CLONING and EXPRESSION OF BENZALDEHYDE LYASE GENE FROM *Pseudomonas fluorescens* BIOVAR I in *Pichia pastoris*

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

BY

ARDA BÜYÜKSUNGUR

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
BIOTECHNOLOGY

AUGUST 2006

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. Fatih Yıldız
Head of Department

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

Prof. Dr. H. Tunçer Özdamar Co-Supervisor	Prof. Dr. Pınar Çalık Supervisor

Examining Committee Members

Prof.Dr. Güzide Çalık	(Ankara University, CHE)	
Prof. Dr. Pınar Çalık	(METU, CHE)	
Prof.Dr. Ufuk Bakır	(METU, CHE)	
Prof.Dr. Zümrüt Begüm Ögel	(METU, FDE)	
Assist. Prof.Dr. A. Elif Erson	(METU, BIO)	

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

Name, Last name : Arda Büyüksungur

Signature :

ABSTRACT

CLONING and EXPRESSION OF BENZALDEHYDE LYASE FROM *Pseudomonas fluorescens* BIOVAR I in *Pichia pastoris*

Büyüksungur, Arda
M.S., Department of Biotechnology
Supervisor: Prof. Dr. Pınar Çalık
Co-Supervisor: Prof. Dr. H. Tunçer Özdamar

August 2006, 144 pages

Benzaldehyde lyase (BAL, EC 4.1.2.38) from *Pseudomonas fluorescens* Biovar I, a thiamine pyrophosphate (ThDP) dependent enzyme, catalyzes the enzymatic kinetic resolution of racemates by C-C bond cleavage and concomitant C-C bond formation. In this study, benzaldehyde lyase gene from *Pseudomonas fluorescens* Biovar I was cloned into *Pichia pastoris*, with the aim of the extracellular production of the enzyme. For this purpose, firstly, PCR amplified *bal* gene was cloned into an integration vector pPICZalphaA. Thereafter the recombinant plasmid pPICZalphaA::*bal* was transformed into *P.pastoris*. Extracellular benzaldehyde lyase enzyme was expressed under the control of the strong AOX promoter and the secretion of the enzyme in the fermentation medium was achieved by means of *S. cerevisiae* alpha factor signal sequence. The recombinant cells were grown for 48-72 hours in solid medium then the cells inoculated in glycerol containing medium. After being separated by centrifugation cells were transferred into methanol containing production

medium. In methanol containing medium cells were grown for 72 h. Starting from t=24 h methanol was added to medium as an inducer of AOX promoter and the carbon source in order to produce BAL in every 24 hour. SDS-PAGE analyses illustrated that extracellular benzaldehyde lyase enzyme produced by the recombinant *P.pastoris* strain had the size of 59 kDa, which is the size of benzaldehyde lyase monomer. FPLC analysis showed that concentration of the tetrameric form of benzaldehyde lyase enzyme, active form, was much less than the monomeric form of the enzyme indicating that the enzyme produced by recombinant *P.pastoris* mostly could not fold into multimeric form in the fermentation medium.

Keywords: Benzaldehyde lyase, Extracellular production, Recombinant *P.pastoris*

ÖZ

Pseudomonas fluorescens BİOVAR I BENZALDEHİT LİYAZ GENİNİN *Pichia pastoris* MİKROORGANİZMASINA KLONLANMASI VE EKSPRESİYONU

Büyüksungur, Arda
Yüksek Lisans, Biyoteknoloji Bölümü
Tez Yöneticisi: Prof. Dr. Pınar Çalık
Ortak Tez Yöneticisi: Prof. Dr. H. Tunçer Özdamar

Ağustos 2006, 144 sayfa

Pseudomonas fluorescens Biovar I mikroorganizmasının benzaldehit liyaz (BAL, EC 4.1.2.38) tiamin difosfat bağımlı enzimi C-C bağlarının yapımı ve yıkımında görevlidir. Bu çalışmada, *Pseudomonas fluorescens* Biovar I benzaldehit liyaz geninin, benzaldehit liyaz enziminin hücredışı üretimi amacı ile *Pichia pastoris* mayasına klonlanması amaçlanmıştır. Bu kapsamda, polimeraz zincir tepkimesiyle çoğaltılan *bal* geni pPICZalphaA entegrasyon vektörüne aktarılmış, ve oluşturulan rekombinant pPICZalphaA::*bal* plazmidi *P.pastoris* hücresine aktarılarak gen *P.pastoris* kromozomuna integre edilmiştir. Hücredışı benzaldehit liyaz enzimi *P.pastoris* hücresinde AOX promotorunun kontrolu altında üretilmiş ve *S. cerevisiae* alpha faktör sinyal dizisi ile hücredışına aktarılmıştır. Rekombinant *P.pastoris* hücresi önce zeosin içeren katı besiyerinde 48-72 saat

çoğalmış ve oluşan tek koloniler gliserol içeren ön besi ortamına aktarılıp 24 saat boyunca çoğalmışlardır. Bu işlemden sonra hücreler santrifüjlenerek ayrılmış ve metanol içeren üretim ortamında 72 st boyunca her 24 saatte bir metanol ile indüklenmişlerdir. SDS-PAGE analizleri rekombinant *P.pastoris* türü tarafından hücredışı üretilen benzaldehit liyaz enzimin monomeri yapısının 59 kDa büyüklüğünde olduğunu göstermiştir. FPLC analizi üretim ortamında enzimin tetramer formunun monomer formundan çok daha az miktar bulunduğunu göstermiştir.

Anahtar Kelimeler: Benzaldehit Liyaz, Hücredışı üretim, Rekombinant *P.pastoris*

To My Family

Seher, H. İnan, Ü. Anıl

ACKNOWLEDGEMENTS

I wish to express my deepest gratitude to my supervisor Prof. Dr. Pınar Çalık for her advice, criticism, insight and especially encouragements throughout the research.

I would also like to express my appreciation to Prof. Dr. H. Tunçer Özdamar for giving me the opportunity to perform part of my experiments in his Laboratories (Ankara University, Chemical Engineering Department, Biochemical Reaction Engineering and Genetic Engineering Laboratories).

I wish to thank to Prof. Dr. Ayhan Sıtkı Demir for his valuable comments and for giving me the opportunity to use his laboratories (METU, Chemistry Department) for TLC and GC-MS analyses.

I also wish to give my special thanks to all my friends in our research group in Industrial Biotechnology and Metabolic Engineering Laboratory and Protein Engineering Laboratory; and members of the research group of Biochemical Reaction Engineering Laboratory in Ankara University. I am especially grateful to Hande Kaya, Nazar İleri, Serdar Güler, Nuriye Korkmaz, Eda Çelik, Vahideh Angardi, A. Hande Ölçeroğlu, Işık Haykır, Birgül Şentürk, Yasemin Demirci, A. Deniz Öztürk, Didem Sutay, Ayşegül Ersayın, Özlem Ak, Selin Aytar, Beril Korkmaz and Aytaç Kocabaş for their invaluable help, cooperation and friendship. I also wish to thank Peruze Ayhan for her guidance and friendship during the experiments I performed in Chemistry Department.

Financial Support provided by TÜBİTAK through the project MAG-104M258 and METU-Research Fund are gratefully acknowledged.

Finally, I would like to express my deepest thanks to my mother Seher Büyüksungur, my father H. İnan Büyüksungur, and my brother Ü. Anıl Büyüksungur for their helping, supporting, encouraging and loving me all through my life.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

NOMENCLATURE

A	Benzaldehyde lyase activity, U cm^{-3}
C_X	Cell concentration, kg dry cell m^{-3}
N	Agitation or shaking rate, min^{-1}
P	Product
PCR	Polymerase Chain Reaction
RE	Restriction Enzyme
S	Substrate
t	Bioreactor cultivation time, h
T	Bioreaction medium temperature, °C
U	One unit of an enzyme

Greek Letters

μ	Specific cell growth rate, h^{-1}
λ	Wavelength, nm

Abbreviations

Ac	Acetic acid
Ala	Alanine
AOX	enzyme alcokol oxidase
BAL	Benzaldehyde lyase
bal	Gene of Benzaldehyde Lyase
BGSC	*Bacillus* Genetic Stock Center
Cys	Cysteine
DO	Dissolved oxygen
DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
EC	Enzyme Commission
FP	Forward Primer
hybrid	Hybrid Gene

Ile	Isoleucine
LIC	Ligation Independent Cloning
Mal	Malic Acid
Met	Methanol
NRRL	Northern Regional Research Center
PCR	Polymerase Chain Reaction
Phe	Phenylalanine
*pre-sub*C	Gene of Signal of Peptide of Serine Alkaline Protease
Pro	Proline
Pyr	Pyruvic Acid
RP	Reverse Primer
SOE	Gene Splicing by Overlap Extension Method
Val	Valine
ThDP	Thiamin diphosphate
Tyr	Tyrosine

CHAPTER 1

INTRODUCTION

Biotechnology is the application of scientific and engineering principles to the processing of materials by biological agents to provide goods and services (OECD Definition). Related with the use of enzymes in chemical industry, for the industrial implementation of many transformations alternative methods in which enzymes are used as the catalyst are available (Madigan et. al, 2003). In this sense, the natural catalysts, enzymes, are now used in a wide range in industrial processes as an alternative to chemical catalysts. Enzymes have a lot of advantages over chemical catalysis:

- Enzymes act in a pH range of about 5-8, typically around 7, and in a temperature range of 20–40^{o}C, preferably at around 30^{o}C. This minimizes problems of undesired side-reactions such as decomposition, isomerization, racemization and rearrangement which are the major challenges in processes where chemical catalysts are used.

- Unlike heavy metals, enzymes are biodegradable and environmentally acceptable.

- Enzymes are very efficient catalysts; typically the rates of enzyme-mediated processes are 10^{8}-10^{10} times higher than those of the corresponding non-enzymatic reactions and this is far above the values that chemical catalysts are capable of achieving (Faber, 2000).

The polypeptide chain of enzyme is folded in such a way that the active site on the surface of the enzyme acts as a keyhole for a specific substrate which results in specificity property of the enzyme. Enzymes show three major types of selectivities:

1. *Chemoselectivity:* Enzyme exhibits specificity on a single type of a functional group.

2. *Regioselectivity and Diastereoselectivity:* Enzyme acts on a functional groups, which are situated in different regions of the same substrate.

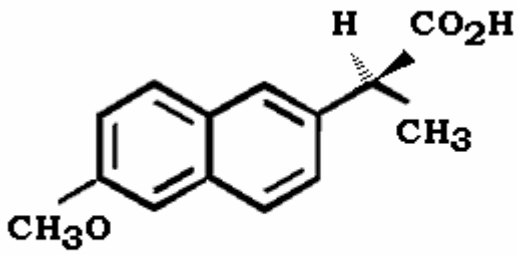
3. *Enantioselectivity:* Certain enzymes can discriminate the differences between the stereoisomer compounds (Scheve, 1984).

Enantioselectivity is the most important property that enzymes display. The synthesis of enantiomerically pure compounds are becoming increasingly more important in the production of pharmaceuticals, agrochemicals (e.g., pesticides, fungicides, herbicides), and flavors (Adam et. al., 1999). Since 1992 both the FDA and the European Committee for Proprietary Medicinal Products have stipulated that the physiological action of each enantiomer of a pharmaceutical product must be individually characterized. In addition to lower dosages and improved efficacy the extension of patent term is also a driving force for pharmaceutical companies to convert a racemic active compound into the enantiomerically pure form.

Enantiomerically pure active compounds are also being used increasingly in the agrochemicals industry. The reasons for this trend are similar to those in the pharmaceutical sector. The targeted synthesis of the respective active enantiomer can improve the economics of the process and lead to reduced quantities applied and thus to reduced environmental impact. (Breuer et. al, 2004).

Practically every known type of synthetic reaction finds its counterpart in enzyme mediated catalysis. Majority of enzymes are highly selective with respect to the chirality of substrate, it is obvious that the enantiomers of a given bioactive compound and cause different biological effects. The effectiveness of a

drug often depends on which enantiomer is used, as does the presence or absence of side effects. For instance, S-naproxen is an important anti-inflammatory drug, while R- enantiomer of naproxen is a liver toxin (Figure 1.1).



(S)- naproxen

Figure 1.1 S- enantiomers of naproxen molecule.

The development of efficient and environmentally acceptable processes for the preparation of enantiomerically pure compounds is essential. Consequently, enantiomerically pure substances can be produced in high enantiomeric excess by utilizing biocatalytic processes since enzymes are chiral materials (Faber, 2000; Hart, 1999).

Enzymes are now widely used for both the biotransformation and synthesis of natural products. Benzaldehyde lyase (BAL, EC 4.1.2.38) is an enzyme, which catalyzes cleavage of the carbon-carbon bond of benzoin to form two benzaldehydes in steriospecific manner and condensation of (R)-benzoin as well. Therefore benzaldehyde lyase enzyme has a potential to produce chiral benzoin, which is used in pharmaceutical industry as a drug intermediate. (Demir et. al. 2001) (Figure 1.2).

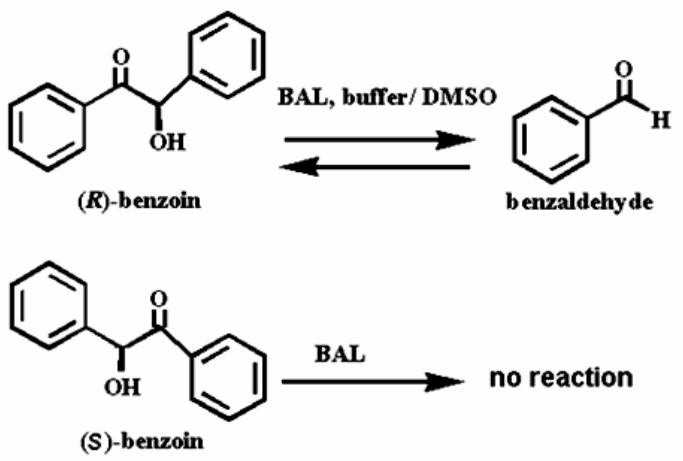


Figure 1.2 R- and S- enantiomers of benzoin molecule.

The steriospecific activity and the ThDP dependency for catalytic activity of the benzaldehyde lyase enzyme were firstly reported by Gonzales and Vicuna (1989). In this study, the enzyme was extracted from *Pseudomonas fluorescens* Biovar I grown in the medium containing benzoin as the sole carbon source.

Hinrichsen et. al. (1994) was determined the nucleotide sequence containing open reading frame encoding benzaldehyde lyase and cloned the gene to an *Escherichia coli* HB101 using vector pUC18. The sequence was later corrected and the substrate specificity was studied with modified pUC18::*bal* plasmid, where hexa histidine was inserted to the C-terminus of the mature gene fragment, by Pohl et. al. (2002).

The reversible catalytic activity of benzaldehyde lyase, where both benzoin and benzaldehyde were substrates, was, firstly, confirmed by Demir et. al. (2001). In this study, the enantioselectivity property of the enzyme, which catalyses both acyloin cleavage and condensation reactions, was proposed; and in further studies the substrate range of the enzyme was expanded (Demir et. al., 2002, 2003, 2004; Dünkelmann et. al., 2002; Sanchez et. al., 2003).

Recently, according to sequence alignments and homology models (Kneen et. al., 2005) and X-ray diffraction analysis (Mosbacher et. al., 2005), the residue similarities between benzaldehyde lyase and other ThDP dependent enzymes, like benzoylformate decarboxylase (BFD) and pyruvate decarboxylase

(PDC), were declared. Together with the detection of active center and reaction geometry, the homotetramer structure of BAL, having a molecular weight of 4 x 58919 Da., provided a significant knowledge about BAL where better approaches towards the substrate selection can be performed.

In a more recent study, Maria et. al. (2005) was investigated the effects of cofactors, cosolvents, and pH on the stability and the activity of the enzyme. The reaction medium containing 30% DMSO in the presence of potassium phosphate buffer with 0.5mM Mg^{+2}, 0.5 mM ThDP and 1mM DTT at pH of 9.5 produced the highest activity and improved the stability of benzaldehyde lyase.

The effects of bioprocess operation parameters to yield the highest cell concentration and benzaldehyde lyase volumetric and specific activities were investigated by Çalık et. al. (2004, 2006); using the defined medium with glucose and $(NH_4)_2HPO_4$ as the sole carbon and nitrogen sources, respectively, the highest BAL activity was obtained at 0.5 vvm, 500min^{-1} as 860 U cm^{-3} with a cell concentration of 2.3 kg m^{-3} at uncontrolled pH of 7.2, using the recombinant *E. coli* K12 carrying pUC18::*bal* plasmid. In this expression system, the gene of *bal* was under the control *trc* promoter, which is a well known hybrid promoter.

Benzaldehyde lyase is being produced as an intracellular product in *E. coli* strains; whereas expression of the *bal* in *Pichia pastoris* may be accomplished as extracellular having an advantage of secretion of target proteins directly to the fermentation broth which leads to a natural separation of the product from cell components which is simplifying downstream processing of the protein.

In the current study, the extracellular production of benzaldehyde lyase enzyme was aimed. For this purpose, yeast *Pichia pastoris* was selected as the host organism. To date, heterologous protein production mostly was achieved by using *E.coli* strains as the host cell. However, there are some occasions where *E.coli* is not the strain of choice and alternative hosts such as *Pichia pastoris* and *Bacillus* species may become attractive. The advantages and the disadvantages of *E.coli*, *Pichia pastoris* and *Bacillus* species as an expression system were summarized in Table 1.1.

Table 1.1 Advantages and disadvantages of *E. coli*, *Bacillus species and Pichia pastoris* as a host microorganism for heterologous protein production (Shuler and Kargı, 2002; Li et. al., 2004; Westers et. al., 2004; Simonen et. al., 1993).

Advantages	Disadvantages
Bacillus species	
• High yield • High growth rate • Capacity to secrete proteins into the extracellular medium(20- 25 mg/ml) • Simple downstream processing of the protein • Much is known about the biochemistry, physiology, and genetics of *B.subtilis* and other species • GRAS(generally regarded as Safe) • Large pH spectrum • Ability to utilize different substrate • Spore formation • Large-scale production systems at high cell densities are easy and usually inexpensive. • *Bacillus subtilis* has been mapped.	• Lack of posttranslational modification • Lack of some strong, Controllable, well-Characterized promoters • Lack of suitable expression vectors • Plasmid instability • Presence of proteases • Occurrence of malfolded proteins
E.coli	
• Has been fully mapped • Easy to handle • Rapid growth rate • Requires an inexpensive-easy to prepare medium • Large yield • Optimum pH:6.0- 7.0	• Glycosylation or post-translational modifications do not occur • Acetate formation • Inclusion body formation • Endotoxin formation
Pichia Pastoris	
• High yield • High productivity • Chemically defined media-simple, inexpensive formulation • Product processing like mammalian cells • Stable production strains • Durability (withstand the adverse cond. of large scale continuous fermentors. This feature makes *Pichia* able to survive unexpected disruptions in fermentation processes). • Low purification cost • High levels of expression of intracellular and secreted proteins • Eukaryotic post-translational modifications • No endotoxin problem • Non-pathogenic • Broad pH range: 3- 7 • Ability of utilizing methanol • Ability of increasing the amount of protein expressed the number of copies of expression cassette increase.	• Potential of ⇒ Proteolysis ⇒ Non-native glycosylation ⇒ Hyper glycosylation • Long time for cell cultivation • Monitoring methanol during a process is very difficult in order to induce AOX1 promoter. • Since methanol is a petrochemical substance, it may be unsuitable for use in the food industry and also storing of this in industrial scale is undesirable because it is a fire hazard.

Pichia pastoris is methylotrophic yeast that can be genetically engineered to express proteins (Higgins and Cregg, 1998). *Pichia pastoris* is suited for foreign protein expression for three key reasons:

1. It can be easily manipulated at the molecular genetic level (e.g. gene targeting, high-frequency DNA transformation, cloning by functional complementation),

2. It can express protein at high levels, intracellularly or extracellularly,

3. It can perform many 'higher eukaryotic' protein modification, such as glycosylation, disulfide-bond formation, and proteolytic processing (Cregg et al., 2000).

Furthermore, the *P. pastoris* itself does not present problems in terms of endotoxin production or viral contamination. *P. pastoris* can be grown to very high cell densities more than 130 kg m^{-3} dry cell weight (Jahicet al., 2002: Wegner, 1990). They also contain a tightly methanol-controlled *alcoholoxidase 1* (*AOX1*) promoter that can be used to drive expression of the foreign gene (Couderc and Baratti, 1980). The first enzyme in the methanol utilization pathway is an AOX enzyme and undetectable in cells grown on glucose, glycerol or ethanol but is dramatically induced when cells are shifted to growth on methanol (Roggenkamp et al., 1984). Heterologous genes can be expressed in *P. pastoris* by placing them under control of the *AOX1* promoter. It has been shown to be simple, easy to scale-up and cost-effective for industrial fermentation (Cregg et al., 1993). The strong promoter, coupled with the high cell density fermentation has allowed production of recombinant product at a very high level; for example Tetanus toxin fragment C at to 12 kg m^{-3} (Clare et al., 1991).

In this study, the strategy of assembly of the expression system for extracellular expression of benzaldehyde lyase enzyme and the production capacity of the recombinant *P.pastoris* in terms of benzaldehyde lyase activity were investigated. In this context, designed pPICZαA:*:bal* plasmid was transfected to *P. pastoris* strain in order to produce benzaldehyde lyase enzyme extracellularly. Thereafter, by using protein analyses techniques extracellular enzyme production was investigated. This study reports the first extracellular production of benzaldehyde lyase in *Pichia pastoris.*

CHAPTER 2

LITERATURE SURVEY

2.1 Enzymes

2.1.1 General Characteristics

Enzymes are proteins that are capable of speeding up chemical reactions. The enzyme neither consumed, nor irreversibly altered during this processes. (Scheve, 1984). Enzymes have extraordinary catalytic power, often far greater than that of synthetic or inorganic catalysts. They have a high degree of specificity for their substrates, they accelerate chemical reactions, and they function in aqueous solutions under very mild conditions of temperature and pH (Nelson, 2004).

All enzymes are proteins, with the exception of a small group of catalytic RNA molecules. Their catalytic activity depends on the integrity of their native protein conformation. If an enzyme is denatured or dissociated into its subunits, catalytic activity is usually lost. If an enzyme is broken down into its component amino acids, its catalytic activity is always destroyed. Thus the primary, secondary, tertiary, and quaternary structures of protein enzymes are essential to their catalytic activity (Nelson, 2004).

Enzymes of high or low specificity can be selected to suit the desired function. The specificity property reduces interference by undesirable substrates and minimizes the problems of unwanted by-products. By-product formation can be a costly inconvenience in industrial processing.

The addition of another process step is always economically undesirable. Enzymes isolated from microorganisms have impressive credentials as catalysts for the synthesis of industrial chemicals. Among the advantages derived from using microbial cells are (Moses and Cape, 1991) :

1) reduced catalyst cost
2) increased stability
3) ease of running multi-catalytic processes
4) decreased time for catalyst production

Some enzymes require no chemical groups for activity other than their amino acid residues. Others require an additional chemical component called a cofactor, either one or more inorganic ions, such as Fe^{+2}, Mg^{+2}, Mn^{+2} (Nelson, 2004).

All enzymes are sensitive to pH and have an optimum pH (or pH range) at which their activity is maximal; at higher or lower pH, activity decreases. Amino acid side chains in the active site may act as weak acids and bases with critical functions that depend on their maintaining a certain state of ionization, and elsewhere in the protein ionized side chains may play an essential role in the interactions that maintain protein structure. Removing a proton from a His residue, for example, might eliminate an ionic interaction essential for stabilizing the active conformation of the enzyme. A less common cause of pH sensitivity is titration of a group on the substrate (Nelson, 2004). pH optima range from acidic, i.e., 1.0, to alkaline, i.e., 10.5 (Kirk and Othmer, 1994).

Enzymes have an optimum temperature at which they work fastest. Usually, enzymes lose activity at quite a low temperature, often slightly above that at which it is typically found. (Bailey, 1986). The temperature optima for enzymes usually lie between 37-47^{o}C (Kirk and Othmer, 1994).

2.1.2 Classification of Enzymes

Enzymes are classified according to a system that was established by the Commission on Enzymes of the International Union of Biochemistry (Atkinson and Mavituna, 1991). This system divides enzymes into six classes, each with subclasses, based on the type of reaction catalyzed (Table 2.1). Each enzyme is

assigned a four-part classification number and a systematic name, which identifies the reaction it catalyzes.(Nelson, 2004)

For example, the EC number of benzaldehyde lyase is EC 4.1.2.38, which catalyzes the cleavage and synthesis of benzoin.

EC 4. Lyases
EC 4.1. Carbon-Carbon Lyases
EC 4.1.2. Aldehyde-Lyases
EC 4.1.2.38. Benzaldehyde lyase

Table 2.1 International classification of enzymes.

No	Class	Type of reaction catalyzed
1	Oxidoreductases	Transfer of electrons
2	Transferases	Group-transfer reactions
3	Hydrolases	Transfer of functional groups to water
4	Lyases	Addition of groups to double bonds or the reverse
5	Isomerases	Transfer of groups within molecules to yield isomeric forms
6	Ligases	Formation of C-C, C-S, C-O, and C-N bonds by condensation reactions coupled to ATP cleavage

2.1.3 Enzyme Activity

The amount of enzyme present or used in a process is difficult to determine in absolute terms (e.g. grams), as its purity is often low and a proportion may be in an inactive, or partially active, state. More relevant parameters are the activity of the enzyme preparation and the activities of any contaminating enzymes. The activity is a measure of enzyme content that is clearly of major interest when the enzyme is to be used in a process. For this reason, enzymes are usually marketed in terms of activity (Chaplin and Bucke, 1990), which was defined by the Commission on Enzymes:

One unit (U) of enzyme activity is defined as the amount which will catalyze the transformation of one micromole of substrate per minute at 25^0C under optimal conditions of measurement.

A comparison of the activity of different enzyme preparations was only possible if the assay procedure is performed exactly in the same way (Faber, 2000). However, these so-termed optimal conditions vary even between laboratories and suppliers (Chaplina and Bucke, 1990). Therefore, the parameters such as temperature, pH, and/or substrate concentration that affects the reaction rates, must be carefully controlled in order to achieve reproducible results (Kirk and Othmer, 1994).

2.2 Benzaldehyde Lyase

Benzaldehyde lyase (BAL, EC 4.1.2.38) from *Pseudomonas fluorescens* Biovar I, which can grow on benzoin as a sole carbon and energy source due to the benzaldehyde lyase activity, was firstly reported by Gonzales et al. (1989). In this study, the purification and the characterization of the enzyme were described and the ability of BAL to catalyze the cleavage of acyloin linkage of benzoin to form benzaldehyde was irreversible described. It was also stated that benzaldehyde lyase requires ThDP for catalytic activity which was maximum at 0.01 mM ThDP, while concentrations higher than 0.5 mM were inhibitory. The activity loss was observed when treated with EDTA which can be restored by addition of 1.0 mM concentrations of $MgCl_2$, $MnSO_4$, or $CaSO_4$. The enzyme showed maximal activity between pH 7.5 and 8.5, whereas it was inactive below pH 6.0. Thereafter, the nucleotide sequence contains open reading frame encoding benzaldehyde lyase and the location of the gene was determined (Hinrichsen et. al., 1994) by cloning the BAL gene into pUC18 plasmid and expressed in *E. coli* HB101. Afterward, the DNA sequence published was corrected and re-submitted; and for easier purification, hexa-histidine tag was fused to the C-terminus of the enzyme and expressed in *E. coli* SG130009 by Pohl et. al. (2002).

Benzaldehyde lyase, an enzyme used for synthesis of enantio pure 2-hydroxyl ketones which are an important class of compounds in natural product and drug synthesis, catalyzes cleavage of the carbon-carbon bond of benzoin to form two benzaldehydes. BAL also catalyzes the reverse acyloin condensation of benzaldehydes resulting in the synthesis of (R)-benzoins. Only one enantiomer of benzoin acts as substrate when racemic mixture of benzoin is reacted with BAL. In other words, only (R)-benzoin is converted into

benzaldehyde through BAL catalysis, while (S)-benzoin gives no reaction at all. The catalytic mechanism of the enzyme is schematized in Figure 2.1 (Demir et. al., 2001).

In further studies, the catalytic activity and substrate range of benzaldehyde lyase like the enantioselective C-C bond cleavage and formation to generate (R)- and (S)-benzoins and (R)-2hydroxypropiophenone ((R)-2HPP) derivatives (Demir et. al., 2001); C-C bond formation from aromatic aldehydes and acetaldehyde (Demir et. al., 2002) and methoxy- and dimethoxy-acetaldehydes (Demir et. al., 2003) on a preparative scale in buffer/DMSO solution; asymmetrical synthesis of mixed benzoins (Dünkelmann et. al., 2002); mixed acyloin condensation between methoxy- substituted benzaldehydes and phenylacetaldehyde (Sanchez et. al., 2003); and hydroxymethylation of aromatic aldehydes with formaldehyde (Demir et. al., 2004) were investigated.

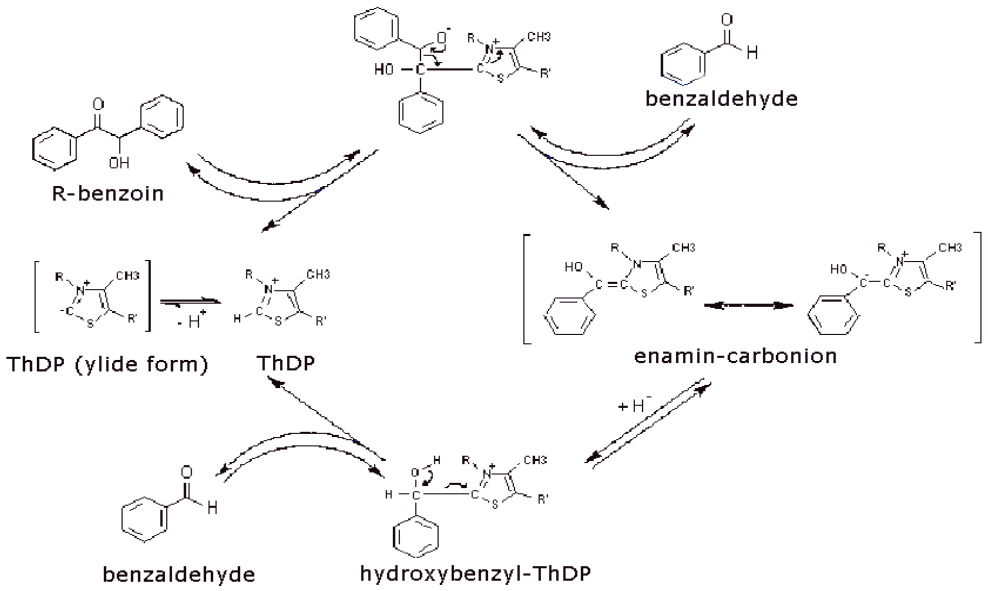


Figure 2.1 Benzaldehyde lyase catalyzed cleavage and synthesis of benzoin. The first step of the catalytic cycle is the attack of ylide form of ThDP on the carbonyl carbon of (R) - benzoin to produce an adduct. The enamine, intermediate product of ThDP dependent enzymes catalyzing the formation of 2-hyroxy ketones, and the first free aldehyde molecule is formed. Protonation of this intermediate then releases the second molecule of aldehyde and restores the cofactor. In the presence of an acceptor aldehyde and enamine intermediate is able to undergo a C-C bond formation reaction.

Recently, the enzyme structure has been modeled to show the residue similarities of ThDP dependent enzymes like benzoylformate decarboxylase (BFD) and pyruvate decarboxylase (PDC). According to sequence alignments and homology models, the active site residues of BAL were identified and it is found that the glutamine residue, Gln113, present in the active site plays an important role wherein replacement with alanine or histidine resulted in 200 fold activity loss (Kneen et. al., 2005).

In the study of Mosbacher et. al (2005), the X-ray structure of benzaldehyde lyase was established in order to show that BAL belongs to a group of closely related ThDP dependent enzymes. The results of X-ray diffraction exposed that BAL is a homotetramer, where each subunit binds to one ThDP molecule using one Mg^{2+} ion. The enzyme is composed of 4 x 563 amino acid residues and has a molecular weight of 4 x 58919 Da. The BAL homotetramer has an overall size of approximately 95x95x75 $Å^3$. No significant structural differences were found between the four crystallographically independent subunits of the tetramer (Figure 2.2). Each subunit consists of the three domains Dom-α, Dom-β and Dom-γ (Fig 2.2). All three domains consist of a central six-stranded parallel β-sheet flanked by a varying number of α-helices. Residues involved in binding of the cofactor ThDP are located at the C-terminal ends of the β -strands of Dom- γ (diphosphates and Mg^{2+}) and of Dom- α′ of a neighboring subunit (pyrimidine moiety). The active centre is defined by the thiazolium ring of ThDP, which sits in a deep pocket opening to the outer surface of the tetramer (Mosbacher et. al., 2005).

The four subunits A, B, C and D form the two tight dimers A–B and C–D around the molecular axis P (Figure 2.2), in which each subunit buries a solvent-accessible surface area of 3270 $Å^2$. The two tight dimers are associated much less tightly around the molecular axes Q and R to form a D_2-symmetric homotetramer. These secondary interfaces bury 1790 $A^{\circ 2}$ per subunit. The tight contact is formed by Dom- α and Dom-γ of subunit A with their counterparts in subunit B. It is stabilized by a large number of hydrogen bonds. The weaker contact results from an association of Dom- α and Dom- β of subunit A with the respective domains of subunit D. It contains only few hydrogen bonds (Mosbacher et. al., 2005).

In all ThDP dependent enzymes ThDP assumes a V-conformation resulting in a close approach between the C_2 atom of the thiazolium ring and the N4' atom of the pyrimidine moiety. A superposition of the cofactors is depicted in Figure 2.3 revealing a remarkable conformational similarity.

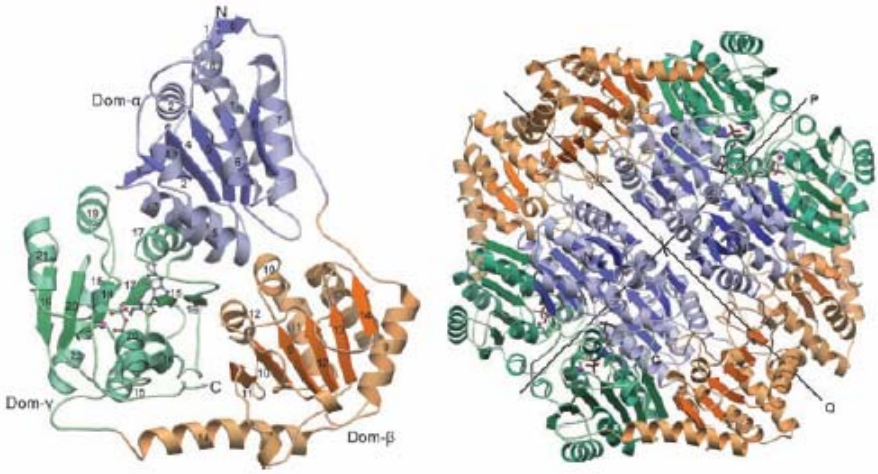


Figure 2.2 Stereo ribbon plot of a BAL subunit composed of the three domains Dom-α (residues 1–183), Dom- β (residues 184–363) and Dom-γ (residues 364–563). The cofactor ThDP is shown as a ball-and-stick model and Mg^{2+} as a sphere. And Stereo ribbon plot of the D_2-symmetric BAL tetramer with the three molecular twofold axes P, Q and R . The tetramer should be described as a dimer of dimers. The tightest interfaces are around axis P. Each tight dimer contains two active centres at its interface. ThDP is shown as a ball-and-stick model(Mosbacher et. al., 2005).

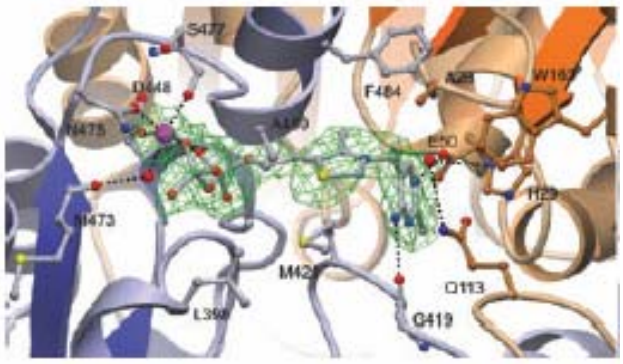


Figure 2.3 Stereo view of ThDP-binding at BAL showing the initial (Fo-Fc)-electron density map of ThDP and Mg2+. The cofactor binds in the typical V-conformation required for catalysis(Mosbacher et. al., 2005).

The diphosphates are tightly bound to the polypeptide of Dom-c using Mg^{2+} as a mediator. The Mg^{2+} ion is octahedrally coordinated to the side-chains of Asp448 and Asn475, to the backbone carbonyl of Ser477, to the diphosphate as well as to a water molecule (Figure 2.3). This binding motif is present in all ThDP-dependent enzymes (Mosbacher et. al., 2005).

Maria et. al. (2005), investigated the effects of cofactors, cosolvents, and pH, on the stability and the activity of the BAL to obtain higher chemical yields for the synthesis of benzoins by carboligation of benzaldehyde derivatives. Among the investigated parameters, it was shown that 30% DMSO content in the presence of potassium phosphate buffer with 0.5mM Mg^{+2}, 0.5 mM ThDP and 1mM DTT, a well known stabilizer of hyrolases, were the optimal for enzyme activity and stability. It was shown that the enzyme prepared in potassium phosphate buffer supplemented with ThDP, Mg^{2+}, DTT as a cofactor and 30% DMSO as a cosolvent leaded to 50% activity loss after 30 h, while the enzyme losses its activity almost completely within 3 h in water. It was also shown that best enzymatic activity appeared when the pH of the reaction medium was increased from 8.0 to 9.5.

The studies in the literature reporting the effects of bioprocess operation parameters on benzaldehyde lyase production from *E. coli* K12 were published by Çalık et. al. (2004, 2006). The host microorganism having the highest benzaldehyde lyase productivity was determined as *E. coli* K12 (ATCC 10798) carrying modified pUC18::*bal* plasmid where inducible hybrid t*rc* promoter is present. Among the investigated media, the highest cell concentration and benzaldehyde lyase activity were obtained as 1.8 kg m^{-3} and 745 U cm^{-3}, respectively, in the medium containing 8.0 kg m^{-3} glucose, 5.0 kg m^{-3} $(NH_4)_2HPO_4$ and the salt solution. Thereafter, the effects of uncontrolled-pH and controlled pH operations and effects of oxygen transfer conditions on benzaldehyde lyase productivity were systematically investigated using the recombinant *E. coli* K12 carrying pUC18::*bal* plasmid. Among the controlled-pH operations, the highest cell concentration and BAL activity were obtained as 2.1 kg m^{-3} and 775 U cm^{-3}, respectively at pH_C 7.0; whereas they were 2.3 kg m^{-3} and 860 U cm^{-3}, respectively at 0.5 vvm, 500 min^{-1} and uncontrolled pH 7.2 condition.

2.3 Microorganisms

In bioprocesses, the selection of host microorganism for production of industrial enzymes is often critical for the success of the process. Potential hosts should give sufficient yields, be able to secrete large amounts of protein, be suitable for industrial fermentations, produce a large cell mass per volume quickly and on cheap media, be considered safe based on historical experience or evaluation by regulatory authorities, and should not produce harmful substances or any other undesirable products (Kirk and Othmer, 1994).

