PRODUCTION OF *Taq* I RESTRICTION AND MODIFICATION ENZYMES BY USING TWO DIFFERENT EXPRESSION SYSTEMS

by

Ebru Toksoy

B.S. in Ch.E., Boğaziçi University, 1993

M.S. in Ch.E., Boğaziçi University, 1995

Submitted to the Institute for Graduate Studies in Science and Engineering in partial fulfillment of the requirements for the degree of

Doctor

of

Philosophy

Boğaziçi University

2000

ACKNOWLEDGMENTS

The research presented in this thesis was conducted at the Chemical Engineering Department of the Boğaziçi University and was supported by TÜBİTAK Research Council TBAG-1533; Boğaziçi University Research Fund Projects 96A0522 and 97A0501 and by the State Planning Organization through project 95K120320.

I would like to express my sincere gratitude and dedicate this thesis to my advisors Prof. Dr. Z. İlsen Önsan and Prof. Dr. Betül Kırdar. Their everlasting encouragement and guidance gave me the power to complete this work. I will always admire their superior personalities and kindness and keep them as my ideal models for the rest of my life.

I would like to thank Prof. Dr. Altan Erarslan, Prof. Dr. Nejat Dalay and Assoc. Prof. Dr. Kutlu Ülgen for the time they have devoted to reading and commenting on my thesis.

Heartfelt thanks are due to Mehmet Mete Altıntaş and Berna Sarıyar for their being more than a friend to me.

I am also indebted to all my friends at ChE Department and to the members of our Biotechnology group who made the hardworking days more than enjoyable.

I wish to express my special thanks to Bilgi Dedeoğlu for technical and moral support and Mine Yediç for her kind help in everything.

I would also like to thank to my little thermophilic enzymes for all the activity they showed me during my experiments, especially at the most desperate moments.

I am especially grateful to my parents and my brother for their continuous support and patience.

A special dedication is to my dearest Hüseyin Öner whose magical love; intelligence and continuous support not only turned this work from just a dream into reality, but also made us a beautiful family.

ABSTRACT

PRODUCTION OF *Taq* I RESTRICTION AND MODIFICATION ENZYMES BY USING TWO DIFFERENT EXPRESSION SYSTEMS

Taq I restriction-modification system has been cloned and expressed using two different expression systems. In the first system, *Taq* I restriction endonuclease was expressed as a fusion protein with the maltose-binding protein. Although the MBP-*Taq* I fusion protein was targeted to the periplasmic space, 80-90 per cent of *Taq* I endonuclease activity was found to be excreted to the growth medium without cell lysis. The fusion protein was also purified via amylose affinity chromatography from the cytoplasm, periplasm and medium of recombinant *E. coli* cells. Co-expression of the *Taq* I methylase gene improved the plasmid stability and the periplasmic and extracellular transport of the fusion protein under controlled bioreactor conditions resulting in 0.6×10^6 U/L extracellular *Taq* I endonuclease activity yield in *E. coli* XL1(pH185, pETMET) culture.

For the second system, *Taq* I endonuclease gene was expressed under the control of the T7 RNA polymerase promoter. Growth and enzyme productivity of the unprotected and methylase protected *E. coli* cultures were analyzed under various fermentation conditions in shake flasks and bioreactor. Co-expression of the methylase gene resulted in higher enzyme production rates for longer periods yielding 250×10^6 U/L *Taq* I endonuclease activity in crude cellular extracts of *E. coli* BL21(DE3)[pTaqR, pMETaq] cells.

ÖZET

Taq I RESTRİKSİYON VE MODİFİKASYON ENZİMLERİNİN İKİ FARKLI EKSPRESYON SİSTEMİ KULLANILARAK ÜRETİLMESİ

Taq I restriksiyon-modifikasyon sistemi, iki farklı ekspresyon sistemi kullanılarak klonlanması ve ekspresyonu sağlanmıştır. Birinci ekspresyon sisteminde, *Taq* I restriksiyon endonükleazı maltoza-bağlanan protein ile füzyon protein oluşturarak üretilmiştir. Oluşturulan MBP-*Taq* I füzyon proteini rekombinant hücrelerin periplazma bölgelerine yönlendirilmiş olmasına rağmen, toplam *Taq* I endonükleaz aktivitesinin yüzde 80-90'ının hücreler parçalanmadan sıvı besi ortamına çıktığı anlaşılmıştır. MBP-*Taq* I füzyon proteini, recombinant *E. coli* hücrelerinin sitoplazma, periplazma ve sıvı besi ortamlarından amiloz afinite kromatografisi yöntemi ile saflaştırılmıştır. Kontrollü fermentör koşullarında, *Taq* I metilaz enziminin koekspresyonu plazmid kararlılığını ve füzyon proteinin periplazma bölgesine ve hücre dışına salgılanmasını arttırmıştır ve *E. coli* XL1(pH185, pETMET) hücrelerinin sıvı besiyerinden $0.6x10^6$ U/L *Taq* I endonükleaz aktivitesi elde edilmiştir.

