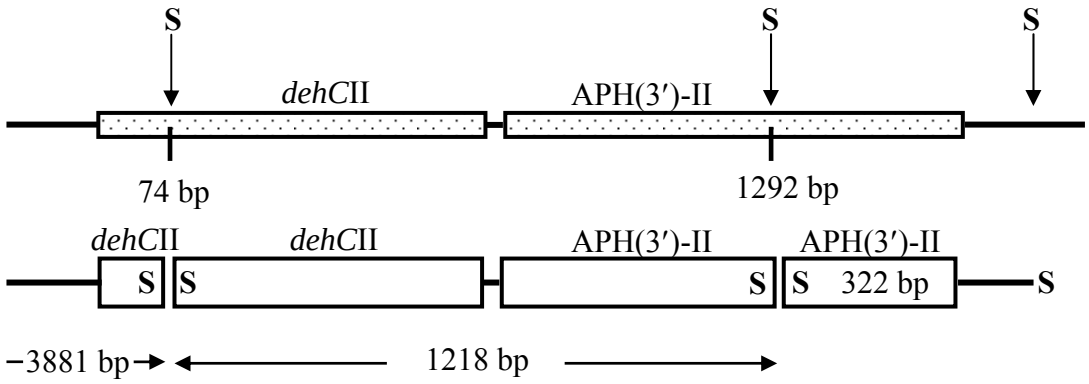
*Ava*II Digestion



*Pvu*II Digestion



*Sph*I Digestion



Overall

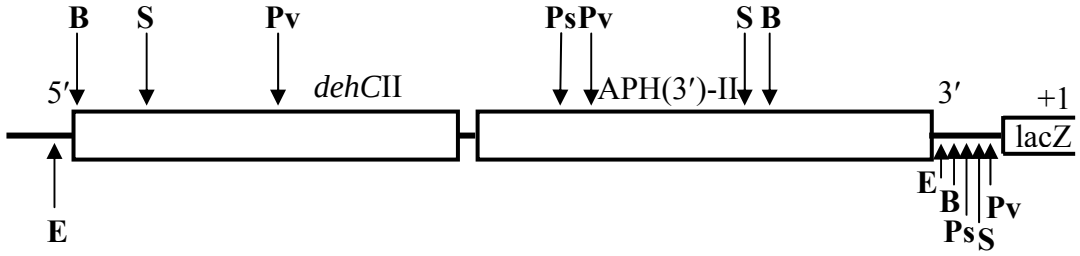


Figure 3.25 continued

*dehC*II-APH(3′)-II fusion construct which was derived from the known sequences of the *Pseudomonas* sp. Strain CBS3 *dehC*II gene and *E. coli* transposon Tn5 APH(3′)-II gene using RestrictionMapper program.

According to these results, *dehC*II-APH(3′)-II fusion construct was successfully cloned in pDrive Cloning Vector.

3.6 Determination of Dehalogenase Activity

Cell-free extracts from the recombinants with CS Large Domain-*dehC*II fusion construct and *dehC*II-APH(3′)-II fusion construct did not display any dehalogenase activity regardless MCA was used as an inducer during growth or not (Figure 3.26).

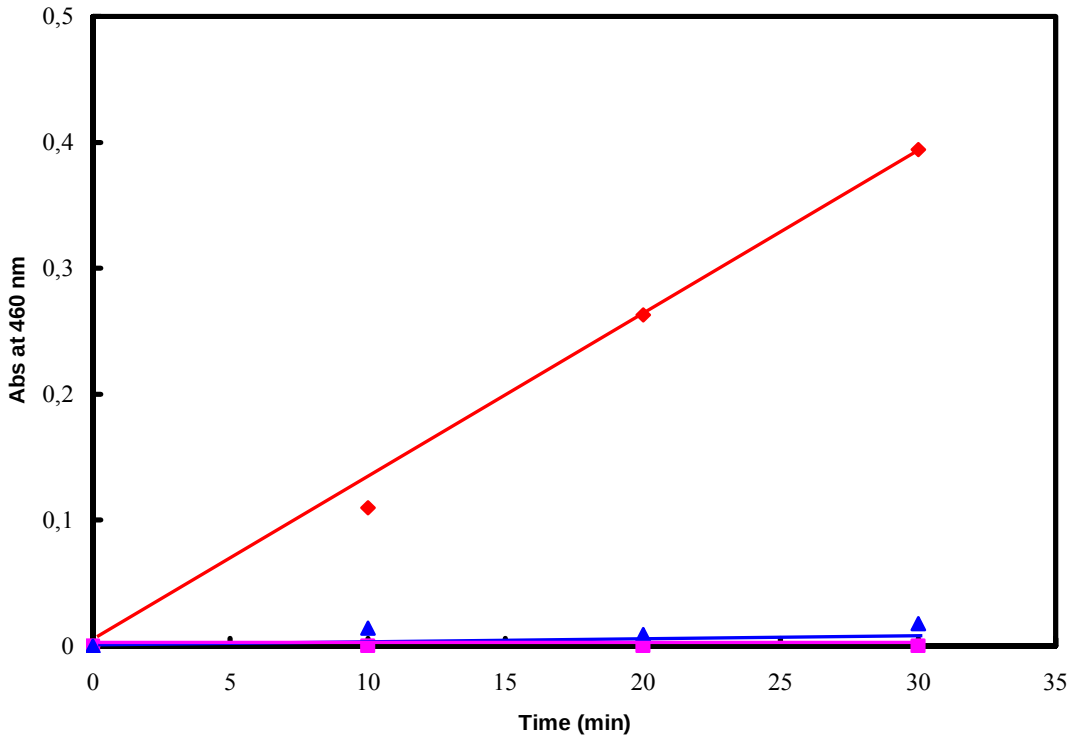


Figure 3.26 Dehalogenase activities of the *E. coli* recombinants containing CS Large Domain-*dehC*II fusion construct (—■—), *dehC*II-APH(3′)-II fusion construct (—▲—) and *E. coli* pUKS 107, which was used as a positive control (—◆—), according to the increase in Cl^- concentration. Cell-free extracts were prepared from the cultures grown in the presence of MCA. Activity was measured using MCA as the substrate in a time dependent manner at 30°C, as described in the Materials and Methods.

3.7 Determination of Neomycin Resistance by Agar-Plate Assay

Neomycin resistance of recombinant *E. coli* strains harboring CS Large Domain-APH(3′)-II and *dehC*II-APH(3′)-II fusions were determined by agar-plate assay as described in the Materials and Methods. *E. coli* pBS-aph and *E. coli* strain TG1 were used as positive and negative controls, respectively. The growth of colonies was observed as induction of antibiotic resistance after 16 hours incubation of the plates at 37°C. *E. coli* strains harboring CS Large Domain-APH(3′)-II and *dehC*II-APH(3′)-II fusions as well as *E. coli* pBS-aph grew successfully on neomycin (upto 200 µg/ml) containing plates while *E. coli* strain TG1 did not display any growth, as expected. This result showed that CS Large Domain-APH(3′)-II and *dehC*II-APH(3′)-II fusion proteins associated APH(3′)-II partner was successfully expressed in *E. coli* (Table 3.4).

Table 3.4 Neomycin resistance of recombinant *E. coli* strains containing CS Large Domain (LD)-APH(3′)-II and *dehC*II-APH(3′)-II fusions with positive control *E. coli* pBS-aph and negative control *E. coli* strain TG1.