Benzaldehyde lyase is naturally produced by wild-type *Pseudomonas fluorescens* (Gonzales et. al., 1989). The gene encoding benzaldehyde lyase was firstly cloned to a prokaryotic microorganism *Escherichia coli* strain by Hinrichsen et. al. (1994), and further studies about the catalysis ability of the enzyme in stereospecific manner were performed by using *E. coli* SG130009 as the host microorganism. In the literature concerning benzaldehyde lyase production, Çalık et. al. (2004, 2006) selected *E. coli* K12 strain as host microorganism among the investigated *E. coli* strains, namely *E. coli* JM109, *E. coli* XL-1 Blue, *E. coli* K12 and *E. coli* GBE 180. An alternative to these, to achieve extracellular production of BAL recombinant *B.subtilis* and *Pichia pastoris* were selected as the host.

2.3.1 The Genius *Bacillus*

The genus, *Bacillus*, are rod-shaped, prokaryotic, gram positive, aerobic or facultative, endospore-forming bacteria. The family's distinguishing feature is the production of endospores, which are highly refractile resting structures formed within the bacterial cells. The ubiquity of *Bacillus* species in nature, the unusual resistances of their endospores to chemical and physical agents, the developmental cycle of endospore formation, the production of antibiotics, the toxicity of their spores and protein crystals for many insects have attracted ongoing interest in the genus (http://www.textbookofbacteriology.net, Last accessed date June 2006).

The genius *Bacillus,* exhibits a wide diversity of physiological abilities; although the majority are mesophilic, there are also psychrophilic and

thermophilic species. Moreover, the morphological divisions of *Bacillus* spp. into (Fogarty, 1990; Sonenshein, 1993).

1) species that produce oval endospores that do not distend the mother cell,
2) species that produce oval endospores that distend the mother cell,
3) species that produce spherical endospores is a useful division of genus.

Like most Gram-positive bacteria the surface of the *Bacillus* is complex and is associated with their properties of adherence, resistance and tactical responses. The vegetative cell surface is a laminated structure that consists of a capsule, a proteinaceous surface layer (S-layer), several layers of peptidoglycan sheeting, and the proteins on the outer surface of the plasma membrane. Walls of gram-positive bacteria are dynamically variable and flexible structures that enclose and protect the underlying cytoplasmic membranes. (http://www.textbookofbacteriology.net, Last accessed date June 2006; Sonenshein, 1993).

In this current study, *B. subtilis* with two protease deficient *apr*$^-$ and *npr*$^-$ strain was selected as the potential producer of benzaldehyde lyase as extracellular.

2.3.1.1 Protein Secretion in *Bacillus* Species

Among the features that *Bacilli* serve, the heterologous protein secretion to the growth medium is the most attractive feature that made the microorganism so popular for the applications in both industry and research (Smith, 1995).

Bacillus subtilis is an attractive host for extracellular protein secretion. The secretory pathway of proteins can be divided into three functional stages: the early stage, involving synthesis of secretory pre-proteins, their interaction with chaperones and binding to the secretory translocase; the second stage, translocation across the cytoplasmic membrane; and the last stage, including removal of the signal peptide, protein refolding and passage through the cell wall (Simonen et. al., 1993).

The proteins secreted are synthesized as precursors with an amino terminal extension, which contains the signal sequence ,or signal peptide. This peptide distinguishes the exported proteins from the cytoplasmic ones and is needed for targeting of the proteins to the export pathway (Simonen et. al., 1993).

The primary structures of amino-terminal signal peptides show similar three distinct domains. The amino-terminal N-domain of signal peptides contains at least one arginine or lysine residue and suggests interacting with the translocation machinery and negatively charged phospholipids in the lipid bilayer of the membrane during translocation. The H-domain, following the N-domain, is formed by a stretch of hydrophobic residues that seem to adopt an α-helical conformation in the membrane. Helix-breaking glycine or proline residues are frequently present in the middle of this hydrophobic core. The latter residues might allow the signal peptide to form a hairpin-like structure that can insert into the membrane. Helix-breaking residues found at the end of the H-domain, are thought to facilitate cleavage by a specific signal peptidase (SPase). The C-domain, following the H-domain, contains the cleavage site for SPase, which removes the signal peptide from the mature part of the secreted protein during or shortly after translocation. The mature part of the protein is thereby released from the membrane and can fold into its native conformation. Finally, the signal peptide is degraded by signal peptide peptidases (SPases) and removed from the membrane (Figure 2.4). Although different amino-terminal signal peptides tend to be quite similar in general structure, apparently small differences between individual signal peptides can cause cleavage by different SPases, export via different pathways, and transport to different destinations (Tjalsma et. al., 2000).

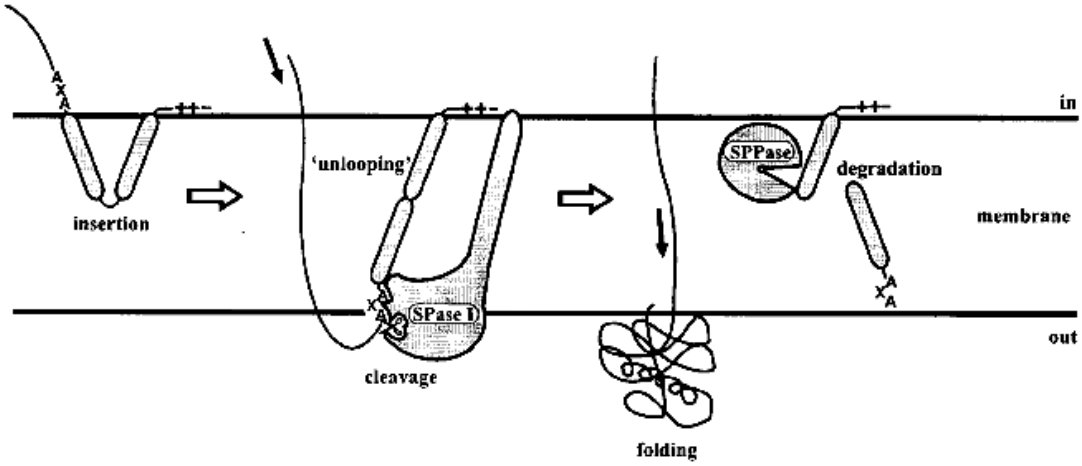


Figure 2.4 Schematic diagram of signal peptide insertion into the cytoplasmic membrane and cleavage by SPase I (Tjalsma et. al, 2000).

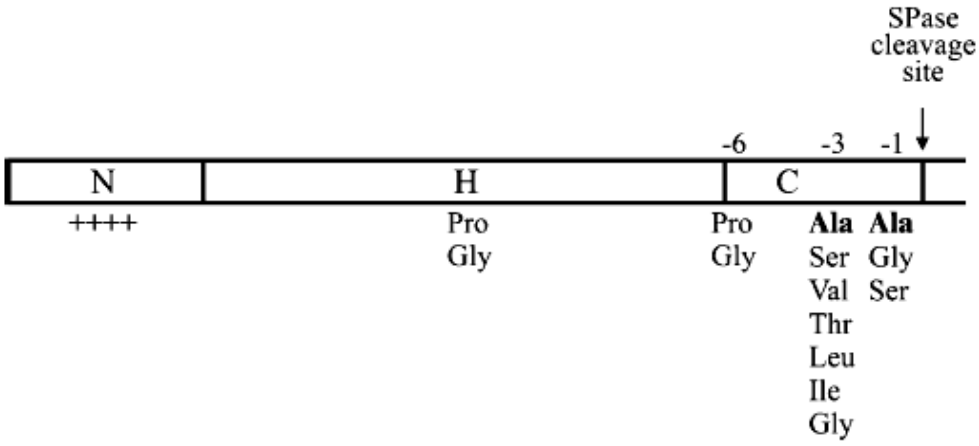


Figure 2.5 Secretory (Sec-type) signal Peptides: positively charged N-terminus (N-region), a central hydrophobic region (H-region), and a polar C-terminal region (C-region) (van Roosmalen et. al.,2004)

Ala-X-Ala is the most commonly observed signal peptidase cleavage sequence in *B. subtilis*. In this group of SPase recognition sequence, small aliphatic residues, often alanine, are preferred at positions -1 and -3 relative to the cleavage site in pre-proteins. For the -1 position, alanine, glysine and serine residues, the -3 position can also accommodate larger residues, such as valine, threonine, leucine and isoleucine (Figure2.5). Especially, an alanine residue at the -1 position relative to the cleavage site seems to be a critical determinant

for SPase I-mediated cleavage of pre-proteins in *B. subtilis* (van Roosmalen et. al., 2004).

Signal peptidases can be a limiting factor for processing of certain precursor proteins (Li et. al., 2004). In attempts to make *Bacillus* species secrete foreign proteins, the joint between *Bacillus* signal peptide and foreign protein is usually made immediately after the signal peptide since extra residues between the mature portion and signal sequence may affect the stability or activity of the protein (Simonen et. al., 1993)

2.3.1.2 Signal Peptide of *Bacillus licheniformis*, *pre-sub*C

The gene encoding subtilisin Carlsberg from *Bacillus licheniformis* NCIB 6816 (*sub*C) was firstly reported by Jacobs et. al. (1985- Accession no: X03341). In this study, the entire nucleotide sequence of the coding sequence as well as 5' and 3' flanking sequences were determined. It was declared that the deduced amino acid sequence reveals an N-terminal signal peptide consisting of 29 residues, a pro-peptide of 76 residues followed by the mature protein comprising 274 residues. Later, the endogenous promoter activity of *sub*C was identified and the 5' and 3' ends of *sub*C transcripts were characterized (Jacobs et. al., 1995). The putative signal sequence of subtilisin Carlsberg has similar properties to other signal sequences belong to the secretory (Sec-type) signal peptides. Jacobs et. al. (1985) has reported that the cleavage site of subC follows the (-3, -1) rule, giving cleavage preferentially after the residues Ala-X-Ala. Based on this rule and other structural considerations, it was predicted that the signal peptidase processing occurs after the residues Ala-Ser-Ala as shown in Table 2.2.

Table 2.2 Signal peptides of *Bacillus* species. Signal peptides were identified as underlined text, positively charged lysine (K) and arginine (R) residues in the N-domain are indicated in bold letters; the hydrophobic H-domain is indicated in gray shading; and signal peptide cleavage site is indicated by arrows (Tjalsma et. al., 2000; Simonen et. al., 1993)

Protein	Species of origin	Signal Peptide
α-Amylase	*B. subtilis*	MFA**KR**F**K**TSLLPLFAGFLLLFHLVLAGPAA**ASA**↓
α-Amylase	*B. amyloliquefaciens*	MIQ**KRKR**TVSFRLVLMCTLLFVSLPITK**TSA**↓
α-Amylase	*B. licheniformis*	M**K**QH**KR**LYARLLPLLFALIFLLPHSA**AAA**↓
Extracellular protease	*B. subtilis*	M**K**NMSC**K**LVVSVTLFFSFLTIGPL**AHA**↓
Subtilisin E	*B. subtilis*	M**R**S**KK**LWISLLFALTLIFTMAFSNMS**VQA**↓
Subtilisin	*B. amyloliquefaciens*	MRG**KK**VWISLLFALALIFTMAFGSTSS**AQA**↓
Subtilisin Carlsberg	***B. licheniformis***	MM**RKK**SFWLGMLTAFMLVFTMAFSDS**ASA**↓
Pectate lyase (Pel)	*B. subtilis*	M**KK**VMLATALFLGLTPAG**ANA**↓
Levansucrase (SacB)	*B. subtilis*	MNI**KK**FA**K**QATVLTFTTALLAGGATQ**AFA**↓

There is no significant evidence for the function of pro-peptide. In the studies using the signal sequence for heterologous protein secretion in *B. subtilis*, only the pre-peptide region is used (Wong et. al., 1986-a&b; Wang et. al., 1988). In the current study, the signal peptide (*pre-sub*C) with and without its own promoter was tried for extracellular production of benzaldehyde lyase.

2.3.1.3 The Vectors for *Bacillus sp.* cloning experiments

pUC series vectors are primarily designed for general cloning and sequencing. They can be used as a sub-cloning vector or can also be used for expression purposes.

In previous studies, the gene encoding benzaldehyde lyase enzyme was purified and cloned to pUC18 cloning vector (Hinrichsen et. al., 1994) and the enantioselective synthesis ability of BAL and the bioprocess operation parameters were investigated with *Escherichia coli* strain using recombinant pUC18::*bal* plasmid (Demir et al., 2001, 2002, 2003 and Çalık et. al., 2004, 2006). The *bal* gene under the control of *trc* promoter was cloned to the multiple cloning site present on the *lacZ* gene fragment and expressed in *E. coli.* In this current study, the pUC19 cloning vector (Figure 2.6) was used for application of sub-cloning.

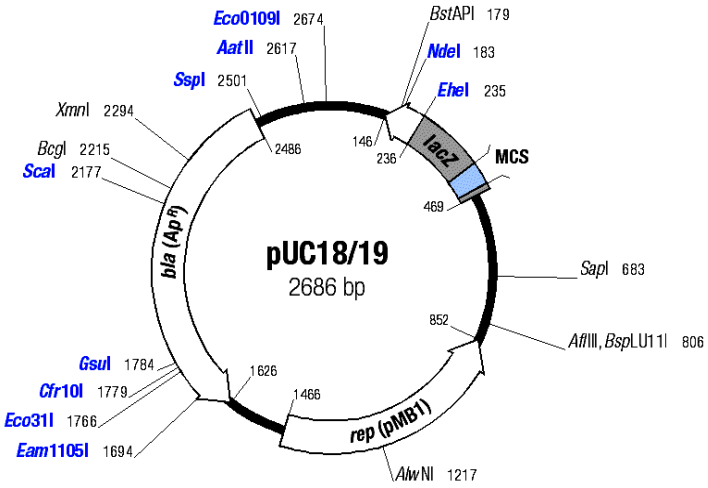


Figure 2.6 pUC18/19 cloning vector (http://www.fermentas.com, Last accessed date June 2006).

The first strategy for the extracellular expression of the benzaldehyde lyase gene in *B.* Subtilis was to clone the hybrid gene in pMK4. The *pre-subC*

sequence was amplified with its own promoter and joined with the *bal* gene by means of SOE method.

The second strategy for the extracellular expression of the benzaldehyde lyase gene in *B.* Subtilis was to clone the hybrid gene in pDG148-*Stu* plasmid by the help of ligation independent cloning method.

The third strategy for the extracellular expression of the benzaldehyde lyase gene in *B.* Subtilis was to clone the hybrid gene in pMUTIN4 and to integrate the hybrid gene into genomic DNA.

2.3.2 *Pichia pastoris*

The methylotrophic yeast *Pichia pastoris* is an eukaryotic microorganism and has a potent expression system with high capability to secrete recombinant proteins to the fermentation medium (Cereghino et al., 2002). Protein concentrations in the culture supernatant can exceed 10 kg m^{-3} (Werten et al., 1999). *P. pastoris* expression system has several advantages mentioned in introduction section, but two are the most compelling. The first is an unusually efficient and tightly regulated promoter from the alcohol oxidase I gene (*AOX1*) that is used to drive the expression of the foreign gene (Cregg et. al., 1988). The *AOX1* promoter is strongly repressed in cells grown on glucose and most other carbon sources, but is induced over 1000-fold when cells are shifted to a medium containing methanol as a sole carbon source (Cereghino et al., 2002). In general, strains are grown initially in a defined medium containing glycerol as its carbon source. During this time, biomass accumulates but heterologous gene expression is fully repressed (Cereghino and Cregg 2000). The second reason is that, physiologically, *P. pastoris* prefers a respiratory rather than a fermentative mode of growth. Fermentation products include ethanol and acetic acid, which quickly reach toxic levels in the high cell density environment of a fermenter with strongly fermentative organisms (Cereghino et al., 2002).

Foreign proteins expressed in *P. pastoris* can be produced either intracellularly or extracellularly. Since *P. pastoris* secretes only low levels of endogenous proteins, the secreted heterologous protein constitutes the vast majority of total protein in the medium (Cereghino and Cregg 2000). Therefore,

directing a heterologous protein to the culture medium can serve as a substantial first step in purification.

Although several different secretion signal sequences have been used successfully for secretion in *P.pastoris*, *S. cerevisiae* α factor signal sequence has been used with the most success (Cereghino and Cregg 2000). The processing of this signal sequence involves three steps. The first is the removal of the *-pre* signal by signal peptidase in the endoplasmic reticulum. Second, Kex2 endopeptidase cleaves between Arg- Lys of the pro leader sequence. This is rapidly followed by cleavage of Glu-Ala repeats by the Ste13 protein (Brake et. al., 1984). In this current study *S. cerevisiae* α factor signal sequence was used for the excretion of benzaldehyde lyase into fermentation medium. The Ste13 protein cleaves the signal peptide at the beginning of the sequence just before the *EcoRI* sequence.

In yeast, proteins destined for the cell surface or for export from the cell are synthesized on and translocated into the endoplasmic reticulum. From there they are transported to the golgi body for processing and packaging into secretory vesicles. Fusion of the secretory vesicles with the plasma membrane then occurs constitutively or in response to an external signal (Rothman, 1992).

P. pastoris is able to add both *O*-linked and *N*-linked carbohydrate moieties to secreted proteins (Cereghino *et. al, 2002*). Eukaryotic cells assemble O-linked saccharide onto the hydroxyl groups of serine and threonine. In mammals, O-linked oligosaccharides are composed of a variety of sugars, including N-acetylgalactosamine, galactose, and sialic acid. In contrast, lower eukaryotes such as *P. pastoris* add O-oligosaccharides composed solely of mannose residues (Cereghino and Cregg, 2000).

In all eukaryotes, N-glycosylation begins in the endoplasmic reticulum with the transfer of a lipid-linked oligosaccharide unit, to asparagine at the recognition sequence Asn-X-Ser/Thr (Goochee et. al., 1991).

Glycosylation of the heterologous proteins might be harmful to the enzyme. Because oligosaccarides' chains can potentially interfere with the folding or function of a foreign protein.

2.3.2.1 Methanol metabolism in *P. pastoris*

The conceptual basis for the *P. pastoris* expression system stems are that some of the enzymes required for methanol metabolism are present at substantial levels only when cells are grown on methanol (Veenhuis et. al., 1983; Egli et. al., 1980). Biochemical studies showed that methanol utilization requires a novel metabolic pathway involving several unique enzymes (Veenhuis et. al., 1983). The enzyme alcohol oxidase (AOX) catalyzes the first step in the methanol utilization pathway, the oxidation of methanol to formaldehyde and hydrogen peroxide. AOX is sequestered within the peroxisome along with catalase, which degrades hydrogen peroxide to oxygen and water. A portion of the formaldehyde generated by AOX leaves the peroxisome and is further oxidized to formate and carbon dioxide by two cytoplasmic dehydrogenases, reactions that are a source of energy for cells growing on methanol. The remaining formaldehyde is assimilated to form cellular constituents by a cyclic pathway that starts with the condensation of formaldehyde with xylulose 5-monophosphate, a reaction catalyzed by a third peroxisomal enzyme dihydroxyacetone synthase (DHAS). The products of this reaction, glyceraldehyde 3-phosphate and dihydroxyacetone, leave the peroxisome and enter a cytoplasmic pathway that regenerates xylulose 5-monophosphate and, for every three cycles, one net molecule of glyceraldehyde 3- phosphate. Two of the methanol pathway enzymes, AOX and DHAS, are present at high levels in cells grown on methanol but are not detectable in cells grown on most other carbon sources (Couderc et. al., 1980;Roggenkamp et. al., 1984). Figure 2.7 illustrates the methanol metabolism in *P. pastoris.*

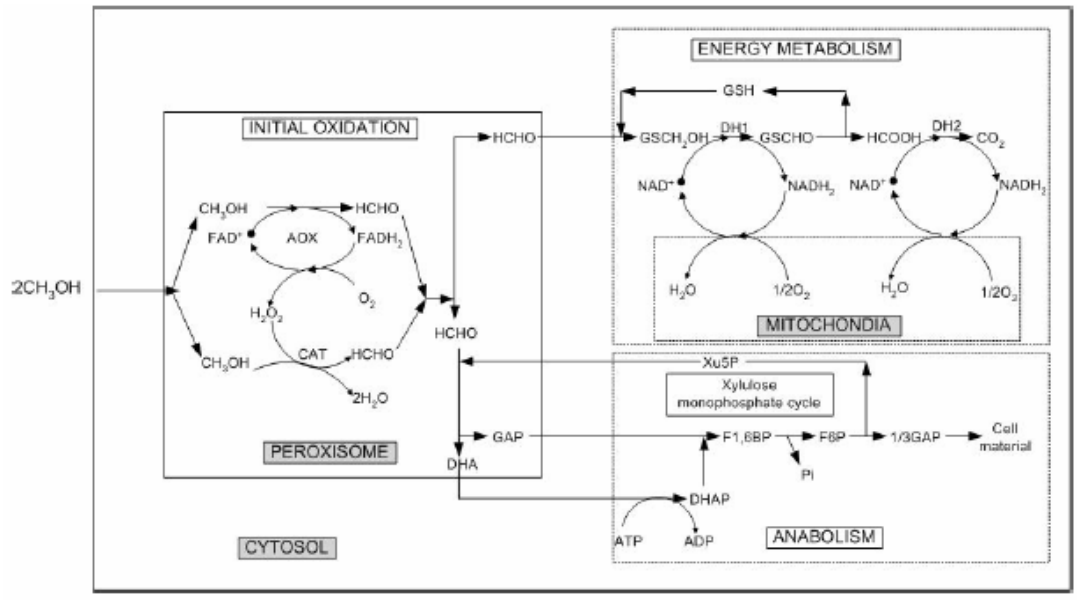


Figure 2.7 Methanol metabolism in *P. pastoris*. AOX, alcohol oxidase; CAT, catalase; DH1, formaldehyde dehydrogenase; DH2, formate dehydrogenase; GSCH2OH, S-hydroxymethylglutathione; GSCHO, S-formylglutathione; GAP, glyceroldehyde-3-phosphate; DHA, dihydroxyacetone; DHAP, dihydroxyacetone phosphate; F1,6BP, fructose-1,6-bisphosphate; F6P, fructose-6-phosphate; Xu5P, xylulose-5-phosphate. (Charoenrat, 2005)

Expression of the AOX1 gene is controlled at the level of transcription (Tschopp et. al., 1987; Gregg et. al., 1989). In methanol-grown cells, about 5% of poly(A) RNA is from AOX1; however, in cells grown on other carbon sources, AOX1 message is undetectable (Cregg et. al., 1988). The regulation of the AOX1 gene appears to involve two mechanisms: a repression/derepression mechanism. In the absence of a repressing carbon source, such as glucose in the medium, does not result in substantial transcription of AOX1. The presence of methanol is essential to induce high levels of transcription (Tschopp et. al., 1987).

2.3.2.2 The Vector pPICZαA

Plasmids most commonly used in recombinant DNA technology are those that are engineered to optimize their use as vectors in DNA cloning have three common region essential for DNA cloning: a replication origin; a marker that permits selection, usually a drug resistance gene; and a region in which

exogenous DNA fragments can be inserted (Lodish et. al., 2003). Most of the *P.pastoris* expression vectors have an expression cassette composed of a fragment from AOX1 composed of the 5' promoter sequences and a second short AOX1-derived fragment with sequences required for transcription termination (Koutz et. al., 1988). Between the promoter and terminator sequences is a site or multiple cloning site (MCS) for insertion of the foreign coding sequence (Cereghino and Cregg, 2000).

In previous studies, the gene encoding benzaldehyde lyase enzyme was purified and cloned to pUC18 cloning vector (Hinrichsen et. al., 1994) and the enantioselective synthesis ability of BAL and the bioprocess operation parameters were investigated with *Escherichia coli* strain using recombinant pUC18::*bal* plasmid (Demir et al., 2001, 2002, 2003 and Çalık et. al., 2004, 2006). The *bal* gene under the control of *trc* promoter was cloned to the multiple cloning site present on the *lacZ* gene fragment and expressed in *E. coli.* In this study, pPICZαA vector (Figure 2.8) was used for extracellularly production of benzaldehyde lyase.

pPICZαA vector has several useful features as a vector such as it integrates into genome such a way that the heterologous gene expressed under the control of AOX1 promoter one of the strongest and most regulatable promoters known (Cereghino & Cregg 1999, 2000); pUC origin to provide high copy replication and growth in *E. coli*; multiple cloning site containing 12 restriction enzyme recognition sequence; AOX1 terminator permitting efficient transcription termination; the bleomycin-related drug Zeocin resistant gene (Invitrogen, catalog V195-20).

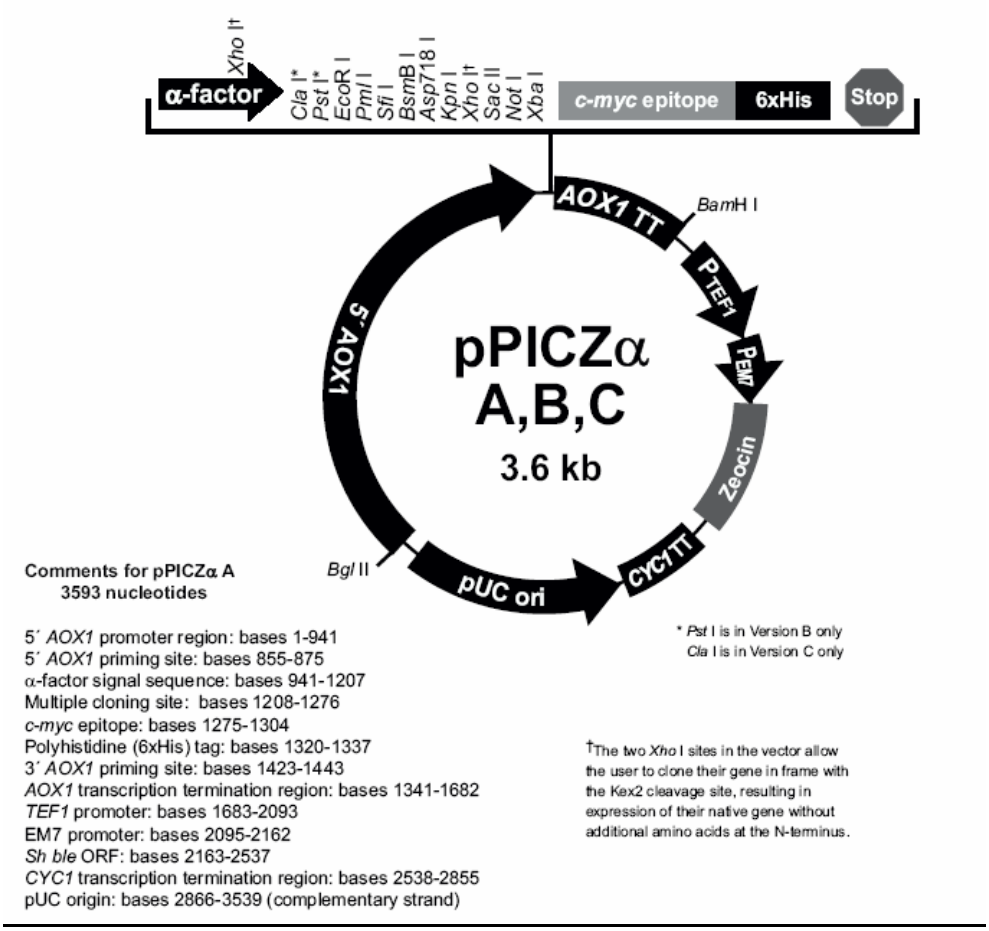


Figure 2.8 pPICZαA vector (Invitrogen, catalog V195-20).

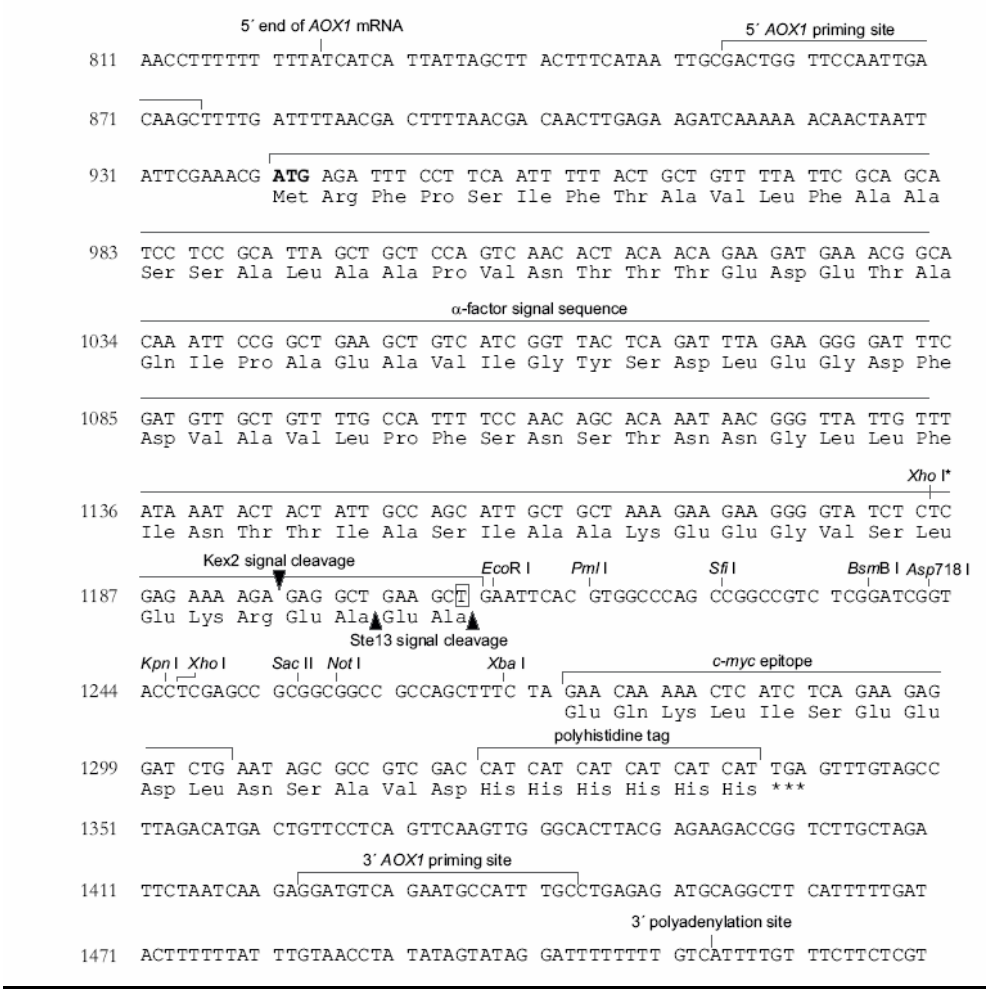


Figure 2.9 Multiple Cloning Site of pPICZαA (Invitrogen, catalog V195-20)

2.4 Genetic Engineering Techniques: Methodology

The techniques of genetic engineering, alternatively known as recombinant DNA technology, brought about a revolution in biotechnology science. Since the industrial enzymes has become more and more important, improving product formation or cellular properties to design and create the optimal biocatalysts through the modification in enzymatic reactions by using recombinant DNA technology lead to the development of genetic engineering techniques. Amplifying the specific DNA regions by Polymerase Chain Reaction (PCR) method, determination of DNA concentration, SDS-Polyacrylamide analysis of secreted protein and restriction digestion are some basic principles that used in recombinant DNA technology.

The genetic engineering techniques, for the development of the r-*P. pastoris* producing the Benzaldehyde lyase enzyme are illustrated in Figure 2.10.

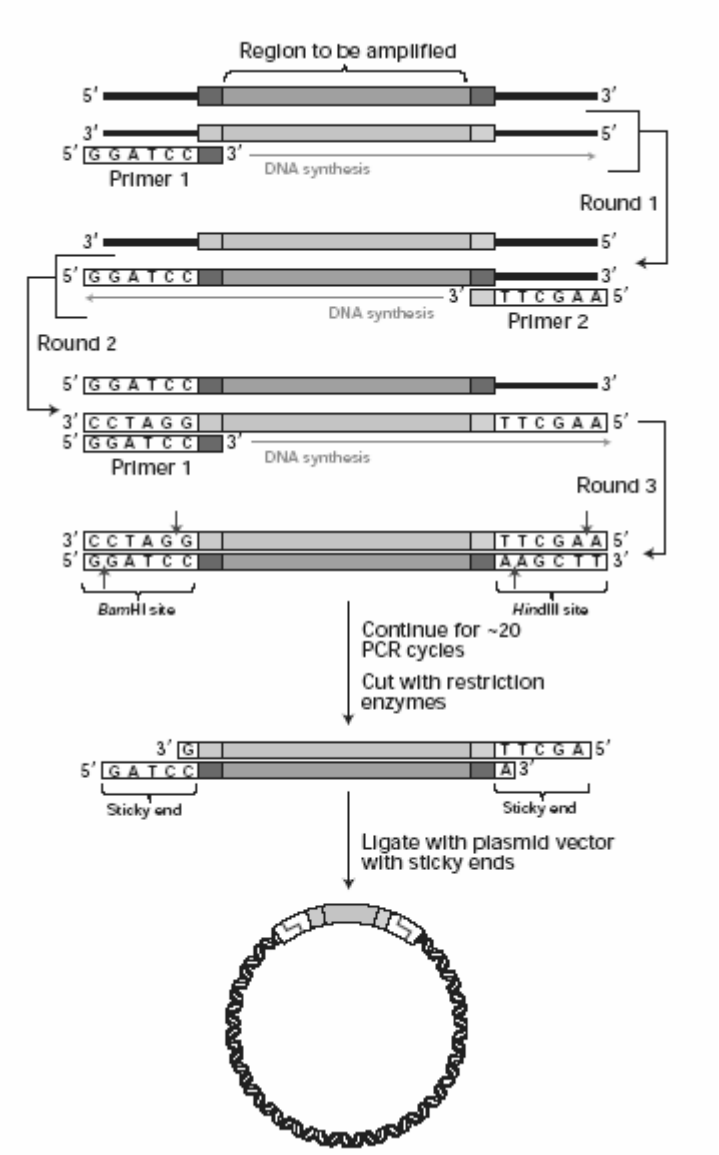


Figure 2.10 Representative scheme of genetic engineering techniques (Lodish et. al., 2003)

2.4.1 Polymerase Chain Reaction (PCR) Method

The PCR involves two oligonucleotide primers, which flank the DNA sequence that is to be amplified. The primers hybridize to opposite strands of the DNA after it has been denatured, and are orientated so that DNA synthesis by the polymerase proceeds through the region between the two primers. The

extension reactions create two doublestranded target regions, each of which can again be denatured ready for a second cycle of hybridization and extension. The third cycle produces two doublestranded molecules that comprise precisely the target region in double-stranded form. By repeated cycles of heat denaturation, primer hybridization and extension, there follows a rapid exponential accumulation of the specific target fragment of DNA (Primrose et. al., 1994). The net result of a PCR is that by the end of n cycles, the reaction contains a theoretical maximum of 2^n double-stranded DNA molecules that are copies of the DNA sequence between the primers (Watson, 1992).

The heat-stable enzyme commonly used is derived from a thermophilic Gram-negative eubacterium, *Thermus aquaticus* (Glazer, 1995). One of these early isolates *Taq* DNA polymerase and its derivatives have a 5′ to 3′ polymerization depended exonuclease activity. For nucleotide incorporation, the enzyme works best at 75-80^{o}C, depending on the target sequence; its polymerase activity is reduced by a factor of 2 at 60^{o}C and by a factor of 10 at 37^{o}C (Sambrook and Russell, 2001).

While the PCR is simple in concept, practically there are a large number of variables which can influence the outcome of the reaction. The template concentration, primers concentrations may effect the reaction.

2.4.1.1 Gene Splicing by Overlap Extension (SOE) Method

Engineering of recombinant DNA molecules mostly depends on standard methods like cutting with proper restriction enzymes and ligating to rejoin the DNA fragments and requires specific nucleotide sequences to be recognized by restriction enzymes. Gene splicing by overlap extension method basing on the methodology of recombination of fragments from the genes in separate polymerase chain reaction was firstly reported by Horton et. al., (1989) and Ho et. al., (1989). This method provides to join of DNA fragments without any need for the restriction enzymes and is especially very useful where precise recombination in frame as in the case of creating fusion proteins is required. Apart from recombination, specific alterations in the sequences can be performed which allows site directed mutagenesis to be performed simultaneously.

The basic scheme of gene splicing by overlap extension is illustrated in Figure 2.11. First, two PCR products are made in separate reactions; primers a and b produce product AB from gene I and primers c and d are used to amplify fragment CD from gene II. DNA segments are depicted as paired anti-parallel strands. Primers b and c have had sequences added to their 5' ends so that the right end of AB matches the sequence at the left end of the CD. When these products are mixed in an SOE reaction, the top strand of AB overlaps with the bottom strand of CD, their 3' ends being oriented toward each other. This allows them to act as primers on one another to make a recombinant product. The other strands, which point in the wrong directions, do not form product and are not necessary to the reaction (Horton et. al, 1993).

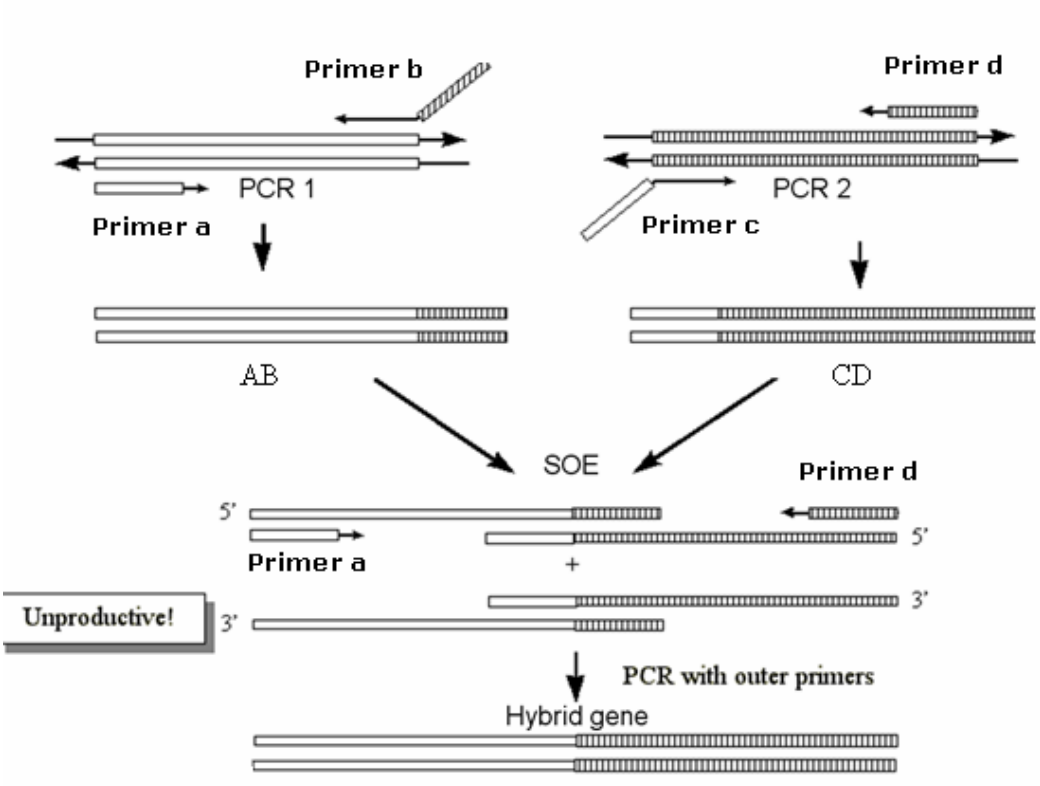


Figure 2.11 Schematic diagram of SOE method (Horton et. al, 1993).

2.4.2 Restriction Enzyme Digestion

These enzymes serve the bacteria in which they occur as protection from foreign DNA. The enzymes cut foreign DNA from outside the helix, but do not digest the host DNA due to characteristic methylation patterns on the DNA (Scragg, 1988).