İkinci sistemde, *Taq* I restriksiyon endonükleaz geninin T7 RNA polimeraz promotörü altında ekspresyonu sağlanmıştır. Korunmayan ve metilaz aktivitesi ile korunan *E. coli* hücrelerinin büyüme özellikleri ve enzim üretimleri, değişik fermentasyon koşullarında çalkalayıcı ve fermentörde incelenmiştir. Metilaz geninin koekspresyonu enzim üretiminin daha hızlı ve uzun süreli olmasını sağlamış ve *E. coli* BL21(DE3)[pTaqR, pMETaq]hücrelerin özütlerinden $250x10^6$ U/L *Taq* I endonükleaz aktivitesi elde edilmiştir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS / ABBREVIATIONS

B	absorbance of the blank
c	concentration (g/L)
D	dilution factor
OD	absorbance
U	units
q_p	specific *Taq* I endonuclease production rate
t	time (hr)
t_{lag}	lag phase period (hr)
x	biomass formation (g/L)
Yp/s	Units of *Taq* I formed per g glucose consumed (g/g)
Yp/x	Units of *Taq* I formed per g cells formed (g/g)
Yx/s	g cells formed per g glucose consumed (g/g)
ΔA	absorbance difference
ε	adsorption coefficient (1/mmol/cm)
μ	specific growth rate (hr^{-1})
μ_{max}	maximum specific growth rate (hr^{-1})
A	absorbance
CFU	colony forming units
DCW	dry cell weight
DO	dissolved oxygen
GSH	reduced glutathione
HIC	hydrophobic interaction chromatography
IA	immobilized antibody
IEC	ion-exchange chromatography
I-IgG	immobilized immunoglobin G
IMAC	immobilized metal affinity chromatography
IS	immobilized substrate

PAGE	polyacrylamide gel electrophoresis
RME	reverse micellar extraction
rpm	rounds per minute
SBD	starch-binding domain
Seph	Sepharose

1. INTRODUCTION

Restriction-modification systems are a bacterium's first line of defense against invading phages. The essential enzymatic components of the system are a modification enzyme that methylates the host DNA, and a restriction endonuclease that recognizes the DNA as a foreigner and cleaves any DNA not carrying the host specific methylation pattern. Restriction endonucleases are valuable tools in various fields of molecular biology and genetics such as DNA mapping, sequencing or recombinant *in vitro* cloning. Therefore, it is desirable to obtain them in large amounts.

Taq I restriction-modification system of the thermophilic eubacterium *Thermus aquaticus* YT1 consists of the two thermostable *Taq* I restriction endonuclease and *Taq* I methylase enzymes. *Taq* I is a type II restriction endonuclease that is highly specific to the double-stranded palindromic 5' TCGA 3' target sequence. On the other hand, *Taq* I methylase enzyme modifies the adenine residue on each strand of the same sequence by transferring a methyl group from its cofactor S-adenosyl-L-methionine. The recognition sequence of these enzymes exists at most of the polymorphic regions of the human genome. Therefore, *Taq* I restriction endonuclease is widely used in molecular diagnosis. Moreover, it is a model enzyme to study DNA-protein interactions. Due to its extensive use in many molecular biology and biotechnology applications, obtaining large quantities of this enzyme is considered to be very important. However, the isolation of *Taq* I restriction endonuclease from its native host *Thermus aquaticus* requires high fermentation costs and results in very low yields upon purification. Therefore, cloning and expression of the *Taq* I restriction-modification system in *Escherichia coli* is considered to be advantageous for the high yield recovery of the *Taq* I restriction endonuclease enzyme.

In the framework of this study, two different expression systems were used for the high-level production of *Taq* I restriction endonuclease in recombinant *E. coli* cells. In the first system, the pH185 plasmid construct was used, where the *Taq* I restriction endonuclease gene was genetically fused to the *malE* gene of *E. coli* encoding the Maltose Binding Protein (MBP) (Özdinler, 1996). In recombinant *E. coli* cells harbouring this plasmid construct, the expressed MBP-*Taq* I fusion protein was targeted to the periplasmic

space. Preliminary shake flask experiments revealed that 80-90 per cent of *Taq* I endonuclease activity was excreted to the extracellular growth medium without cell lysis, which was very beneficial for the subsequent downstream processing purposes (Toksoy *et al.*, 1999).