	Genotypes of Recombinants		Positive control	Negative control
	CS LD-APH(3′)-II	*dehC*II-APH(3′)-II	APH(3′)-II	TG1
Only Ampicillin (50 µg/ml)	+	+	+	–
Only Neomycin (100 µg/ml)	+	+	+	–
Neo + Amp (20 µg/ml) (50 µg/ml)	+	+	+	–
Neo + Amp (50 µg/ml) (50 µg/ml)	+	+	+	–
Neo + Amp (100 µg/ml) (50 µg/ml)	+	+	+	–
Neo + Amp (200 µg/ml) (50 µg/ml)	+	+	+	–

3.8 APH(3′)-II Assay

The APH(3′)-II activities in the extracts of the recombinant *E. coli* strains containing the CS Large Domain-APH(3′)-II fusion construct and *dehC*II-APH(3′)-II fusion construct were measured in a pyruvate kinase/lactate dehydrogenase coupled assay system in a continuous fashion, as described in the Materials and Methods. In these experiments, cell-free extract of *E. coli* pBS-aph and *E. coli* strain TG1 were used as positive and negative controls, respectively. Cell-free extracts of the recombinants with CS Large Domain-APH(3′)-II fusion construct and *dehC*II-APH(3′)-II fusion construct were displayed an APH(3′)-II activity of 3.31 µmoles/ml/min and 1.32 µmoles/ml/min, respectively. APH(3′)-II activity of the extract from *E. coli* pBS-aph

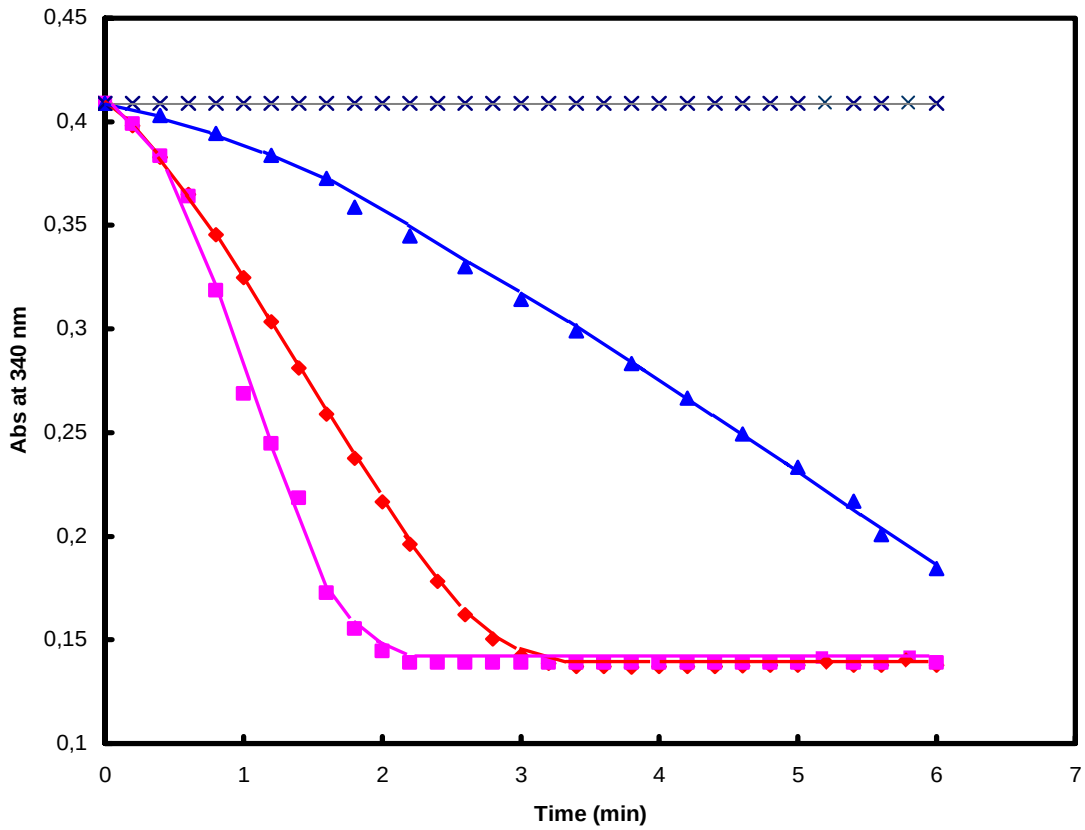


Figure 3.27 APH(3′)-II activity of the CS Large Domain-APH(3′)-II fusion construct (—◆—), *dehC*II-APH(3′)-II fusion construct (—▲—), *E. coli* pBS-aph, as positive control (—■—) and *E. coli* strain TG1, as negative control (—×—), using a pyruvate kinase/lactate dehydrogenase coupled assay system. The resulting decrease was monitored at 33°C, as described in the Materials and Methods.

(positive control) was found as 4.82 µmoles/ml/min. The activity curves are shown in the Figure 3.27.

3.9 Thermostability

To determine the thermostability of the APH(3′)-II activity associated with CS Large Domain-APH(3′)-II fusion construct, cell-free extract of the respective recombinant *E. coli* strain was incubated for 10 minutes at 60°C and the remaining activity was measured. APH(3′)-II activities measured before and after heat treatment were 3.31 µmoles/ml/min and 0.31 µmoles/ml/min, respectively. Thus, the APH(3′)-II activity in the extract of the recombinant strain with CS Large Domain-APH(3′)-II retained 9.4% of the original activity, after 10 minutes at 60°C. The same heat treatment totally abolished the APH(3′)-II activity in the extract of *E. coli* pBS-aph strain (Figure 3.28).

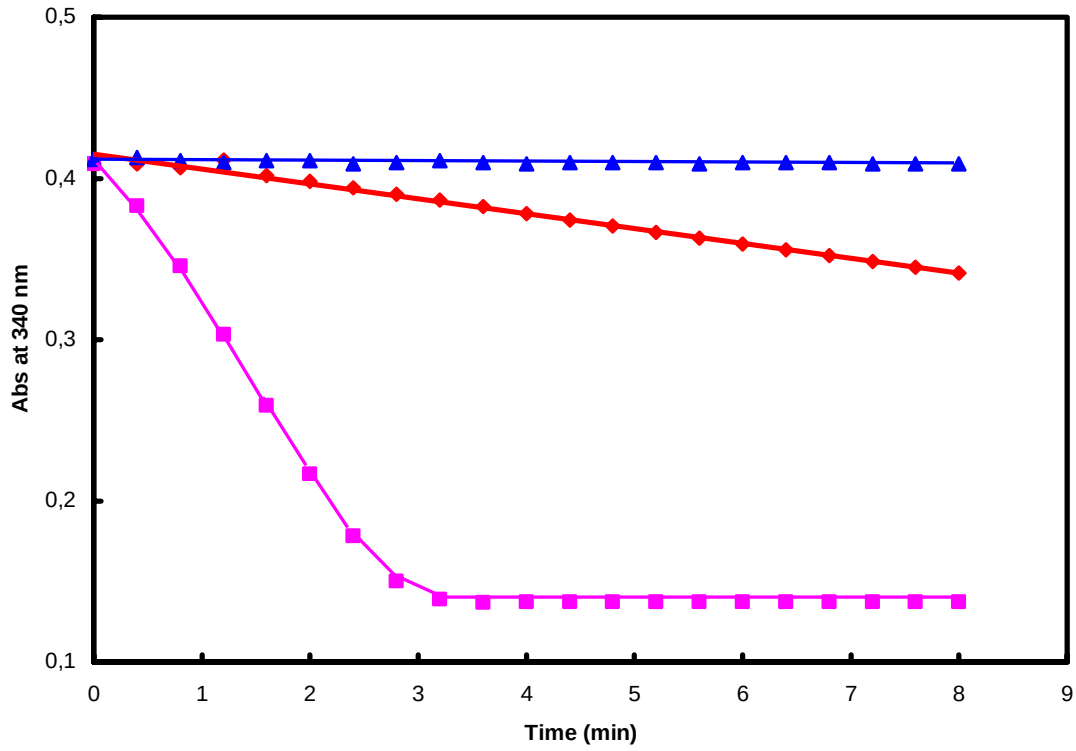


Figure 3.28 APH(3′)-II activities of the CS Large Domain-APH(3′)-II fusion construct before (—■—) and after (—◆—) the heat treatment together with *E. coli* pBS-aph strain (—▲—), as a positive control. Samples were incubated for 10 minutes at 60°C. Remaining activities were measured under standard assay condition.