Restriction enzymes are of several types; the most useful for cloning, the Type II restriction enzymes recognize specific sequences, usually 4-8 bp in length, and cut DNA molecules within these sequences (Kirk and Othmer, 1994). For example the restriction enzyme called *EcoR*I recognizes the double-stranded six-nucleotide sequence of GAATTC and cleavages each strand between the G and A residues. Cleavage of the DNA at this site produces complimentary single-stranded tails called as sticky ends which can later anneal with complementary single-stranded tails on other DNA fragments. Some restriction enzymes cut straight across both chains to form blunt ends.

The restriction enzymes used in this study and their recognition sequences are listed in Table 2.3.

Table 2.3 Recognition sites and cleavage points of restriction enzymes

Enzyme	Target site
*EcoR*I	5'-G^AATTCC-3'
*Sac*I	5'-GAGCT^C-3'
XbaI	5'-T^CTAGA-3'
*Pst*I	5'-CTGCA^G-3'

The DNA fragment of interest cut by a restriction enzyme leads to single stranded tails, sticky ends, which have a tendency to anneal with the complementary strand present in the reaction mixture. The addition of vector DNA cut open by the same restriction enzyme results in the annealing of the foreign DNA to the complementary ends of the cut vector.

2.5 Ligation Reaction

The phosphodiester bonds missing between the attached strands is covalently bond by DNA ligase. T4 DNA ligase seals single-stranded nicks between adjacent nucleotides in a duplex DNA chain (Olivera *et al.* 1968, Gumport and Lehman 1971). The T4 enzyme requires ATP, while the *E. coli* enzyme requires NAD^+. The cofactor is split and forms an enzyme– AMP complex. The complex binds to the nick, which must expose a 5′ phosphate and 3′ OH group, and makes a covalent bond in the phosphodiester chain. When termini created by a restriction endonuclease that creates cohesive ends associate, the joint has nicks a few base pairs apart in opposite strands. DNA ligase can then repair these nicks to form an intact duplex. This reaction, performed *in vitro* (Primrose et. al., 1994).

The ligation reaction is the rate limiting step in genetic engineering techniques since this reaction requires the cohesive ends of foreign DNA and open plasmid DNA to attach in correct orientation and anneal while preventing the relegation of opened vector DNA. Therefore, optimum ligation reaction conditions should be determined by both paying attention to foreign DNA and plasmid DNA concentrations. The insert:vector ratio is also very important parameter in ligation reaction.

2.6 Transformation

After ligation, the mixture bearing the desired vector-donor combinations is then moved into the recipient or host cell. In most cases this is done by transformation (Schuler and Kargı, 2002). A very simple, moderately efficient transformation procedure for use with *E. coli* involves resuspending log-phase cells in ice-cold 0.1 M calcium chloride at about 10^{10} cells ml^{-1} and keeping them on ice for about 30 min. Plasmid DNA (0. 1 µg) is then added to a small aliquot (0.2 ml) of these now *competent* cells, and the incubation on ice continued for a further 30 min, followed by a heat shock of 90 s at 42°C. The cells are then usually transferred to nutrient medium and incubated for some time (30 min to 1 h) to allow phenotypic properties conferred by the plasmid to be expressed, e.g. antibiotic resistance commonly used as a selectable marker

for plasmid containing cells. Finally the cells are plated out on selective medium (Primrose et. al., 1994).

The pPICZαA vector do not have Blue-White selection property. Therefore the selection of the true transformants must be done by isolating the plasmid of all the colonies.

2.7 SDS-Polyacrylamide Gel Electrophoresis of Proteins

Electrophoretic separation of proteins is most commonly performed in *polyacrylamide gels.* When a mixture of proteins is applied to a gel and an electric current is applied, smaller proteins migrate faster through the gel than do larger proteins. Gels are cast between a pair of glass plates by polymerizing a solution of acrylamide monomers into polyacrylamide chains and simultaneously cross-linking the chains into a semisolid matrix. The pore size of a gel can be varied by adjusting the concentrations of polyacrylamide and the cross-linking reagent. The rate at which a protein moves through a gel is influenced by the gel's pore size and the strength of the electric field. By suitable adjustment of these parameters, proteins of widely varying sizes can be separated. In the most powerful technique for resolving protein mixtures, proteins are exposed to the ionic detergent SDS (sodium dodecylsulfate) before and during gel electrophoresis (Figure 2.12). SDS denatures proteins, causing multimeric proteins to dissociate into their subunits, and all polypeptide chains are forced into extended conformations with similar charge:mass ratios. SDS treatment thus eliminates the effect of differences in shape, and so chain length, which corresponds to mass, is the sole determinant of the migration rate of proteins in SDS-polyacrylamide electrophoresis. Even chains that differ in molecular weight by less than 10 percent can be separated by this technique (Lodish et. al., 2003).

SDS-PAGE analyses are used to detect the production of the proteins by microorganisms.

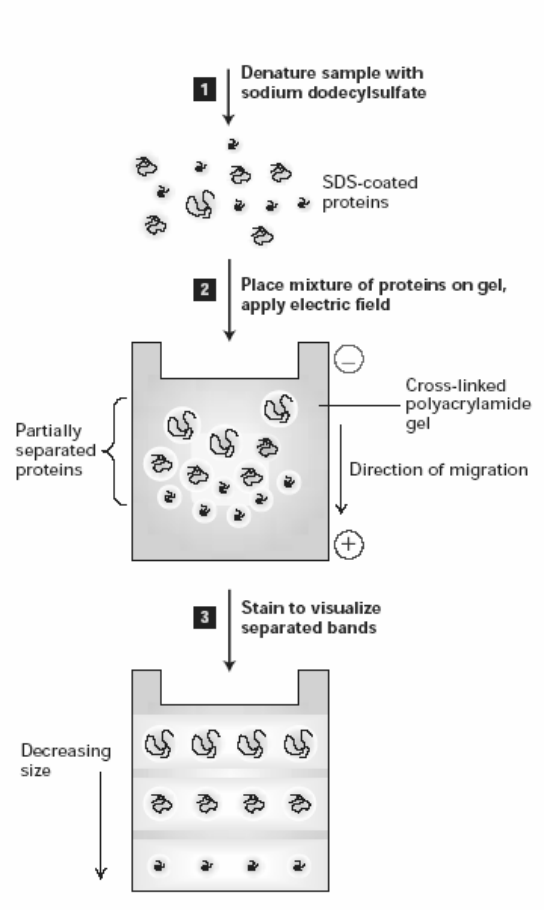


Figure 2.12 Representative scheme of SDS-PAGE (Lodish et. al., 2003)

CHAPTER 3

MATERIALS AND METHODS

3.1 Chemicals

All chemicals were analytical grade, and obtained from Sigma Ltd., Difco Laboratories, Fluka Ltd. and Merck Ltd.

3.2 The Microorganism and Plasmids

Bacillus subtilis and *Pichia pastoris* were used to produce benzaldehyde lyase enzyme extracellularly. *E.coli* BL21 (DE3) pLysS carrying *pRSETA::bal* microorganism was used as the control producer. These three microorganism experiments were given in following sections.

The microorganisms are stored in the microbanks (PRO-LAB), by inoculating young colonial growth into cyropreservative fluid present in the vial. After providing the adsorption of microorganisms into the porous beads, excess cryopreservative was aspirated and inoculated cyrovial was stored at -55^{o}C.

3.3 Procedure for Benzaldehyde Lyase Production in Recombinant *Escherichia coli* strain

Benzaldehyde lyase (BAL, EC 4.1.2.38) was produced from *E.coli* BL21 (DE3) pLysS carrying *pRSETA::bal* plasmid (Kaya, 2006). It was first inoculated into solid medium, then transferred into precultivation medium and finally inoculated into the

production medium. The selective antibiotics were added to the mediums. The final concentrations of antibiotics for recombinant *E.coli* BL21 (DE3) pLysS cells are given in Table 3.1. The details were given in the following sections.

Table 3.1 The final concentration of antibiotics in the mediums for recombinant *E.coli* BL21 (DE3) pLysS cells

Antibiotics	Concentration, kg m^{-3}
Chloramphenicol	0.035
Ampicillin	0.100

3.3.1 Solid Medium for *Escherichia coli* strain

Luria-Bertani (LB) agar was used as the solid medium. Microorganisms were received as very few drops from the microbank stored at -55°C and inoculated onto the LB agar slants in sterile conditions, and incubated at 37 ^{0}C for 10 hours. According to the antibiotic resistance ability of the recombinant microorganism, ampicillin and chloramphenicol were added into the solid medium Table 3.2; the amounts of antibiotics are given in Table 3.1.

Table 3.2 Composition of the reference solid medium for *E.coli* BL21 carrying *pRSETA::bal* plasmid

Compound	**Concentration, kg m^{-3}**
Soytryptone	10.0
Yeast Extract	5.0
NaCl	10.0
Agar	15.0
Ampicilin	0.100
Chloramphenicol	0.035

3.3.2 Precultivation Medium for *Escherichia coli* strain

Microorganisms grown in the solid medium were inoculated into precultivation medium and incubated at 37ºC and N=200 min^{-1} for 14 h. Experiments were conducted in agitation and heating rate controlled orbital shakers, using air-filtered Erlenmeyer flasks 150 ml in size that had working volume capacities of 33 ml. LB medium was used as the precultivation medium for biomass production, and its constituents is given in Table 3.3. According to the antibiotic resistance ability of the recombinant microorganism, ampicillin and chloramphenicol were added in to solid medium the amounts of antibiotics are given in Table 3.1.

Table 3.3 Composition of the reference precultivation medium for *E.coli* BL21

Compound	Concentration, kg m-3
Soytryptone	10.0
Yeast Extract	5.0
NaCl	10.0
Ampicilin	0.100
Chloramphenicol	0.035

3.3.3 Production Medium for *Escherichia coli* strain

After incubation in the precultivation medium for 12 hours, microorganisms were inoculated into the production medium, contained in 150 ml volume bioreactor, with 1/10 inoculation ratio. In agitation and heating rate controlled orbital shakers, incubation took place at T=37 ^{0}C and N= 200 min^{-1} for 12 hours. At the 4th hour of the production period, isopropyl β-D-thiogalactopyranoside (IPTG) to obtain a final concentration of 1 mM was added. According to the antibiotic resistance ability of the recombinant microorganism, ampicillin and chloramphenicol were added in to solid medium the amounts of antibiotics are given in Table 3.1. The contents of the production medium are shown in Table 3.4.

Table 3.4 Composition of the reference BAL production medium for *E.coli* BL21

Compound	Concentration, kg m^{-3}
Glucose	20.0
Na_2HPO_4	6.7
KH_2PO_4	3.1
NaCl	0.5
$MgSO_4.7H_2O$	0.5
$(NH_4)_2HPO_4$	11.8
Ampicilin	0.100
Chloramphenicol	0.035

3.4 Procedure for Benzaldehyde Lyase Production in Recombinant *Bacillus sp.*

Procedure and the solid medium for *Bacillus sp.* were the same with *E.coli* strain. The microorganisms used in *Bacillus sp.* cloning studies were given in Table 3.5.

Table 3.5 Strains and plasmids used in *Bacillus sp.* cloning studies.

Genus	Species	Strain	Genotype/ plasmid	Source
Escherichia	*coli*	JM109	Wild type	Yanish-Peron et. al (1985)
Escherichia	*coli*	XL1-Blue	pUC19	Yanish-Peron et. al. (1985)
Escherichia	*coli*	pUC19::*hybrid*	pUC19::*hybrid*	This study
Bacillus	*subtilis*	1A751	apr-, npr-	BGSC
Bacillus	*licheniformis*	Wild type	Wild type	DSM1969
Bacillus	*licheniformis*	pMK4::*hybrid*	pMK4::*hybrid*	This study
Escherichia	*coli*	pMK4::*hybrid*	pMK4::*hybrid*	This study

Antibiotics used for recombinant *E.coli* strain and recombinant *Bacillus sp*. strains are given in Table 3.6.

Table 3.6 The final concentration of antibiotics in the mediums for recombinant *E.coli* strain and recombinant *Bacillus sp.*

Antibiotic	Concentration, kg m^{-3}
Ampicillin for *E.coli* pMK4::*hybrid*	0.100
Chloramphenicol for *Bacillus licheniformis* pMK4::*hybrid*	0.035

3.4.1 The Precultivation Medium for *Bacillus sp.*

The recombinant *Bacillus* species, grown in the solid medium, were inoculated into precultivation medium and incubated at 37ºC and N=200 min^{-1} for 12 h in agitation and heating rate controlled orbital shakers (B.Braun, Certomat BS-T) using air-filtered Erlenmeyer flasks 150 ml in size that had working volume capacities of 33 ml.

Table 3.7 The composition of the precultivation medium of *Bacillus* species.

Compound	Concentration, kg m^{-3}
Soytryptone	15.0
Peptone	5.0
$MnSO_4.2H_2O$	1.0×10^{-2}
Na_2HPO_4	0.25
$CaCl_2$	0.10
Chloramphenicol	0.035

3.4.2 The Production Medium for *Bacillus sp.*

Recombinant *Bacillus* species were inoculated in production medium whose compositions were given in Table 3.8 (Çalık et. al., 2003-a).

Table 3.8 The composition of the production medium for *Bacillus* species.

Macronutrients	Concentration, kg m^{-3}
Glucose	10.0
$(NH_4)_2HPO_4$	4.71
KH_2PO_4	2.0
Na_2HPO_4	4.3×10^{-2}
NaH_2PO_4	5.63
$MgSO_2$	1.0×10^{-5}
Chloramphenicol	0.035

3.5 The Microorganism and Plasmids for *P.pastoris* studies

For extracellular production of benzaldehyde lyase, microorganisms and plasmids used were listed in Table 3.9.

Table 3.9 Strains and plasmids used for *P.pastoris* experiments.

Genus	**Species**	**Strain**	**Genotype/plasmid**	**Source**
Escherichia	*coli*	JM109	Wild type	Yanish-Peron et. al (1985)
Escherichia	*coli*	TOP10-pPICZαA	pPICZαA	Invitrogen (USA)
Escherichia	*coli*	pPICZαA*::bal*	pPICZαA::*bal*	This study
Pichia	*pastoris*	X-33	wild type	Invitrogen (USA)
Pichia	*pastoris*	IBMEG-Bal1	AOX1::pPICZαA-bal	This study

3.5.1 The Solid Medium

The recombinant *E. coli* JM109 and TOP10-pPICZαA strains stored on agar slants (Table 3.10) at 4ºC or stored on microbanks at -55°C, were inoculated onto the freshly prepared agar slants under sterile conditions, and were incubated at 37ºC overnight. Recombinant *Pichia pastoris* strains were also stored both on agar slants (Table 3.11) and stored at -55°C but were incubated at 30^0C for 48-72 hours. According to the antibiotic resistance ability of the microorganisms, antibiotic was added to the agars after steam sterilization at 121^oC for 20 minutes proper. Amounts of the antibiotics added to the medium are stated in Table 3.12.

Table 3.10 The composition of the solid medium (LSLB) for *E.coli* strains.

Compound	Concentration, kg m^{-3}
Yeast extract	5.0
Soytryptone	10.0
NaCl	5.0
Agar	15.0
pH	7.0

Table 3.11 The composition of the solid medium for *P.pastoris* strains.

Compound	Concentration, kg m^{-3}
Yeast extract	10.0
Peptone	20.0
Glucose	20.0
Agar	20.0
pH	7.0

Table 3.12 The final concentration of antibiotics in the mediums.

Antibiotic	Concentration, kg m^{-3}
Zeocin for *E.coli* pPICZαA*::bal*	0.025
Zeocin for *P.pastoris* IBMEG-Bal1	0.100
Chloramphenicol for *P.pastoris* IBMEG-Bal1	0.035

3.5.2 The Precultivation Medium

The recombinant *P.pastoris* strain, grown in the solid medium, was inoculated into precultivation medium and incubated at 30ºC and N=225 min^{-1} for 24 h in agitation and heating rate controlled orbital shakers (B.Braun, Certomat BS-1) using air-filtered Erlenmeyer flasks 150 ml in size that had working volume capacities of 10 ml. The composition of precultivation medium was given in Table 3.13. The selective antibiotics, zeocin or chloramphenicol, were added to the precultivation medium in amounts stated in Table 3.12 after sterilization.

Table 3.13 The composition of the precultivation medium of *P.pastoris* strain (BMGY).

Compound	Concentration, kg m^{-3}
Yeast extract	10.0
Peptone	20.0
Potassium phosphate buffer pH 6.0	0.1 M
YNB	13.4
Biotin	4×10^{-5}
Glycerol	10.0

3.5.3 The Production Medium

The recombinant *P.pastoris* strain inoculated in precultivation medium was grown for 24 h. The cells were harvested by centrifugation at 4000 rpm, 10 min^{-1} at $+4^0C$ and resuspended in BMMY production medium. The recombinant cells incubated at 30^0C and N=225 min^{-1} for 72 h using air-filtered, baffled Erlenmeyer flasks 250 ml in size that had working volume capacities of 50 ml. Every 24 h, 0.5% methanol was added to the production medium. The composition of production medium was given in Table 3.14. The selective antibiotics, zeocin or chloramphenicol, were added to the production medium in amounts stated in Table 3.12 after sterilization.

Table 3.14 The composition of the production medium of *P.pastoris* strain (BMMY)

Compound	Concentration, kg m^{-3}
Yeast extract	10.0
Peptone	20.0
Potassium phosphate buffer pH 6.0	0.1 M
YNB	13.4
Biotin	4×10^{-5}
Methanol	10.0

3.6 Analysis

Throughout the bioprocesses, samples were taken at characteristic cultivation times. After determining the cell concentration, the medium was centrifuged at 4000 min^{-1} for 10 min at 4^oC to precipitate the cells. In recombinant *P.pastoris* strains, supernatant was used to determine benzaldehyde lyase activity, SDS-PAGE and TLC analyses.

3.6.1 Cell Concentration

Cell concentrations based on dry weights were measured with a UV-Vis spectrophotometer (Thermo Spectronic, Heλiosα).

3.6.2 Benzaldehyde Lyase Activity

Benzaldehyde lyase activity was determined by measuring the conversion of benzoin into benzaldehyde. Samples from the culture broth were harvested by centrifugation (Sigma 1-15) at 4000 min^{-1} for 10 min at 4°C. After removing the cell debris, supernatant was directly added to the substrate solution. Fresh substrate solutions were prepared daily by mixing 0.5 volume activity buffer [40 mM Tris-HCl (pH=8.0), 0.02 mM ThDP, 0.2 mM $MgCl_2$ buffer], with 0.35 volume stock benzoin (0.1mM) and 0.15 volume 15%PEG solution with a final concentrations of 0.035 mM benzoin, 20 mM Tris-HCl (pH=8.0), 0.01 mM ThDP, 0.1 mM $MgCl_2$, 7.5% PEG and incubated at 37°C. Stock benzoin solution (0.1mM) was prepared in 15% PEG solution and used after overnight incubation at room temperature. The conversion reaction was carried out at 37°C.

Reaction was started by the addition of 10 µl of samples to 3 cm^3 of substrate solution and further incubated for 10 seconds. Because of the light absorbing behavior of yeast extract (at 250 nm) used in the production medium of *P.pastoris*, and the accuracy of the detection same volumes of samples were added to the blank solution. Enzymatic activity was monitored spectrophotometrically at 250 nm by following the change in absorbance in ten seconds. The substrate benzoin and the product benzaldehyde have similar absorbance maxima. However, since two molecules of benzaldehyde are produced per each molecule of benzoin cleaved, the following formula was used to relate absorbance change to product formed (Gonzalez and Vicuna, 1989):

$$\text{Nanomoles of product formed} = [2(A_f - A_i)/(\varepsilon_s - 2\varepsilon_p)]*10^6$$

With ε_s and ε_p being molar extinction coefficients of the substrate and the product, respectively. $A_f - A_i$ is the change in optical density during the reaction time, measured at 250 nm. One unit of enzymatic activity was defined as the

amount of enzyme that catalyzes the cleavage of benzoin into one nanomoles of benzaldehyde at 37^{o}C and pH 8.0 in one second (Çalık et. al., 2004).

Benzaldehyde lyase activity was also determined by the conversion of benzaldehyde into (R)-benzoin. Samples from the culture broth were harvested by centrifugation (Z323K, Hermle) at 4000 min^{-1} for 10 min at 4^{o}C. After removing the cell debris, chemical compounds were directly added to the supernatant. The final concentration of 1 mM ThDP, 0.5 mM $MgSO_4$, 30% DMSO and incubated at 37^{o}C for 24 h (Maria et. al., 2005). Enzymatic activity was followed by TLC analysis.

3.7 Ultrafiltration

Concentration and desalting of production medium (supernatant) was achieved by ultrafiltration using 50ml-400 ml stirred cells (Amicon), holding a flat sheet membrane with a diameter of 44.5 mm for 50 ml cells, and 76 mm for 400 ml cells. Effective membrane areas were 13.4 cm^2 and 41.8 cm^2, respectively, for 44.5 and 76 mm diameter regenerated cellulose membranes (Amicom, Millipore), with molecular weight cut of (MWCO) 10 000 Da and 30 000 Da. The system was operated at cold room (2-8^{0}C) temperature with an applied pressure of P=3 bars using nitrogen gas. Applied stirring rate was N= 400 rpm. The process was carried until at least 5 fold concentrated solution was obtained.

3.8 Thin Layer Chromatographic Analysis

To determine the benzoin condensation reaction TLC analyses were done. TLC analysis was carried out on silica gel F_{254} gel plates (Merck). Samples were loaded on the plates and developed in Ethyl acetate: Chloroform [3:1 (vol/vol)]. When the mobile phase had moved an appropriate distance, stationary phase was removed, the mobile phase is rapidly dried, and the zones are detected under ultraviolet (UV) light because both benzaldehyde and benzoin are naturally absorb UV light. The retention factor, or R_f, is defined as the distance migrated by the solute divided by the distance migrated by the solvent (Sherma and Fried 2003).

$$Rf = \frac{\text{Distance moved by the solute}}{\text{Distance moved by the mobile phase}}$$

When benzoin biotransformation reaction detected at TLC results, samples were further analyzed by GC-MS.

3.9 Gas Chromatography-Mass Spectrometry (GC-MS) Analysis

GC-MS analysis was used to detect the chemical compounds in reaction media (section 3.6). Analysis was carried out using ThermoQuest Finnigan multi Mass (EI, 70eV) (San Jose, CA, USA). Samples were purified and concentrated with Ethyl acetate extraction followed by a silica column chromatography. The amount of time that a compound is retained in the GC column is known as the retention time. The retention time can aid in differentiating some compounds but it is not a reliable factor all alone. Therefore, by comparing the retention time and mass spectrum, which was identified with Mass Spectrometry, of the chemical compounds molecules were identified.

Table 3.15 GC-MS Oven parameters

Time	Temperature	Ramp rate
2 min	60^{o}C	-
36 min	-	5^{o}C min^{-1}
12 min	240^{o}C	-

Once the sample solution is introduced into the GC inlet it is vaporized immediately because of the temperature and swept onto the column by Helium. The used flow rate was Splitless flow, Constant flow (1ml min^{-1}).

3.10 Fast Performance Liquid Chromatography (FPLC)

To determine molecular weights of the proteins produced extracellularly by the recombinant *P.pastoris* strain FPLC analysis was done. The recombinant *P.pastoris* cells inoculated in the reference production medium were harvested at 24^{th} h by centrifugation and supernatant was used for further analysis. Samples were freeze-dried (Hetosicc Freeze Dryer Type CD25, Heto, Birkerød, Denmark) for 7 h at -50^0C under 0.5 mbar pressure then sample was stored at -20^0C. Thereafter, sample was solved in 1 ml mobile phase (Appendix A) and 25x concentration was achieved. The sample then analyzed in Fast Performance Liquid Chromatography (Prostar Varian) with Hiload 16/60 superdex 200pg column (16/600mm)(Amersham Biosciences). Flow-rate was 0.5 ml min^{-1} and detection was carried out with an UV detector at 280nm.The process was done at $+4^0$C.The mobile phase used was PBS buffer.

3.11 Genetic Engineering Techniques

3.11.1 Materials

3.11.1.1 Enzymes, Kits, and Molecular Size Markers

Taq DNA polymerase, *Pfu* DNA polymerase, Ribonuclease A (DNase and protease free), T4 DNA ligase, dNTP mixture, and restriction enzymes (EcoRI and XbaI) and their buffers were purchased from MBI Fermentas.

QIAGEN Plasmid Purification Kit, QIAguick PCR purification Kit and QIAexpress Ni-NTA Spin Columns were obtained from QIAGEN Inc. Gene Elution Kit was purchased from GeneMark Molecular Biology Tools.

Lambda DNA/HindIII Marker and 6X Loading Dye were from MBI Fermentas.

3.11.1.2 Buffers and Solutions

All buffers and stock solutions listed in Appendix A were prepared with distilled water (dH_2O). The sterilization of solutions was performed either by

autoclaving at 121°C for 20 min or by filter sterilization through 0.20 µm filters (Sartorius).

3.11.2 Determination of DNA Concentration

The concentration of DNA fragments after restriction digestion, PCR amplification or concentration of isolated plasmids after digestion or purification; or DNA molecules treated with any manipulation were analyzed by gel electrophoresis with 0.8-1.7% (w/v) agarose gels according to the weight of the DNA fragment and 1XTBE buffer (for DNA fragments bigger than 1500 bp 0.8% (w/v) agarose gels were used). DNA samples of 10-20µl, mixed with 1/5 volume of 6X loading dye (MBI Fermentas) were applied to the gel which was supplemented by ethidium bromide (Sigma-10 mg/ml) with a final concentration of 0.8µL/ml. At the end of the electrophoresis, bands were visualized with a UV transilluminator and gel photographs were taken using gel imaging and documentation system (UVP BioImaging System, and Hamamatsu Digital CCD Camera).

The molecular weights (Appendix G) and the concentrations of the DNA fragments analyzed were determined by referring to Labworks image acquisition and analysis software (UVP BioImaging System)

3.11.3 PCR Amplification of Target Genes

Primers were designed in accordance with the sequences of *bal* gene (Appendix B). Nucleotide sequences of *bal* gene (Accession no: AX349268) was obtained from National Institutes of Health (USA), National Library of Medicine, National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/ BLAST, Last accessed date June 2006). Restriction enzyme recognition sites were determined (Appendix D) by the help of Restriction Mapper web-page of USA Molecular Biology Resources (http://www.restrictionmapper.org, Last accessed date June 2006). After determination of proper restriction enzyme recognition sites for *bal* gene, primers were designed for the cloning into appropriate plasmids selected according to the following rules:

• Primers should be 17 to 30 nucleotides in length.

• A GC content of about 50% is ideal. For primers with a low GC content, it is desirable to choose a long primer so as to avoid a low melting temperature.

• Sequences with long runs (i.e. more than three or four) of a single nucleotide should be avoided.

• Primers with significant secondary structure are undesirable.

• There should be no complementarity between the two primers.
(Primrose et. al., 1994)

The possibility of dimer formation and self-complimentarity of primers and melting temperature, ΔG, ΔH and ΔS values were checked with a computer program (NAR) and they were illustrated in Appendix C. Designed primers were synthesized in Thermo Hybaid GmbH (Germany) laboratories. The sequences and abbreviations of designed primers were demonstrated in Table 3.16 and Table 3.17.

Table 3.16 Primers designed for *Bacillus sp.* cloning experiments.

Name	Sequence	Targeted Gene Fragment
pMK4-pre-subC Forward Primer	5' GCGCAATTCGCGGTCTATTCATACTTTCG 3'	*pre-sub*C and *hybrid*
pDG- pre-subC Forward Primer	5′ AAGGAGGAAGCAGGTATGATGAGGAAAAAGAGTT 3′	*pre-sub*C and *hybrid*
pDG-bal Reverse Primer	5′ GACACGCACGAGGTTTAGTGATGGTGATGGTG 3′	*bal* and *hybrid*
SOE Primer 1	5' CCTGTAATCATCGCCATAGCAGAAGCGGAATCG 3′	*pre-sub*C
SOE Primer 2	5′ CGATTCCGCTTCTGCTATGGCGATGATTACAGGC 3′	*bal*
bal Reverse Primer	5′ GCGAGCTCGCTTAGTGATGGTGATGGTGATG 3′	*pre-sub*C and *hybrid*

Table 3.17 Primers designed for amplification of desired gene fragments for *P.pastoris* experiments.

Name	Sequence	Targeted Gene Fragment
EcoRI-bal Forward Primer	5' GGGAATTCATGGCGATGATTACAGGC 3'	*bal*
XbaI-bal Reverse Primer	5′GGTCTAGATTAGTGATGGTGATGGTGATG3′	*bal*
AOX1 Forward Primer	5′ GACTGGTTCCAATTGACAAGC 3′	*AOX1*
AOX1 Reverse Primer	5' GCAAATGGCATTCTGACATCC 3′	*AOX1*

PCR amplification was carried out with thermal cycling machine (Techgene, Flexigene). The cycles and the contents of the reaction mixture of 50µl final volume were arranged as follows:

Table 3.18 PCR process parameters

1 cycle	T_1 = 94°C,	3-4 min
30 cycle	T_1 = 94°C,	1 min
	T_2 = 55-60°C,	1 min
	T_3 = 72°C,	0.25-2.5 min
1 cycle	T_3 = 72°C,	5-10 min
	T_4 = 4°C,	5 min

Table 3.19 Components of reaction mixture of PCR

10XPCR Buffer (with Mg^{++})	5µl
dNTPs (1mM)	10µl
Forward Primer (10µM)	1µl
Reverse Primer (10µM)	1µl
Template DNA (0.01-1µg)	1-5 µl
dH2O	up to 49 µl
DNA polymerase	2.5 U

3.11.4 Purification of PCR products

The purification of PCR products were performed by using QIAquick PCR Purification Kit according to manufacturer's recommendations. 1 volume of PCR reaction sample was mixed with 5 volumes of PB buffer and placed to QIAquick spin column. After centrifugation, column was washed with PE twice and DNA molecules were eluted in proper amount of water.

3.11.5 Ligation Reaction

PCR amplified gene was cloned into suitable expression vector from the sticky ends occurred after restriction digestion with altering the gene/vector molar ratio between 3 and 5.

The amount of insert DNA to be added to the reaction mixture was calculated such that insert:vector ratio of 1:3 was achieved, as given in the equation.

$$100 \text{ ng vector} \times \frac{\text{Size of insert (bp)}}{\text{Size of vector(bp)}} \times \frac{3}{1} = \text{amount of insert (ng)}$$

The ligation reactions were performed by incubating the reaction mixture of 10µl final volume at 16°C for 16h. The composition of the ligation reaction mixture was arranged as follows:

Table 3.20 Ligation reaction components

10X ligation buffer	1µl
Insert DNA	45-75 ng
Double digested vector DNA	100 ng
T4 DNA ligase	1µl
Sterile dH_2O	Up to 10 µl

3.11.6 Restriction Enzyme Digestion Reaction

Restriction digestion of the genes of interests were performed by incubating DNA fragments with proper restriction enzymes (REs) and specified buffers, of 20µl final volume at 37°C for overnight. The composition of the restriction digestion reaction mixtures was arranged as given in Tables 3.21, 3.22, 3.23, 3.24. After restriction digestion, the reaction was ended by incubating the mixture at 65°C for the enzymes.

Table 3.21 Components of reaction mixture of restriction digestion with *EcoR*I RE.

Components	Amounts
DNA fragment	~ 45-200 ng
EcoRI RE (10U/ µl)	1 µl
10X Buffer *EcoRI*	2 µl
dH_2O	up to 19 µl

Table 3.22 Components of reaction mixture of restriction digestion with *Xba*I RE.

Components	Amounts
DNA fragment	~ 45-200 ng
XbaI RE (10U/ µl)	1 µl
10X Buffer Tango	2 µl
dH_2O	up to 19 µl

Table 3.23 Components of reaction mixture of restriction digestion with P*st*I RE.

Components	Amounts
DNA fragment	~ 45-200 ng
PstI RE (10U/ µl)	1 µl
10X Buffer Tango	2 µl
dH_2O	up to 19 µl

Table 3.24 Components of reaction mixture of restriction digestion with *BamHI* RE.

Components	Amounts
DNA fragment	~ 45-200 ng
*BamH*I RE (10U/ µl)	2.5 µl
Buffer *BamH*I	2 µl
dH_2O	up to 19 µl

Table 3.25 Components of reaction mixture of restriction digestion with *SacI* RE.

Components	Amounts
DNA fragment	~ 45-200 ng
*Sac*I RE (10U/ µl)	2 µl
1X Buffer *Sac*I^{+}	2 µl
dH_2O	up to 19 µl

3.11.7 DNA Sequencing

The DNA sequencing is performed by automatic DNA sequencers (Microsynth GmbH, Switzerland) by using the primers designed to control the insertion of target genes. And the primer for AOX1 gene primers supplied by Invitrogen (USA). For *Bacillus* experiments insert was controlled by using both *pre-subC* and *bal* genes primers.

3.11.8 Transformation of Plasmid DNA by $CaCl_2$ Method to *E.coli*

1. Incubate *Escherichia coli* overnight in LB-solid medium at 37^{o}C,
2. Pick a single colony from a selective plate and inoculate a starter culture of 5 ml LB medium. Grow for 12h (overnight) at 37^{o}C, with vigorous shaking (~200 rpm)
3. Transfer 1ml sample from the precultivation medium to an 100 ml- LB medium and incubate at 37^{o}C and 250 min^{-1} for 3.5 hours,
4. Transfer 10 ml of broth into 30 ml sterile polypropylene tubes; and place on ice for 10 minutes,
5. Separate the microorganisms by centrifugation at 4000 min^{-1}, 4^{o}C for 10 minutes,
6. Separate the supernatant; let the cells dry on a paper tissue for 1 minute,
7. Add 1 ml of 0.1 M $CaCl_2$ solution onto the cells and make a complete solution by vortex, set on ice for 10 minutes,
8. Precipitate the microorganisms by centrifugation at 4000 min^{-1}, 4^{o}C for 10 minutes,

9. Separate the supernatant; let the cells dry on a paper tissue for 1 minute,
10. Add 200 µl of 0.1M $CaCl_2$ solution onto the cells and make a complete solution by vortex, set on ice for 10 minutes,
11. Transfer 200 µl of solution to an eppendorf tube, and add 0.1 µg of plasmid DNA to this solution. Incubate on ice for 30 minutes,
12. Apply heat-shock to the solution at 42°C for 90 seconds and quickly place the tube on ice for 1 minute,
13. Transfer the cell suspension to sterile culture tubes containing 800 µl of LB medium without antibiotics and incubate at 37°C for 45 minutes with shaking at 140 min^{-1} to recover cells,
14. Transfer 250 µl of the cultured cells onto the center of LSLB plate containing the desired antibiotic. Immediately spread the cells over the entire surface of the LB plate using a sterile, bent glass rod.
15. Invert the plates and incubate at 37°C overnight. Selected colonies should be visible in 14-24 hours (Sambrook, 2001).

3.11.9 Isolation of Plasmid DNA

1. Pick a single colony from a selective plate and inoculate a starter culture of 30 ml LB medium. Grow for 12h (overnight) at 37°C, with vigorous shaking (~200 rpm)
2. Pour 1ml of culture into microfuge tube and centrifuge at 13200 min^{-1}, 4°C, for 30 s,
3. Remove the supernatant and add again 1ml of culture and repeat the centrifugation step,
4. Remove the supernatant and take off all fluid by micropipette; place the tube on ice,
5. Resuspend the bacterial pellet in 100 µl of ice-cold alkaline lysis solution I one by vigorous vortexing. Make sure that the bacterial pellet is completely dispersed in alkaline lysis solution I,
6. Add 200 µl of freshly prepared alkaline lysis solution II to each bacterial suspension. Close the tube tightly, and mix the content by inverting the tube gently 5 times and store at room temperature for 5 minutes,

7. Add 150 µl of ice-cold alkaline lysis solution III. Close the tube and disperse alkaline lysis solution III through the viscous bacterial lysate by inverting the tube 5 times. Store the tube on ice for 10 minutes,
8. Centrifuge the bacterial lysate at 13200 min^{-1}, 4°C, for 10 minutes. Transfer the supernatant to a fresh tube.
9. Add 1/10 volumes of NaAc and 2 volumes of EtOH. Mix the solution by inverting and then allow the mixture to stand for at least 10 minutes at -20°C,
10. Collect the precipitated plasmid DNA by centrifugation at 13200 min^{-1}, 4°C, for 10 minutes.
11. Remove the supernatant gently and stand the tube in an inverted position on a paper towel to allow all of the fluid to drain away.
12. Dissolve the plasmid DNA in suitable amount of dH_2O and store the solution at -20°C (Sambrook, 2001).

3.11.10 Purification of Plasmid DNA

Plasmid purification was carried out by using QIAGEN Minipreps Plasmid Purification Kit according to the procedure provided by manufacturer. Bacterial cells incubated overnight were harvested by centrifugation and resuspended in P1 buffer containing RNase A. After cell lysis in P2 buffer, genomic DNA, proteins, cell debris and SDS were precipitated by adding P3 buffer to the mixture and incubated on ice until lyasate become less viscous. Sample was centrifuged to remove the precipitate and supernatant loaded into equilibrated OIAGEN-tip and allowed to enter the resin by gravity. After washing with QC buffer twice, DNA was eluted with QF buffer into clean microcentrifuge tubes. 0.7 volumes isopropanol was added on each sample to precipitate the DNA and centrifuged for 30 min. After washing pellet with 70% ethanol, each pellet was dissolved in proper amount of water.

3.11.11 Chromosomal DNA isolation from *Bacillus* Species

Chromosomal DNA of *B. licheniformis* DSM 1969 was isolated according to the method given by Posprech and Neumann (1995) with some modifications.

1. Pick a single colony from plate and inoculate a starter culture of 30 ml LB medium. Grow for 12h (overnight) at 30°C, with vigorous shaking (~200 rpm),
2. Harvest bacterial cells by centrifugation at 3000g, 4°C, for 10 min after removing the supernatant, resuspend the bacterial pellet in 5 ml SET buffer,
3. Add lyzosim (1mg/ml) and incubate at 37°C for 60 min to lysate the bacterial cell wall,
4. Add 1/10 volume of 10%SDS onto lysate, mix and immediately add proteinase-K (0.5mg/ml) over the mixture. Incubate the mixture at 4°C for 2 h with gently shaking,
5. Add 1/3 volume of 5M NaCl and equal volume of chloroform and incubate at room temperature for 30 min. with gently shaking,
6. Centrifuge the two-phase mixture at 4500g for 15 min to assure the separation of each phase. Then, take the water phase carefully with micropipette,
7. Precipitate chromosomal DNA by adding equal volume of room temperature isopropanol. Mix and centrifuge immediately at 4500g for 15 min. Wash DNA pellet with 70% ethanol and centrifuge at 4500g for 15 min. Carefully decant the supernatant without disturbing the pellet.
8. Redissolve the DNA in a suitable volume of water.

3.11.12 Natural Transformation of Plasmid DNA into *Bacillus* Species

Transformation into *Bacillus* species was done according to the method described by Özçelik, 2003, with some modifications.

1. Pick a single colony from plate and inoculate a starter culture of 5 ml LB medium. Grow for 12h (overnight) at 37°C, with vigorous shaking (~180 rpm),
2. Transfer 250 µl sample from the precultivation medium to an 50 ml LB medium and incubate at 37 °C and 180 min^{-1} for 3 hours,
3. Harvest bacterial cell by centrifugation at 4000 min^{-1}, 4°C for 15 minutes,

4. Wash separated cells with 3ml 1 mM HEPES (pH=7.0) buffer and with cold electrophoration buffer (25% PEG and 0.1M mannitol) twice. Keep cells cold during washing processes.
5. Resuspend washed cells in 1/200 volume of cold electrophoration buffer (~250 µl) and keep on ice for 10 min.
6. Proceed natural transformation by adding 3 µl (~ 100ng) plasmid DNA into 40 µl competent cells.
7. Transfer transformants into sterile culture tubes containing 200 µl of LB medium without antibiotics and incubate at 37°C for 3 h with shaking at 180 min^{-1} to recover cells,
8. Spread the cells over the entire surface of the LB plate containing proper antibiotic using a sterile, bent glass rod,
9. Invert the plates and incubate at 37°C overnight. Selected colonies should be visible in 14-24 hours.