In order to protect cellular DNA from harmful effects of endonuclease expression, *Taq* I methylase gene was amplified by PCR from *Thermus aquaticus* bacterial genome by using specifically designed primers and cloned within the tetracycline gene of the pBR322 vector. In *E. coli* cells harboring the resulting pMETaq plasmid construct, *Taq* I methylase gene was expressed constitutively under the control of the tetracycline resistance gene promoter. In order to select the recombinants with an antibiotic other than ampicillin, pETMET plasmid was constructed by subcloning the *Taq* I methylase gene together with the tetracycline promoter region into the pET28a+ vector which contained the kanamycin resistance gene.

A series of experiments were conducted in shake flasks and New Brunswick BioFlo III batch/continuous fermentor in order to analyze the effects of host strain, induction conditions, medium composition and methylase protection on the plasmid stability, cell viability, production and distribution of *Taq* I endonuclease activity in the three cellular compartments, i.e. cytoplasm, periplasm and medium, of the *E. coli* cells.

For the second expression system, *Taq* I restriction endonuclease gene was amplified by PCR from *Thermus aquaticus* bacterial genome by using specifically designed primers and cloned into the pET28a+ vector. In unprotected *E. coli* BL21(DE3)[pTaqR] and methylase protected *E. coli* BL21(DE3)[pTaqR, pMETaq] lysogens, *Taq* I restriction endonuclease was expressed under the control of the strong bacteriophage T7 RNA polymerase promoter. Both cultures were compared in terms of their growth characteristics and *Taq* I endonuclease productivity under various fermentation conditions in batch shake flasks containing complex and chemically defined media. Effect of methylase protection on growth kinetics, enzyme productivity, substrate utilization, acetic acid excretion, cell viability and plasmid stability of these *E. coli* cells over-producing *Taq* I restriction endonuclease, was also analyzed in chemically defined medium under controlled bioreactor conditions.

A detailed literature survey about the theoretical background is given in Chapter 2. Chemicals, buffers, solutions, various laboratory equipment used in conducting the experimental work and the experimental methods applied during this study are described in Chapter 3 and Chapter 4, respectively. The results of this study are presented and discussed in Chapter 5. Chapter 6 contains the major conclusions that can be drawn from this study and the recommendations for future work.

2. THEORETICAL BACKGROUND

2.1. Recombinant Protein Expression in *Escherichia coli*

One of the practical aims of genetic engineering is to force a host cell to produce an optimal level of one or more desirable proteins. Genetically engineered cells are used for the production of large amounts of a recombinant protein or they are used for the production of intermediate enzymes that facilitate the synthesis of an end product in the host cell. In the latter case, genetic engineering techniques are employed, for example, to produce an enzyme that speeds up the limiting reaction in the production of a desired product (Georgiou, 1988).

In a commercial process, the overall objective is to maximize the yield of a desired product and to increase the scale of production if possible. However, the overproduction of a recombinant protein usually disrupts the coordination of metabolic processes and may ultimately result in cell death, and is, hence, unfavorable. Therefore, the maximization problem of the chemical engineer becomes an optimization problem in which the physiology of the genetically engineered cells must be taken into account (Georgiou, 1988).

Recombinant proteins can be produced in high levels in various expression systems, and the choice depends on (a) cell growth characteristics; (b) expression levels; (c) intracellular and extracellular expression; (d) posttranslational modifications; (e) biological activity of the protein when expressed with that specific expression system; (f) degree of authenticity required and regulatory issues in the production of therapeutic proteins (Makrides, 1996).

The choice of the host for the expression of a recombinant protein is the most important step in the selection of the expression system. For large-scale productions, the host must propagate fast with simple nutritional requirements. It must not be pathogenic and must not produce any toxins that may contaminate the product. The cells should be able to grow to high cell densities to maximize protein production (Georgiou, 1988).