CHAPTER IV

DISCUSSION

Comparative structural analysis of the citrate synthases isolated from psychorophiles, mesophiles, thermophiles and hyperthermophiles have highlighted many structural adaptations that confer on protein thermostability and pointed out the structural trends that correlate with their differing thermostabilities. The presence of ion-pair networks, which is common to almost all the hyperthermophilic proteins and improved hydrophobic packing, compactness, additional hydrogen bonds are the measures observed in these proteins. Investigations to search the contribution made by the subunit interaction in the thermostability and the thermoactivity between thermophilic *Thermoplasma acidophilum* citrate synthase and hyperthermophilic *Pyrococcus furiosus* citrate synthase in which chimeric mutants were created by swapping the large and small domains of the proteins' subunits showed that the origin of the large domains primarily dictated the thermal stability of the chimeric enzymes (Arnott *et al.*, 2000).

In this study, a novel fusion strategy was followed in order to increase the thermostability of two mesophilic enzymes encoded by *Pseudomonas* sp. strain CBS3 2-haloalkanoic acid dehalogenase II gene and *E. coli* transposon Tn5 aminogylcoside 3´-phosphotransferase II gene. Today, gene fusion technology not only circumvents many of the problems inherit in the use of *E. coli* for expression, folding, purification, secretion and detection of recombinant proteins but also provides an increasing number of applications in biochemistry, biotechnology and immunology including display of proteins on surfaces of cells and phages, development of subunit vaccine as well as increase the *in vivo* half-lives of therapeutic gene products (Nygren *et al.*, 1994; Stahl and Nygren, 1997). This study

presents a novel application of the gene fusion strategy, to enhance the thermostability of mesophilic enzymes, by tagging with a stability domain from a thermophilic enzyme.

First fusion strategy we have employed was a sequence-dependent standard method that relied on the ligation of the *Thermoplasma acidophilum* CS large domain to the *Pseudomonas* sp. Strain CBS3 *dehC*II gene through their complementary sticky ends (of *Bam*HI). This method did not yield satisfactory results due to the drawbacks such as absence of blue/white screening for the recombinant colonies and highly favored self-ligation of the vector itself. Moreover, the strategy was time-consuming and labour intensive, therefore it was replaced by an alternative strategy called Gene splicing by overlap extension, "Gene SOEing".

Gene SOEing is a PCR-based approach described by Horton and co-workers (1989) and does not depend on the occurrence of restriction enzyme recognition sequences at the recombination site. Instead, genes from different sources that are to be generated are combined in a subsequent fusion reaction in which overlapping sequences anneal, allowing the 3′ overlap of each strand to serve as a primer for the 3′ extension of the complementary strand. The process is achieved by adding a complementary sequence to the 5′-end one the primers that become incorporated into the end of the product molecule and overlap to the other gene fragment.

Thermoplasma volcanium citrate synthase gene has recently been cloned, expressed, sequenced and purified in our laboratory. Results of the amino acid sequence alignment of the *Tp. acidophilum* citrate synthase with *Tp. volcanium* citrate synthase revealed that there is a high conservation between the citrate synthase sequences of the two *Thermoplasma* species with a sequence homology of 87%. Referring this alignment *Tp. volcanium*'s large domain was determined and used as the fusion partner during the construction experiments by Gene SOEing.

Designing of the primers were done consulting to the amino acid sequences of CS,

*dehC*II and APH(3′)-II. To construct the CS Large Domain-*dehC*II fusion the CS-Rv-1 primer was designed to add the homologous sequence to the 3′-end of the resulting fragment. To construct the CS Large Domain-APH(3′)-II fusion again the CS-Rv-2 primer was designed to add the homologous sequence to the 3′-end of the resulting fragment. To construct the *dehC*II-APH(3′)-II fusion the *dehC*II-Rv-2 primer was designed to add the homologous sequence to the 3′-end of the resulting fragment. So these special Rv primers include two regions: The priming region, found at the 3′-end of the oligo, contained the sequences which allow it act as a primer on its template and the overlap region, found at the 5′-end of the oligo, contained sequences which overlap with the 3′-end of the fragment to be joined. Since these primers are too long (37 bp for CS Large Domain-*dehC*II fusion construct, 39 bp for CS Large Domain-APH(3′)-II fusion construct and 40 bp for *dehC*II-APH(3′)-II fusion construct) (See Appendix D), it was possible that some excess Rv primers left after purification that was done before the SOE reactions. Therefore, with addition of the flanking primers, templates of the gene fragments that are to be joined were produced in the SOE reactions. These minor products appeared in agarose gel photos of SOE reactions and it was clear that SOE reactions were not hindered by the presence of them since bands of fusion constructs were the major products. SOE reactions resulted in the amplification of 1722 bp long CS Large Domain-*dehC*II fusion protein, 1750 bp long CS Large Domain-APH(3′)-II fusion protein and 1586 bp long *dehC*II-APH(3′)-II fusion protein.

Pfu polymerase was preferred in the PCR reactions because of its 3′→5′ exonuclease activity and relatively blunt ended fragments were produced after PCR. *Taq* and other non-proofreading DNA polymerases add single A overhangs at each end of PCR products which is an unwanted situation for the next step of the strategy. SOE reactions were carried out using *Taq* polymerase since an A/U complementation-based strategy was used for cloning. PCR amplified fusion constructs, 1722 bp long CS Large Domain-*dehC*II fusion, 1750 bp long CS Large Domain-APH(3′)-II fusion and 1586 bp long *dehC*II-APH(3′)-II fusion, were directly cloned into the pDrive

Cloning vector by A/U ligation. The restriction analysis of the fusion constructs when judged by referring to the restriction maps of the participant genes, revealed that we have constructed and cloned 3 gene fusions successfully (Figure 4.1, 4.2 and 4.3).

Experiments revealed that Gene SOEing strategy allowed greater flexibility than direct ligation strategy by decreasing the dependence on restriction sites and provided much more straightforward results in generation of the fusion proteins. It was proved to be more convenient and successful as compared to the restriction enzyme-based direct ligation approach.

Recombinant *E. coli* strains harboring CS Large Domain-APH(3′)-II fusion construct and *dehC*II-APH(3′)-II fusion construct displayed considerable APH(3′)-II activity, indicating successful expression of the APH(3′)-II gene in these constructs. This result was also confirmed by agar-plate assay. However, we could not detect any dehalogenase activity for CS Large Domain-*dehC*II fusion construct and *dehC*II-APH(3′)-II fusion construct. Transcription of the dehalogenase gene of the CS Large Domain-*dehC*II fusion was under the control of the promoter region of the citrate synthase. On the other hand, the *dehC*II-APH(3′)-II fusion construct was under the control of its own dehalogenase promoter. The absence of dehalogenase activity associated with these two constructs could be due to requirement for some additional genetic elements like a permease or transport protein or a regulatory protein for expression (Greer *et al.*, 1989; Janssen *et al.*, 1989). Another explanation could be inappropriate folding of the *Pseudomonas* associated *dehC*II in the form of fusion protein in *E. coli*.

Also, expression of the *dehC*II gene could be studied by placing the fusion constructs downstream of strong *E. coli* promoters such as *lac* or *tac*. The *Pseudomonas cepacia* MBA42-haloacid halidohydrolase IVa (*hdl* IVa) gene has a 56% amino acid sequence similarity with *dehC*II and its expression was increased 10-fold when inserted in the correct orientation with respect to the *tac* promoter of the expression vector pBTac1 (Murdiyatmo *et al.*, 1992).