3.11.13 Transfection of *Pichia pastoris*

Transfection of *P. pastoris* was performed using LiCl method according to manufacturer's instructions (Invitrogen, catalog V195-20).

YPD plate was inoculated with *Pichia pastoris* X-33 and incubated for 48 h in 30°C incubator. 5 ml YPD was inoculated with a single colony and grown overnight to saturation in 30°C shaker. 50 ml culture of YPD was inoculated using preculture, to an initial OD_{600} of approximately 0.1 and incubated at 30°C with shaking to an OD_{600} of 0.8 to 1.0 (approximately 10^8 cells/ml; 6-7 h). During this period, the plasmid DNA to be integrated into the genome had to be digested at a single site. Therefore, pPICZαA*::bal* plasmid was digested with *Sac* I at 37°C for overnight in Buffer *Sac* I. Full digestion was verified by running 2 µl sample on agarose gel and then purified. The concentration of the plasmid DNA was adjusted to 0.1-0.2 µg/µl and verified by Agarose gel electrophoresis. When the OD_{600} reached 0.8 - 1.0, the cells were harvested at 4000xg for 5 min, washed with 25 ml of sterile water, and centrifuged at 1500xg for 10 min at room temperature. The cell pellet was resuspended in 1 ml of 100 mM LiCl, transferred to a 1.5 ml microcentrifuge tube, centrifuged at maximum speed for 15 sec. LiCl was removed with a pipet and the

cells were resuspended in 400 µl of 100 mM LiCl. For each transformation, 50 µl of the cell suspension was dispensed into a 1.5 ml microcentrifuge tube, immediately centrifuged at maximum speed for 15 sec and LiCl was removed with a pipet. To each tube for transformation, 240 µl of 50% PEG, 36 µl of 1 M LiCl, 10 µl of 5 mg/ml single-stranded DNA and 5-10 µg plasmid DNA in 50 µl sterile water were added in the order given and vortexed vigorously until the cell pellet is completely mixed. The tube was incubated at 30°C for 30 min without shaking, then heat shocked in a water bath at 42°C for 25 min. The cells were pelleted by centrifugation at 6000 rpm for 15 sec, gently resuspended in 1 ml of YPD and incubated at 30°C with shaking. After 2 h of incubation, 25-100 µl was spread on YPD + Zeocin plates and incubated for 2-3 days at 30°C.

3.11.14 Isolation of Genomic DNA from Yeast

The isolation of genomic DNA from yeast was performed according to the method described by Burke (2000) with slight modifications.

- 10 ml yeast culture was grown to saturation in YPD at 30° C in 50 ml Falcon tube.
- Cells were collected by centrifugation at 5 000 rpm for 6 min.
- The cells were resuspended in 0.5 ml of dH_2O, transferred to a 1.5 ml microfuge tube and collected by centrifugation at 13200 rpm for 2 min.
- The supernatant was decanted and the pellet was vortexed in residual supernatant.
- 200µl yeast lysis solution was added and mixed by inversion to ensure lysis of cells and then 200µl phenol:chloroform:isoamyl alcohol (25:24:1) and 0.3 g acid-washed glass beads were added.
- The tube was wrapped with Parafilm and vortexed for 3-4 min.
- 0.2 ml of TE (pH 8.0) was added and centrifuged at 13200 rpm for 5min in a microfuge.
- The aqueous layer was transferred to a fresh Eppendorf tube.
- 1 ml of 100 % EtOH was added and mixed by inversion.
- The tube was centrifuged at 13200 rpm for 2 min and the supernatant was discarded.

- The pellet was resuspended in 0.4 ml of TE and 3 µl of a 10 mg/ml solution of RNase A.
- The solution was incubated for 10 min at 37 °C and 10 µl of 4 M ammonium acetate plus 1 ml of 100 % EtOH were added and mixed by inversion.
- The DNA was pelleted by centrifugation at 13200 rpm for 5 min in a microfuge and the supernatant was discarded. The pellet was air-dried and resuspended in 50 µl of sterile dH_2O.
- 10 µl was used for each sample to be analyzed by southern blotting. This corresponds to approximately 2-4 µg of genomic DNA.

3.11.15 DNA Extraction from Agarose Gel

DNA fragments analyzed with gel electrophoresis were extracted from the agarose gel by using Gel Elusion Kit (GeneMark). After electrophoresis desired DNA bands were cut and gel slices (up to 350 mg) incubated in 400 µl Binding Solution at 60°C for 15-30 minutes were loaded into spin columns and centrifuged at maximum speed. Columns washed with Washing Solution twice were transferred into new sterilized microcentrifuge tubes and DNA molecules were eluted in proper amounts of water.

3.11.16 Purification of 6xHis-tagged proteins

The purification of 6xHis-tagged Proteins was performed by using QIAexpress Ni-NTA spin columns according to manufacturer's recommendations. 6x-His-tagged benzaldehyde lyase proteins were purified from *P.pastoris* under native conditions. Fermentation media were centrifuged at 4000rpm for 10 min and lysis buffer was added to the supernatant. Then the mixture was loaded into previously equilibrated Ni-NTA spin column. Centrifugation was performed at 700g for 3-4 min to ensure the effective binding of lysate with resin. Ni-NTA spin column was washed with 600 µl wash buffer twice with an open lid to ensure that the centrifugation step is completed after 2 min. Pellet proteins were eluted with 200 µl elution buffer.

3.11.16.2 Purification of His-Tag by BD TALON Resin

1. Thoroughly resuspend the BD TALON Resin.
2. Immediately transfer the required amount of resin suspension to a sterile tube that will accommodate 10–20 times the resin bed volume.
3. Centrifuge at 700 rpm for 2 min to pellet the resin.
4. Remove and discard the supernatant.
5. Add 10 bed volumes of 1X Equilibration/Wash Buffer and mix briefly to pre-equilibrate the resin.
6. Recentrifuge at 700 rpm for 2 min to pellet the resin. Discard the supernatant.
7. Repeat Steps 5 and 6.
8. Add the clarified sample from Section VI.A, B or C to the resin.
9. Gently agitate at room temperature for 20 min on a platform shaker to allow the polyhistidine-tagged protein to bind the resin.
10. Centrifuge at 700 rpm for 5 min.
11. Carefully remove as much supernatant as possible without disturbing the resin pellet.
12. Wash the resin by adding 10–20 bed volumes of 1X Equilibration/Wash Buffer. Gently agitate the suspension at room temperature for 10 min on a platform shaker to promote thorough washing.
13. Centrifuge at 700 rpm for 5 min.
14. Remove and discard the supernatant.
15. Repeat Steps 12–14.
16. Add one bed volume of the 1X Equilibration/Wash Buffer to the resin, and resuspend by vortexing.
17. Transfer the resin to a 2-ml gravity-flow column with an end-cap in place, and allow the resin to settle out of suspension.
18. Remove the end-cap, and allow the buffer to drain until it reaches the top of the resin bed, making sure no air bubbles are trapped in the resin bed.
19. Wash column once with 5 bed volumes of 1X Equilibration/Wash Buffer.
20. [Optional]: If necessary, repeat Step 19 under more stringent conditions using 5–10 mM imidazole in 1X Equilibration/Wash Buffer.
21. Elute the polyhistidine-tagged protein by adding 5 bed volumes of Elution Buffer to the column. Collect the eluate in 500-µl fractions.

3.11.17 SDS Page

Proteins secreted were analyzed with SDS-PAGE according to the method described by Laemmli, et. al. (1970), with some modifications. Gels were stained with silver staining method after electrophoretic run was completed using the procedure of Blum *et al*. (1987). Silver staining method was performed in 6 steps including fixing, washing with 50% ethanol, pretreatment, impregnation, developing and stopping. Preparation of reagents and the procedure of silver staining method were given in Appendix E.

After staining, the gel was photographed. The molecular weights of enzymes were determined by measuring the migration distance and by comparing them with molecular weight markers.

The relative mobility (R_f) of each protein was determined by dividing its migration distance from the top of gel to the center of the protein band by the migration distance of the tracking dye from the top of the gel.

The equation of relative mobility (R_f) was given as:

$$R_f = \frac{\text{Distance migrated by protein}}{\text{Distance migrated by tracking dye}}$$

3.11.17.1 Pouring SDS-polyacrlamide Gels

1. Clean the glasses with ethanol. And assemble the glass plates according to the manufacturer's instructions. To check whether the glasses are properly sealed, pour distilled water between glasses. If water level does not decrease, glasses are properly sealed. Then, pour out the water and let the glasses to dry.
2. In an Erlenmeyer flask, prepare appropriate volume of solutions containing the desired concentration of monomer solution for 12% separating gel, using the values given in Appendix E. Mix the solutions in order shown.

Polymerization will begin as soon as the NNN'N'-Tetramethylethylenediamine (TEMED) and 10% (w/v) ammonium persulfate (APS) have been added.

3. Swirl the mixture rapidly and immediately pour the solution into the gap between the glass plates. Leave sufficient space for the stacking gel. Add some water to overlay the monomer solution and leave the gel in a vertical position until polymerization is completed.
4. After 30 min, pour off the water and dry the area above the separating gel with filter paper before pouring the stacking gel Place a comb in the gel sandwich and tilt it so that the teeth are at a slight (~10°) angle. This will prevent air from being trapped under the comb teeth while the monomer solutions are poured. Allow the gel to polymerize 30-45 minutes.

3.11.17.2 Preparation of Samples and Running the Gel

1. While stacking gel is polymerizing, prepare samples by diluting at least 1:1 with sample buffer and heated at 95°C for 5 minutes.
2. After polymerization is complete (30 min), mount the gel in electrophoresis apparatus and fill the reservoir with running buffer.
3. Load up 20 µl of each sample into the wells and start running with 30 mA. After the dye front has moved into the separating gel increase the applied current. The usual run time is approximately 45 minutes. This electrophoresis cell is for rapid separation and is not recommended for runs over 60 minutes long.

3.11.17.3 Staining SDS-Polyacrylamide Gels with Coomassie Brilliant Blue

1. After running is completed, immerse the gel in 5 volumes of staining solution and place on a slowly rotating platform for 4h at room temperature.
2. Remove the stain and save it for future use. Destain the gel by soaking it in the methanol:acetic acid solution without the dye on a slowly rotating platform for 4-8h, changing the destaining solution 3-4 times. After destining store the gel in dH_2O.

3.12.15.4 Staining SDS-Polyacrylamide Gels with Silver Salts

The gels were silver stained using the procedure of Blum *et* al. (1987).

Table 3.26 Procedure for silver staining

	STEP	SOLUTION	TIME OF TREATMENT	COMMENTS
1	Fixing	Fixer	≥ 1 hr	Overnight incubation is all right
2	Washing	50% Ethanol	3 x 20 min	Should be fresh
3	Pre-treatment	Pretreatment Solution	1 min	Should be fresh
4	Rinse	Distilled water	3 x 20 sec	Time should be exact
5	Impregnate	Silver Nitrate Solution	20 min	
6	Rinse	Distilled water	2 x 20 sec	Time should be exact
7	Developing	Developing Solution	~ 5 min	After a few minutes add some distilled water to proceed the reaction slowly. Time should be determined by observation of color development
8	Wash	Distilled water	2 x 2 min	
9	Stop	Stop Solution	≥ 10 min	The gels can be kept in this solution overnight

CHAPTER 4

RESULTS AND DISCUSSION

Benzaldehyde lyase enzyme catalyzes both cleavage and condensation of only (R)-Benzoin. Therefore it is possible to use this enzyme in preparation of pure (R)-Benzoin. In this study, benzaldehyde lyase gene from *Pseudomonas fluorescens* Biovar I was cloned into *B.subtilis* and *P. pastoris*, with the aim of the extracellular production of the enzyme. Thereafter, the extracellular production of benzaldehyde lyase by recombinant *P. pastoris* was analyzed by SDS-PAGE and FPLC, and the products of the reaction catalyzed by the enzyme were analyzed with TLC and GC-MS analyses.

4.1 *Bacillus sp.* cloning experiments

For expression of the extracellular benzaldehyde lyase enzyme in *Bacillus subtilis*, three different plasmids were used. Although three different plasmids were used, the cloning strategy was the same, for the extracellular expression in *Bacillus subtilis; the subC* gene signal peptide with and without the promoter from *Bacillus licheniformis* was associated with the *bal* gene by the help of gene splicing by overlap extension (SOE) Method.

4.1.1 Primer Design for Generation of *bal*, *pre-sub*C and *hybrid* Genes

The gene encoding subtilisin Carlsberg from *Bacillus licheniformis* NCIB 6816 (*sub*C) was firstly reported by Jacobs et. al. (1985- Accession no: X03341). The signal peptide of serine alkaline protease, *pre-sub*C gene, was obtained from the chromosomal DNA of *B. licheniformis* DSM 1969 (Çalık et. al., 2003-b) and recombinant pUC18::*bal* plasmid was used as the template sequence (Pohl, et. al., 2002-Accession no: AX349268) for amplification of *bal* gene.

Complete DNA sequences of the templates were attained from National Institutes of Health (USA), National Library of Medicine, National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/BLAST, Last accessed date June 2006) and nucleotide sequences were illustrated in Appendix E.

4.1.2 Vector pMK4 for *Bacillus sp.* cloning experiments

The first strategy for the extracellular expression of the benzaldehyde lyase gene in *B. subtilis* was to clone the hybrid gene in pMK4. The *pre-subC* sequence was amplified with its own promoter and joined with the *bal* gene by means of SOE method.

For the gene cloning in *Bacillus* species, among the constructed vectors, *E. coli*/*B. subtilis* shuttle vector pMK4 was selected to clone the *subC* gene with its own promoter (Figure 4.1). In this designed vector, seven unique restriction sites flanked by two transcriptional terminators, ampicillin resistant gene for selection in *E. coli* and chloramphenicol resistant gene for selection in *B. subtilis* and the *LacZ* gene for detection of the gene insertion, are available (Sullivan et al., 1984).

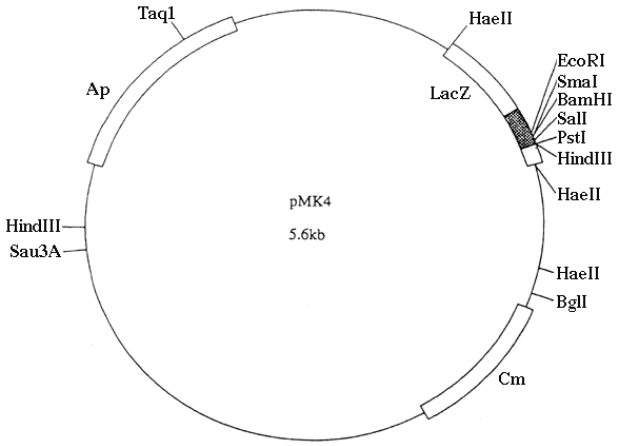


Figure 4.1 pMK4 expression vector. (Sullivan et al. 1984),

Four different primers, two for *pre-sub*C gene, others for *bal* gene were used to create the *hybrid* gene. At the beginning of the forward primer of *pre-sub*C, the *EcoRI* restriction enzyme recognition sequence was added. As it can be seen from Figure 4.3, extra sequences were added to the 5' end of oligonucleotide primers 2 and 3 in order to provide overlapping ends to the amplified fragments. By this way, in the following denaturation and reannealing processes, strand from two fragments will act as primers on each other and extension of overlap by DNA polymerase will result in the recombinant hybrid gene. Similarly, to obtain the mature protein gene encoding *bal* from the donor plasmid, the forward primer with complimentary strand to *pre-sub*C and *bal* reverse primer designed previously with *Sac*I recognition sequence were used (Figure 4.2).

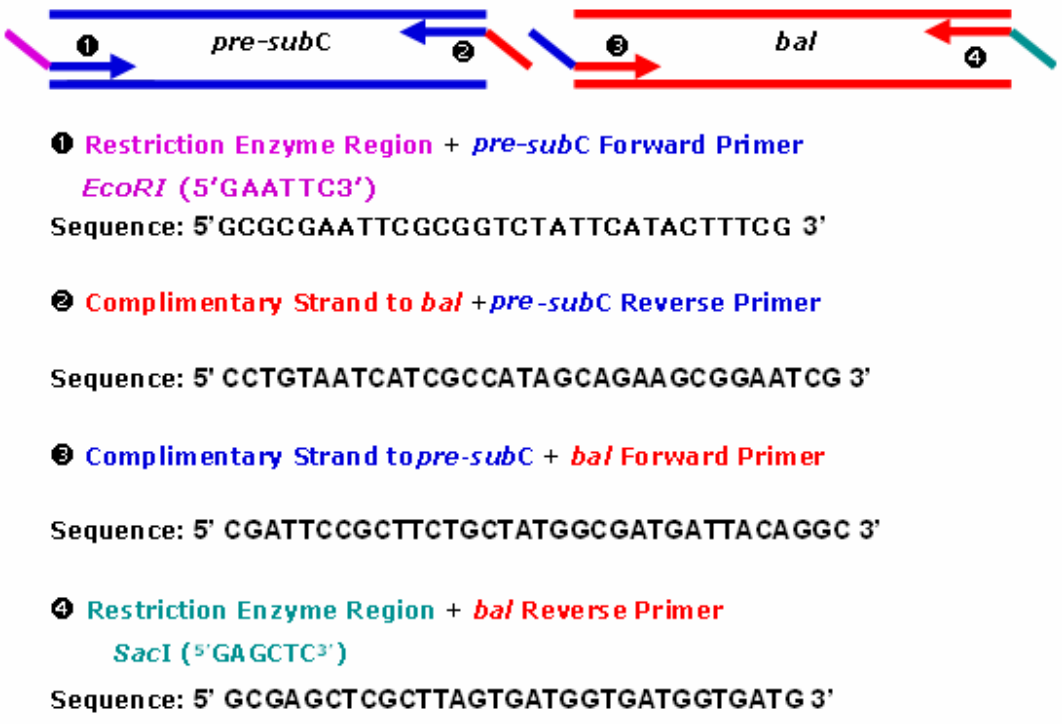
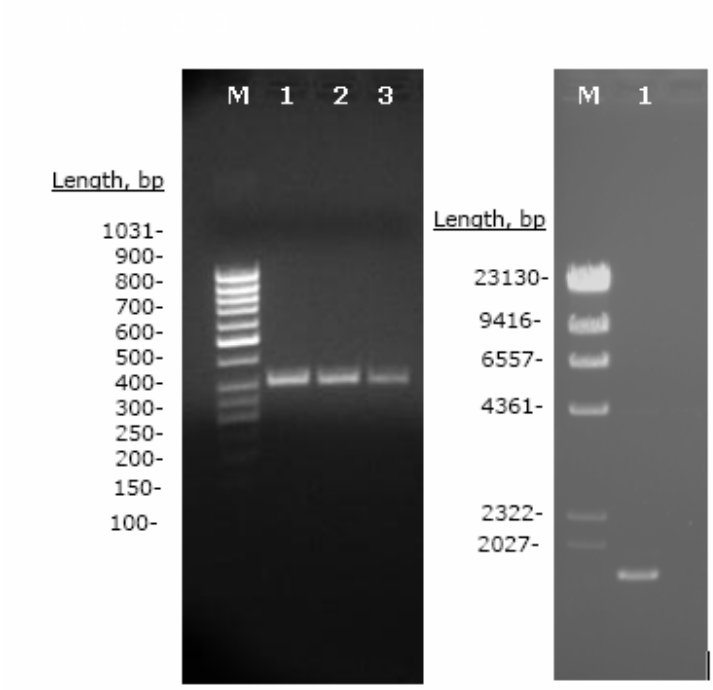


Figure 4.2 Schematic illustration of primer design for cloning pMK4 to express in *Bacillus* species. 1) *EcoRI* RE sequence was associated in front of pre-subC Forward Primer as a single-stranded tail. pre-subC Forward Primer is complimentary to anti-sense strand of the *pre-subC* gene; 2) Overlapping extension complimentary to bal anti-sense strand was added to 5' end of pre-subC Reverse Primer, which is complementary to sense strand of pre-subC gene; 3) Overlapping extension complimentary to pre-subC sense strand was added to 5' end of bal Forward Primer, which is complementary to anti-sense strand of bal gene; 4) *Sac*I RE sequence was associated in front of *bal* Reverse Primer as a singe-stranded tail. bal Reverse Primer is complimentary to sense strand of the *bal* gene.

4.1.2.1 Amplification of *bal*, *pre-sub*C and *hybrid* Genes by Polymerase Chain Reaction (PCR) and SOE Method

In the genetic applications performed for cloning in pMK4 to express the enzyme extracellularly in *Bacillus subtilis*, the signal sequence of subtilisin Carlsberg gene present on genomic DNA of *Bacillus licheniformis* DSM 1969, *pre-sub*C, with *EcoRI* recognition sequence (final length=315bp) was isolated as demonstrated in Figure 4.3-a. In the second step the mature protein gene encoding BAL with *Sac*I

sequence (final length=1739bp) was obtained from recombinant pUC18::*bal* plasmid (Figure 4.3-b).



(a) (b)

Figure 4.3 Gel electrophoresis image of *pre-sub*C and *bal* genes for cloning into pMK4. (a) M: Low Range Marker; 1-3: *pre-sub*C gene amplified with *Pfu* DNA polymerase; (b) M: λ DNA/HindIII Marker; 1: *bal* gene amplified with *Pfu* DNA polymerase.

These two sequences were associated with gene splicing by overlap extension method (SOE). Complimentary primers, *pre-sub*C Reverse Primer and *bal* Forward Primer, were used to generate two DNA fragments having overlapping ends. These fragments were combined in a subsequent fusion reaction in which the overlapping ends anneal, allowing the 3' overlap of each strand to serve as a primer for the 3' extension of the complementary strand (Ho et. al., 1989). The resulting fusion product was further be amplified in 3rd PCR (Figure 4.4) with a final length of 2051bp.

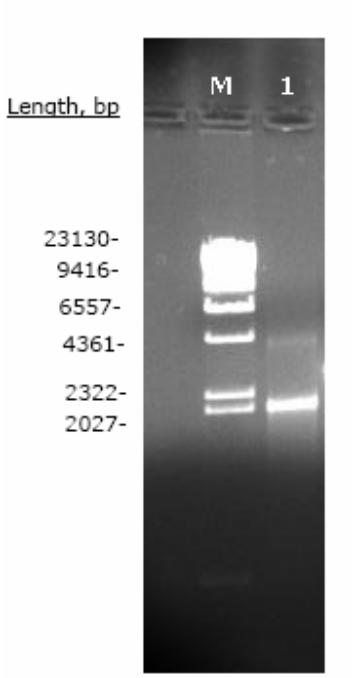


Figure 4.4 Gel electrophoresis image of *hybrid* gene for cloning into pMK4. M: λ DNA/HindIII Marker; 1: *hybrid* gene amplified with *Taq* DNA polymerase.

The annealing temperature and time were determined in accordance with the length of each DNA fragment and PCR components and parameters were illustrated in Tables 4.1 and 4.2 *pre-sub*C and *bal* genes were amplified with proofreading *Pfu* DNA polymerase not to allow any additional nucleotide insertion between the signal peptide and the mature part of *bal* gene since frame shift occurred due to additional residues can alter the 3-D conformation of the expressed protein or decrease the efficiency of signal peptidase act during cleavage of signal peptide from mature part (Simonen et. al., 1993).

Taq DNA polymerase, on the other hand, has an activity to add 3' A-overhangs to each end of the PCR product and thus not recommended for precise combination of two DNA fragments. *Hybrid* gene was produced with *Taq* DNA polymerase since it was previously shown that *Taq* DNA polymerase has superior efficiency in terms of yield and 3' A-overhangs does not cause any frame shift for fused *hybrid* gene.

The annealing temperature of pre-*sub*C, *bal* and *hybrid* genes were kept at 60^{o}C for accurate binding of primers. But for the SOE reaction annealing temperature was chosen at 55^{o}C.

Table 4.1 PCR process parameters for amplification of *pre-sub*C, *bal* and *hybrid* genes.

		Genes to be amplified		
		*pre-sub*C gene	*bal* gene	*hybrid* gene
1 cycle	T_1 = 94^{o}C,	3 min.	3 min.	3 min.
30 cycle	T_1 = 94^{o}C,	1 min.	1 min.	1 min.
	T_2 = 60^{o}C,	1 min.	1 min.	1 min.
	T_3 = 72^{o}C,	20 s.	2 min.	2.5 min.
1 cycle	T_3 = 72^{o}C,	10 min.	10 min.	10 min.
	T_4 = 4^{o}C,	5 min.	5 min.	5 min.

Table 4.2 Components of reaction mixture of PCR for amplification of *pre-sub*C, *bal* and *hybrid* genes

	Genes to be amplified		
	*pre-sub*C gene	*bal* gene	*hybrid* gene
Components	Amount	Amount	Amount
10XPCR Buffer	5µl	5µl	5µl
dNTPs (1mM)	10µl	10µl	10µl
FP (10µM)	pre-subC FP: 0.5µl	BAL FB: 0.5µl	pre-subC FP: 1µl
RP (10µM)	pre-subC RP: 0.5µl	BAL RP: 0.5µl	BAL RP: 1µl
Template DNA	~3 µg cDNA	~60 ng pUC18::*bal*	~30 ng *pre-sub*C + ~30 ng *bal* genes
dH_2O	up to 49 µl	up to 49 µl	up to 49 µl
DNA polymerase	2.5 U *Pfu*	2.5 U *Pfu*	2.5 U *Taq*

Starting with pure template DNA can eliminate any kind of contamination especially when fresh PCR products were directly used as template since reaction mixture can contain many components like salts, proteins, metals and organic solvents which can inhibit the polymerase efficiency and reduce PCR yield. Also having lots of redundant substances in the reaction mixture can mask the required amplifiable templates resulting in reduced yield. Therefore, *pre-sub*C and *bal* genes were purified with PCR Purification Kit (QIAquick) before utilized as template DNA.

4.1.2.2 Ligation of *hybrid* Gene into pUC19 Cloning Vector and Transformation in *E. coli* XL1-Blue Strain

PCR amplified and purified *hybrid* gene and isolated pUC19 vector were digested with *Sac*I and *EcoRI* restriction enzymes in reaction mixture of 20 µl final volume, subsequently, for overnight at 37°C. After restriction digestion, reaction mixtures were kept at 85°C for 20 min. For thermal inactivation of the enzymes. Between each subsequent restriction digestion, DNA fragments were purified by precipitating with NaAc(3M) and EtOH(100%). Thereafter, digested and purified *hybrid* gene was cloned into purified linear pUC19 cloning vector treated with the same restriction enzymes from the multiple cloning site present on *LacZ* gene fragment. Purified gene fragments were associated at 16°C for 16 h by keeping the Gen/Vector ratio at 3:1.

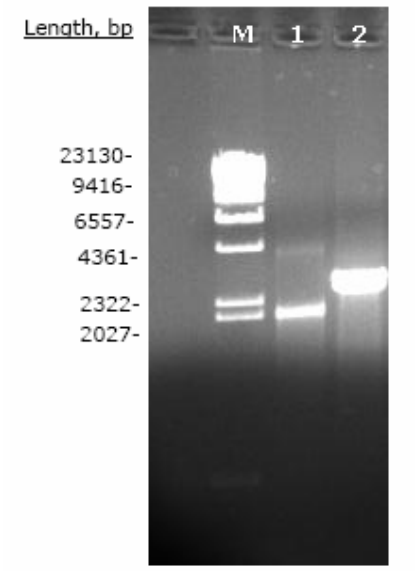


Figure 4.5 Gel electrophoresis image of *hybrid* gene and digested pUC19 plasmid. M: λ DNA/HindIII Marker; 1: *hybrid* gene digested with *EcoRI* and *SacI* REs; 2: *pUC19*plasmid digested with *EcoRI* and *SacI* REs

After transformation of ligation product into *E. coli* XL1-Blue strain by $CaCl_2$ method, transformants were incubated on LB-agar medium supplemented with 100 µg/ml ampicillin, IPTG and X-GAL for 12-18 h. Putative colonies were selected according to the blue-white screening method and incubated overnight in LB Broth supplemented with 100 µg/ml ampicillin at 37°C for plasmid DNA isolation.

The assembly of the recombinant molecule, pUC19::*hybrid*, was schematically illustrated in Figure 4.6.

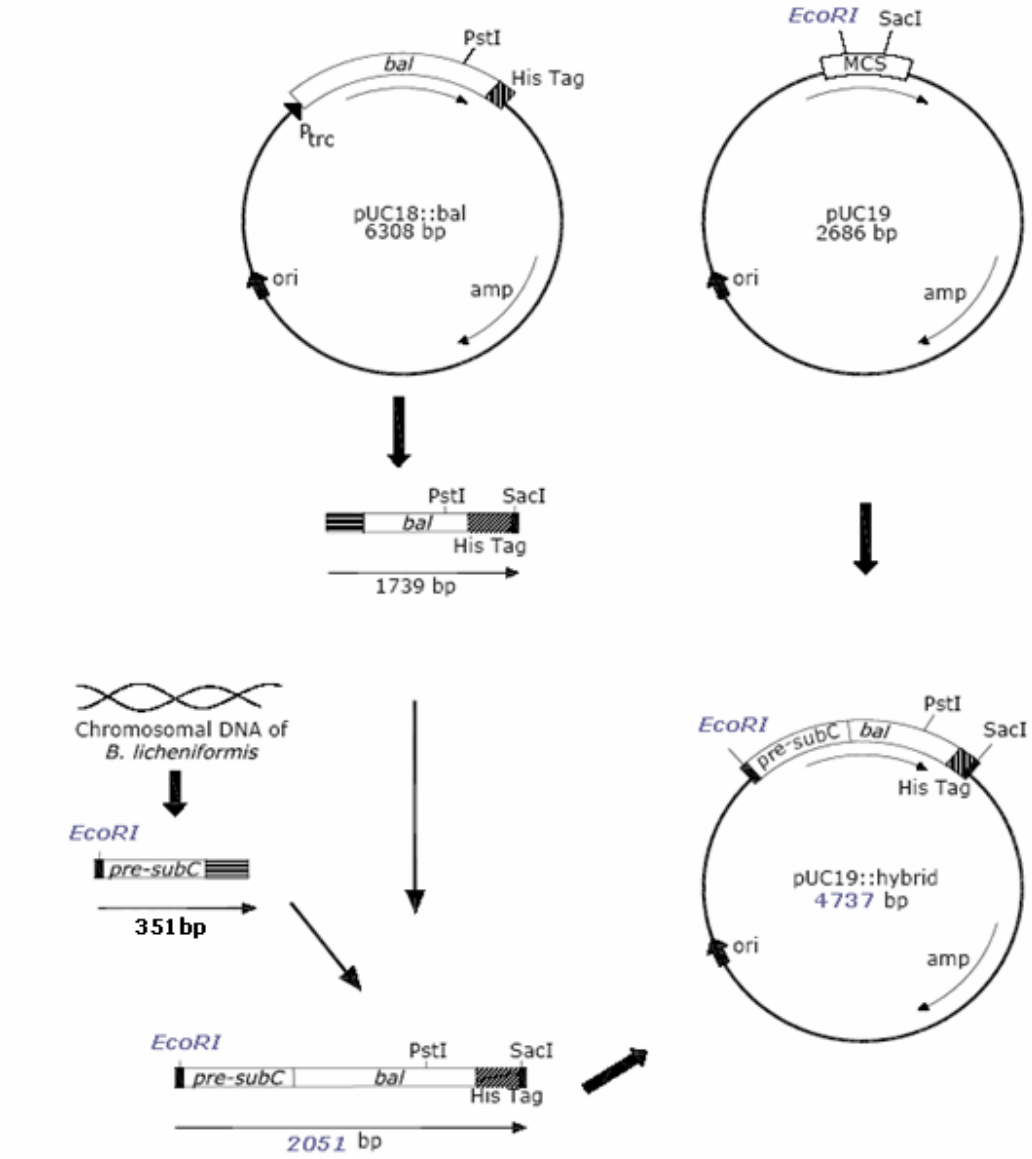


Figure 4.6 Scheme for the construction of recombinant pUC19::*hybrid*. 351 bp DNA fragment of signal peptide of serine alkaline protease with *EcoRI* extension and complementary sequence to *bal* gene (horizontally hatched region) was associated to 1739 bp DNA fragment of *bal* gene having *Sac*I end and complementary sequence to *pre-sub*C gene leading to 2051 bp DNA fragment of *hybrid* gene. *Hybrid* gene *(pre-subC::bal)* with *Sac*I and *EcoRI* extensions was cloned into pUC19 plasmid from *Sac*I and *EcoRI* restriction sites. The insertion and transcription directions were indicated by arrows; Cross hatched region indicates the location of poly-histidine tag used to the N-terminus of *bal* gene; P, promoter; MCS, multiple cloning site.

Putative recombinant pUC19::*hybrid* plasmids were analyzed by restriction digestion to confirm the insertion of *hybrid* gene to the cloning plasmid. Putative recombinant plasmids were treated, firstly, with *Sac*I and *EcoRI* restriction enzymes and a gene portion in the same length of *hybrid* gene was observed. Digested putative recombinant plasmids were further digested with *Pst*I restriction enzyme, which is known to cut the *hybrid* gene from the middle leading to two gene fragments of 1273bp and 778bp. The gel electrophoresis results verified that *hybrid* gene was cloned into pUC19 cloning vector in correct orientation.

4.1.2.3 Ligation of *hybrid* Gene into *pMK4* Expression Vector and Transformation in *Bacillus* Species

The *hybrid* gene sub-cloned into pUC19 cloning vector was isolated from r-pUC19 plasmid by subsequent digestion of the plasmid with *Sac*I and *EcoRI* and purified from agarose gel. Similarly, isolated *pMK4 E. coli/B. subtilis* shuttle vector was cut with the same REs and purified by Gene Elusion Kit.

Digested *hybrid* gene was ligated to purified linear *pMK4* plasmid from the cohesive ends by the help of T4 DNA ligase and reaction was performed at 16^0C for 16h. Ligation mixture was transformed into *E. coli* XL1-Blue by $CaCl_2$ method and putative colonies carrying r-*pMK4* plasmid with ampicillin resistance gene were picked from solid medium and isolated for further control experiments.

The assembly of the recombinant molecule, *pMK4*::*hybrid*, was schematically illustrated in Figure 4.7.

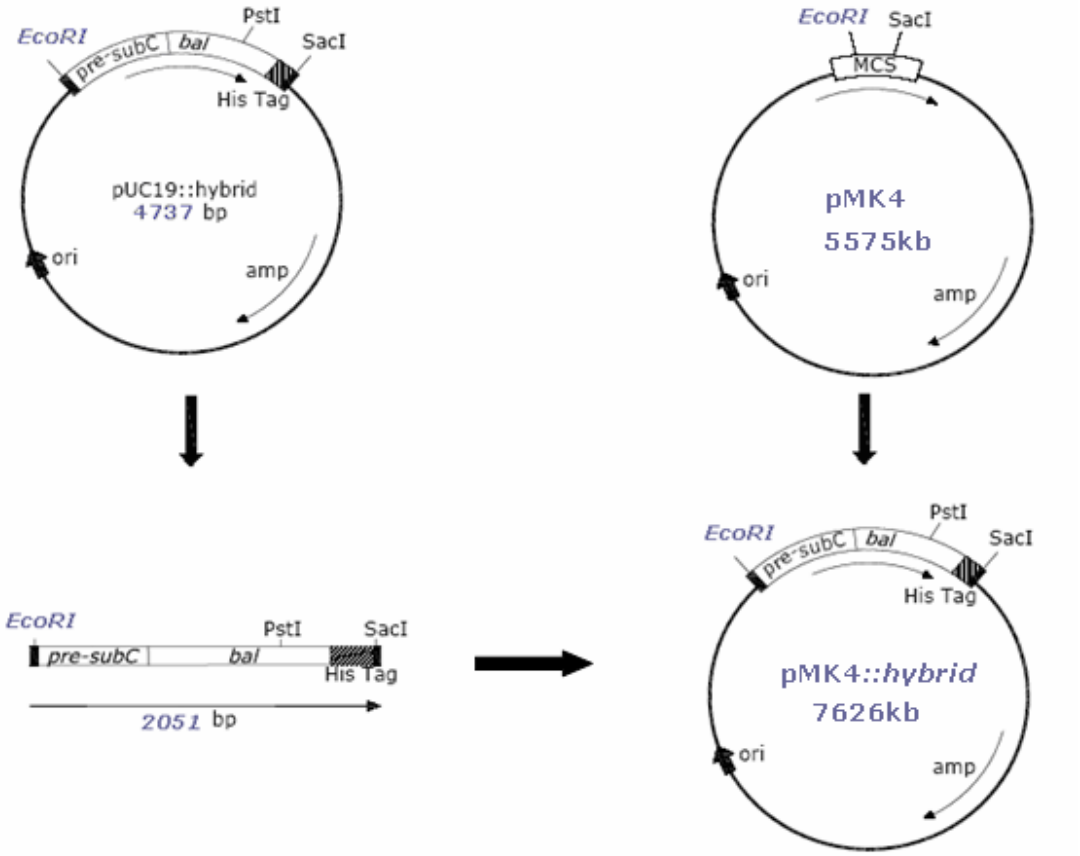
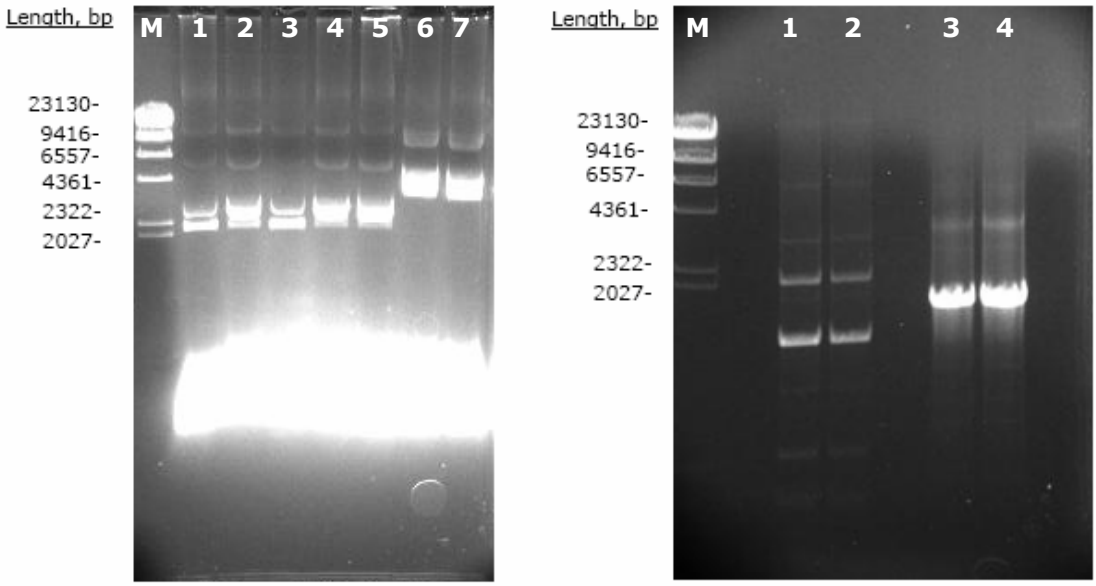


Figure 4.7 Scheme for the construction of *pMK4*::*hybrid* for BAL production in *Bacillus subtilis*. 2051 bp DNA fragment of f *hybrid* gene (*pre-sub*C::*bal*) with *Sac*I and *EcoRI* endings was cloned into *pMK4* plasmid from *Sac*I and *EcoRI* restriction sites. The insertion and transcription directions were indicated by arrows; Cross hatched region indicates the location of poly-histidine tag used to the C-terminus of *bal* gene; P, promoter; MCS, multiple cloning site.