Escherichia coli is one of the most widely used hosts for the production of heterologous proteins and its genetics are far better characterized than those of any other microorganism which makes it easier to model, predict and use in large scale productions (Baneyx, 1999). *Escherichia coli* is also versatile because there are many high expression vectors available for it. However, there are some drawbacks to this host, too. It is unable to perform posttranslational modifications required for eukaryotic proteins, it cannot facilitate disulfide bond formation necessary for the three dimensional conformation of some proteins and some proteins cannot assume their correct three dimensional structure that is essential for biological activity and instead precipitate inside the cell forming inactive protein aggregates called inclusion bodies. It cannot secrete the protein into the culture medium, which would be beneficial from a biochemical engineering point for the simple purification of the target protein. The stability and translational efficiency of mRNA, the ease of protein folding, degradation of the protein by host cell proteases, differences of codon usage between the foreign gene and *Escherichia coli* and the toxicity of the protein to the host determine the efficiency of expression of a protein in *Escherichia coli* (Makrides, 1996).

2.1.1. Selection of Host Strain

The choice of strain for expression of recombinant protein depends on the target protein, the expression vector, the fermentation conditions and the particular aim of overproduction. For example, for proteins that are highly susceptible to proteolysis, it may be necessary to use protease deficient strains. When it is aimed to secrete the target protein to the culture medium, periplasmic "leaky" strains may be employed (Miller *et al.,* 1998). The genotype of the host is a useful guide in selection of the best strain for expression of a specific gene.

JM109 is a frequently used strain. It carries the $lacI^q$ gene, strongly repressing the *lac* promoter and derivatives such as the *tac* promoter until induction. This is useful in preventing leaky expression of toxic recombinant proteins. However, this strain has been shown to be inferior to BL21 strain with regard to low acetate concentration maintenance (Shiloach *et al.,* 1996).

TB1 has been reported to be the most widely used strain with pMAL expression vector. It is an α-complementing host, allowing selection of recombinant plasmids harboring an insert. XL1 strain is often used for high quality plasmid DNA purifications. It carries the gene *endAl*, preventing nonspecific activity of endonuclease 1 on DNA. It also carries the *recA* gene and hence homologous recombination is abolished increasing plasmid stability. This is particularly desirable when working with sequences containing direct repeats larger than 50 base pairs. Moreover, it is an α-complementing host (Sambrook *et al.*, 1989).

ER2508 is a protease deficient strain constructed at the laboratories of New England BioLabs (USA). This strain is defective in its *lon* gene encoding the La protease and hence this strain is most frequently used when proteolysis is a problem. The *mal*B gene deletion includes the *mal*E gene, so this strain does not produce any chromosomal MBP. This is beneficial when MBP is expressed by the expression vector as a fusion partner to a target protein. All the MBP produced is part of the fusion partner, increasing the efficiency of amylose column chromatographic purification.

2.1.2. Transcriptional Regulation

The gene coding for a desired protein acts as a template for the production of that protein. The gene may be inserted into the host chromosomal DNA for expression; however, the amount of mRNA determines the rate of expression of a protein. Therefore, if the number of copies of the target gene in the host can be increased, the rate of product formation will also be increased. To achieve this aim, the gene is inserted into high copy number plasmids that are propagated from one generation to the next and exist in several copies per cell. Plasmids containing a desired foreign gene and designed specifically to produce a high level of product synthesis are called expression vectors (Georgiou, 1988).

An ideal plasmid should have a low molecular weight for easier handling and higher copy number, and should have single sites for a large number of restriction endonucleases. The expression vector should also harbor a gene that confers a phenotypic trait to simplify selection and propagation of the recombinant colonies. This is usually an antibiotic resistance gene. Hence, only those cells harboring the plasmid, i.e. recombinant cells, can

grow in the presence of antibiotic supplemented to the medium (Old and Primrose, 1989). The essential components of an expression vector are the promoter, the ribosome binding site, the transcription terminator and the origin of replication.

The gene of interest must be placed under the control of an *Escherichia coli* promoter. Two types of promoters exist that initiate transcription: constitutive and regulated. Constitutive promoters are always turned on and the rate of mRNA synthesis is not modulated. On the other hand, a regulated promoter controls the level of mRNA synthesis and hence production of the protein under its control. Repressible or inducible promoters may be employed to control gene expression. The inducible promoters are usually preferred because they are induced by the addition of a substrate or by temperature shifts; however, repressible promoters require the dilution of the repressor from the growth medium by changing the culture medium, which is not practical (Georgiou, 1988).