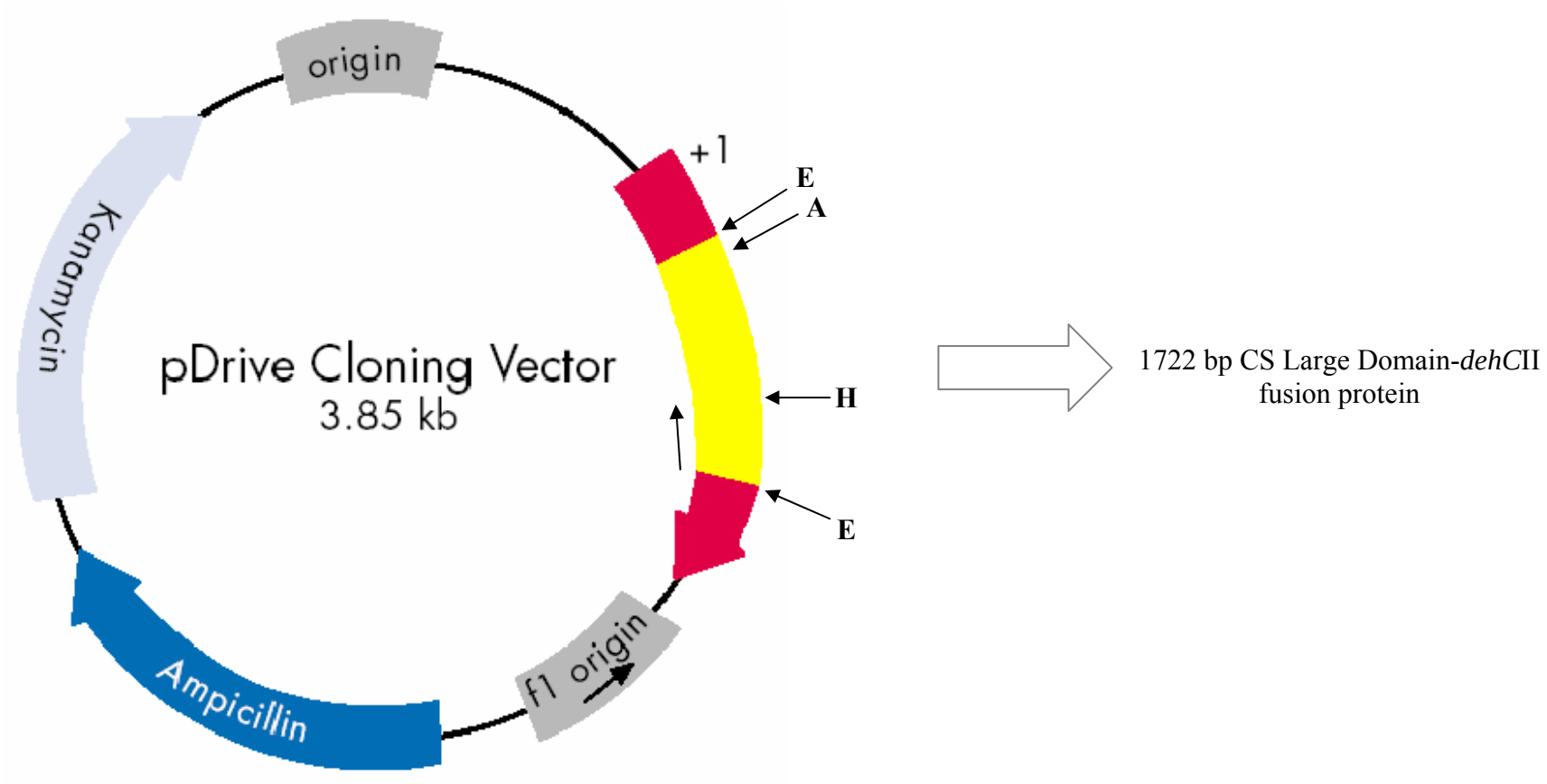


Figure 4.1 Recombinant pDrive Vector containing cloned CS Large Domain-*dehC*II fusion gene. **E:** *Eco*RI, **H:** *Hin*dIII, **A:** *Ava*II.

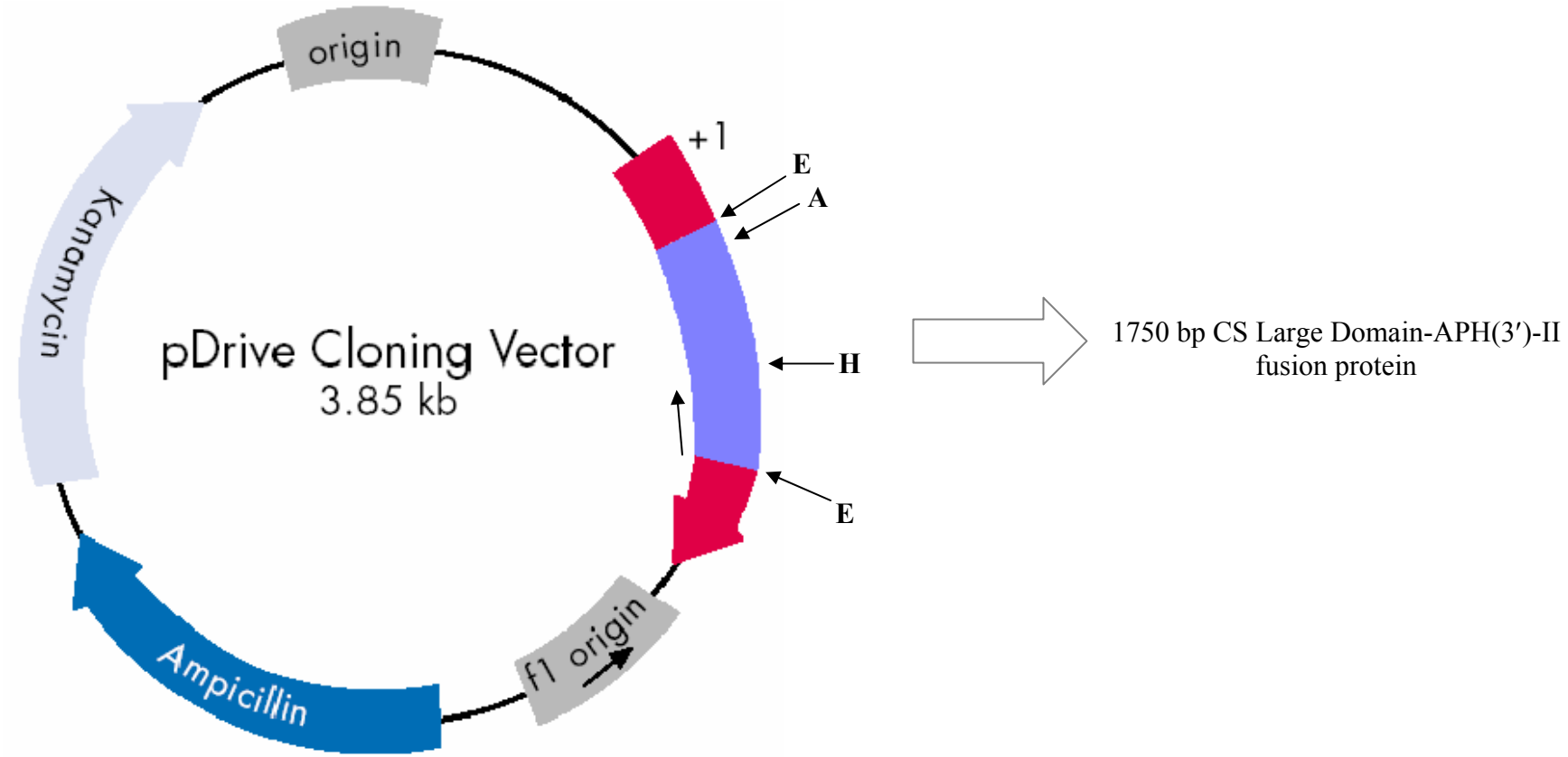


Figure 4.2 Recombinant pDrive Vector containing cloned CS Large Domain-APH(3′)-II fusion gene. **E:** *Eco*RI, **H:** *Hin*dIII, **A:** *Ava*II.