Similar tests performed with r-pUC19 plasmid were carried out to ensure the insertion of the gene in correct orientation. Putative recombinant *pMK4*::*hybrid* plasmids were explored to RE digestion and PCR by being used as template DNA. The gel electrophoresis image (Figure 4.8) shows that cloning was successful.



(a) (b)

Figure 4.8 Control of the transformation. (a) M: λ DNA/HindIII Marker; 1-7: Putative recombinant colonies. (b) M: λ DNA/HindIII Marker; 1: Control PCR for *bal* gene of the first putative colony; 2: Control PCR for *bal* gene of the second putative colony; 3: Control PCR for *hybrid* gene of the first putativecolony; 4: Control PCR for *hybrid* gene of the second putative colony.

The DNA sequence of insert DNA was further controlled by automatic DNA sequencers (Microsynth GmbH, Switzerland) and results revealed that cloning was not successful. There was an eight nucleotides deletion between 590^{th}-597^{th} bases in the *bal* gene. After deletion a frameshift occurred and *the bal* gene sequence had changed. The changed nucleotide sequence and aminoacid sequence is given in Table 4.3.

Table 4.3 The *bal* gene sequence and aminoacid sequence after deletion and frameshift.

```
   1 - ATGGCGATGATTACAGGCGGCGAACTGGTTGTTCGCACCCTAATAAAGGCTGGGGTCGAA - 60
     - M  A  M  I  T  G  G  E  L  V  V  R  T  L  I  K  A  G  V  E
  61 - CATCTGTTCGGCCTGCACGGCGCGCATATCGATACGATTTTTCAAGCCTGTCTCGATCAT - 120
     - H  L  F  G  L  H  G  A  H  I  D  T  I  F  Q  A  C  L  D  H
 121 - GATGTGCCGATCATCGACACCCGCCATGAGGCCGCCGCAGGGCATGCGGCCGAGGGCTAT - 180
     - D  V  P  I  I  D  T  R  H  E  A  A  A  G  H  A  A  E  G  Y
 181 - GCCCGCGCTGGCGCCAAGCTGGGCGTGGCGCTGGTCACGGCGGGCGGGGGATTTACCAAT - 240
     - A  R  A  G  A  K  L  G  V  A  L  V  T  A  G  G  G  F  T  N
 241 - GCGGTCACGCCCATTGCCAACGCTTGGCTGGATCGCACGCCGGTGCTCTTCCTCACCGGA - 300
     - A  V  T  P  I  A  N  A  W  L  D  R  T  P  V  L  F  L  T  G
 301 - TCGGGCGCGCTGCGTGATGATGAAACCAACACGTTGCAGGCGGGGATTGATCAGGTCGCC - 360
     - S  G  A  L  R  D  D  E  T  N  T  L  Q  A  G  I  D  Q  V  A
 361 - ATGGCGGCGCCCATTACCAAATGGGCGCATCGGGTGATGGCAACCGAGCATATCCCACGG - 420
     - M  A  A  P  I  T  K  W  A  H  R  V  M  A  T  E  H  I  P  R
 421 - CTGGTGATGCAGGCGATCCGCGCCGCGTTGAGCGCGCCACGCGGGCCGGTGTTGCTGGAT - 480
     - L  V  M  Q  A  I  R  A  A  L  S  A  P  R  G  P  V  L  L  D
 481 - CTGCCGTGGGATATTCTGATGAACCAGATTGATGAGGATAGCGTCATTATCCCCGATCTG - 540
     - L  P  W  D  I  L  M  N  Q  I  D  E  D  S  V  I  I  P  D  L
 541 - GTCTTGTCCGCGCATGGGGCCAGACCCGACCCTGCCGATCTGGATCAGGCTTTTGCGCAA - 600
     - V  L  S  A  H  G  A  R  P  D  P  A  D  L  D  Q  A  F  A  Q
 601 - GGCGGAGCGGCCGGTCATCGTGCTCGGCTCAGAAGCCTCGCGGACAGCGCGCAAGACGGC - 660
     - G  G  A  A  G  H  R  A  R  L  R  S  L  A  D  S  A  Q  D  G
 661 - GCTTAGCGCCTTCGTGGCGGCGACTGGCGTGCCGGTGTTTGCCGATTATGAAGGGCTAAG - 720
     - A  *  R  L  R  G  G  D  W  R  A  G  V  C  R  L  *  R  A  K
 721 - CATGCTCTCGGGGCTGCCCGATGCTATGCGGGGCGGGCTGGTGCAAAACCTCTATTCTTT - 780
     - H  A  L  G  A  A  R  C  Y  A  G  R  A  G  A  K  P  L  F  F
 781 - TGCCAAAGCCGATGCCGCGCCAGATCTCGTGCTGATGCTGGGGGCGCGCTTTGGCCTTAA - 840
     - C  Q  S  R  C  R  A  R  S  R  A  D  A  G  G  A  L  W  P  *
 841 - CACCGGGCATGGATCTGGGCAGTTGATCCCCCATAGCGCGCAGGTCATTCAGGTCGACCC - 900
     - H  R  A  W  I  W  A  V  D  P  P  *  R  A  G  H  S  G  R  P
 901 - TGATGCCTGCGAGCTGGGACGCCTGCAGGGCATCGCTCTGGGCATTGTGGCCGATGTGGG - 960
     - *  C  L  R  A  G  T  P  A  G  H  R  S  G  H  C  G  R  C  G
 961 - TGGGACCATCGAGGCTTTGGCGCAGGCCACCGCGCAAGATGCGGCTTGGCCGGATCGCGG - 1020
     - W  D  H  R  G  F  G  A  G  H  R  A  R  C  G  L  A  G  S  R
1021 - CGACTGGTGCGCCAAAGTGACGGATCTGGCGCAAGAGCGCTATGCCAGCATCGCTGCGAA - 1080
     - R  L  V  R  Q  S  D  G  S  G  A  R  A  L  C  Q  H  R  C  E
1081 - ATCGAGCAGCGAGCATGCGCTCCACCCCTTTCACGCCTCGCAGGTCATTGCCAAACACGT - 1140
     - I  E  Q  R  A  C  A  P  P  L  S  R  L  A  G  H  C  Q  T  R
1141 - CGATGCAGGGGTGACGGTGGTAGCGGATGGTGCGCTGACCTATCTCTGGCTGTCCGAAGT - 1200
     - R  C  R  G  D  G  G  S  G  W  C  A  D  L  S  L  A  V  R  S
1201 - GATGAGCCGCGTGAAACCCGGCGGTTTTCTCTGCCACGGCTATCTAGGCTCGATGGGCGT - 1260
     - D  E  P  R  E  T  R  R  F  S  L  P  R  L  S  R  L  D  G  R
1261 - GGGCTTCGGCACGGCGCTGGGCGCGCAAGTGGCCGATCTTGAAGCAGGCCGCCGCACGAT - 1320
     - G  L  R  H  G  A  G  R  A  S  G  R  S  *  S  R  P  P  H  D
1321 - CCTTGTGACCGGCGATGGCTCGGTGGGCTATAGCATCGGTGAATTTGATACGCTGGTGCG - 1380
     - P  C  D  R  R  W  L  G  G  L  *  H  R  *  I  *  Y  A  G  A
1381 - CAAACAATTGCCGCTGATCGTCATCATCATGAACAACCAAAGCTGGGGGGCGACATTGCA - 1440
     - Q  T  I  A  A  D  R  H  H  H  E  Q  P  K  L  G  G  D  I  A
1441 - TTTCCAGCAATTGGCCGTCGGCCCCAATCGCGTGACGGGCACCCGTTTGGAAAATGGCTC - 1500
     - F  P  A  I  G  R  R  P  Q  S  R  D  G  H  P  F  G  K  W  L
1501 - CTATCACGGGGTGGCCGCCGCCTTTGGCGCGGATGGCTATCATGTCGACAGTGTGGAGAG - 1560
     - L  S  R  G  G  R  R  L  W  R  G  W  L  S  C  R  Q  C  G  E
1561 - CTTTTCTGCGGCTCTGGCCCAAGCGCTCGCCCATAATCGCCCCGCCTGCATCAATGTCGC - 1620
     - L  F  C  G  S  G  P  S  A  R  P  *  S  P  R  L  H  Q  C  R
1621 - GGTCGCGCTCGATCCGATCCCGCCCGAAGAACTCATTCTGATCGGCATGGACCCCTTCGG - 1680
     - G  R  A  R  S  D  P  A  R  R  T  H  S  D  R  H  G  P  L  R
1681 - ATCTCATCACCATCACCATCACTAA - 1705
     - I  S  S  P  S  P  S  L  X
```

As can be seen in the Table 4.3 the sequence was changed in such a way that the new sequence has many stop codons, which lead to incomplete translation of the bal enzyme. Although the sequence changed; the benzoin condensation reaction and the spectrophotometric activity were investigated for the mutant enzyme but it showed no benzaldehyde lyase activity.

4.1.3 Vector pDG148-*Stu* for *Bacillus sp.* cloning experiments

The second strategy for the extracellular expression of the benzaldehyde lyase gene in *B.* Subtilis was to clone the hybrid gene in pDG148-*Stu* plasmid by the help of ligation independent cloning method. In this cloning strategy was to clone the *pre-subC* gene without its own promoter. pDG148-*Stu* plasmid derives from pDG148 and combines the main property of this vector; conditional expression of the gene in response to isopropylL-D-thiogalactopyranoside, with (i) rapid orientated cloning by a ligation independent procedure and (ii) a ribosome binding site of high translational efficiency(Joseph et. al, 2001).

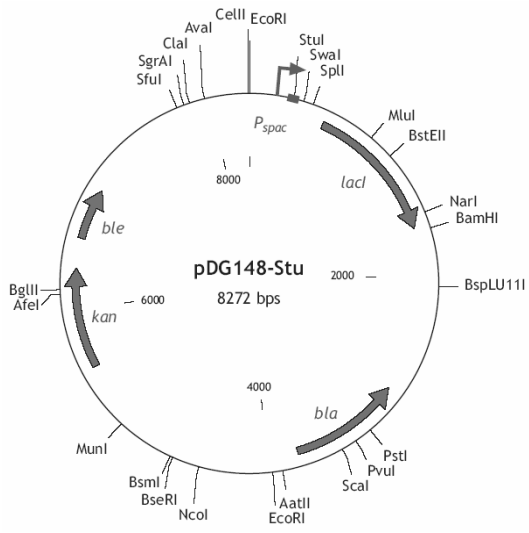


Figure 4.9 Schematic image of pDG148-*Stu* vector (http://www.bgsc.org, Last accessed date June 2006).

The assembly of the recombinant molecule, *pDG148-Stu*::*hybrid*, was schematically illustrated in Figure 4.10.

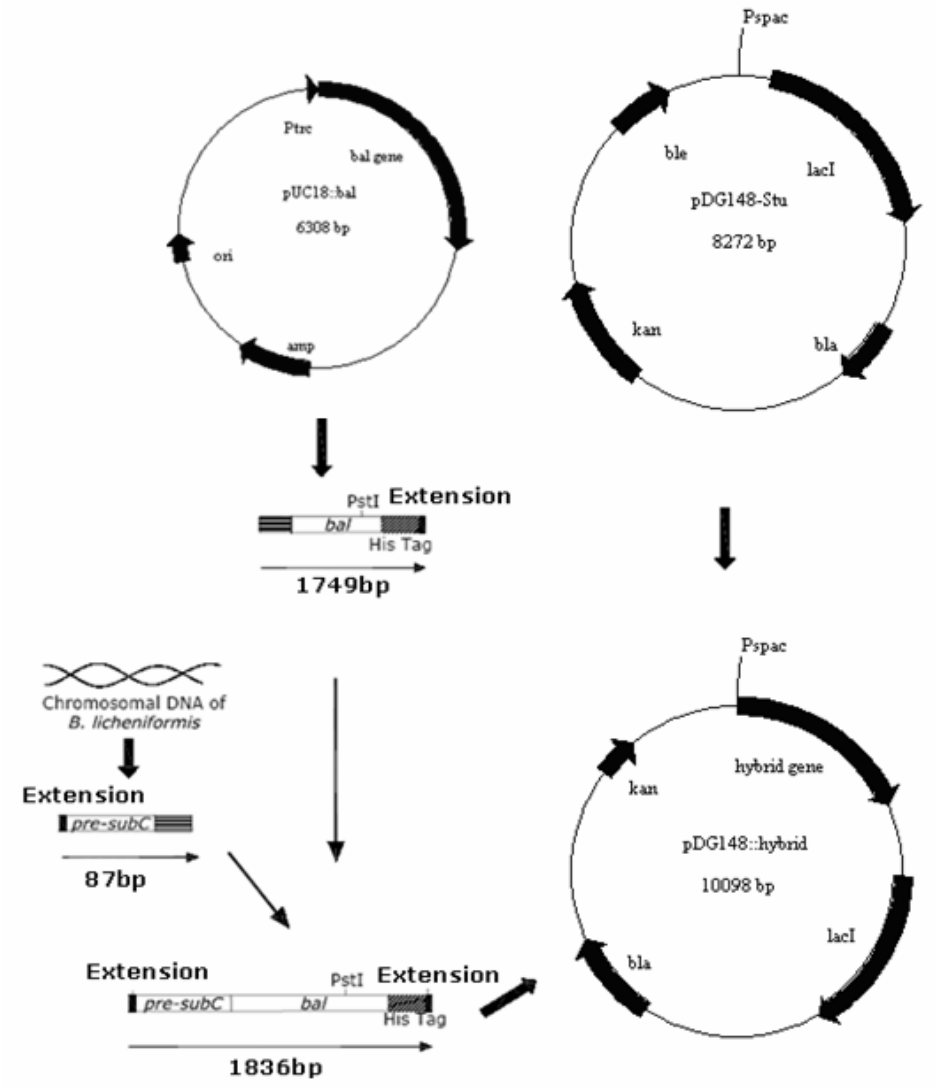


Figure 4.10 Scheme for the construction of *pDG148-Stu*::*hybrid* for BAL production in *Bacillus subtilis*.

The general cloning procedure for LIC is described in Figure 4.11. The cloning procedure was to amplify the hybrid gene with specifically designed first and fourth primers, where the pre-subC forward primer had 5'-AAGGAGGAAGCAGGT-3' extra nucleotides followed by start codon, ATG, and the reverse primer had the extension 5'-GACACGCACGAGGT-3' (Joseph et. al, 2001). The outer primers did not have any

restriction enzyme recognition sites. Because the method does not need any restriction enzyme. The SOE primers were the same of the primer, which were used for the pMK4 cloning experiments.

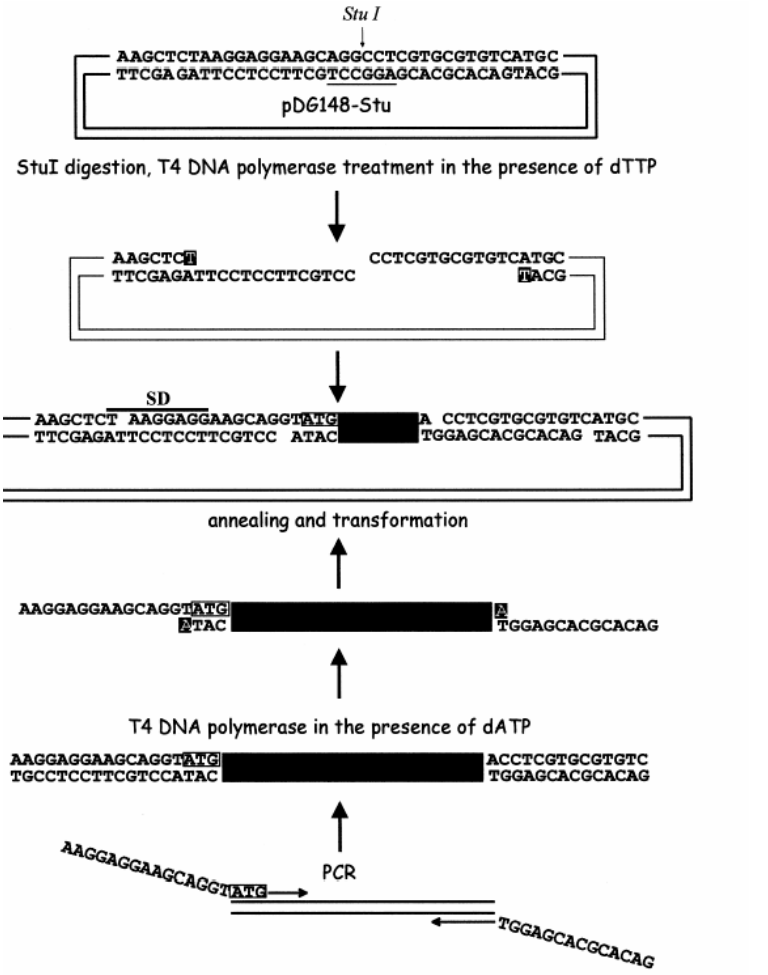


Figure 4.11 Schematic procedures for ligation independent cloning (Joseph et. al., 2001).

The procedure does not require the use of the restriction enzymes, T4 DNA ligase. After digestion with *StuI* enzyme the vector treated with the T4 DNA polymerase in the presence of dTTP. The PCR product was treated with the T4 DNA polymerase in the presence of dATP. The 3'-terminal sequences were removed by the action of the T4 DNA polymerase in the presence of dNTP, leading to fragments with 5'-extending single-stranded tails. Therefore the first and fourth primers had to have the extra nucleotides. After treatment with the T4 DNA polymerase, the vector and the PCR product were flanked by 13 and 14bp complementary overhangs. Then

they were mixed, the overhangs were thought to be annealed and the gapped circle directly transformed into *E.coli* (Joseph et. al, 2001).

The strategy for the cloning was that only the signal sequence of the *pre-subC* gene was associated with the *bal* gene no other sequences were included. Therefore only 87 nucleotides (29 aminoacids) of the *subC* gene were associated with the *bal* gene by the help of SOE method. The designed primers are shown in Figure 4.12.

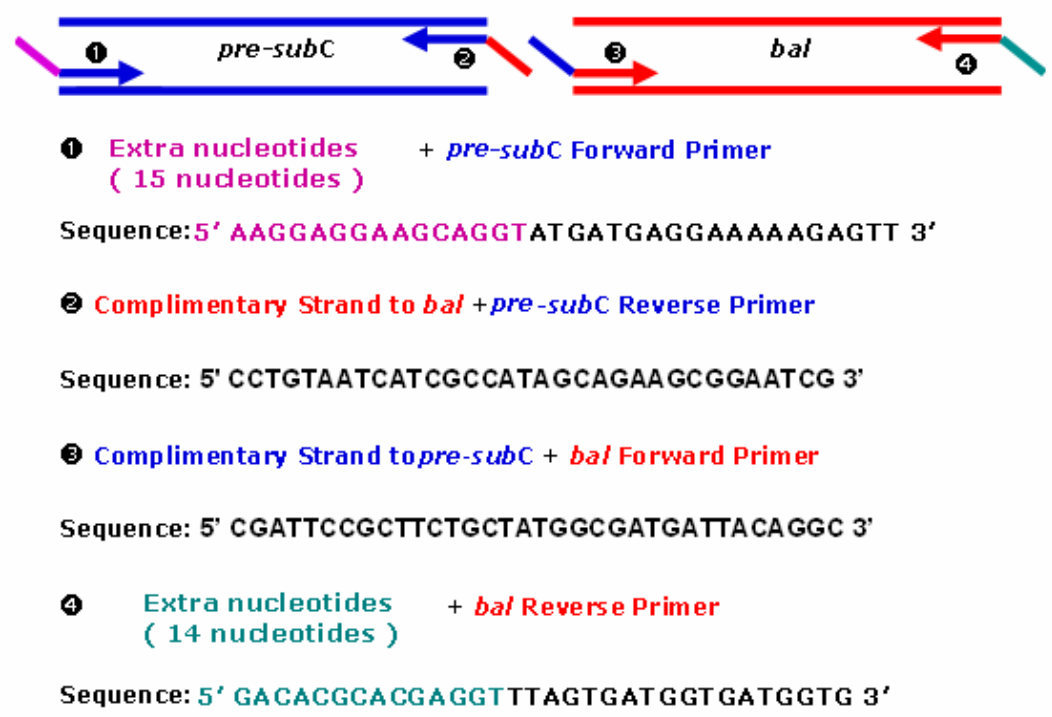


Figure 4.12 Schematic illustration of primer design for cloning pDG148-*Stu* to express in *Bacillus* species. 1) *15 nucleotides* were associated in front of *pre-subC* Forward Primer. *pre-subC* Forward Primer is complimentary to anti-sense strand of the pre-subC gene; 2) Overlapping extension complimentary to *bal* anti-sense strand was added to 5' end of *pre-subC* Reverse Primer, which is complementary to sense strand of pre-subC gene; 3) Overlapping extension complimentary to pre-subC sense strand was added to 5' end of *bal* Forward Primer, which is complementary to anti-sense strand of bal gene; 4) *14 nucleotides* were associated in front of *bal* Reverse Primer as a singe-stranded tail. *bal* Reverse Primer is complimentary to sense strand of the *bal* gene.

PCR process parameters and the PCR components for amplification of *pre-subC, bal* and *hybrid* genes are given in table 4.4 and table 4.5.

Table 4.4 PCR process parameters for amplification of *pre-sub*C, *bal* and *hybrid* genes for pDG148-*Stu* cloning experiments.

		Genes to be amplified		
		*pre-sub*C gene	*bal* gene	*hybrid* gene
1 cycle	T_1 = 94^{o}C,	3 min.	3 min.	3 min.
30 cycle	T_1 = 94^{o}C,	1 min.	1 min.	1 min.
	T_2 = 60^{o}C,	1 min.	1 min.	1 min.
	T_3 = 72^{o}C,	10 s.	2 min.	2 min.
1 cycle	T_3 = 72^{o}C,	10 min.	10 min.	10 min.
	T_4 = 4^{o}C,	5 min.	5 min.	5 min.

Table 4.5 Components of reaction mixture of PCR for amplification of *pre-sub*C, *bal* and *hybrid* genes for pDG148-*Stu* cloning experiments.

	Genes to be amplified		
	*pre-sub*C gene	*bal* gene	*hybrid* gene
Components	Amount	Amount	Amount
10XPCR Buffer	5µl	5µl	5µl
dNTPs (1mM)	10µl	10µl	10µl
FP (10µM)	pre-subC FP: 0.5µl	BAL FB: 0.5µl	pre-subC FP: 1µl
RP (10µM)	pre-subC RP: 0.5µl	BAL RP: 0.5µl	BAL RP: 1µl
Template DNA	~3 µg cDNA	~60 ng pUC18::*bal*	~30 ng *pre-sub*C + ~30 ng *bal* genes
dH_2O	up to 49 µl	up to 49 µl	up to 49 µl
DNA polymerase	2.5 U *Pfu*	2.5 U *Pfu*	2.5 U *Taq*

In figure 4.13 the PCR product hybrid gene and the *pDG148-Stu* plasmid were shown. In figure, the hybrid was extracted from agarose gel and the plasmid, which was digested with *StuI* enzyme and extracted from agarose gel. These DNA fragments were directly treated with T4 DNA polymerase and mixed with each others. Then EDTA was added to the mixture to stop the T4 DNA polymerase activity. Thereafter the gapped circle directly transformed into *E.coli* cells.

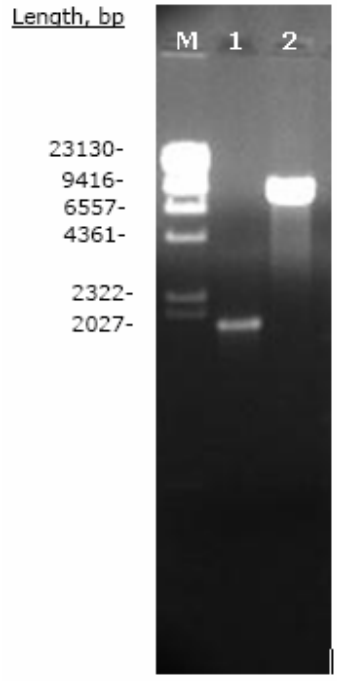


Figure4.13 Gel electrophoresis image of the pDG148-Stu plasmid digested with *StuI* and the hybrid gene. M: λ DNA/HindIII Marker; 1: hybrid gene; 2: pDG148-Stu plasmid digested with *StuI.*

Although the all process parameters for ligation independent cloning such as; different T4 DNA polymerase quantity, different insert and vector quantities were tried and the different insert:vector ratios were tried, the cloning experiments were not successful.

It can be concluded that the ligation independent cloning procedure is not an efficient method.

4.1.4 Vector pMUTIN4 for *Bacillus sp.* cloning experiments

The third strategy for the extracellular expression of the benzaldehyde lyase gene in *B. subtilis* was to clone the hybrid gene in pMUTIN4 and to integrate the hybrid gene into genomic DNA. pMUTIN4 plasmid is an integration vector but it integrates the genome in the cloned gene (Vagner et. al., 1998).

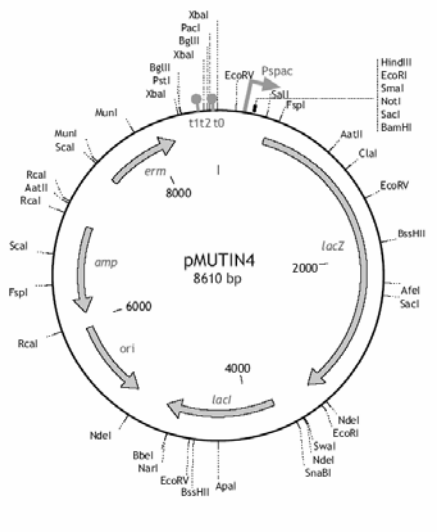
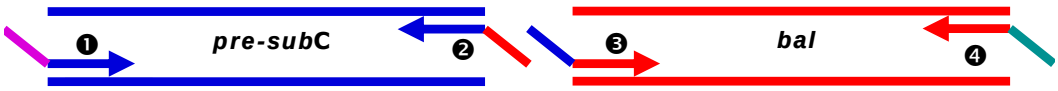


Figure 4.14 The scheme of pMUTIN4 plasmid (http://www.bgsc.org, Last accessed date June 2006).

The cloning strategy was to clone the *pre-subC* gene into pMUTIN4 plasmid without its own promoter. The primers were designed in such a way that the *pre-subC* sequence was amplified with its own ribosomal binding sit but without its promoter. Therefore *pre-subC* sequence was amplified without its own promoter and joined with the *bal* gene by means of SOE method.

The designed primers are shown in Figure 4.15.



❶ **Restriction Enzyme Region + *pre-subC* Forward Primer**
***BamHI* (5'GGATCC3')**
Sequence: 5' GCGGATCCGCAGCAATCTCCTGTCATTCG 3'

❷ **Complimentary Strand to *bal* + *pre-subC* Reverse Primer**

Sequence: 5' CCTGTAATCATCGCCATAGCAGAAGCGGAATCG 3'

❸ **Complimentary Strand to *pre-subC* + *bal* Forward Primer**

Sequence: 5' CGATTCCGCTTCTGCTATGGCGATGATTACAGGC 3'

❹ **Restriction Enzyme Region + *bal* Reverse Primer**
***SacI* (5'GAGCTC3')**
Sequence: 5' GCGAGCTCGCTTAGTGATGGTGATGGTGATG 3'

Figure 4.15 Schematic illustration of primer design for expression in *Bacillus* species. 1) *BamH*I RE sequence was associated in front of *pre-subC* Forward Primer as a single-stranded tail. *pre-subC* Forward Primer is complimentary to anti-sense strand of the *bal* gene; 2) Overlapping extension complimentary to *bal* anti-sense strand was added to 5' end of *pre-subC* Reverse Primer, which is complementary to sense strand of pre-subC gene; 3) Overlapping extension complimentary to *pre-subC* sense strand was added to 5' end of *bal* Forward Primer, which is complementary to anti-sense strand of *bal* gene; 4) *Sac*I RE sequence was associated in front of *bal* Reverse Primer as a singe-stranded tail. *bal* Reverse Primer is complimentary to sense strand of the *bal* gene.

In this cloning study, the strategy was that after digestion with *BamHI* and *SacI* restriction enzyme the pMUTIN4 plasmid lost its 1994 bp fragment and in that region the hybrid gene tried to be cloned. The resulting plasmid is shown in Figure 4.16.

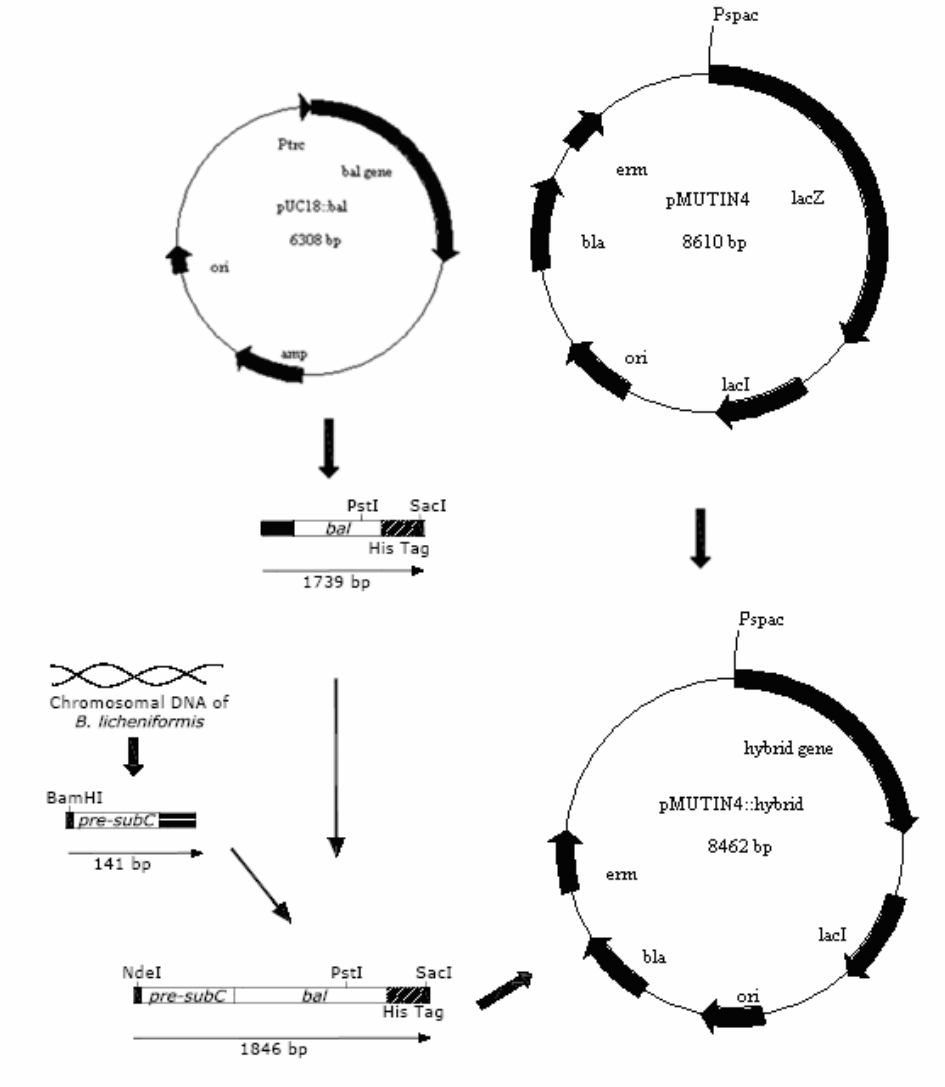


Figure4.16 Scheme of recombinant pMUTIN4 plasmid.

In this third cloning study, the signal sequence of subtilisin Carlsberg gene present on chromosomal DNA (cDNA) of *Bacillus licheniformis* DSM 1969, *pre-sub*C, with *BamH*I recognition sequence (final length=141bp) was isolated as demonstrated in Figure 4.17-a. In the second step the mature protein gene encoding BAL with *Sac*I sequence (final length=1739bp) was obtained from recombinant pUC18::*bal* plasmid (Figure 4.17-b). And these two fragments were

joined by means of SOE method. The resulting fusion product was further be amplified in 3rd PCR (Figure 4.19) with a final length of 1846bp.

Table 4.6 PCR process parameters for amplification of *pre-sub*C, *bal* and *hybrid* genes for pMUTIN4 cloning experiments.

		Genes to be amplified		
		*pre-sub*C gene	*bal* gene	*hybrid* gene
1 cycle	T_1 = 94°C,	2 min.	3 min.	3 min.
30 cycle	T_1 = 94°C,	1 min.	1 min.	1 min.
	T_2 = 60°C,	1 min.	1 min.	1 min.
	T_3 = 72°C,	15 s.	2 min.	2.5 min.
1 cycle	T_3 = 72°C,	5 min.	10 min.	10 min.
	T_4 = 4°C,	5 min.	5 min.	5 min.

Table 4.7 Components of reaction mixture of PCR for amplification of *pre-sub*C, *bal* and *hybrid* genes for pMUTIN4 cloning experiments.

	Genes to be amplified		
	*pre-sub*C gene	*bal* gene	*hybrid* gene
Components	Amount	Amount	Amount
10XPCR Buffer	5µl	5µl	5µl
dNTPs (1mM)	10µl	10µl	10µl
FP (10µM)	pre-subC FP: 1µl	BAL FB: 1µl	pre-subC FP: 1µl
RP (10µM)	pre-subC RP: 1µl	BAL RP: 1µl	BAL RP: 1µl
Template DNA	~3 µg cDNA	~60 ng pUC18::*bal*	~30 ng *pre-sub*C + ~30 ng *bal* genes
dH_2O	up to 49 µl	up to 49 µl	up to 49 µl
DNA polymerase	2.5 U *Pfu*	2.5 U *Pfu*	2.5 U *Taq*

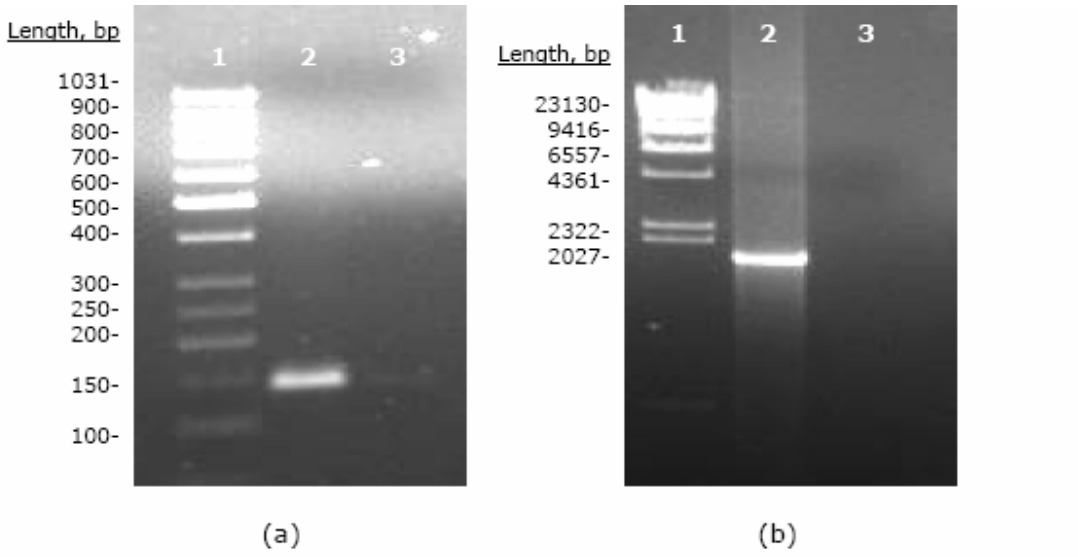


Figure 4.17 Gel electrophoresis image of *pre-sub*C and *bal* genes. (a) 1: Low Range Marker; 2: *pre-sub*C gene amplified with *Pfu* DNA polymerase; (b) 1: λ DNA/HindIII Marker; 2: *bal* gene amplified with *Taq* DNA polymerase.

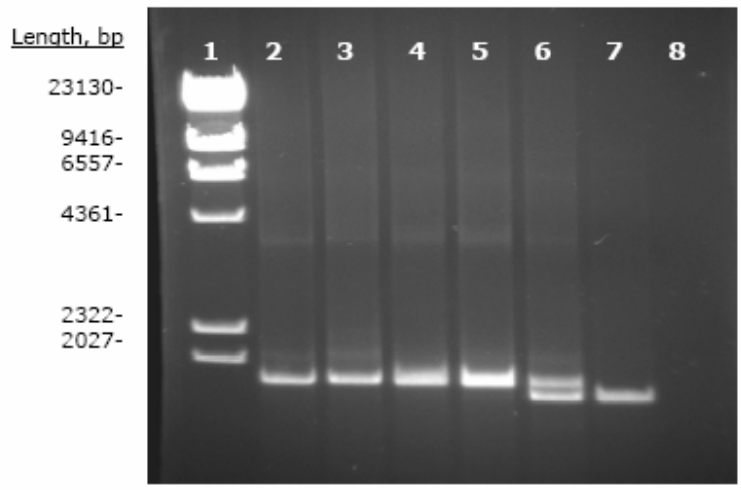


Figure 4.18 Gel electrophoresis image of *hybrid* gene. 1: λ DNA/HindIII Marker; 2-5: *hybrid* gene amplified with *Taq* DNA polymerase; 6: *hybrid* + *bal* genes for control; 7: Positive control (*bal* gene) control.

For the pMUTIN4 cloning experiments; if the cloning had been successful, the hybrid gene would have transcribed under the control of *subC* gene promoter. Because the plasmid integrates into genome in the cloned sequence.

Although all the cloning parameters were investigated pMUTIN4 cloning experiments were unsuccessful.

4.2 Development of recombinant *P.pastoris* carrying benzaldehyde lyase gene

The research program for the cloning of *bal gene* and extracellularly expression of benzaldehyde lyase in *Pichia pastoris* was carried out mainly in two parts. Firstly, the gene encoding benzaldehyde lyase (*bal*) was cloned into vector pPICZαA, which carries the α-Factor pre-pro peptide. Thereafter the recombinant vector was introduced and expressed in *P.pastoris*. The research plan for the development of the recombinant microorganism producing extracellular benzaldehyde lyase is summarized in Figure 4.19.