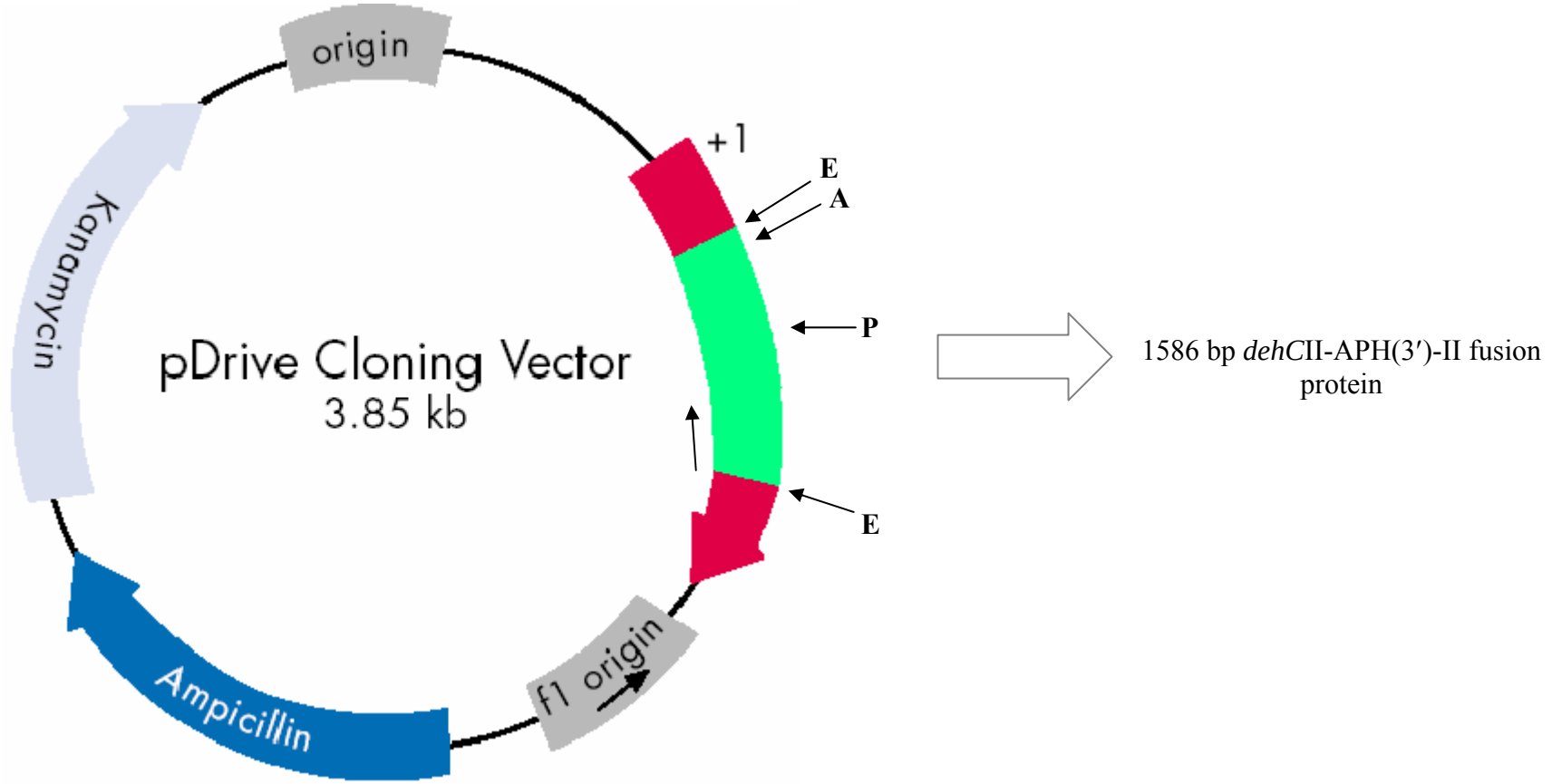


Figure 4.3 Recombinant pDrive Vector containing cloned *dehC*II-APH(3′)-II fusion gene. **E:** *Eco*RI, **A:** *Ava*II**, P:** *Pst*I**.**

Heat treatment results showed that 9.4% of the APH(3′)-II activity for CS Large Domain-APH(3′)-II fusion construct was retained at 60ºC after 10 minutes. This result is important for enhancement of heat stability of the mesophilic enzymes by incorporation of stability domains from thermostable enzymes, through gene fusion strategy. This result suggested that only the incorporation of CS Large Domain would be enough for thermostability. This data was also in correlation with the Arnott and co-workers (2000) study. They demonstrated that the large domains of the domain swapped chimeric mutants (CS(PF-ta) and CS(TA-pf)) comprised the inter-subunit contacts within the dimers and was the main determinant of the enzymes' thermostability properties where the small, flexible domain of the *Thermoplasma* enzyme did not confer any significant loss of thermostability when incorporated into the *Pyrococcus* enzyme and, conversely, the small domain of the *Pf*CS did not increase the thermostability of the less stable *Ta*CS.

CHAPTER V

CONCLUSION

1. In this study we have followed two strategies for construction of the fusion proteins. A) Direct ligation of *Pseudomonas dehC*II gene to *Tp. acidophilum* CS Large Domain B) Gene SOEing. Since it was time and labour intensive, first strategy was replaced by more convenient PCR-based Gene SOEing strategy.

2. Three fusion proteins were constructed using Gene SOEing strategy. PCR amplified 922 bp CS Large Domain fragment and 800 bp *dehC*II fragment were fused at an annealing temperature of 67ºC to yield the PCR amplified CS Large Domain-*dehC*II fusion protein which was 1722 bp long. PCR amplified 922 bp CS Large Domain fragment and 828 bp APH(3′)-II fragment were fused at an annealing temperature of 65ºC to yield the PCR amplified CS Large Domain-APH(3′)-II fusion protein which was 1750 bp long. PCR amplified 758 bp *dehC*II fragment and 828 bp *dehC*II fragment were fused at an annealing temperature of 60ºC to yield the PCR amplified *dehC*II-APH(3′)-II fusion protein which was 1586 bp long.

3. PCR amplified three fusion genes were cloned in *E. coli*. Cloning was confirmed by restriction analysis of isolated plasmids from recombinant colonies.

4. CS Large Domain-APH(3′)-II fusion protein and *dehC*II-APH(3′)-II fusion protein associated APH(3′)-II partner was successfully expressed in *E. coli* and this enzyme was still active at 60ºC after 10 minutes. Thus, this is the first report for the enhancement of thermostable enzyme by gene fusion approach. CS Large Domain-*dehC*II and *dehC*II-APH(3′)-II fusions displayed no dehalogenase activity which require further investigation.

REFERENCES

Arnott MA, Micheal RA, Thompson CR, Hough DW, Danson MJ: Thermostability and thermoactivity of citrate synthases from the thermophilic archaea, *Thermoplasma acidophilum* and *Pyrococcus furiosus*. *Journal of Molecular Biology* 2000, 304:657-668.

Baneyx F: Recombinant protein expression in *Escherichia coli. Current Opinion in Biotechnology* 1999, 10:411-421.

Baquero F and Blázquez J: Evolution of antibiotic resistance. *Trends in Ecology and Evolution* 1997, 12:482-487.

Beck E, Ludwig G, Auerswald EA, Reiss B, Schaller H: Nucleotide sequence and exact localization of the neomycin phosphotransferase gene from transposon Tn5. *Gene* 1982, 19:327-336.

Bell GS, Russell RJM, Connaris H, Hough DW, Danson MJ, Taylor GL: Stepwise adaptations of citrate synthase to survival at life's extremes. *European Journal of Biochemistry* 2002, 269:6250-6260.

Carter AP, Clemons WM, Broderson DE, Morgan-Warren RJ, Wimberly BT, Ramakrishnan V: Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* 2000, 407:340-348.

Chalfie M, Tu Y, Eurkirshen G, Ward WW, Prasher DC: Green fluorescent protein as a marker gene expression. *Science* 1994, 263:802-805.

Charbit A, Molla A, Saurin W, Hofnung M: Versatility of a vector for expressing foreign polypeptides at the surface of Gram-negative bacteria. *Gene* 1988, 70:181-189.

Chong S, Shao Y, Paulus H, Benner J, Perler FB, Xu M-Q: Protein splicing involving the *Saccharomyces cerevisiae* VMA intein. *Journal of Biological Chemistry* 1996, 271:22159-22168.

Cronan JE: Biotination of proteins *in vivo. Journal of Biological Chemistry* 1990, 265:10327-10333.

Cull MG, Miller JF, Schatz PJ: Screening for receptor ligands using large libraries of peptides linked to the C terminus of the *lac* repressor. *PNAS* 1992, 89:1865-1869.

Danson MJ and Hough DW: Structure, function and stability of enzymes from the Archaea. *Trends in Microbiology* 1998, 6:307-313.

Davies J, Wright GD. Bacterial resistance to aminoglycoside antibiotics. *Trends in Microbiology* 1997, 5:234-240.

Duplay P, Bedouelle H, Fowler A, Zabin I, Saurin W, Hofnung M: Sequences of the *malE* gene and of its product, the maltose-binding protein of *Escherichia coli* K12. *Journal of Biological Chemistry* 1984, 259:10606-10613.

Erduran İ, Kocabıyık S: Amino acid substitutions in the subunit interface enhancing thermostability of *Thermoplasma acidophilum* citrate synthase. *Biochemical and Biophysical Research Communications* 1998, 249:566-571.

Fetzner S: Bacterial dehalogenation. *Applied Microbiology and Biotechnology* 1998, 50:633-657.

Francisco JA, Earhart CF, Georgiou G: Transport and anchoring of β-lactamase to the external surface of *Escherichia coli. PNAS* 1992, 89:2713-2717.

Gerike U, Hough DW, Russell NJ, Dyall-Smith ML, Danson MJ: Citrate synthase and 2-methylcitrate synthase: structural, functional and evolutionary relationships. *Microbiology* 1998, 144:929-935.

Greenwood J, Willis AE, Perham RN: Multiple display of foreign peptides on a filamentous bacteriophage. Peptides from *Plasmodium falciparum* circumsporozoite protein as antigens. *Journal of Molecular Biology* 1991, 220:821-827.

Greer CW, Beaumier D, Samson R: Application of on-line sensors during growth of the dichloroethane degrading bacterium *Xanthobacter autotrophicus*. *Journal of Biotechnology* 1989, 12:261-274.

Guan C-D, Li P, Riggs PD, Inouye H: Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. *Gene* 1988, 67:21-30.

Hardman D: Biotransformation of halogenated compounds. *Critical Reviews in Biotechnology* 1991, 11:1-40.

Hill KE, Marchesi JR, Weightman AJ: Investigations of two evolutionary unrelated halocarboxylic acid dehalogenase gene families. *Journal of Bacteriology* 1999, 181:2535-2547.

Hochuli E, Bannwarth W, Döbeli H, Gentz R, Stüber D: Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. *Bio/Technology* 1988, 6:1321-1325.

Hopp TP, Prickett KS, Price VL, Libby RT, March CJ, Cerretti DP, Urdal DL, Conlon PJ: A short polypeptide marker sequence useful for recombinant protein identification and purification. *Bio/Technology* 1988, 6:1204-1210.

Horton RM, Hunt HD, Ho SN, Pullen JK, Pease LR: Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* 1989, 77:61-68.

Hough DW, Danson MJ: Extremozymes. *Current Opinion in Biotechnology* 1999, 3:39-46.

Iwasaki I, Utsumi S, Hagino K, Ozawa T: A new spectrophotometric method for the determination of small amounts of chloride using the mercuric thiocyanate method. *Bulletin of the Chemical Society of Japan.* 1956, 29:860-864.

Janssen DB, Pries F, van der Ploeg J, Kazemier B, Terpstra P, Witholt B: Cloning of 1,2-Dichloroethane degrading genes of GJ10 and expression and sequencing of the *dhlA* gene. *Journal of Bacteriology* 1989, 171:6791-6799.

Janssen DB, Oppentocht JE, Poelarends GJ: Microbial dehalogenation. *Current Opinion in Biotechnology* 2001, 12:254-258.

Karp M, Oker-Blom C: A streptavidin-luciferase fusion protein: comparisons and applications. *Molecular Engineering* 1999, 16:101-104.

Kasher MS, Wakulchik M, Cook JA, Smith MC: One-step purification of recombinant human Papillomavirus type 16 E7 oncoprotein and its binding to the retinoblastoma gene product. *BioTechniques* 1993, 14:630-641.

Kawasaki H, Tsuda K, Matsushita I, Tonomura K: Lack of homology between two haloacetate genes encoded on a plasmid from *Moraxella* sp. strain B. *Journal of General Microbiology* 1992, 138:1317-1323.

Kocabıyık S, Erduran İ: The effect of valine substitution for glycine in the dimer interface of citrate synthase from *Thermoplasma acidophilum* on stability and activity. *Biochemical and Biophysical Research Communications* 2000, 275:460-465.

LaVallie ER, DiBlasio EA, Kovacic S, Grant KL, Schendel PF, McCoy JM: A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. *Bio/Technology* 1993, 11:187-193.

LaVallie ER, McCoy JM: Gene fusion expression systems in *Escherichia coli*. *Current Opinion in Biotechnology* 1995, 6:501-506.

Liu JQ, Kurihara T, Hasan AKMQ, Nardi-Dei V, Koshikawa H, Esaki N, Soda K: Purification and characterization of thermostable and nonthermostable 2-haloacid dehalogenases with different stereospecificities from *Pseudomonas* sp. Strain YL *Applied and Environmental Microbiology* 1994, 60:2389-2393.

Ljungquist C, Breitholtz A, Brink-Nilsson H, Moks T, Uhlén M, Nilsson B: Immobilization and affinity purification of recombinant proteins using histidine peptide fusions. *European Journal of Biochemistry* 1989, 186:563-569.

Lu Z, DiBlasio-Smith EA, Grant KL, Warne NW, LaVallie ER, Collins-Racie LA, Follettie MT, Williamson MJ, McCoy JM: Histidine patch thioredoxins. *Journal of Biological Chemistry* 1996, 271:5059-5065.

Motosugi K, Esaki N, Soda K: Purification and properties of a new enzyme, DL-2-haloacid dehalogenase, from *Pseudomonas* sp. *Journal of Bacteriology* 1982, 150:522-527.

Motosugi K, Esaki N, Soda K: Enzymatic preparation of D- and L-lactic acid from racemic 2-chloropropionic acid. *Biotechnology and Bioengineering* 1984, 26:805-806.

Murdiyatmo U, Asmara W, Tsang SH, Baines AJ, Bull AT, Hardman DJ: Molecular biology of the 2-halo-acid halidohydrolase IVa from *Pseudomonas cepacia* MBA4. *Biological Journal* 1992, 284:87-93.