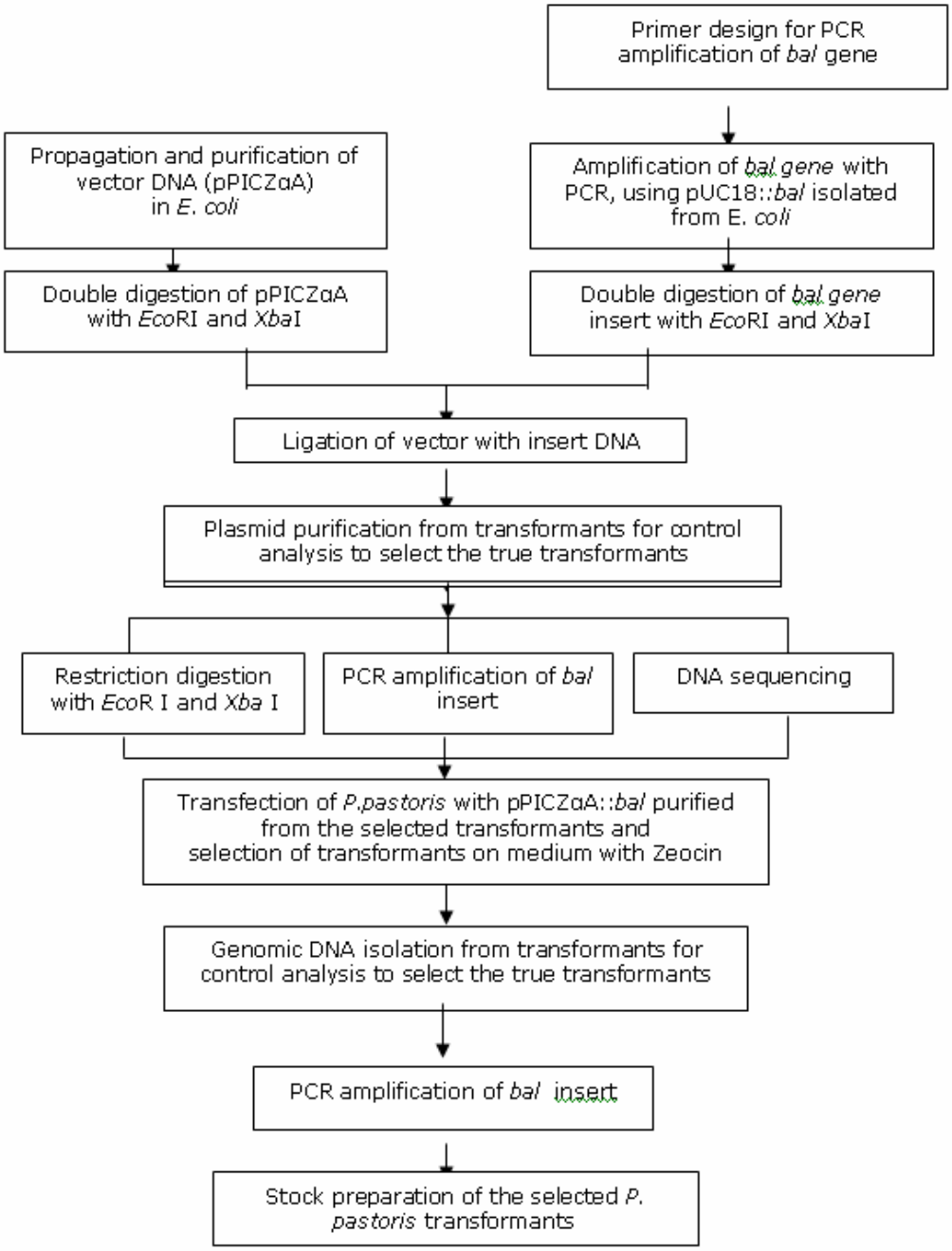


Figure 4.19 Flowchart of the research plan, for the development of the r-*P. pastoris* producing extracellular benzaldehyde lyase.

4.2.1 Primer Design for Generation of *bal* Gene

Recombinant pUC18::*bal* plasmid was used as the template sequence for amplification of *bal* gene, (Appendix B), using the nucleotide sequence defined by Pohl et. al. (2002).

For the amplification of *bal* gene, two primers were designed, one of which is *EcoRI-bal* forward, where *EcoRI* restriction enzyme recognition sequence was added; and the other *XbaI-bal* reverse, where *XbaI* restriction enzyme recognition sequence was added to the *bal* gene sequence. Restriction enzymes were chosen in such a way that they were not digesting the *bal* gene and their location in the vector suited the design. After digestion with the enzymes the vector lost its 63 nucleotides and the *bal* gene was integrated in that gap. Primers were designed according to the rules given in section 3.11.3; and the assembly of primers designed was illustrated in Figure 4.21. The designed primer sequences are also shown in *bal* gene sequence in Appendix B. Designed primers (Figure 4.20) were synthesized in Thermo Hybaid GmbH (Germany) laboratories.

The restriction enzymes' recognition sequences that were added to the 5' ends were selected from those of not cutting the *bal* gene according to map of sites for restriction endonucleases obtained from Restriction Mapper web-page of USA Molecular Biology Resources (http://www.restrictionmapper.org, Last accessed date June 2006) by introducing the gene sequence of *bal* (Pohl et. al., 2002). The entire list of restriction enzymes both cutting and noncutting the *bal* gene were demonstrated in Appendix D and Table A.2.

4.2.2 Amplification of *bal* Gene by Polymerase Chain Reaction (PCR)

After synthesis of proper single-stranded primers, *bal* gene with *EcoR*I and *Xba*I restriction enzyme extensions (final length=1730 bp) was amplified by polymerase chain reaction (PCR). The PCR involves two oligonucleotide primers, *EcoRI-bal* Forward Primer and *XbaI-bal* Reverse Primer, which flank the DNA sequence that is to be amplified. The primers hybridize to opposite strands of the DNA after it has been denatured, and are orientated so that DNA synthesis by the polymerase proceeds through the region between the two primers. The extension

reactions create two double-stranded target regions, each of which can again be denatured ready for a second cycle of hybridization and extension. The third cycle produces two double-stranded molecules that comprise precisely the target region in double-stranded form.

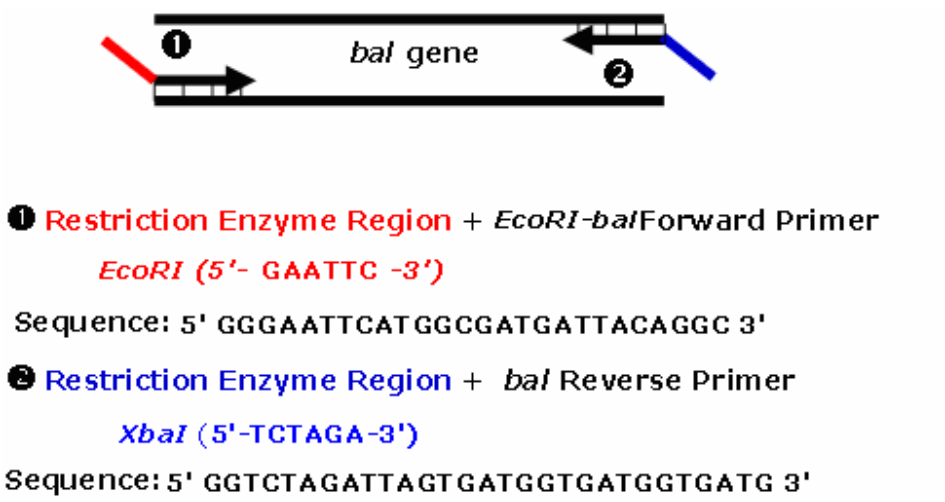


Figure 4.20 Schematic illustration of primers designed for expression in *P.pastoris* strains. 1) *EcoR*I RE sequence was associated in front of *bal* Forward Primer as a single-stranded tail. *bal* Forward Primer is complimentary to anti-sense strand of the *bal* gene; 2) *XbaI* RE sequence was associated in front of *bal* Reverse Primer as a singe-stranded tail *bal* Reverse Primer is complimentary to sense strand of the *bal* gene.

The annealing temperature and time are central variables in determining the specificity of a PCR. Hence, temperatures and times used vary depending on the sequences to be amplified. Time of amplification of each cycle, which can be calculated from the formula of 1 min per kb, was determined as 2 min to allow enough time for full amplification. The annealing temperature should be between the range of $(T_m\text{-}5) \leq T_a \leq (T_m\text{+}12)$. Melting temperature of both primers were arranged as 60.7°C (Appendix C), and thus the finest annealing temperature was determined as 60°C.

The parameters and the final compositions of the polymerase chain reaction mixture were given in Tables 4.8. As *Pfu* DNA polymerase having 3'=>5' exonuclease (proofreading) activity, was used for the accuracy of the *bal* gene.

In principle, physical and chemical components of PCR can be modified to produce a potential increase in yield, specificity and sensitivity. In all PCR experiments, control reactions, i.e., negative control which contains all reaction components except the template DNA in order to visualize whether any contamination occurred; or positive control which contains PCR product of previously amplified *bal* gene in order to make precise comparison with cloned gene, were included in gel electrophoresis analysis.

The template DNA of pUC18::*bal* plasmid was isolated and purified by the Plasmid Purification Kit (QIAGEN) before it was utilized as template, to remove RNA, proteins, and chemicals remaining after isolation, since these materials can capture Mg^{2+} ions present in PCR buffer that inhibit the DNA polymerase enzyme and reduce PCR yield. Figure 4.21 illustrates the uncut form of the template pUC18::*bal* plasmid.

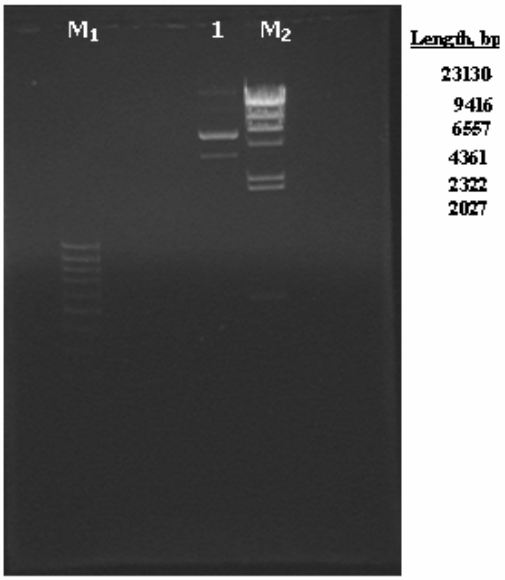


Figure 4.21 Agarose gel electrophoresis image of pUC18::*bal* plasmid. M_1: 50bp DNA ladder; M_2: λ DNA/HindIII Marker; 1: Uncut form of PUC18::*bal* plasmid

Table 4.8 PCR process parameters and components of reaction mixture of PCR for amplification of *bal* gene

Target DNA	PCR conditions	PCR Reaction Mixture Composition:
bal gene	1 cycle T_1 = 94°C, 4 min 30cycle T_1 = 94°C, 1 min; T_2 = 60°C,1 min; T_3 = 72°C, 2 min 1 cycle T_4 = 72°C,10min; T_5 = 4°C, 5 min	10XPCR Buffer 5µl dNTPs (1mM) 10µl FP (10µM) 1µl RP (10µM) 1µl Template DNA ~60ng pUC18::*bal* 1µl dH_2O 31µl *Pfu* DNA polymerase (2.5 U/µl) 1µl

The PCR results are shown in Figure 4.22.

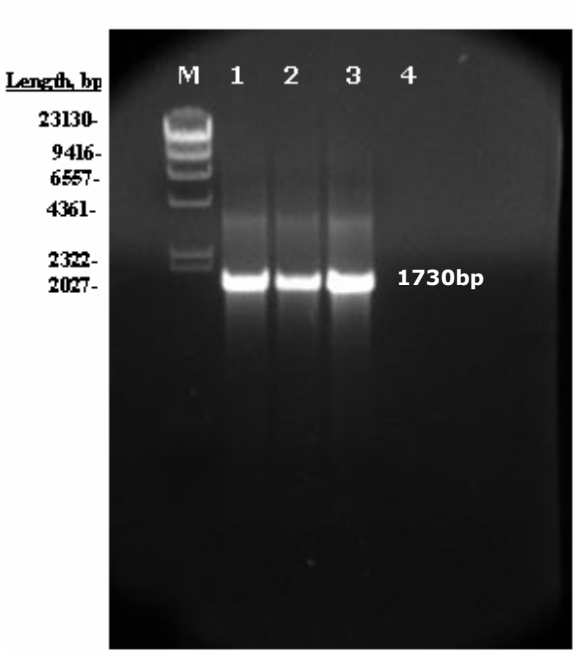


Figure 4.22 Agarose gel electrophoresis image of *bal* gene. M: λDNA/HindIII Marker; 1-3: BAL gene amplified with *Pfu* DNA polymerase; 4: Negative control (No template).

Figure 4.22 illustrates PCR results, after amplification with designed primers the *bal* gene was seen 1730bp in length, which was the expected size of the gene after extension.

4.1.3 Restriction Enzyme Digestion and Ligation

The PCR amplified 1.7kb *bal* gene purified by the PCR Purification Kit (QIAquick), and the pPICZαA vector isolated by the Plasmid Purification Kit (QIAGEN), were digested with *EcoRI* and *Xba*I restriction enzymes, separately, in specified buffers (Tables 4.9 and 4.10) of 20µl final volume at 37^{o}C for overnight. The reaction mixtures were kept at 65^{o}C for 20 min for the thermal inactivation of the enzymes used at the end of each restriction digestion. Between subsequent digestion reactions, processed *bal* gene and linear pPICZαA vector were purified by treating with NaAC (3 M) and EtOH (100%). After separating the fluid by centrifugation from the precipitates, both of the DNA fragments were dissolved in suitable amount of water. After digested with both restriction enzymes DNA fragments were extracted from agarose gels (Section 3.11.13).

Table 4.9 Components of reaction mixture of restriction digestion of *bal* gene and pPICZαA vector with *EcoRI* RE.

	bal gene	pPICZαA
Components	Amounts	Amounts
DNA fragment	~ 150 ng	~ 150 ng
EcoRI RE (10U/ µl)	1µl	1µl
10X Buffer *EcoRI*	2µl	2µl
dH_2O	up to 19µl	up to 19µl

Table 4.10 Components of reaction mixture of restriction digestion of *bal* gene and pPICZαA vector with *XbaI* RE.

	bal gene	pPICZαA
Components	Amounts	Amounts
DNA fragment	~ 150 ng	~ 150 ng
XbaI RE (10U/ µl)	1µl	1µl
10X Buffer Tango	2µl	2µl
dH_2O	up to 19µl	up to 19µl

The ligation reaction is the crucial step in genetic engineering techniques since this reaction requires the cohesive ends of foreign DNA and open plasmid DNA to bind in correct orientation. Therefore the ligation reaction was done by altering the gene/vector molar ratio between 3 and 5.

The amount of insert DNA to be added to the reaction mixture was calculated such that insert: vector ratio of 1:3 was achieved, as given in the equation.

$$100 \text{ ng vector} \times \frac{1.7 \text{ bp (}bal\text{ gene)}}{3.5 \text{ bp (vector)}} \times \frac{3}{1} = 145.1 \text{ ng } bal \text{ gene}$$

The ligation reactions were performed by incubating the reaction mixture of 10 µl final volume at 16°C for 16h.

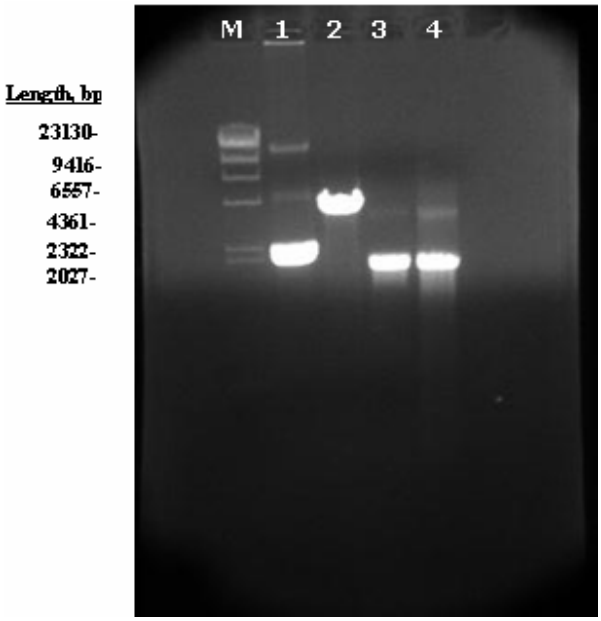


Figure 4.23 Agarose gel electrophoresis image of *bal* gene and pPICZαA vector after restriction digestion; M: λDNA/HindIII Marker; 1: undigested pPICZαA vector 2: pPICZαA vector digested with *EcoRI* and *XbaI*; 3: *bal* gene digested with *EcoRI* and *XbaI* and purified. 4: *bal* gene digested with *EcoRI* and *XbaI.*

Table 4.11 pPICZαA::bal ligation reaction conditions for 1:3 ratio.

10X ligation buffer	: 1µl
Insert DNA (1730 bp)	: 145 ng
Double digested vector DNA (3536 bp)	: 100 ng
T4 DNA ligase	: 1µl
Sterile dH_2O	: to 10 µl

The assembly of the recombinant plasmid, pPICZαA::*bal,* was schematically illustrated in Figure 4.24.

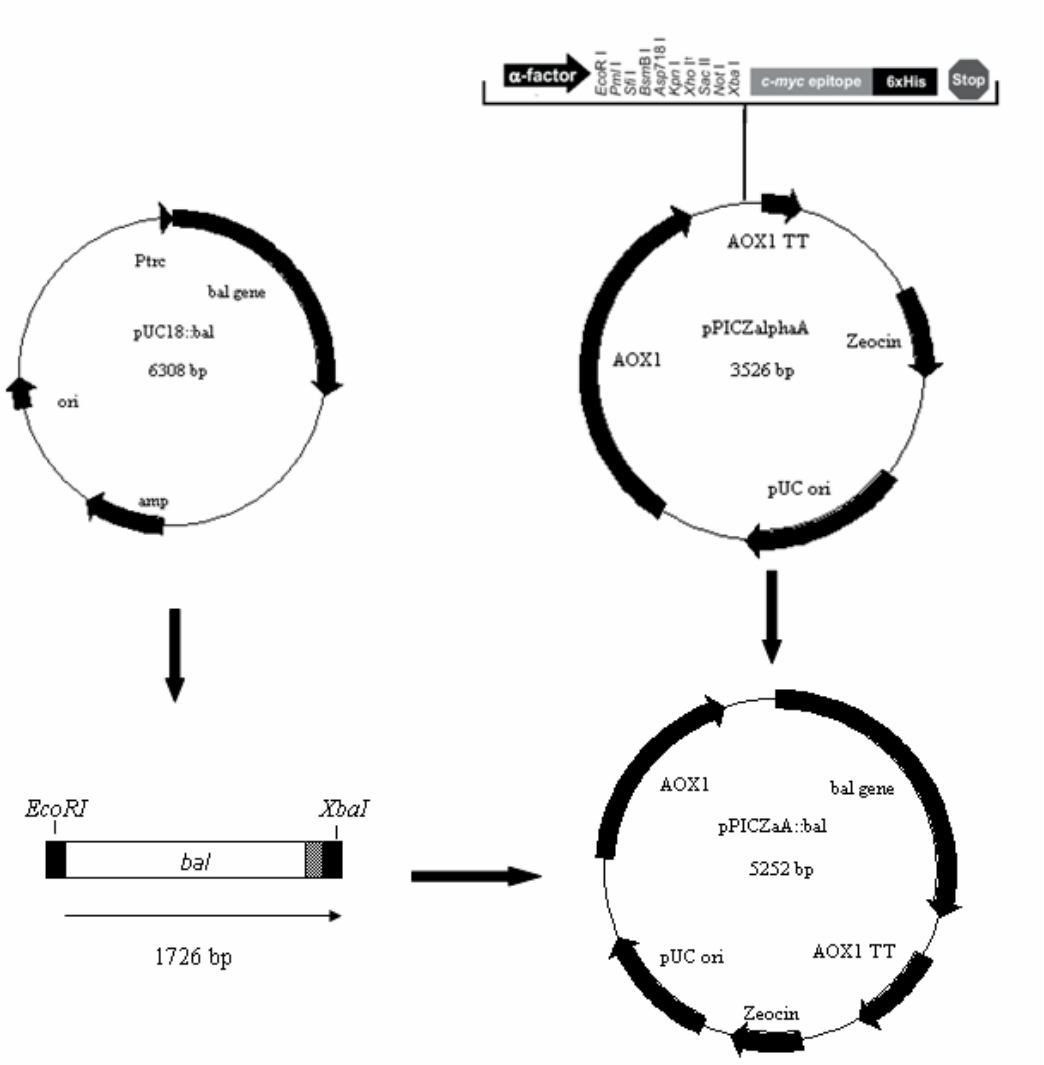


Figure 4.24 Scheme for the construction of the expression system for BAL production in *P.pastoris* strain. 1726 bp *bal* gene digested with *EcoRI* and *XbaI* restriction enzymes was cloned into pPICZαA vector from *EcoRI* and *XbaI* restriction sites. The transcription directions were indicated by arrows; Cross hatched region indicates the poly-his tag of the *bal* gene; TT, transcription termination region; ori, origin of replication in *E.coli*

4.2.4 Transformation of *E.coli* cells with pPICZαA::*bal* and Selection of the Transformants

To select and propagate the ligation products, *E. coli* JM109 strain was transformed by the $CaCl_2$ method described in section 3.12.8 and fresh transformants were grown on LSLB-agar + Zeocin (0.025 kg m^{-3}) for 12-18 h. Thirty two colonies were selected and further incubated for plasmid DNA isolation. The putative recombinant plasmids of three colonies were detected that they had molecular weights bigger than pPICZαA vector (Figure 4.25). These plasmids were plated for short term storage and further tests.

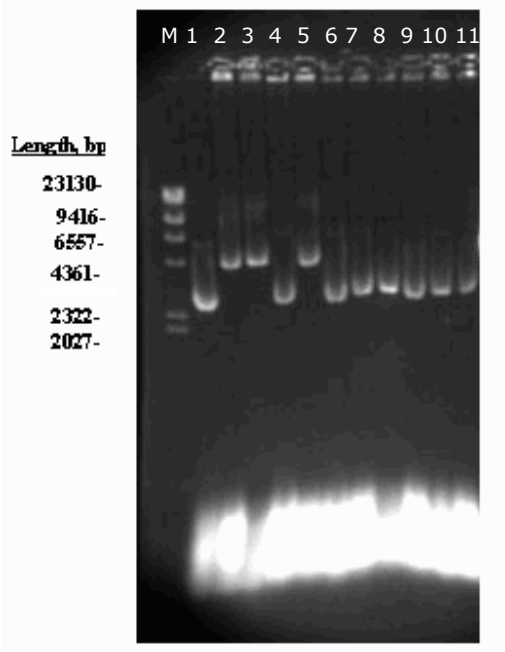


Figure 4.25 The putative recombinant plasmids M: λ DNA/HindIII Marker; 2-3-5: Putative recombinant plasmids

The putative recombinant plasmids were first controlled with subsequent restriction digestion of putative recombinant plasmids with *EcoRI, XbaI* and *Pst*I restriction enzymes. Components of restriction digestion of *bal* gene by restriction enzymes were given in Table 4.12. *Pst*I is cutting the *bal* gene from the 938th nucleotide and thus remaining part which is 778bp in length was expected to be observed in the reaction medium.

Table 4.12 Components of reaction mixture of restriction digestion of putative recombinant pPICZαA vector with *EcoRI, XbaI* and *Pst*I REs.

Components	Amounts	Components	Amounts	Components	Amounts
DNA fragment	~ 120 ng	DNA fragment	~ 120 ng	DNA fragment	~ 120 ng
*Xba*I RE (10U/ µl)	1.5µl	*Pst*I RE (10U/µl)	1.5µl	*EcoR*I RE (10U/ µl)	1.5µl
10X Buffer Tango	2µl	10X Buffer Tango	2µl	Buffer *EcoR*I	2µl
dH_2O	up to 19µl	dH_2O	up to 19µl	dH_2O	up to 19µl

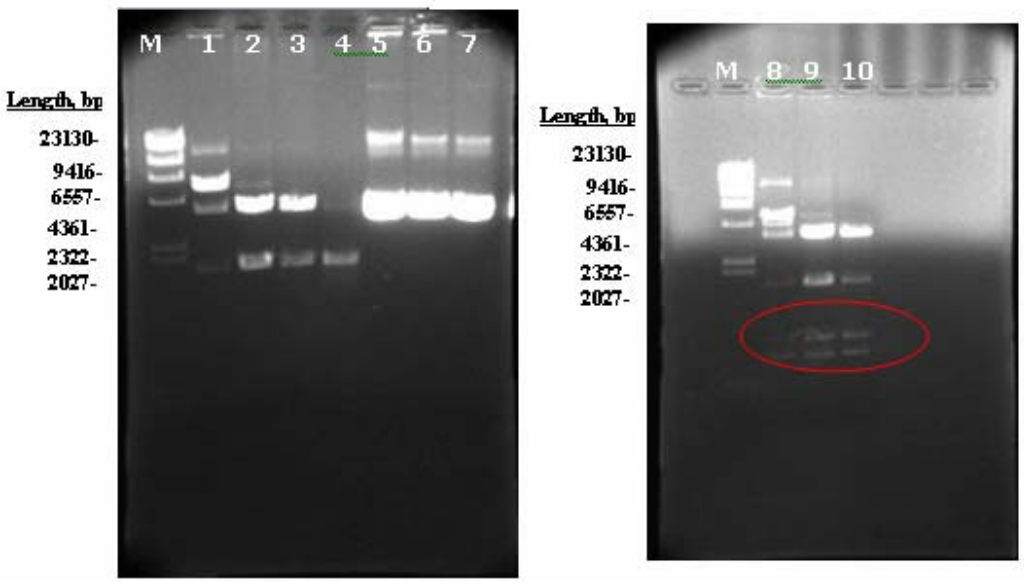


Figure 4.26 Agarose gel electrophoresis image of pPICZαA::*bal* plasmid after restriction digestion. M: λDNA/HindIII Marker; 1: First putative pPICZαA::*bal* plasmid digested with both *EcoRI* and *XbaI* REs; 2: Second putative pPICZαA::*bal* plasmid digested with both *EcoRI* and *XbaI* REs; 3: Third putative pPICZαA::*bal* plasmid digested with both *EcoRI* and *XbaI* REs; 4: *bal* gene amplified from *pUC18::bal* plasmid; 5-7: Undigested putative pPICZαA::*bal* plasmids; M: λDNA/HindIII Marker; 8-10: Putative pPICZαA::*bal* plasmids digested with *EcoRI*, XbaI *and PstI* REs;

The *EcoRI* and *XbaI* digestion reaction results were expected same bp in length as *bal* gene; 1730 bp. The gel electrophoresis results prove that *bal* gene was inserted into the pPICZαA plasmid (Figure 4.26).

The putative recombinant plasmids were further checked by PCR for the presence of the insert. The primers and expected PCR products listed in Table 4.13. Agarose gel electrophoresis image (Figure 4.27) showed that *bal* gene is present in these plasmids.

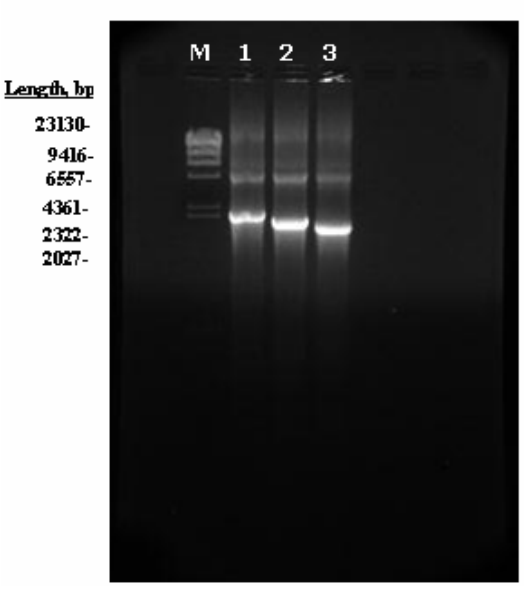


Figure 4.27 Agarose gel electrophoresis image of *PCR products of* pPICZαA::*bal* plasmid; M: λDNA/HindIII Marker; 1: PCR product *of* pPICZαA::*bal* plasmid with 5'-AOX1 forward primer and *XbaI-bal* reverse primer 2: PCR product *of* pPICZαA::*bal* plasmid with 3'-AOX1 reverse primer and *EcoRI-bal* forward primer 3: PCR product *of* pPICZαA::*bal* plasmid with *EcoRI-bal* forward primer and *XbaI-bal* reverse primer (bal gene).

Table 4.13 Primers and expected PCR products length

Forward primer	Reverse primer	Expected length
5' AOX primer	*bal* gene primer	2085
bal gene primer	3' AOX primer	1899
bal gene primer	*bal* gene primer	1726

The DNA sequence of insert DNA was further controlled by automatic DNA sequencers, (Microsynth GmbH, Switzerland) using AOX primers. The results proved that cloning was successful and the inserted sequence was same with the *bal* gene.

4.2.5 Transformation of *P.pastoris* cells with pPICZαA::*bal* plasmid

The pPICZαA vector is an integration vector but has to be linearized at its AOX promoter region so that it can be integrated into the genome from the AOX locus. *Sac*I restriction enzyme was chosen as a single-cutter for linearization of pPICZαA::*bal* plasmid. *pPICZαA::bal* was plasmid digested by *Sac*I and full digestion was verified by agarose gel electrophoresis. Then the digestion product was purified by the extraction from agarose gel. The concentration of the digested plasmid solution was adjusted 0.1-0.2 µg/µl for efficient transfection.

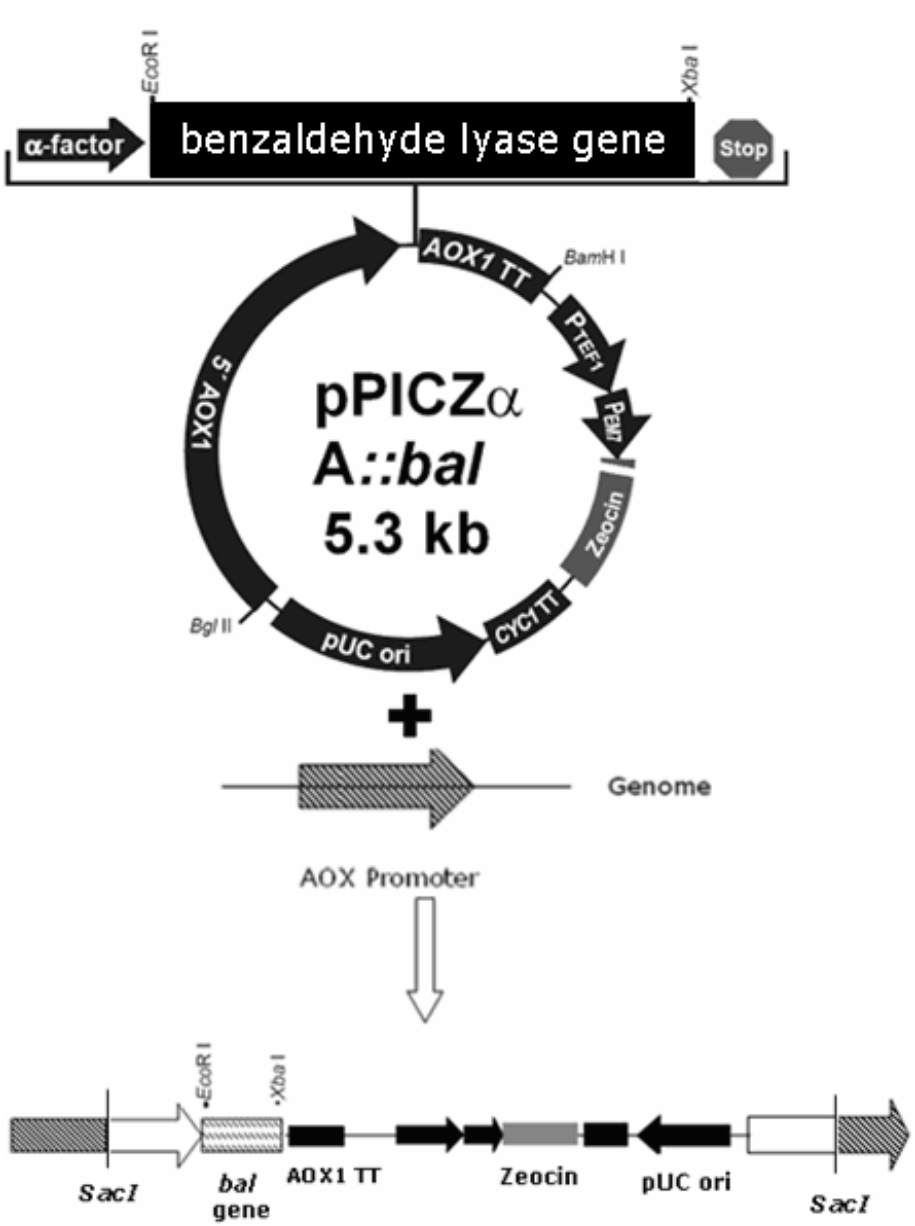


Figure 4.28 Schematic representation of pPICZαA::*bal* integration into *P. pastoris* genome. The recombinant plasmid was digested with *Sac*I from AOX promoter region, yielding a linearized plasmid having homologous regions to the AOX promoter in the genome at both ends. After integration of the plasmid to the genome, there are two functional copies of the AOX promoter in the genome

Transfection was performed as explained in Section 3.11.11. After 60 h of incubation at 30^0C, six single colonies were selected for further controls. The selected colonies were inoculated onto YPD + Zeocin (0.100 kg m^{-3}) plates for short term storage and then inoculated into 10 ml YPD + Zeocin (Appendix A) medium. Their genomic DNA was isolated to be used in controls by PCR.

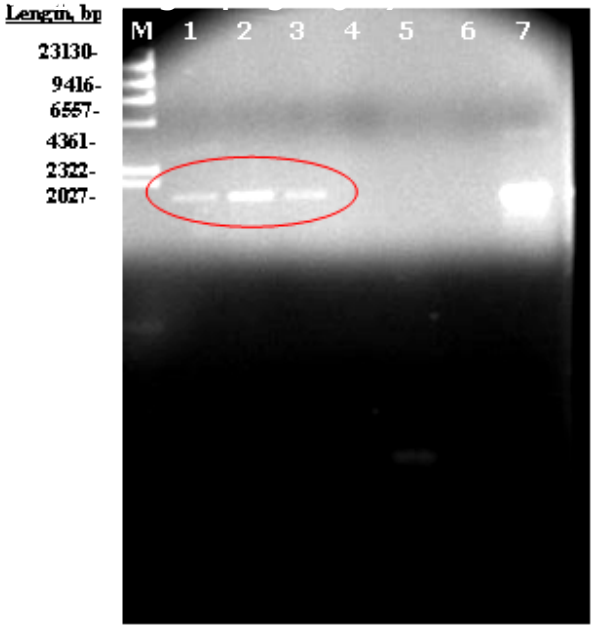


Figure 4.29 Agarose gel electrophoresis image of PCR products *of bal gene insert* from genomic DNA *P.pastoris transformants;* M: λDNA/HindIII Marker; 1: PCR product of first colony; 2: PCR product of second colony 3: PCR product of third colony; 4: PCR product of fourth colony; 5: PCR product of fifth colony; 6: PCR product of sixth colony; 7: bal gene as positive control

Among six putative recombinant colonies only three of them showed positive results in PCR reaction. These three colonies were compared according to their cell concentration and benzaldehyde lyase production capacity, and then one of them was chosen as a producer of extracellular benzaldehyde lyase enzyme.

4.3 Enzyme Production and Activity Analysis

4.3.1 SDS-PAGE Analysis of Benzaldehyde Lyase Excreted

The extracellular benzaldehyde lyase enzyme produced by the recombinant *P.pastoris* cells was analyzed with SDS-PAGE. The recombinant *P.pastoris* cells inoculated in the reference production medium (section 3.5.2) were harvested at 24^{th} h of the fermentation by centrifugation and supernatant was used for further analysis.

After the production, benzaldehyde lyase was partially purified and concentrated by ultrafiltration (Section 3.7). Since the monomer form of benzaldehyde lyase enzyme is 59 kDa, 10 kDa and 30 kDa ultrafiltration membranes were used to loss of the enzyme. Regenerated cellulose was preferred rather than polyethersulfone membranes, because they give the highest possible retention with the lowest possible adsorption of the protein.

The protein containing solution was kept at 2-8°C to prevent proteolytic degradation. In addition to this, ultrafiltration experiments were carried out in the cold room (4°C) and the samples were kept on ice during the whole process.

The concentrated benzaldehyde lyase proteins containing a polyhistidine tag extension were purified under native conditions with QIAexpress Nİ-NTA spin columns (Figure 4.30) and with cobalt-based metal affinity resins (BD Talon) (Figure 4.31).

Analytical electrophoresis of benzaldehyde lyase was carried out in polyacrylamide gel by dissociating the protein into its individual polypeptide subunits by treating mixture with reducing agent SDS followed by heat exposure. The enzyme is composed of 4 x 563 amino acid residues and has a molecular weight of 4 x 58919 Da. (Mosbacher et. al., 2005). After electrophoresis, separated proteins were stained with silver salts. The enzyme migrated in sodium dodecyl sulfate- containing polyacrylamide gel as a single polypeptide with a molecular weight of ~59 kDa matching with molecular weight observed for previously utilized *E. coli* K12 strain carrying r-pUC18::*bal* plasmid (Çalık et. al., 2004).

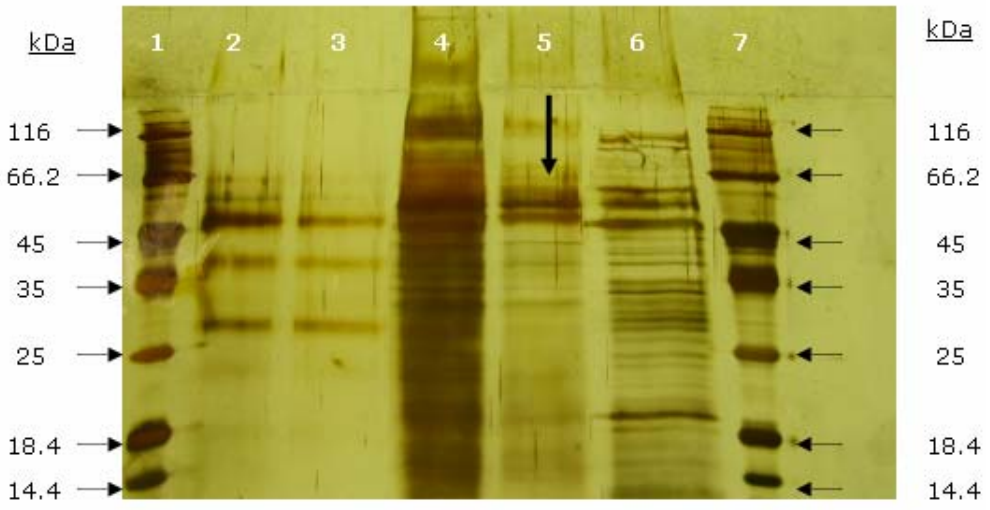


Figure 4.30 SDS-PAGE analysis of recombinant *P.pastoris* extracellular products: 1: Protein Marker; 2: crude extract; 3: crude extract; 4: concentrated supernatant; 5: purified benzaldehyde lyase produced from *P.pastoris;* 6: second elution of spin column; 7: Protein Marker

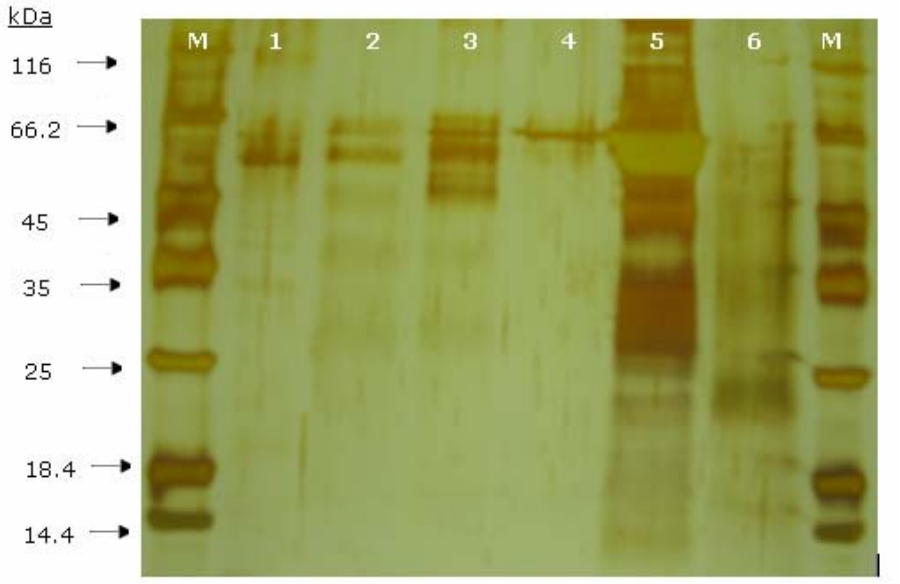


Figure 4.31 SDS-PAGE image of benzaldehyde lyase enzyme: M: Protein Marker; 1: crude extract; 2: BD Talon purified benzaldehyde lyase produced from *P.pastoris*; 3: second elution of purified benzaldehyde lyase produced from *P.pastoris;* 4: third elution of BD Talon; 5: purified benzaldehyde lyase from *E.coli* K12 strain carrying pUC18*::bal;* 7: Control untransformed *P.pastoris; M*: Protein Marker

SDS-PAGE results illustrated that benzaldehyde lyase enzyme was cloned into *P.pastoris* and the recombinant *P.pastoris* produced the enzyme extracellularly. As SDS-PAGE analyses determines the monomeric forms of the proteins not the multimeric forms of the protein to determine the multimeric form of the enzyme in the supernatant FPLC analysis was done.