Müller R, Lingens F: Microbial degradation of halogenated hydrocarbons: A biological solution to pollution problems? *Angewandte Chemie. International Edition in English* 1986, 25:779-789.

Nardi-Dei V, Kurihara T, Park C, Miyagi M, Tsunasawa S, Soda K, Esaki N: DL-2-

Haloacid dehalogenase from Pseudomonas sp. 113 is a new class of dehalogenase catalyzing hydrolytic dehalogenation not involving enzyme-substrate ester intermediate. *Journal of Biological Chemistry* 1999, 274:20977-20981.

Nilsson B, Abrahmsén L: Fusions to *Staphylococcus* protein A. *Methods in Enzymology* 1990, 185:144-161.

Nilsson J, Ståhl S, Lundeberg J, Uhlén M, Nygren P-Å: Affinity fusion strategies for detection, purification, and immobilization of recombinant proteins. *Protein Expression and Purification* 1997, 11:1-16.

Nurizzo D, Shewry SC, Perlin MH, Brown SA, Dholakia JN, Fucks RL, Deva T, Baker EN, Smith CA: The crystal structure of aminoglycoside-3′-phosho-transferase-IIa, an enzyme responsible for antibiotic resistance. *Journal of Molecular Biology* 2003, 327:491-506.

Nygren P-Å, Ståhl S, Uhlén M: Engineering proteins to facilitate bioprocessing. *Trends in Biotechnology* 1994, 12:184-188.

Panagiotidis CA, Silverstein SJ: pALEX, a dual-tag prokaryotic expression vector for the purification pf full-length proteins. *Gene* 1995, 164:45-47.

Pang BCM and Tsang JSM: Mutagenic analysis of the conserved residues in dehalogenase IVa of *Burholderia cepacia* MBA4. *FEMS Microbiology Letters* 2001, 204:135-140.

Porath J, Carlsson J, Olsson I, Belfrage G: Metal chelate affinity chromatography, anew approach to protein fractionation. *Nature* 1975, 258:598-599.

Pryor KD, Leiting B: High-level expression of soluble protein in *Escherichia coli* using a his_6-tag and maltose-binding-protein double-affinity fusion system. *Protein Expression and Purification* 1997, 10:309-319.

Remington SJ, Wiegand G, Huber R: Crystallographic refinement and atomic models of two different forms of citrate synthase at 2.7 and 1.7 Å resolution. *Journal of Molecular Biology* 1982, 158:111-152.

Remington SJ: Structure and mechanism of citrate synthase. *Current Topics in Cellular Regulation* 1992, 33:209-229.

Russell RJM, Danson MJ, Hough DW, Taylor GL: The crystal structure of citrate synthase from the thermophilic archaeon, *Thermoplasma acidophilum*. *Structure* 1994, 2:1157-1167.

Russell RJM, Taylor GL: Engineering thermostability: lessons from thermophilic proteins. *Current Opinion in Biotechnology* 1995, 6:370-374.

Russell RJM, Ferguson JMC, Hough DW, Danson MJ, Taylor GL: The crystal structure of citrate synthase from the hyperthermophilic archaeon *Pyrococcus furiosus* at 1.9 Å resolution. *Biochemistry* 1997, 36:9983-9994.

Russell RJM, Gerike U, Hough DW, Danson MJ, Taylor GL: Structural adaptations of the cold-active citrate synthase from an Antarctic bacterium. *Structure* 1998, 6:351-361.

Schatz PJ: Use of peptide libraries to map the substrate specificity of a peptide modifying enzyme: A 13 residue consensus peptide specifies biotinylation in *Escherichia coli. Bio/Technology* 1993, 11:1138-1143.

Schneider B, Müller R, Frank R, Lingens F: Complete nucleotide sequences and comparisons of the structural genes of two 2-haloalkanoic acid dehalogenases from Pseudomonas sp. strain CBS3. *Journal of Bacteriology* 1991, 173:1530-1535.

Scott JK, Smith GP: Searching for peptide ligands with an epitope library. *Science* 1990, 249:386-390.

Sharrocks AD: A T7 expression vector for producing N- and C-terminal fusion proteins with glutathione *S*-transferase. *Gene* 1994, 138:105-108.

Smith JM, Harrison K, Colby J: Purification and characterization of D-2-haloacid dehalogenase from *Pseudomonas putida* strain AJ1/23. *Journal of General Microbiology* 1990, 136:881-886.

Smith DB, Johnson KS: Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione *S*-transferase. *Gene* 1988, 67:31-40.

Smith LD, Stevenson KJ, Hough DW, Danson MJ: Citrate synthase from the thermophilic archaebacteria *Thermoplasma acidophilum* and *Sulfolobus acidocaldarius*. *FEBS* 1987, 225:277-281.

Smith MC, Furman TC, Ingolia TD, Pidgeon C: Chelating peptide-immobilized metal ion affinity chromatography. *Journal of Biological Chemistry* 1988, 263:7211-7215.

Ståhl S, Nygren P-Å: The use of gene fusions to protein A and protein G in immunobiology and biotechnology. *Pathologie Biologie* 1997, 45:66-67.

Stathopoulos C, Georgiou G, Earhart CF: Characterization of *Escherichia coli* expressing an Lpp′OmpA(46-159)-PhoA fusion protein localized in the outer membrane. *Applied Microbiology and Biotechnology* 1996, 45:112-119.

Stevens RC: Design of high-throughput methods of protein production for structural biology. *Structure* 2000, 8:R177-R185.

Sutherland KJ, Henneke CM, Towner P, Hough DW, Danson MJ: Citrate synthase from the thermophilic archaebacterium *Thermoplasma acidophilum*. Cloning and sequencing of the gene. *European Journal of Biochemistry* 1990, 194:839-844.

Swanson P: Dehalogenases applied to industrial-scale biocatalysis. *Current Opinion in Biotechnology* 1999, 10:365-369.

Uhlén M, Moks T: Gene fusions for purpose of expression: An introduction. *Methods in Enzymology* 1990, 185:129-143.

Van Der Ploeg J, Hall G van, Janssen DB: Characterization of the haloacid dehalogenase from *Xanthobacter autrophicus* GJ10 and sequencing of the *dhlB* gene. *Journal of Bacteriology* 1991, 173:7925-7933.

Van Dyke MW, Sirito M, Sawadogo M: Single-step purification of bacterially expressed polypeptides containing an oligo-histidine domain. *Gene* 1992, 111:99-104.

Walsh C: Molecular mechanism that confer antibacterial drug resistance. *Nature* 2000, 406:775-781.

Weightman AJ, Weightman AL, Slater JH: Stereospecificity of 2-monochloropropionic dehalogenation by the two dehalogenases of *Pseudomonas putida* PP3: evidence for two different dehalogenation mechanism. *Journal of General Microbiology* 1982, 128:1755-1762.

Winter J, Neubauer P, Glockshuber R, Rudolph R: Increased production of human proinsulin in the periplasmic space of *Escherichia coli* by fusion to DsbA. *Journal of Biotechnology* 2000, 84:175-185.

APPENDIX A

BUFFERS AND SOLUTIONS

LB Medium

10 g tryptone

5 g yeast extract

10 g NaCl

TAE Buffer

0.04 M Tris-acetate

0.001 M EDTA, pH 8.0

Gel Loading Buffer

0.25% Bromophenol blue

40% Sucrose

TE Buffer

10 mM Tris HCl

1 mM EDTA, pH 7.0

Buffer B

3 M $NaCH_3COO$, pH 5.5

TMND Buffer

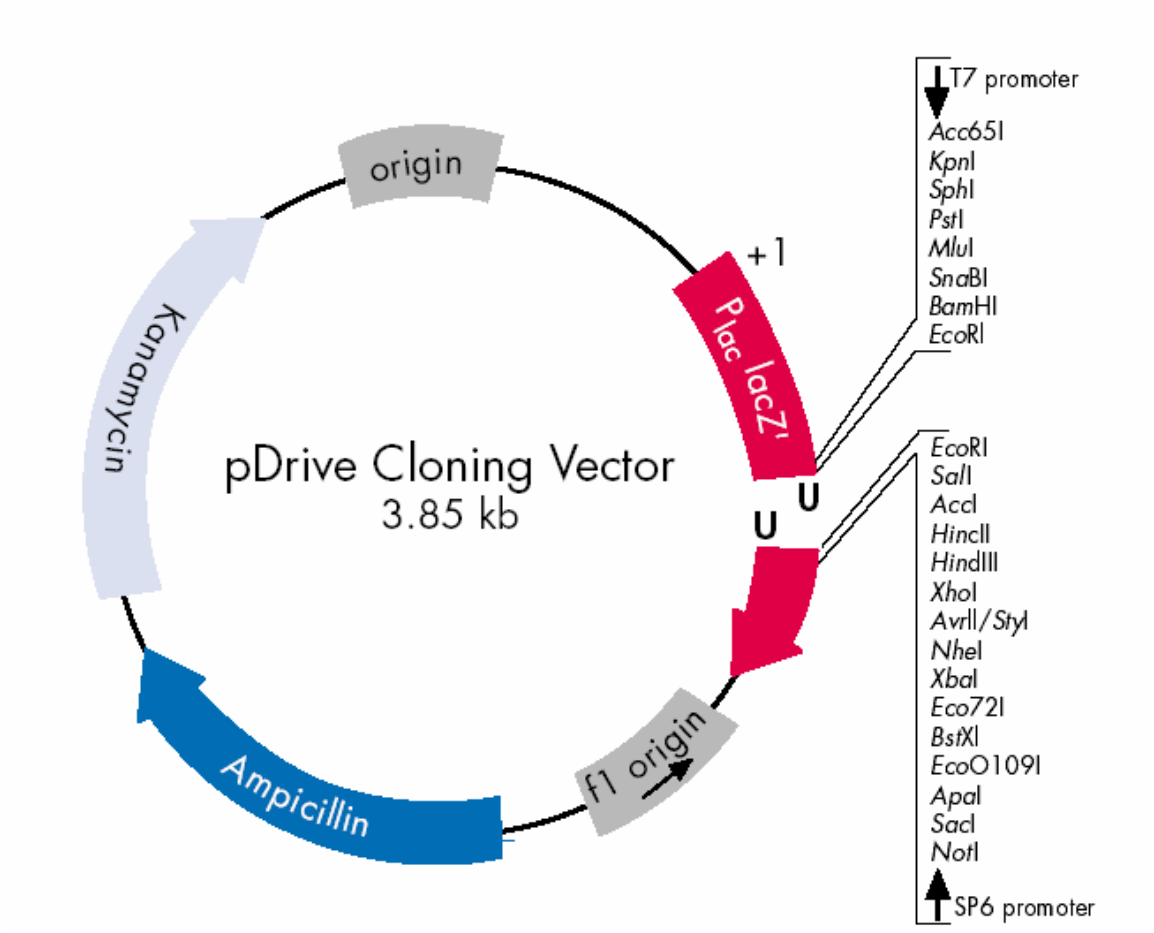
20 mM Tris.HCl pH 7.4

1.1 mM DTT

8.6 mM $MgCl_2.6H_2O$

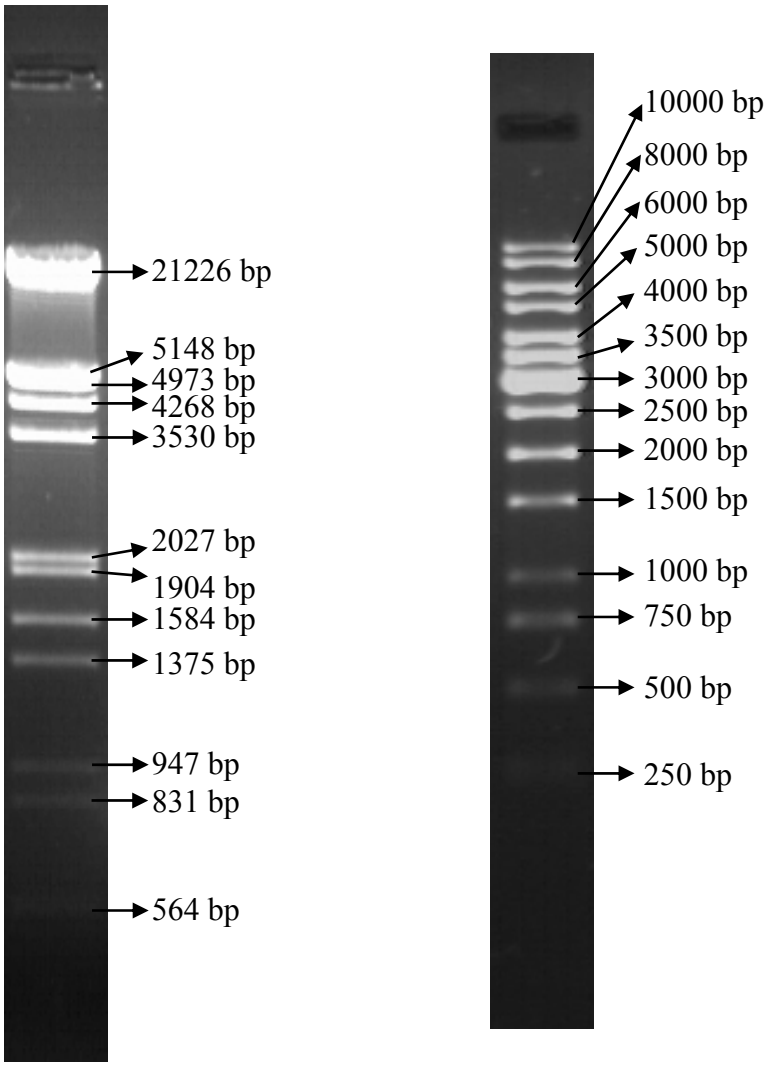
APPENDIX B

pDRIVE CLONING VECTOR



APPENDIX C

MOLECULAR SIZE MARKERS



(A) (B)

(A) Lambda DNA/*Eco*RI+*Hin*dIII, molecular size marker (MBI Fermentas AB, Vilnius, Lithuania)

(B) Gene Ruler 1 kb DNA Ladder (MBI Fermentas AB, Vilnius, Lithuania)

APPENDIX D

Nucleotide sequence of *Tp. volcanium* citrate synthase (CS) (above) and *Pseudomonas* sp. Strain CBS3 2-haloalkanoic acid dehalogenase II (*dehC*II) (bottom) with marked primer sequences. CS-Fr-1 is indicated in red color while CS-Rv-1 is displayed in frame and the overlapping sequence is marked with blue color. *dehC*II-Fr-1 and *dehC*II-Rv-1 gene are pointed out in pink color.

Nucleotide sequence of *Tp. volcanium* CS (above) and *E. coli* transposon Tn5 APH(3´)-II gene (bottom) with marked primer sequences. CS-Fr-2 primer is indicated in red color while CS-Rv-2 primer is displayed in frame and the overlapping sequence is showed in blue color. APH-Fr-1 and APH-Rv-1 primers are pointed out in pink color.

Nucleotide sequence of *Pseudomonas* sp. Strain CBS3 *dehC*II (above) and *E. coli* transposon Tn5 APH(3′)-II gene (bottom) with marked primer sequences. *dehC*II-Fr-2 primer is indicated in red color while *dehC*II-Rv-2 primer is displayed in frame and the overlapping sequence is showed in blue color. APH-Fr-2 and APH-Rv-2 primers are pointed out in pink color.