SDS-PAGE results showed that his-tag purification was not a 100% efficient method, spin columns not only purify benzaldehyde lyase enzyme but also some other proteins, which should have histidine residues in their structure. In Figure 4.32 it can be seen that after purification with spin columns the eluted enzyme lose its concentration in elution gradually and the other bands became more concentrated. Thus one of enzymes produced by recombinant *P.pastoris* might be a protease. In the literature protease problem of *P.pastoris* was also reported. (Cereghino and Cregg, 2000).

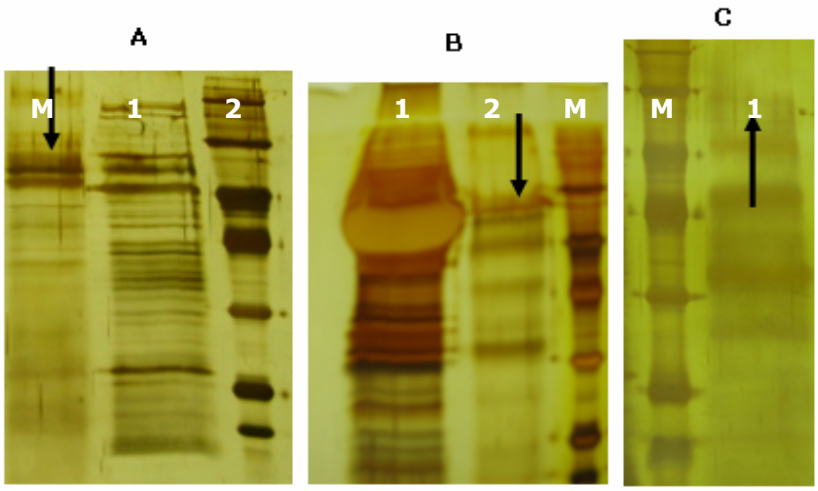


Figure 4.32 SDS-PAGE images of proteolytic activity. A gel: 1: First day of the purification; 2: Second elution of spin columns; M: Protein Marker, B gel: 1: *E.coli* BL21 cells produced and purified BAL enzyme; 2: Second day of the purification; M: Protein Marker C gel: M: Protein Marker; 1: Third day of the purification

Figure 4.32 illustrates the putative proteolytic activity after purification of the enzyme. After purification the BAL band was lost its concentration while the other bands increased in concentration.

4.3.2 FPLC Analysis of Protein Produced by *P.pastoris*

To determine the extracellular proteins produced by recombinant *P.pastoris* cells, fermentation broth was further analyzed with FPLC. The recombinant *P.pastoris* cells inoculated in the reference production medium were harvested at 24^{th} h by centrifugation and supernatant was used. The first step of process after production was the freeze-drying of the supernatant (Section 3.10). Freeze-drying process was done approximately 7 h, until the sample's weight became constant. Thereafter, the sample was solved in 1 ml mobile phase (PBS Buffer) which resulted in 25 fold concentrated fermentation broth. Finally sample was filtered through 0.20 µm filters (Sartorius) and loaded to the Hiload 16/60 superdex 200pg column that separate proteins according to their molecular weight. Flow-rate was 0.5 ml min^{-1} and detection was carried out with an UV detector at 280nm. The process was carried out at $+4^0$C.

Table 4.14 Molecular Weights of Discrete Proteins of Protein Marker for FPLC and elution times

	Molecular Weight(kDa)	Time (min)
Ferritin	440	103
Albumin	67	140
Riba MA	13.7	198

Their elution times are illustrated in Figure 4.33 in red circles.

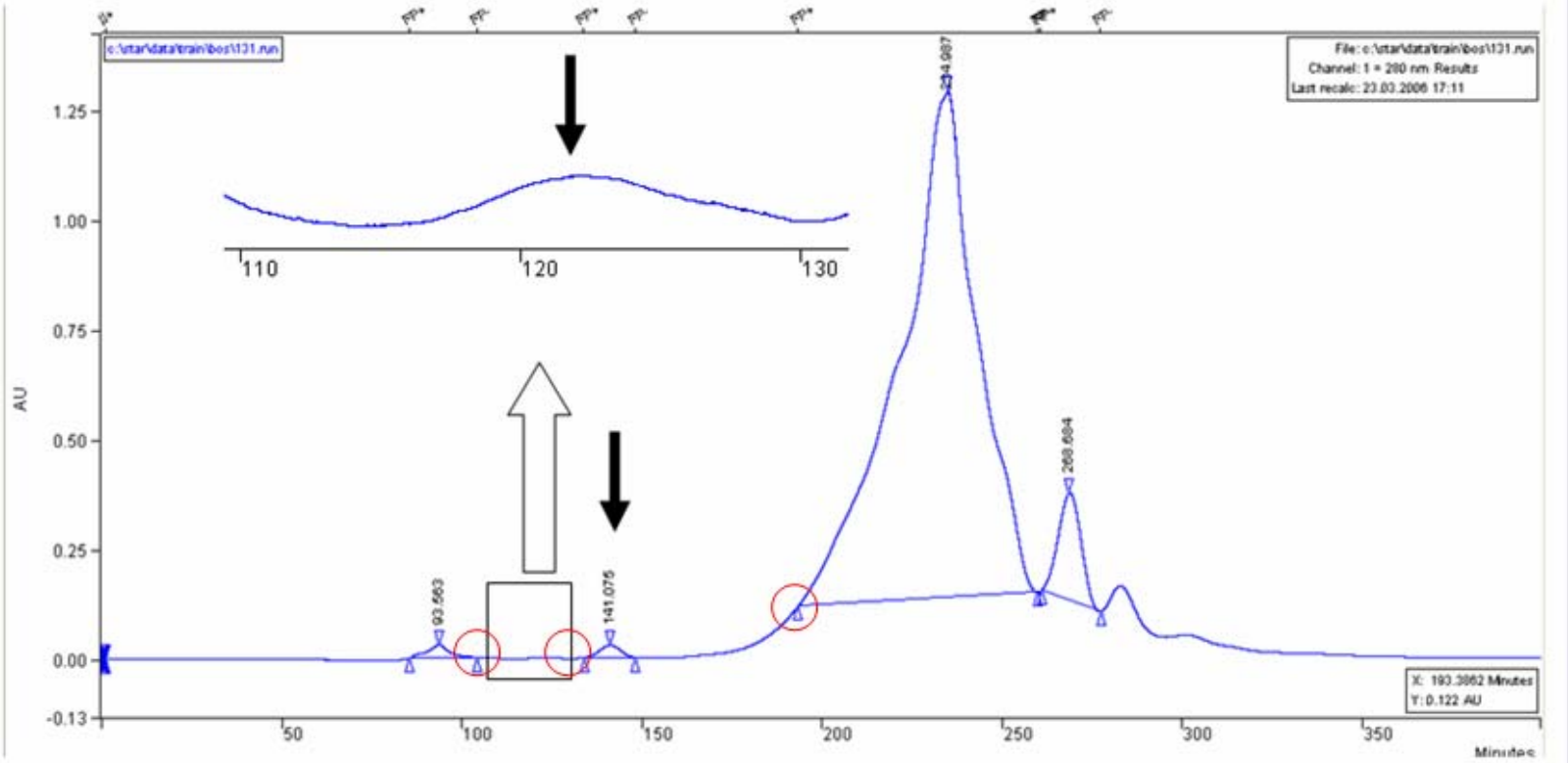


Figure 4.33 FPLC Chromatogram of the protein produced by recombinant *P.pastoris* strain with Hiload 16/60 superdex 200pg column at +4°C, 0.5 ml min^{-1}.Sample was withdrawn at 24th h of production medium

Figure 4.33 illustrates that the peak at 141 min belongs to monomer benzaldehyde lyase enzyme and the second peak, which eluted around 115 min belongs to tetramer benzaldehyde lyase enzyme. FPLC result showed that the benzaldehyde lyase enzyme produced extracellularly by recombinant *P.pastoris* as mainly monomer, thus extracellular benzaldehyde lyase enzyme could not gain its active tetramer form in the fermentation medium, as the concentration of tetramer form was very low.

4.3.3 Thin Layer Chromatography (TLC) Analysis of Benzoin Condensation Reaction

Thin layer chromatography was used to follow the condensation reaction of benzaldehyde to benzoin (Section 3.6). Samples were taken from reaction mixtures at t= 1 h and t=24 h of the reaction. Benzaldehyde, benzoin and reaction products were analyzed by TLC on F_{254} gel plates (Merck) and developed in Ethyl acetate: Chloroform [3:1 (vol/vol)].

Benzoin condensation reaction products were analyzed by TLC analysis. In Table 4.15 the reaction components are shown. The second condensation reaction components were the optimum concentrations of the reaction components found by Maria et.al., 2005.

Table 4.15 Components for benzoin condensation reaction

Component	First	Second
ThDP*	1 mM	1 mM
$Mg2^{+}$	0.5 mM	0.5 mM
DMSO**	20%	30%
Benzaldehyde	90 mM	100mM

*ThDP: Thiamin diphosphate

**DMSO: Dimethyl sulfoxide

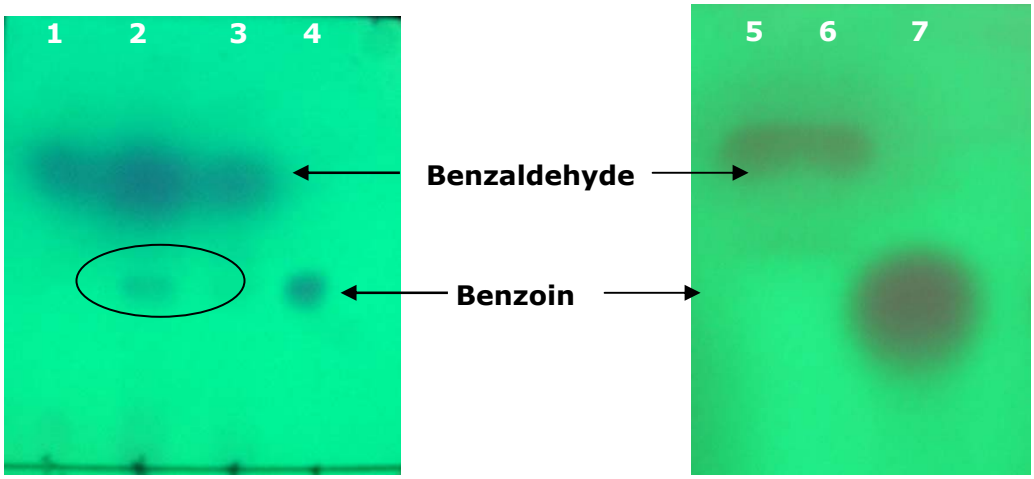


Figure 4.34 TLC results of biotransformation reaction. 1: Benzaldehyde standard; 2: First reaction mixture; 3: Second reaction mixture; 4: Benzoin standard; 5: Benzaldehyde standard; 6: Untransformed *P.pastoris*; 7: Benzoin standard

In Figure 4.34, the positions of the spots were seen as viewed under 254 nm UV light after the run. Benzaldehyde standard, benzoin standard and reaction products can be seen In Figure 4.34. On the other hand, no product was detected by TLC at t= 1 h of the reaction.

Figure 4.34 illustrates that the extracellular benzaldehyde lyase enzyme reacted with its substrate benzoin. And the optimum benzaldehyde concentration found by Maria et.al., 2005 might not be the optimum concentration for extracellular enzyme. Thus, concentration of the substrate, benzaldehyde lyase, is indeed important, as expected. The samples were further analyzed with GC-MS to prove the TLC results.

4.3.4 GC-MS Analysis of Benzoin Condensation Reaction

The Samples, which was used in TLC analysis, were further analyzed in GC-MS analysis. The reaction medium was extracted with Ethyl acetate (4X 50ml). The organic phase was washed several times with water, to remove residual DMSO. After drying the organic layer with $MgSO_4$ and removing the solvent under reduced pressure, the products were purified with a silica column. The initial temperature 60^0C was constant during 2 min, and then a slope of 5^oC

min^{-1} was set until 240^0C. Thereafter the temperature 240^0C was constant for 12 min. Determination of configuration was done by comparing to both relative retention times and mass spectrum of the compounds.

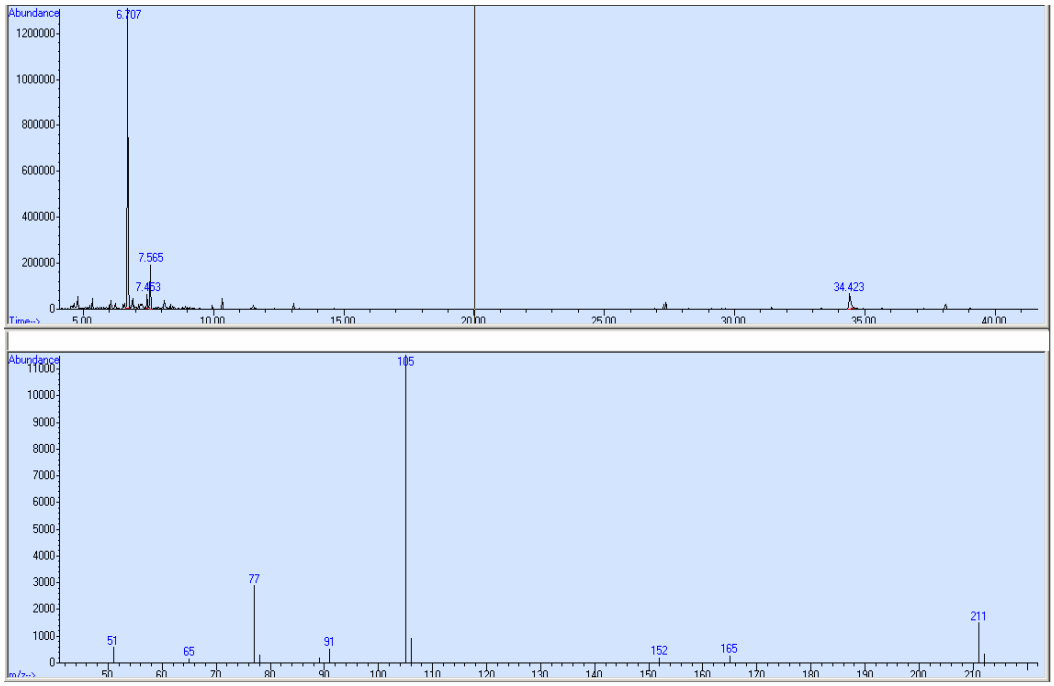


Figure 4.35 GC-MS Chromatogram of benzoin transformation for the first benzoin condensation reaction

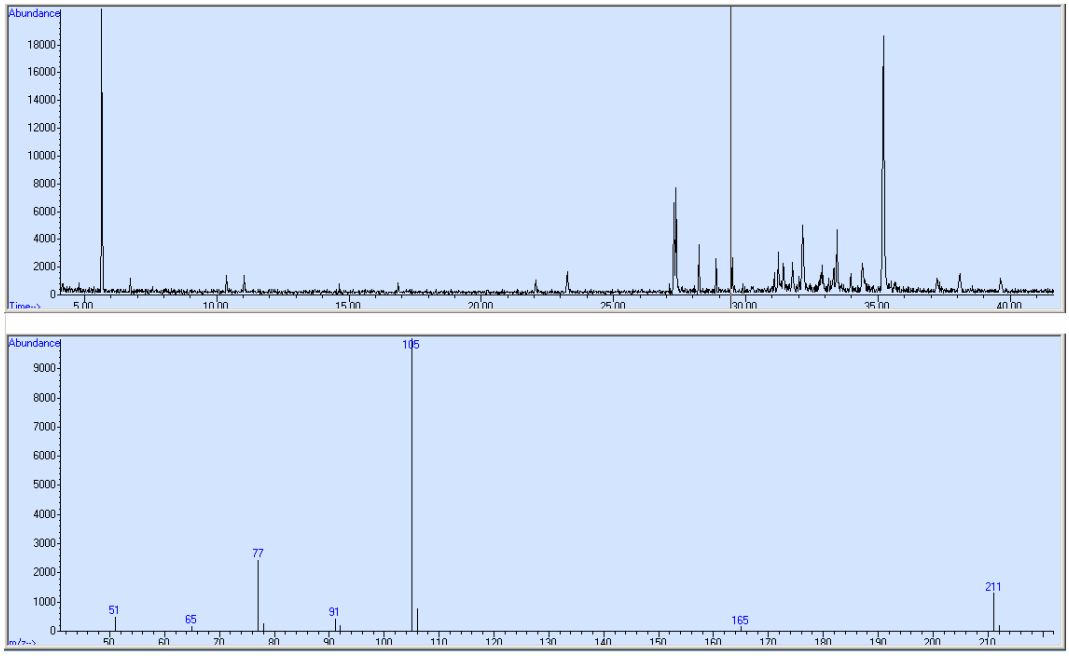


Figure 4.36 GC-MS Chromatogram of benzoin transformation for the second benzoin condensation reaction

Figure 4.35 and Figure 4.36 illustrate the retention times of benzoin and benzaldehyde, also their mass spectrum eluted from Zebron column (60 m). The 211, 195, 181,165, 135 and 119 peaks are the molecular spectrum of benzoin. Benzoin was eluted from the column at 28-34 m. The 105, 90, 77 and 65 peaks are the molecular spectrum of benzaldehyde. Benzaldehyde was eluted from the column at 18-25 m. As observed from the figure, benzoin condensation reaction took place but benzaldehyde molecules are not completely converted to benzoin.

GC-MS results and TLC results for the condensation of benzoin showed positive results. The extracellular benzaldehyde lyase enzyme reacted with benzaldehyde and converted it into benzoin.

4.3.5 Benzaldehyde Production

The recombinant *P.pastoris* strain was inoculated in YPD+Zeocin agar plates for 48-72 h at 30^0C. Then a single colony was chosen and inoculated in BMGY precultivation medium at 30^0C and N=225 min^{-1} for 24 h. The cells were harvested by centrifugation at 4000 rpm, 10 min^{-1} at $+4^0$C and re-suspended in BMMY production medium. The cells were incubated 30^0C and N=225 min^{-1} for 72 h. Every 24h in BMMY medium methanol induction was done. But for most of processes production finished after 24 h.

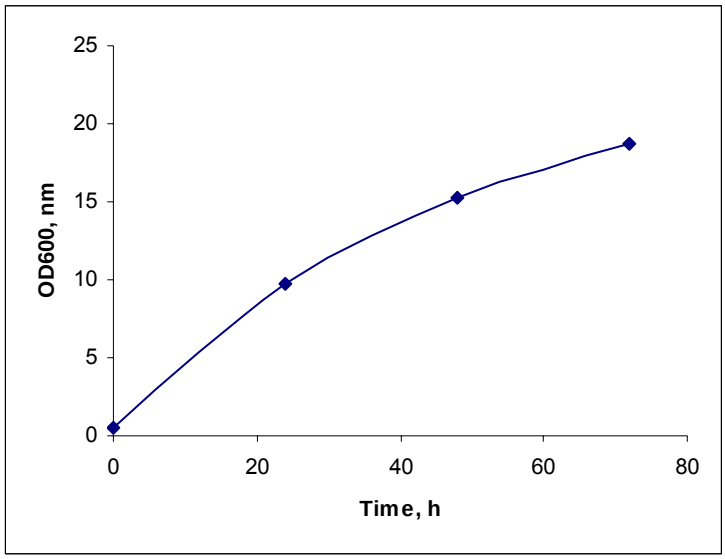


Figure 4.37 Optical densities at 600 nm of the BMMY medium during production phase of *P. pastoris* cultivation. Every 24 h methanol induction

4.3.6 Benzaldehyde Lyase Activity

According to the research program for the extracellular production of benzaldehyde lyase, the recombinant pPICZαA::*bal* plasmid was transfected into *P.pastoris* strain and expression capacity was investigated. The cells were separated with centrifugation and supernatant was used for analysis of SDS-PAGE, TLC analyses and activity assay.

Benzaldehyde lyase activity, determined by measuring the conversion of benzoin into benzaldehyde, was 470 U cm^{-3} at 24^{th} in the BMMY production medium. This result was obtained by modified activity assay explained in section 3.6.2.

The recombinant enzyme showed activity just after collected from the fermentation broth, which detected with spectrophotometer; however, in a short time it loses its activity. On the other hand, in SDS-PAGE analyses, the monomer benzaldehyde lyase enzyme detected for all the processes. These results reflect that the extracellular benzaldehyde lyase enzyme was produced by recombinant *P.pastoris* but it loses its tetrameric form in a few hours.

As previously stated, Mosbacher et. al. (2005) was studied the X-ray structure of benzaldehyde lyase to show that BAL belongs to a group of closely related ThDP dependent enzymes. According to the results, it was proposed that BAL is a homotetramer, where each subunit is composed of 4 x 563 amino acid residues and has a molecular weight of 4 x 58919 Da. The FPLC results (section 4.2.2) illustrated the monomeric form of the enzyme much higher then the tetrameric form of the enzyme. This result reflects that either the enzyme lost its tetrameric form or could not fold into tetrameric form or this may be the most crucial reason for observing less BAL activity in the fermentation broth, where volume of the bulk liquid is incomparable with the volume of the cell. Therefore, coalescence probability of each monomer in such a huge environment, even if all other factors like salts, by-products produced or other chemicals present in the medium are assumed to be negligible, to form tetramer structure may be the limiting factor on BAL activity. On the other hand, in the study of Jungo et. al.,

2006 the homotetramer avidin protein cloned in *P.pastoris* strain and the extracellularly produced protein showed activity. But avidin is already a glycoprotein and has a very strong affinity for biotin, which avidin bind, with a kD (dissociation constant) approximate to 10^{-15} M^{-1}, the highest known between any ligand and a protein (Bayer et.al., 1980). Each subunit of avidin has a molecular weight of 17 kDa and the active form (tetrameric form) of recombinant avidin is 65.4 kDa (Zocchi et.al., 2003). The tetrameric form of avidin has almost same molecular weight of monomers of benzaldehyde lyase.

In the study of Maria et. al. (2005), it was shown that the active form of BAL in homotetramer structure losses its activity almost completely within 3 h in water, and hence, carboligation reactions were performed in potassium phosphate buffer supplemented with ThDP, Mg^{2+}, DTT as a cofactor and 30% DMSO as a cosolvent. In this respect, the production rate of the enzyme by *P.pastoris* may be much smaller than the inactivation rate of enzyme since enzyme losses its active form within 3 h in aqueous solutions like fermentation broth. Furthermore, since inactive forms of BAL, monomers, are less stable and sensible, expressed amount, may directly degraded by proteases.

P. pastoris which is the most commonly used and studied yeast for foreign protein production. But *P. pastoris* glycosylates the heterologous proteins (Section 2.3.1). Benzaldehyde lyase enzyme has two possible regions for the N-linked glycosylation, which are 474.-475.-476. and 501.-502.-503. amimoacids. In the study of Mosbacher et. al. (2005) was studied the X-ray structure of benzaldehyde lyase. According to the results each subunit binds one ThDP molecule using one Mg^{+2} ion. And according to the results of this study, 474. aminoacid asparagine and 476. amino acid serine have roles in the ThDP binding domain. Glycolysation may occur at this side and it might prevent the ThDP binding on the enzyme. Therefore the enzyme might not react with benzaldehyde properly. But the enzyme showed activity for both volumetric and condensation reaction. So glycosylation may be occurred in the enzyme sequence but it does not interfere the ThDP binding. All possible glycosylation sites listed in Appendix F.

Although expression of heterologous proteins can be done in shake-flask culture, protein levels are typically much higher in fermenter cultures. Because

the mineral media for *P. pastoris* (containing only glycerol or methanol, biotin, salts and trace elements) are economical and well defined, this yeast is nearly ideal for large-scale production of heterologous proteins in fermenters. Most importantly, it is only in fermenters, where parameters such as pH, aeration and carbon source feed rate can be controlled, that it is possible to achieve ultra-high cell densities (>100 kg m^{-3} dry cell weight; >400 kg m^{-3} wet cell weight) (Cereghino et. al.,2002). In this current study, all processes were done in shake-flask culture therefore the enough enzyme production might not be reached for the monomer concentration to form tetrameric enzyme.

CHAPTER 5

CONCLUSION

In this study, extracellularly benzaldehyde lyase production capacity in recombinant *P.pastoris* was aimed. For this purpose, the gene encoding BAL with poly-His tag extension was cloned and expressed in *P.pastoris* strain under the control of strong AOX promoter; and extracellular benzaldehyde lyase production was investigated in the laboratory scale bioreactors. In this context the following conclusions were drawn:

1. For extracellular expression of BAL in *Pichia pastoris*, *bal* forward and *bal* reverse primers for amplification of *bal* gene were designed. The PCR product cloned into pPICZαA vector, *E. coli*/ *Pichia pastoris* shuttle vector after restriction digestion with proper restriction enzymes. Recombinant pPICZαA*::bal* plasmid was transfected into *Pichia pastoris* for extracellular expression.

2. Benzaldehyde lyase volumetric activity, determined by measuring the conversion of benzoin into benzaldehyde, was 470 U cm^{-3} at 24 h in the BMMY production medium; T=30^0C and N=225 min^{-1}, V=50 cm^3.

3. Benzaldehyde lyase volumetric activity, determined by measuring the conversion of benzoin into benzaldehyde, was measured zero at 48 h in the BMMY production medium. This result indicated that the fermentation medium contaminated with *Bacillus species*. After that chloramphenicol added to the precultivation and the production medium. But more importantly the result illustrates that the new adapted benzaldehyde lyase volumetric activity method detect conversion properly without influenced with yeast extract in the fermentation medium.

4. The SDS-PAGE results for all the processes illustrated that after purification with His-Tag column or BD-Talons, the monomers of the benzaldehyde lyase enzyme were detected in polyacrylamide gels. Suggesting that the extracellular enzyme was produced by the recombinant *P.pastoris* strain. And SDS-PAGE results also showed that the His-Tag column did not only purify the BAL enzyme but also purify the putative proteases.

5. TLC analysis results, done at 24^{th} of the production medium, illustrated that the extracellular benzaldehyde lyase enzyme reacted with the benzoin and convert it into benzaldehyde molecules.

6. GC-MS results illustrated that the benzoin condensation reaction was taken place, benzaldehyde turn into benzoin molecule by means of the extracellular enzyme. Therefore GC-MS and TLC analyses results showed that extracellular benzaldehyde lyase enzyme produced by recombinant *Pichia pastoris* could fold into its active tetramer form in the fermentation medium.

7. The FPLC results showed that the tetrameric form of the benzaldehyde lyase enzyme was much less than the monomeric form of the enzyme. This could reflect both the monomeric structures could not fold into multimeric form of the enzyme in the fermentation medium or the active tetrameric form splited into its

monomeric forms because the enzyme loses its almost all activity in three hours.

8. The highest benzaldehyde lyase volumetric activity was obtained where the highest cell concentration was obtained. This result indicates that, extracellular benzaldehyde lyase production capacity could be enhanced by increasing the cell concentration by using the pilot scale bioreactor.

9. Cloning experiments for the production of benzaldehyde lyase enzyme extracellularly by *Bacillus subtilis;* three different plasmids were used. Only *pMK4* cloning experiments were successful, but DNA sequencing revealed that eight nucleotides were missing in the *bal* gene and this deletion caused the frameshift, which changed the aminoacid sequence. Therefore all the *Bacillus subtilis* cloning experiments were unsuccessful.

10. *Bacillus subtilis* cloning experiments with pDG148-*Stu* plasmid showed that ligation independent cloning is not an efficient cloning method.

11. On the basis of the results of the benzaldehyde lyase volumetric activity, SDS-PAGE, FPLC, TLC, GC-MS assays; for extracellular benzaldehyde lyase production by *P.pastoris*, the following strategies can be proposed:

    - For the homomultimeric enzymes, intracellular production should be chosen to improve the folding of the enzyme.

    - The enzyme loses almost its activity in 3 h. This may be the result of that monomers separates each-other in time. The situation should be investigated by SDS-PAGE and FPLC analyses.

- For determination of tetrameric structure the Native-PAGE experiments should be done.

- His-Tag tail should be removed from the enzyme and the activity of the enzyme should be investigated after that.

- The stability of the enzyme should be increased by means of changing amino acids.

REFERENCES

Adam, W., Lazarus, M., Möller, C., Schreier, P., 1999. Biocatalytic Synthesis of Optically Active α-Oxyfunctionalized Carbonyl Compounds, Accounts of Chemical Research, 32, 837-845.

AgileBio, http://www.justbio.com, Last accessed date june 2006.

Atkinson, B., Mavituna, F., 1991. Biochemical Engineering and Biotechnology Handbook, 2nd ed., Macmillan Publishers Ltd., England.

Bailey, E. J., Ollis, F. D., 1986. Biochemical Engineering Fundamentals, 2nd ed., McGraw-Hill, Inc., Singapore.

Bayer, E. A. and Wilchek, M., 1980. The use of the avidin-biotin complex in molecular biology. Methods Biochem. Anal. 26, 1-45.

Breuer, M., Ditrich, K., Habicer, T., Hauer, B., Kebeler, M., Stürmer, R. And Zeliski, M., 2004. Industrial Methods for the Production of Optically Active Intermetiates, Angew. Chem. Int. Ed., 43, 788-824

Burke, D., Dawson, D., Stearns, T. *Methods in Yeast Genetics: 2000 Edition*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA.

Charoenrat, T., 2005. Process Design for Production of Thai Rosewood β-glucosidase in *Pichia pastoris*, Licentiate Thesis.

Çalık, P., Kalender, N., Özdamar, T. H., 2003-a. Overexpression of Serine Alkaline Protease Encoding Gene in *Bacillus* Species: Performance Analyses, Enzyme and Microbial Tech., 33, 967-974.

Çalık, P., Tomlin, G. C., Oliver, S. G., Özdamar, T. H., 2003-b. Overexpression of Serine Alkaline Protease Gene in *Bacillus licheniformis* and Its Impact on Metabolic Reaction Network, Enzyme and Microbial Tech., 32, 706-720.

Çalık, P., Yılgör, P., Ayhan, P., Demir, A. S., 2004. Oxygen Transfer Effects on Recombinant Benzaldehyde Lyase Production, Chemical Engineering Science, 59, 5075-5083.

Çalık, P., Yılgör, P., Demir, A. S., 2006. Influence Of Controlled-pH and Uncontrolled-pH Operations on Recombinant Benzaldehyde Lyase Production by *Escherichia coli*, Enzyme and Microbial Technology, 38, 617-627.

Clare, J.J., Rayment, F.B., Ballantine, S.P., Sreekrishna, K. and Romanos M.A. 1991. High level expression; of tetanus toxin fragment C in *Pichia pastoris* strains containing mulyiple tandem integration of the gene. Bio/Technology. 9: 455-460.

Cereghino,J.L. and Cregg, J.M. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*.2000. FEMS Microbiology Reviews 24 45-66.

Cereghino, G.P, Cereghino, J. L., Ilgen, C., Cregg, J.M.. 2002. Production of recombinant proteins in fermenter cultures of the yeast Pichia pastoris. Curr Opin Biotechnol 13:329–332.

Couderc, R., and Baratti, J., 1980. Oxidation of methanol by the yeast Pichia pastoris: purification and properties of alcohol oxidase. Agric. Biol. Chem. 44: 2279-2289.

Cregg, J.M., Madden, K.R. 1988. Development of the methylotrophic yeast, *Pichia pastoris*, as a host system for the production of foreign proteins. *Dev Ind Microbiol* 29:33-41.

Cregg, J.M., Madden, K.R., Barringer, K.J., Thill, G.P. and Stillman, C.A. 1989 Functional characterization of the two alcohol oxidase genes from the yeast Pichia pastoris. Mol. Cell. Biol. 9, 1316-1323.

Cregg, J.M., Vedvick, T.S. and Raschke W.C. 1993. Recent advances in the expression of foreign genes in *Pichia pastoris*. Bio/Technology. 11: 905-910.

Chaplin, M. F., Bucke, C., 1990. Enzyme Technology, Cambridge University Press, USA.

Couderc, R. and Baratti, J. 1980. Oxidation of methanol by the yeast Pichia pastoris: purification and properties of alcohol oxidase. Agric. Biol. Chem. 44, 2279-2289.

Demir, A. S., Pohl, M., Janzen, E., Muller, M., 2001. Enantioselective Synthesis of Hydroxy Ketones Through Cleavage and Formation of Acyloin Linkage, J. Chem. Soc., Perkin Trans., 633-635.

Demir, A. S., Şeşenoğlu Ö., Eren, E., Hosrik, B., Pohl, M., Janzen, E., Kolter, D., Feldmann, R., Dünkelmann, P., Müller, M., 2002. Enantioselective Synthesis of α-Hydroxy Ketones via Benzaldehyde Lyase-Catalyzed C-C Bond Formation Reaction, Advanced Synthesis &Catalysis. 344(1), 96-103.

Demir, A. S., Şeşenoğlu, Ö., Dünkelmann, P., Muller, M., 2003. Benzaldehyde Lyase-Catalyzed Enantioselective Carboligation of Aromatic Aldehydes with Mono- and Dimethoxy Acetaldehyde, Organicletters, 5, 2047-2050.

Demir, A. S., Ayhan, P., Igdir, C. A., Duygu, N. A., 2004. Enzyme Catalyzed Hydroxymethylation of Aromatic Aldehydes with Formaldehyde. Synthesis of Hydroxyacetophenones and (S) – Benzoins, Tetrahedron, 60, 6509-6512.

Doran, P. M., 1995. Bioprocess Engineering Principles, Academic Press, USA.

Dünkelmann, P., Kolter-Jung, D., Nitsche, A., Demir, A. S., 2002. Development of a Donor-Acceptor Concept for Enzymatic Cross-Coupling Reactions of Aldehydes: the First Asymmetric Cross-Benzoin Condensation, J. Am. Chem. Soc., 124, 12084-12085.

Egli, T., van Dijken, J.P., Veenhuis, M., Harder, W. and Fiechter, A. 1980. Methanol metabolism in yeasts: regulation of the synthesis of catabolic enzymes. Arch. Microbiol. 124, 115-121.

Fermentas Life Sciences, http://www.fermentas.com, Last accessed date, june 2006.

Faber, K., 2000. Biotransformations in Organic Chemistry, 4th Ed., Springer, Germany.

Fogarty, M. W., Kelly, T. C., 1990. Microbial Enzymes and Biotechnology, 2nd Ed., Elsevier Science Publishers Ltd., ENGLAND.

Glazer, A. N., Nikaido, H., 1995. Microbial Biotechnology: Fundamentals of Applied Biotechnology, W.H. Freeman and Company, USA.

Goochee, C.F., Gramer, M.J., Andersen, D.C., Bahr, J.B. and Rasmussen, J.R. 1991. The oligosaccharides of glycoproteins: bioprocess factors aiecting oligosaccharide structure and their eiect on glycoprotein properties. Biotechnology (NY) 9, 1347-1355.

Gonzalez, B., Vicuna, R., 1989. Benzaldehyde Lyase, a Novel Thiamine PP_i Requiring Enzyme, from *Pseudomonas fluorescens* Biovar I, J. Bacteriol., 171, 2401-2405.

Hart., H., Craine, L. E., Hart, D. J., 1999. Organic Chemistry: A Short Course, Houghton Mifflin Company, USA.

Higgins, D.R. and Cregg, J.M. 1998. Methods in Molecular Biology: Pichia Protocols. Humana Press. New Jersey USA.

Hinrichsen, P., Gomez, I., Vicuna, R., 1994. Cloning and Sequencing of the Gene Encoding Benzaldehyde Lyase *Pseudomonas fluorescens* Biovar I., Gene, 144, 137-138.

Ho, N. S, Hunt, D. H., Horton, M. R., Pullen, K. J., Pease, R. L., 1989. Site-Directed Mutagenesis by Overlap Extension Using the Polymerase Chain Reaction, Gene, 77, 51-59.

Horton, M. R., Ho, N. S., Pullen, K. J., Hunt, D. H., Cai, Z., Pease, R. L., 1993. Gene Splicing by Overlap Extension, Methods in Enzymology, 217, 270-279.

Horton, M. R., Hunt, D. H., Ho, N. S., Pullen, K. J., Pease, R. L., 1989. Engineering Hybrid Genes Witout the Use of Restricition Enzymes: Gene Splicing by Overlap Extension, Gene, 77, 61-68.

Jacobs, M., 1995. Expression of the subtilisin Carlsberg-Encoding Gene in *Bacillus licheniformis* and *Bacillus subtilis*, Gene, 152, 69-74.

Jacobs, M., Eliasson, M., Uhlen, M., Flock, J. I., 1985. Cloning, Sequencing and Expression of subtilisin Carlsberg from *Bacillus licheniformis*, Nucleic Acids Research, 13, 8913-8926.

Jahic, M., Rotticci-Mulder, J.C., Martinelle, M., Hult, K. and Enfors, S-O. 2002. Modeling of growth and energy metabolism of *Pichia pastoris* producing a fusion protein. Bioprocess Biosyst Eng. 24: 385-393.

Joseph, P., Fantino, J.R., Herbaud, M.L., Denizot, F., 2001, Rapid orientated cloning in a shuttle vector allowing modulated gene expression in *Bacillus subtilis.* Fems Mic. Let., 205, 91-97

Kazusa DNA Research Institute, http://www.kazusa.or.jp/codon/, Last accessed: June 2006

Khatri, N., Hoffmann, F., 2005. Impact of methanol concentration on secreted protein production in oxygen-limited cultures of recombinant *Pichia pastoris.,* Biotechnology and Bioengineering. 93: 871-879

Kirk, R. E., Othmer, D. F., 1994. Encyclopedia of Chemical Technology, 4th Ed., the Interscience Encyclopedia Inc., New York.

Kneen, M. M., Pogozheva, D. I., Kenyon, L. G., Mcleish, J. M., 2005. Exploring the Active Site of Benzaldehyde Lyase by Modeling and Mutagenesis, Biochemica Et Biophysica Acta, 1753, 263-271.

Koutz, P., Davis, G.R., Stillman, C., Barringer, K., Cregg, J. And Thill, G. 1989. Structural comparison of the Pichia pastoris alcohol oxidase genes. Yeast 5, 167-177.

Laemmli, U. K., 1970. Cleavage of Structural Proteins During the Assembly of the Head of Bacteriophage T4, Nature, 227, 680-685.

Lodish, H., Berk, A., Zipursky, S. L., Matsudaira, P., Baltimore, D., Darnell, J., 2003. Molecular Cell Biology, 5th Ed., W.H. Freeman Co., New York.

Li, W., Zhou, X., Lu, P., 2004. Bottlenecks in the Expression and Secretion of Heterologous Proteins in *Bacillus subtilis,* Research in Microbiology, 155, 605-610.

Madigan, M.T., Martinko, J., Parker, J., 2002. Brock Biology of Microorganisms. 10 edition. Prentice Hall; Saddle River USA

Maria, D. D. P., Stillger, T., Pohl, M., Wallert, S., Et Al., 2005. Preparative Enantioselective Synthesis of Benzoins and (R)-2-Hydroxy-1-Phenylpropanone Using Benzaldeyhyde Lyase, Journal of Molecular Catalysis B: Enzymatic, 38, 43-47.

Mosbacher, G. T., Mueller, M., Schulz, E. G., 2005. Structure and Mechanism of the ThDP-Dependent Benzaldehyde Lyase from *Pseudomonas fluorescens*, FEBS Journal, 272, 6067-6076.

Moses, V., Cape, R. E., 1991. Biotechnology: The Science and the Business, Harwood Academic Publishers, Switzerland

Nelson, D. L., Cox M. M. 2004 Lehninger's Principles of Biochemistry Fourth Edition W.H. Freeman New York USA

Nielsen, J., Villadsen, J., 1994. Bioreaction Engineering Principles, Plenum Press, New York.

Özcelik, İ. Ş., 2003. Genetik Mühendisliği ve Metabolik Yolizi Mühendisliği Teknikleriyle L-Fenilalanin Üretimi İçin Biyoproses Geliştirilmesi, Doktora Tezi, Ankara Üniversitesi, Ankara.

Pohl, M., Mueller, M., Demir, A. S., 2002. Nucleotide Sequence Encoding Benzaldehyde Lyase and Method for the Stereoselective Synthesis of 2-Hyroxyketone, Patent: WO0202753-A.

Posprech, A., Neumann, B., 1995. A Versatile Quick-Prep of Genomic DNA From Gram Positive Bacteria. Trends in Genetics, 11, 217-218.

Primrose, S., Twyman,R. 2002. Principles of Gene Manipulation, 6th edition. Blackwell Publishing USA

Roggenkamp, R., Janowicz, Z., Stanikowski, B. and Hollenberg, C.P. 1984. Biosynthesis and regulation of the peroxisomal methanol oxidase from the methylotrophic yeast Hansenula polymorpha. Mol. Gen. Genet. 194, 489-493.

Sambrook, J., Russell, D. W., 2001. Molecular Cloning: A Laboratory Manual Vol. 1, 2, 3, 3rd ed., Cold Spring Harbor, USA.

Sanchez-Gonzales, M., Rosazza, P. N. J., 2003. Mixed Aromatic Acyloin Condensations with Recombinant Benzaldehyde Lyase: Synthesis of α-Hydroxydihydrochalcones and Related α-Hydroxy Ketones, Adv. Synth. Catal., 345, 819-824.

Scheve, L. G., 1984. Elements of Biochemistry, Allyn and Bacon. Inc., USA.

Shiloach, J., Kaufman, J., Guillard, S., Fass, R., 1996. Effect of Glucose Supply Strategy on Acetate Accumulation, Growth, and Recombinant Protein Production by *Escherichia coli* BL21 (Λde3) and *Escherichia coli* JM109, Biotechnology and Bioengineering, 49, 421-428.

Shuler, M. L., Kargi, F., 2002. Bioprocess Engineering: Basic Concepts, 2nd Ed., Prentice Hall Inc., USA.

Simonen, M., Palva, I., 1993. Protein Secretion in *Bacillus s*pecies, Microbiological Reviews, 57, 109-137.

Smith, A., 1995. Gene Expression in Recombinant Microorganisms, 1st Ed., Marcel Dekker Inc., USA.

Smith, Wood, 1991. Molecular Biology and Biotechnology, Chapman & Hall Limited, Hong Kong.

Sonenshein, A. L., 1993. *Bacillus subtilis* and Other Gram-Positive Bacteria, American Society for Microbiology, USA.

Sullivan, M.A., Yasbin, R.E., Young, F.E., 1984. New shuttle vectors for Bacillus subtilis and Escherichia coli which allow rapid detection of inserted fragments, Gene, 29, 21-26.

Stephanopoulos, G. N., Aristidou, A. A., Nielsen, J., 1998. Metabolic Engineering: Principles and Methodologies, Academic Press, UK.

Tjalsma, H., Bolhuis, A., Jongbloed, J. D. H., Bron, S., Van Dijl, J. M., 2000. Signal Peptide-Dependent Protein Transport in *Bacillus subtilis*: a Genome-Based Survey of the Secretome, Microbiology and Molecular Biology Reviews, 64, 515-547.

Todar's Online Textbook of Bacteriology, http://www.textbookofbacteriology.net, Last accessed date june 2006.

Tschopp, J.F., Brust, P.F., Cregg, J.M., Stillman, C.A. and Gingeras, T.R. 1987. Expression of the LacZ gene from two methanol-regulated promoters in Pichia pastoris. Nucleic Acids Res. 15, 3859-3876.

Vagner, V., Dervyn, E. and Erlich, S.D., 1998, Avector for systematic gene inactivation in *Bacillus subitilis*, Microbiol. 144, 3097-3104.

Van Roosmalen, M. L., Geukens, N., Jongloed, J. D. H., Tjalsma, H., Dubois, J-Y. F., Bron, S., Van Dijl, J. M., Anne, J., 2004. Type I Signal Peptidases of Gram-Positive Bacteria, Biochemica Et Biophysica Acta, 1694, 279-297.

Veenhuis, M., van Dijken, J.P. and Harder, W. 1983 The significance of peroxisomes in the metabolism of one-carbon compounds in yeast. Adv. Microb. Physiol. 24, 182.

Wang, L. F., Wong, L. S., Lee, G. S., Kalyan, K. N., Hung, P. P., Hilliker, S., Doi, H. R., 1988. Expression and Secretion of Human Atrial Natriuretic α-Factor in *Bacillus subtilis* Using the subtilisin Signal Peptide, Gene, 69, 39-47.

Watson, J. D., Gilman, M., Witkowski, J., Zoller, M., 1992. Recombinant DNA, 2nd Ed., Scientific American, Inc., USA.

Wegner, G., 1990. Emerging application of methelotrophic yeast. FEMS Microbiol. Rev. 87: 279-284.

Werten, M.W.T., van den Bosch, TJ., Wind, R.D., Mooibroek, H., DeWolf, F.A. 1999. High-yield secretion of recombinant gelatins by Pichia pastoris. Yeast 15:1087–1096.

Westers, L., Westers, H., Quax, J. W., 2004. *Bacillus subtilis* as Cell Factory for pharmaceutical Proteins: a Biotechnological Approach to Optimize the Host Organism, Biochimica Et Biophysica Acta, 1694, 299-310.

Wong, S. L., Doi, H. R., 1986-a. Determination of the Signal Peptidase Cleavage Site in the Preprosubtilisin of *Bacillus subtilis*, the Journal of Biological Chemistry, 261, 10176-10181.

Wong, S. L., Kawamura, F., Doi, H. R., 1986-b. Use of the *Bacillus subtilis* Subtilisin Signal Peptide for Efficient Secretion of TEM β-Lactamase During Growth, Journal of Bacteriology, 1005-1009.

Wong, S., L., 1995. Advances in the Use of *Bacillus subtilis* for the Expression and Secretion of Heterologous Proteins, Current Opinion in Biotech., 6, 517-522.

Zocchi, A., Marya Jobe, A., Neuhaus, J.M. and Ward, T.R. 2003. Expression and purification of a recombinant avidin with a lowered isoelectric point in Pichia pastoris. *Protein Expression and Purification, 32*, 167-174

APPENDIX A

PREPARATION OF BUFFERS AND SOLUTIONS USED IN EXPERIMENTS

LB

Soytryptone	10 kg m^{-3}
Yeast exract	5 kg m^{-3}
NaCl	10 kg m^{-3}

LBA

Soytryptone	10 kg m^{-3}
Yeast exract	5 kg m^{-3}
NaCl	10 kg m^{-3}
Agar	15 kg m^{-3}

ALKALINE LYSIS SOLUTION I

Glucose	50 mM
Tris-HCl (pH=8.0)	25 mM
EDTA	10 mM

ALKALINE LYSIS SOLUTION II

NaOH	0.2 N
SDS	% 1

ALKALINE LYSIS SOLUTION III

Potassium Acetate	5 M
Acetic Acid	11.5 (v/v)

SET

NaCI	75 mM
EDTA	25 mM

10X TBE

Tris	108 kg m^{-3}
Boric Acid	55 kg m^{-3}
EDTA	9.3 kg m^{-3}

TSE

Tris HCl, pH= 8.0	10 mM
NaCl	300 mM
EDTA	10 mM

SOLUTION A

Tris HCl, pH= 8.1	10 mM
EDTA	10 mM
NaCl	50 mM
Sacchorose	8% (w/v)

SOLUTION B

SDS	1% (w/v)
NaOH	0.2 M

SOLUTION C

Potassium Acetate	5 M
Acetate Acid	60 ml
dH_2O	Up to 100 ml

Antibiotics	Stock Solutions	
	Concentration	Storage
Zeocin	100 mg/ml in dH_2O	-20°C
Chloramphenicol	35 mg/ml in ethanol	-20°C
Ampcillin	100 mg/ml in dH_2O	-20 °C

1x Equilibration / wash buffer for His-Tag purification	50 mM sodium phosphate buffer, pH 7.0; 300 mM NaCl.
1x Elution buffer for His-Tag purification	50 mM sodium phosphate buffer, pH 7.0; 300 mM NaCl; 150 mM Imidazole.

PBS buffer

8.0 g NaCl, 0.2 g Kcl, 1.44 g $NaHPO_4$, 0.24 g K_2HPO_4 were dissolved in dH_20 and the volume made up to 1000 ml. The pH was adjusted at pH:7.0

1 M potassium phosphate, pH 6.0

56.48 g KH_2PO_4, 14.8 g K_2HPO_4 was dissolved in dH_20 and the volume made upto 500 ml. The pH was controlled. The buffer was autoclaved and stored at room temperature.

20x YNB Stock solution

17 g Yeast Nitrogen Base without amino acids, 50 g $(NH_4)_2SO_4$ was dissolved in dH_20 and the volume made upto 500 ml. The solution was autoclaved, aliquoted into 50 ml Falcon® tubes and stored at room temperature in dark.

3 M Sodium acetate, pH 5.2	24.6 g sodium acetate was dissolved in 80 ml dH_2O and the pH was adjusted to 5.2 with 3M acetic acid. The buffer was filter sterilized and stored at 2-8°C.
Yeast Lysis Solution	2 % Triton X-100, 1% SDS, 100 mM NaCl, 10 mM Tris-Cl-pH8.0, 1mM Na_2EDTA. The solution was autoclaved and stored at room temperature.
50 mM potassium acetate, pH 5.5	0.049 g potassium acetate, dissolved in 8.5 ml dH_2O, titrated with 10N glacial acetic acid to pH 5.5 and make up to 10 ml with dH_2O. Autoclaved and store at room temperature.
TE Buffer, pH 8.0	1 ml of 1M Tris-Cl (pH 8.0), 200 µl of 0.5 M EDTA (pH 8.0) was added to dH_2O and the volume was made up to 100 ml. The buffer was autoclaved and stored at room temperature.
0.125 M (or 0.5 M) EDTA, pH 8.0	4.65 g (or 18.61 g) Ethylenediaminetetra acetic acid disodium salt dihydrate was was dissolved in 80 ml dH_2O. NaOH was added until EDTA was dissolved. The final pH was further adjusted to pH 8.0 and the final volume was adjusted to 100 ml. The buffer was autoclaved and stored at room temperature.

APPENDIX B

GENE SEQUENCE OF *pre-sub*C GENE:

(Accession No:X03341)

Gcggtctattcatactttcgaactgaacatttttctaaaacagttattaataaccaaaaaatttaaa
ttggtcctccaaaaaaataggcctaccatataattcatttttttttctataataaattaacagaataat
tggaatagattatattatccttctatttaaattattctgaataaagaggaggagagtgagtaatg*at
gaggaaaaagagtttttggcttgggatgctgacggccttcatgctcgtgttcacgatggcattcagcg
attccgcttctgctgctcaaccggcgaaaaatgttgaaaaggattatattgtcggatttaagtcagga
gtgaaaaccgcatctgtcaaaaaggacatcatcaaagagagcggcggaaaagtggacaagcagtttag
aatcatcaacgcggcaaaagcgaagctagacaaagaagcgcttaaggaagtcaaaaatgatccggatg
tcgcttatgtggaagaggatcatgtggcccatgccttg

* Indicates the start codon.

GENE SEQUENCE OF *bal* GENE INCLUDING C-TERMINAL POLY-HIS TAG EXTENSION:

(Accession No:A349268)

atggcgatgattacaggc*ggcgaactggttgttcgcaccctaataaaggctggggtcgaacatctgt
tcggcctgcacggcgcgcatatcgatacgatttttcaagcctgtctcgatcatgatgtgccgatcatc
gacacccgccatgaggccgccgcagggcatgcggccgagggctatgcccgcgctggcgccaagctggg
cgtggcgctggtcacggcgggcgggggatttaccaatgcggtcacgcccattgccaacgcttggctgg
atcgcacgccggtgctcttcctcaccggatcgggcgcgctgcgtgatgatgaaaccaacacgttgcag
gcggggattgatcaggtcgccatggcggcgcccattaccaaatgggcgcatcgggtgatggcaaccga
gcatatcccacggctggtgatgcaggcgatccgcgccgcgttgagcgcgccacgcgggccggtgttgc
tggatctgccgtgggatattctgatgaaccagattgatgaggatagcgtcattatccccgatctggtc
ttgtccgcgcatggggccagacccgaccctgccgatctggatcaggctctcgcgcttttgcgcaaggc
ggagcggccggtcatcgtgctcggctcagaagcctcgcggacagcgcgcaagacggcgcttagcgcct
tcgtggcggcgactggcgtgccggtgtttgccgattatgaagggctaagcatgctctcggggctgccc
gatgctatgcggggcgggctggtgcaaaacctctattcttttgccaaagccgatgccgcgccagatct
cgtgctgatgctgggggcgcgctttggccttaacaccgggcatggatctgggcagttgatcccccata
gcgcgcaggtcattcaggtcgaccctgatgcctgcgagctgggacgcctgcagggcatcgctctgggc
attgtggccgatgtgggtgggaccatcgaggctttggcgcaggccaccgcgcaagatgcggcttggcc
ggatcgcggcgactggtgcgccaaagtgacggatctggcgcaagagcgctatgccagcatcgctgcga
aatcgagcagcgagcatgcgctccacccctttcacgcctcgcaggtcattgccaaacacgtcgatgca
ggggtgacggtggtagcggatggtgcgctgacctatctctggctgtccgaagtgatgagccgcgtgaa
acccggcggttttctctgccacggctatctaggctcgatgggcgtgggcttcggcacggcgctgggcg
cgcaagtggccgatcttgaagcaggccgccgcacgatccttgtgaccggcgatggctcggtgggctat
agcatcggtgaatttgatacgctggtgcgcaaacaattgccgctgatcgtcatcatcatgaacaacca
aagctgggggcgacattgcatttccagcaattggccgtcggccccaatcgcgtgacgggcacccgtt
tggaaaatggctcctatcacggggtggccgccgcctttggcgcggatggctatcatgtcgacagtgtg
gagagcttttctgcggctctggcccaagcgctcgcccataatcgccccgcctgcatcaatgtcgcggt
cgcgctcgatccgatcccgcccgaagaactcattctgatcggcatggacccccttcggatct**catcacc
atcaccatcactaa***

* Indicates the primer binding sequences.

APPENDIX C

THERMODYNAMIC PROPERTIES OF DESIGNED PRIMERS TOGETHER WITH DIMER AND SELF-COMPLIMENTARY FORMATION AFFINITIES

Table A.1 Thermodynamic properties of designed primers.

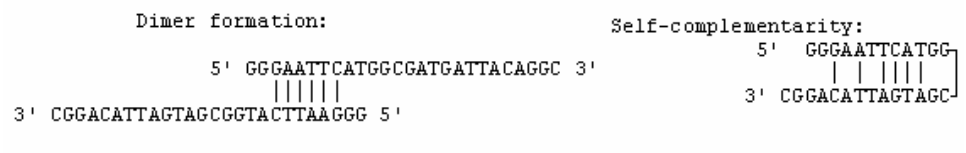
Name	Length	GC%	T_m (°C)	ΔG kcal/mol	ΔH kcal/mol	ΔS eu
BAL Reverse Primer	26	44.8	60.7	-46.7	-205.7	-592.3
BAL Forward Primer	29	50	60.7	-43.4	-200.7	-701.8

The self complimentary and dimer formation affinities are given below:

***EcoRI-bal* forward Primer:**

Sequence: 5' GGGAATTCATGGCGATGATTACAGGC 3'

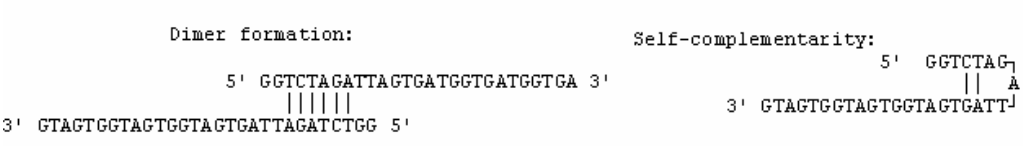
EcoRI RE Sequence



***XbaI-bal* Reverse Primer:**

Sequence: 5' GGTCTAGATTAGTGATGGTGATGGTGA 3'

XbaI RE Sequence



APPENDIX D

MAP OF SITES FOR RESTRICTION ENDONUCLEASES

Restriction enzymes that are cutting the *bal* gene fragment are listed in Table A.2; whereas that of those noncutting the *bal* gene as follows:

Noncutter restriction enzymes of *bal* gene: AarI, AatII, AclI, AflII, AgeI, AloI, AlwNI, ApaI, ApaLI, AscI, AsuII, AvrII, BalI, BamHI, BbvCI, BciVI, BfiI, BplI, Bpu10I, BsaAI, BsaXI, BseRI, BsmI, Bsp1407I, BspLU11I, BstEII, BstXI, BtsI, DraII, DrdI, Eam1105I, Eco31I, Eco57I, Eco57MI, EcoRI, EcoRII, EcoRV, Esp3I, FalI, FseI, FspAI, GsuI, Hin4I, HindIII, HpaI, KpnI, MluI, MmeI, NaeI, NdeI, NheI, NotI, NruI, PI-PspI, PI-SceI, PacI, PfoI, PleI, PmaCI, PmeI, PpiI, PpuMI, PshAI, PsiI, PsrI, PvuI, PvuII, RsrII, SacI, SacII, SanDI, ScaI, SexAI, SfiI, SgfI, SmaI, SmlI, SnaBI, SpeI, SrfI, SspI, StuI, SwaI, TaqII, TatI, TfiI, Tth111I, VspI, XbaI, XhoI

Table A.2 Restriction enzymes that of cuting *bal* gene.

Name	Sequence	Site Length	Overhang	Frequency	Cut Positions
BsrBI	CCGCTC	6	blunt	1	615
BtrI	CACGTC	6	blunt	1	1146
OliI	CACNNNNGTG	6	blunt	1	280
XmnI	GAANNNNTTC	6	blunt	1	1661
ApoI	RAATTY	6	five_prime	1	1369
AvaI	CYCGRG	6	five_prime	1	735
BclI	TGATCA	6	five_prime	1	348
BglII	AGATCT	6	five_prime	1	810
BsmAI	GTCTC	5	five_prime	1	115
ClaI	ATCGAT	6	five_prime	1	89
EcoNI	CCTNNNNNAGG	6	five_prime	1	43
Ksp632I	CTCTTC	6	five_prime	1	292
NcoI	CCATGG	6	five_prime	1	359
SapI	GCTCTTC	7	five_prime	1	292
SgrAI	CRCCGGYG	8	five_prime	1	279
StyI	CCWWGG	6	five_prime	1	359
BseMII	CTCAG	5	three_prime	1	650
BseSI	GKGCMC	6	three_prime	1	1489
BsgI	GTGCAG	6	three_prime	1	58
DraIII	CACNNNGTG	6	three_prime	1	1518
EciI	GGCGGA	6	three_prime	1	625
PstI	CTGCAG	6	three_prime	1	935
Sse8387I	CCTGCAGG	8	three_prime	1	935

Table A.2, continued

Name	Sequence	Site Length	Overhang	Frequency	Cut Positions
TspGWI	ACGGA	5	three_prime	1	1063
TspRI	CASTG	5	three_prime	1	1563
Eco47III	AGCGCT	6	blunt	2	1066, 1592
HindII	GTYRAC	6	blunt	2	903, 1554
AccI	GTMKAC	6	five_prime	2	902, 1553
AflIII	ACRYGT	6	five_prime	2	329, 1143
AvaII	GGWCC	5	five_prime	2	971, 1677
BspHI	TCATGA	6	five_prime	2	117, 1415
BspMI	ACCTGC	6	five_prime	2	879, 1119
FokI	GGATG	5	five_prime	2	1186, 1552
HgaI	GACGC	5	five_prime	2	510, 935
MfeI	CAATTG	6	five_prime	2	1393, 1456
NarI	GGCGCC	6	five_prime	2	191, 367
SalI	GTCGAC	6	five_prime	2	901, 1552
BaeI	ACNNNNGTAYC	7	three_prime	2	1367, 1400
Hpy99I	CGWCG	5	three_prime	2	1150, 1468
MboII	GAAGA	5	three_prime	2	279, 1666
PflMI	CCANNNNNTGG	6	three_prime	2	263, 1458
BsaBI	GATNNNNATC	6	blunt	3	581, 1408, 1544
MslI	CAYNNNNRTG	6	blunt	3	280, 1148, 1414
AcyI	GRCGYC	6	five_prime	3	191, 367, 927
BsrDI	GCAATG	6	three_prime	3	251, 1133, 1441
BsrI	ACTGG	5	three_prime	3	29, 696, 1036
NspI	RCATGY	6	three_prime	3	166, 732, 1105
SduI	GDGCHC	6	three_prime	3	287, 632, 1489
SphI	GCATGC	6	three_prime	3	166, 732, 1105
XcmI	CCANNNNNNNNNTGG	6	three_prime	3	201, 264, 981
BbvI	GCAGC	5	five_prime	4	296, 728, 1068, 1106
TseI	GCWGC	5	five_prime	4	309, 741, 1081, 1094
TspDTI	ATGAA	5	three_prime	4	335, 514, 731, 1432
Cfr10I	RCCGGY	6	five_prime	5	279, 465, 618, 699, 1336
XhoII	RGATCY	6	five_prime	5	477, 810, 859, 1050, 1687

Table A.2, continued

Name	Sequence	Site Length	Overhang	Frequency	Cut Positions
BglI	GCCNNNNNGGC	6	three_prime	5	77, 364, 462, 682, 935
BseYI	CCCAGC	6	five_prime	6	49, 198, 825, 921, 1284, 1430
Tsp45I	GTSAC	5	five_prime	6	213, 243, 1044, 1158, 1332, 1479
BccI	CCATC	5	five_prime	7	390, 982, 1168, 1254, 1336, 1534, 1706
BsePI	GCGCGC	6	five_prime	7	80, 305, 452, 655, 832, 884, 1289
CfrI	YGGCCR	6	five_prime	7	167, 616, 956, 1015, 1298, 1460, 1520
HphI	GGTGA	5	three_prime	7	285, 405, 435, 1170, 1378, 1687, 1693
BcgI	CGANNNNNNTGC	6	three_prime	8	160, 194, 678, 712, 892, 926, 1605, 1639
HaeII	RGCGCY	6	three_prime	8	194, 211, 370, 670, 677, 1068, 1285, 1594
FauI	CCCGC	5	five_prime	10	148, 190, 213, 217, 333, 454, 749, 754, 1617, 1655
SfaNI	GCATC	5	five_prime	12	396, 416, 738, 789, 812, 900, 947, 996, 1085, 1140, 1370, 1625
TauI	GCSGC	5	three_prime	18	20, 154, 157, 169, 367, 445, 618, 688, 805, 1012, 1028, 1217, 1319, 1322, 1401, 1525, 1528, 1579

APPENDIX E

PREPARATION OF SEPARATION AND STACKING GEL FOR SDS-POLYACRYLAMIDE GEL ELECTROPHORESIS

	Stacking Gel	Seperating Gel		
	5%	7.5%	10%	12%
30% acrylamide mix	1.67 ml	2.5 ml	3.33 ml	4 ml
dH_2O	5.68 ml	4.85 ml	4.05 ml	3.35 ml
1.5 M Tris-HCl, pH 8.8	-	2.5 ml	2.5 ml	2.5 ml
0.5 M Tris-HCl, pH 6.8	2.5 ml	-	-	-
10% (w/v) SDS	100 µl	100 µl	100 µl	100 µl
10% (w/v) ammonium persulfate	60 µl	50 µl	50 µl	50 µl
TEMED	15 µl	10 µl	10 µl	10 µl

Preparation of Materials Used for Staining of SDS-Polyacrylamide Gels with Coomassie Brilliant Blue

A. Staining Solution:
Dissolve 0.25g Coomassie Brilliant Blue in 100 ml methanol: acetic acid solution.
B. Methanol: Acetic Acid Solution:
Combine 900 ml of methanol:H_2O (500 ml of methanol and 400 ml of H_2O) and 100 ml of glacial acetic acid.

Preparation of Materials Used for Staining of SDS-Polyacrylamide Gels with Silver Salts

A. Fixer
Mix 150 ml methanol + 36 ml acetic acid + 150 µl 37% formaldehyde and complete to 300 ml with distilled water. This solution can be used several times.

B. 50% Ethanol

Mix 600 ml pure ethanol + 600 ml distilled water. This solution should always be prepared freshly.

C. Pretreatment Solution

Dissolve 0.08 g sodium thiosulphate ($Na_2S_2O_3.5H_2O$) in 400 ml distilled water by mixing with a glass rod. Take 8 ml and set aside for further use in developing solution preparation.

D. Silver Nitrate Solution

Dissolve 0.8 g silver nitrate in 400 ml distilled water and add 300 µl 37% formaldehyde.

E. Developing Solution

Dissolve 9 g potassium carbonate in 400 ml distilled water. Add 8 ml from pretreatment solution and 300 µl 37% formaldehyde.

F. Stop Solution

Mix 200 ml methanol + 48 ml acetic acid and complete to 400 ml with distilled water.

APPENDIX F

POSSIBLE GLYCOSYLATION SITES FOR BENZALDEHYDE LYASE ENZYME

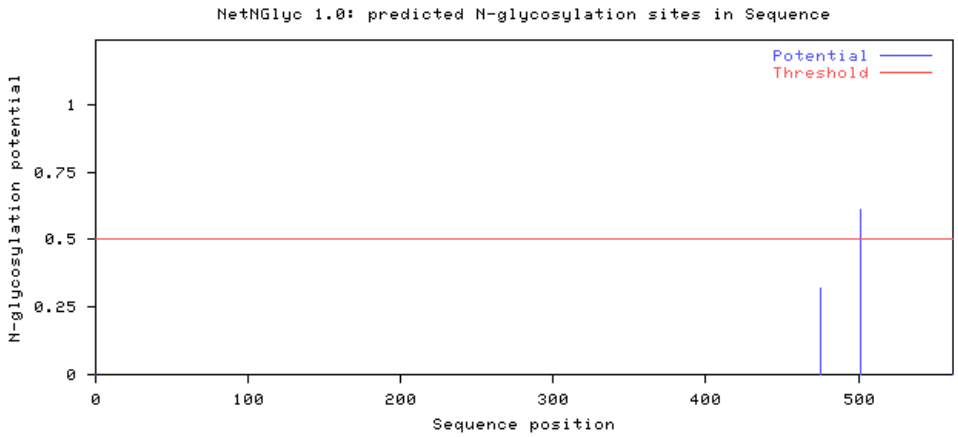


Figure A.1 Molecular weights of discrete proteins of Protein Molecular Marker in kDa (http://www.cbs.dtu.dk/services/NetOGlyc/).

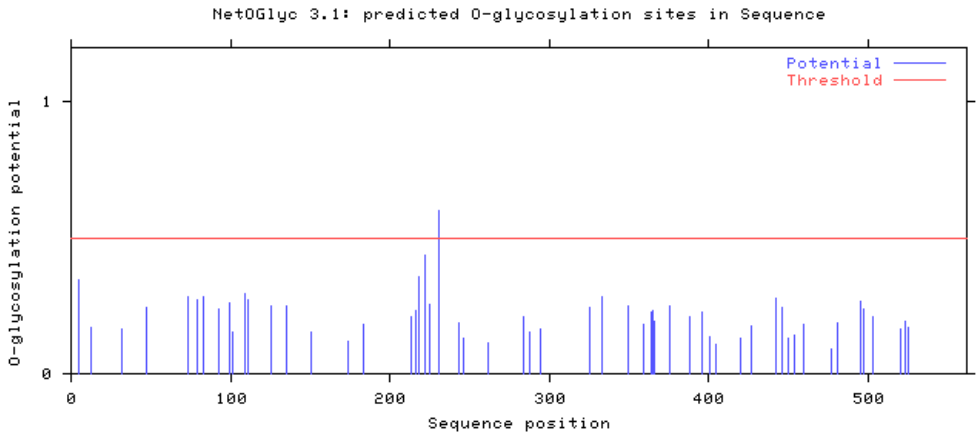


Figure A.2 Molecular weights of discrete proteins of Protein Molecular Marker in kDa (http://www.cbs.dtu.dk/services/NetOGlyc/).

APPENDIX G

.

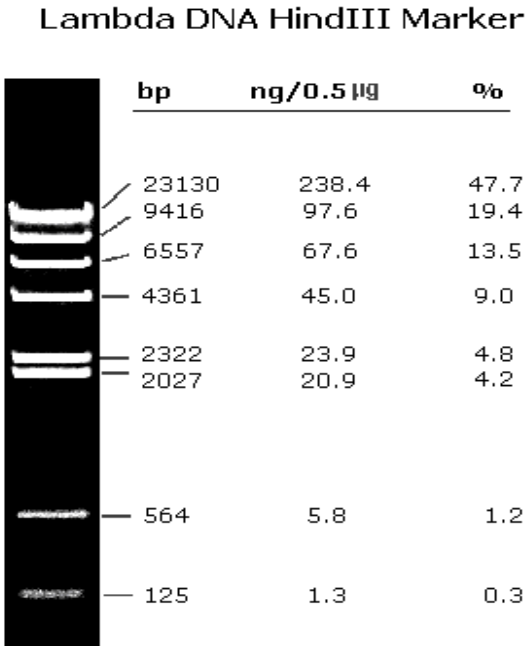


Figure A.3 Discrete DNA fragments of Lambda DNA HindIII Marker in bp.

Protein	Source	Approx. MW (kDa)
β-galactosidase	*E.coli*	116.0
Bovine serum albumin	Bovine plasma	66.2
Ovalbumin	Chicken egg white	45.0
Lactate dehydrogenase	Porcine muscle	35.0
Restriction endonuclease *Bsp*981	*E.coli*	25.0
β-lactoglobulin	Bovine milk	18.4
Lysozyme	Chicken egg white	14.4

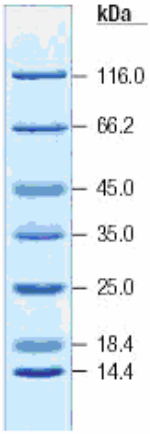


Figure A.4 Molecular weights of discrete proteins of Protein Molecular Marker in kDa

APPENDIX H

PREPERATION OF BUFFERS FOR 6XHIS-TAGGED PROTEIN PURIFICATION

Lysis buffer (1 liter)

50 mM NaH_2PO_4	6.90 g $NaH_2PO_4.H_20$ (MW 137.99 g/mol)
300 mM NaCl	17.54 g NaCl (MW 58.44 g/mol)
10 mM imidazole	0.68 g imidazole (MW 68.08 g/mol)
Adjust pH to 8.0 using NaOH	

Wash buffer (1 liter)

50 mM NaH_2PO_4	6.90 g $NaH_2PO_4.H_20$ (MW 137.99 g/mol)
300 mM NaCl	17.54 g NaCl (MW 58.44 g/mol)
20 mM imidazole	1.36 g imidazole (MW 68.08 g/mol)
Adjust pH to 8.0 using NaOH	

Elution buffer (1 liter)

50 mM NaH_2PO_4	6.90 g $NaH_2PO_4.H_20$ (MW 137.99 g/mol)
300 mM NaCl	17.54 g NaCl (MW 58.44 g/mol)
250 mM imidazole	17.00 g imidazole (MW 68.08 g/mol)
Adjust pH to 8.0 using NaOH	

APPENDIX I

Table A.3 Benzaldehyde lyase enzyme amino acid sequence after secretion into medium

```
   1 - GAATTCATGGCGATGATTACAGGCGGCGAACTGGTTGTTCGCACCCTAATAAAGGCTGGG - 60
     - E  F  M  A  M  I  T  G  G  E  L  V  V  R  T  L  I  K  A  G
  61 - GTCGAACATCTGTTCGGCCTGCACGGCGCGCATATCGATACGATTTTTCAAGCCTGTCTC - 120
     - V  E  H  L  F  G  L  H  G  A  H  I  D  T  I  F  Q  A  C  L
 121 - GATCATGATGTGCCGATCATCGACACCCGCCATGAGGCCGCCGCAGGGCATGCGGCCGAG - 180
     - D  H  D  V  P  I  I  D  T  R  H  E  A  A  A  G  H  A  A  E
 181 - GGCTATGCCCGCGCTGGCGCCAAGCTGGGCGTGGCGCTGGTCACGGCGGGCGGGGGATTT - 240
     - G  Y  A  R  A  G  A  K  L  G  V  A  L  V  T  A  G  G  G  F
 241 - ACCAATGCGGTCACGCCCATTGCCAACGCTTGGCTGGATCGCACGCCCGGTGCTCTTCCTC - 300
     - T  N  A  V  T  P  I  A  N  A  W  L  D  R  T  P  V  L  F  L
 301 - ACCGGATCGGGCGCGCTGCGTGATGATGAAACCAACACGTTGCAGGCGGGGATTGATCAG - 360
     - T  G  S  G  A  L  R  D  D  E  T  N  T  L  Q  A  G  I  D  Q
 361 - GTCGCCATGGCGGCGCCCATTACCAAATGGGCGCATCGGGTGATGGCAACCGAGCATATC - 420
     - V  A  M  A  A  P  I  T  K  W  A  H  R  V  M  A  T  E  H  I
 421 - CCACGGCTGGTGATGCAGGCGATCCGCGCCGCCGTTGAGCGCGCCACGCGGGCCGGTGTTG - 480
     - P  R  L  V  M  Q  A  I  R  A  A  L  S  A  P  R  G  P  V  L
 481 - CTGGATCTGCCGTGGGATATTCTGATGAACCAGATTGATGAGGATAGCGTCATTATCCCC - 540
     - L  D  L  P  W  D  I  L  M  N  Q  I  D  E  D  S  V  I  I  P
 541 - GATCTGGTCTTGTCCGCGCATGGGGCCAGACCCGACCCTGCCGATCTGGATCAGGCTCTC - 600
     - D  L  V  L  S  A  H  G  A  R  P  D  P  A  D  L  D  Q  A  L
 601 - GCGCTTTTGCGCAAGGCGGAGCGGCCGGTCATCGTGCTCGGCTCAGAAGCCTCGCGGACA - 660
     - A  L  L  R  K  A  E  R  P  V  I  V  L  G  S  E  A  S  R  T
 661 - GCGCGCAAGACGGCGCTTAGCGCCTTCGTGGCGGCGACTGGCGTGCCGGTGTTTGCCGAT - 720
     - A  R  K  T  A  L  S  A  F  V  A  A  T  G  V  P  V  F  A  D
 721 - TATGAAGGGCTAAGCATGCTCTCGGGGCTGCCCGATGCTATGCGGGGCGGGCTGGTGCAA - 780
     - Y  E  G  L  S  M  L  S  G  L  P  D  A  M  R  G  G  L  V  Q
 781 - AACCTCTATTCTTTTGCCAAAGCCGATGCCGCGCCAGATCTCGTGCTGATGCTGGGGGCG - 840
     - N  L  Y  S  F  A  K  A  D  A  A  P  D  L  V  L  M  L  G  A
 841 - CGCTTTGGCCTTAACACCGGGCATGGATCTGGGCAGTTGATCCCCCATAGCGCGCAGGTC - 900
     - R  F  G  L  N  T  G  H  G  S  G  Q  L  I  P  H  S  A  Q  V
 901 - ATTCAGGTCGACCCTGATGCCTGCGAGCTGGGACGCCTGCAGGGCATCGCTCTGGGCATT - 960
     - I  Q  V  D  P  D  A  C  E  L  G  R  L  Q  G  I  A  L  G  I
 961 - GTGGCCGATGTGGGTGGGACCATCGAGGCTTTGGCGCAGGCCACCGCGCAAGATGCGGCT - 1020
     - V  A  D  V  G  G  T  I  E  A  L  A  Q  A  T  A  Q  D  A  A
1021 - TGGCCGGATCGCGGCGACTGGTGCGCCAAAGTGACGGATCTGGCGCAAGAGCGCTATGCC - 1080
     - W  P  D  R  G  D  W  C  A  K  V  T  D  L  A  Q  E  R  Y  A
1081 - AGCATCGCTGCGAAATCGAGCAGCGAGCATGCGCTCCACCCCTTTCACGCCTCGCAGGTC - 1140
     - S  I  A  A  K  S  S  S  E  H  A  L  H  P  F  H  A  S  Q  V
1141 - ATTGCCAAACACGTCGATGCAGGGGTGACGGTGGTAGCGGATGGTGCGCTGACCTATCTC - 1200
     - I  A  K  H  V  D  A  G  V  T  V  V  A  D  G  A  L  T  Y  L
1201 - TGGCTGTCCGAAGTGATGAGCCGCGTGAAACCCGGCGGTTTTCTCTGCCACGGCTATCTA - 1260
     - W  L  S  E  V  M  S  R  V  K  P  G  G  F  L  C  H  G  Y  L
1261 - GGCTCGATGGGCGTGGGCTTCGGCACGGCGCTGGGCGCGCAAGTGGCCGATCTTGAAGCA - 1320
     - G  S  M  G  V  G  F  G  T  A  L  G  A  Q  V  A  D  L  E  A
1321 - GGCCGCCGCACGATCCTTGTGACCGGCGATGGCTCGGTGGGCTATAGCATCGGTGAATTT - 1380
     - G  R  R  T  I  L  V  T  G  D  G  S  V  G  Y  S  I  G  E  F
1381 - GATACGCTGGTGCGCAAACAATTGCCGCTGATCGTCATCATCATGAACAACCAAAGCTGG - 1440
     - D  T  L  V  R  K  Q  L  P  L  I  V  I  I  M  N  N  Q  S  W
1441 - GGGGCGACATTGCATTTCCAGCAATTGGCCGTCGGCCCCAATCGCGTGACGGGCACCCGT - 1500
     - G  A  T  L  H  F  Q  Q  L  A  V  G  P  N  R  V  T  G  T  R
1501 - TTGGAAAATGGCTCCTATCACGGGGTGGCCGCCGCCTTTGGCGCGGATGGCTATCATGTC - 1560
     - L  E  N  G  S  Y  H  G  V  A  A  A  F  G  A  D  G  Y  H  V
1561 - GACAGTGTGGAGAGCTTTTCTGCGGCTCTGGCCCAAGCGCTCGCCCATAATCGCCCCGCC - 1620
     - D  S  V  E  S  F  S  A  A  L  A  Q  A  L  A  H  N  R  P  A
1621 - TGCATCAATGTCGCGGTCGCGCTCGATCCGATCCCGCCCGAAGAACTCATTCTGATCGGC - 1680
     - C  I  N  V  A  V  A  L  D  P  I  P  P  E  E  L  I  L  I  G
1681 - ATGGACCCCTTCGGATCTCATCACCATCACCATCACTAA - 1719
     - M  D  P  F  G  S  H  H  H  H  H  H  *  X
```

(http://www.justbio.com)