There are basically four desirable features of a regulated promoter. First, it should be strong. Thus, the yield of the target protein can be as high as 30 per cent of the total cellular protein. It must have a low basal expression level, meaning it must be tightly regulated and not "leaky". Leaky protein expression before induction of the cells may result in the accumulation of the usually toxic recombinant protein, causing cell death due to excess toxicity or metabolic burden. Plasmids also lose stability as a result of expression of the foreign protein, and hence a decrease in growth rate occurs and recombinant protein production may drop significantly. The promoter must be easily transferable to other *Escherichia coli* strains to study expression in different strains. Finally, its induction must be simple and cost effective. For laboratory research, strong promoters such as *tac* or *trc* are usually employed that may be induced by the addition of IPTG. However, IPTG is both toxic and very expensive, making it unsuitable for large-scale use. For large-scale production, promoters induced by thermal induction or chemical inducers are used (Makrides, 1996). Some of the promoters used for the high level expression of genes in *Escherichia coli* are listed in Table 2.1 (Weickert *et al.*, 1996).

Table 2.1. A few promoters used for the high level expression of genes in *Escherichia coli* with their advantages and disadvantages.

Promoter	Induction	Characteristics
lac	IPTG, thermal	Low level expression (about 5 per cent), leaky expression
trp	Trp starvation, indole acrylic acid	Leaky expression
*pho*A	Phosphate starvation	Limited media options
*rec*A	Nalidixic acid	
tac	IPTG, thermal, lactose	High level expression, induction results in cell death, leaky expression
trc	IPTG	
T7*lac*	IPTG	Very high level expression, leaky expression
cad	acid pH	

Improvements have been made in promoter control that increase efficiency of transcription and provide tighter control at a lower cost. A temperature sensitive mutant *lac*I gene that encodes a thermosensitive *lac* repressor is a recent development with which *lac* based promoters can be induced thermally with a tight regulation. However, although thermal induction is cost effective and simple, it may result in the increased production of proteases, which degrade the product. The development of feed strategies for tight low-cost control of *trp* promoter in large scale fermentation and the development of thermally regulated, runaway replication plasmids that allow more efficient repression of the *trp* promoter are noteworthy improvements in promoter control (Weickert *et al.*, 1996). The *cspA* is a cold-shock promoter, which is tightly repressed at and above 37°C and compares

favorably to the *tac* promoter for the expression of an aggregation-prone fusion protein at reduced temperatures (Baneyx, 1999).

Among the nutritionally inducible promoters such as *phoA* and *trp*, the arabinose promoter uses the inexpensive sugar L-arabinose as an inducer and is somewhat weaker than the *tac* promoter. However, the *ara*B promoter can both be tightly repressed and modulated over a wide range of inducer conditions which is advantageous for the control of gene expression. The *ara*B expression systems could be used with other promoters to independently regulate expression of two or more recombinant proteins in *Escherichia coli* (Baneyx, 1999). The regulatory region of the *cadA* gene was used to construct a pH inducible expression system. This is advantageous since bulk chemical inducers such as acids and bases can be employed. Another advantage of pH induction over temperature induction is that high temperatures induce formation of inclusion bodies and heat shock proteins including *lon* protease, which result in degradation (Makrides, 1996).

The transcription terminator is located downstream of the target gene and is a signal to terminate transcription. Its presence prevents the transcription and translation of unnecessarily long gene segments which otherwise would increase the metabolic burden to the host. Moreover, undesirable secondary structures may form in the transcript due to overlong transcripts, which would reduce efficient translation. Finally, essential genes such as the ampicillin resistance gene or the origin of replication may be read through if termination is not present or is weak and hence they may not be transcribed and translated efficiently. If, for example, the origin of replication is read through, plasmid copy number decreases remarkably (Old and Primrose, 1989). Similarly, a transcription terminator placed upstream of the promoter will inhibit transcription through the promoter unless the promoter is induced and thus minimize background transcription. It also serves as a protective element composed of stem loop structures, protecting the mRNA from degradation, enhancing its stability and subsequently increasing the level of protein production. The use of a reversible transcription termination system derived from phage λ to control leaky expression of T7 RNA polymerase and use of tandem transcription terminators that prevent read through transcription into vector sequences are some of the improvements in the transcription termination (Makrides, 1996).

2.1.3. Plasmid Stability

The vector contains the origin of replication, which determines the average number of plasmid molecules per cell, which is also called the plasmid copy number (Georgiou, 1988). Plasmid stability is the ability of plasmid harboring cells to maintain the structure of the plasmid with a high copy number. The batch culture studies show that the recombinant plasmid suffers from segregational and structural instabilities. Structural instability is the result of insertions or rearrangements within the plasmid or the loss of some plasmid material by deletions (Old and Primrose, 1989). Segregational instability is due to the defective partitioning of plasmids resulting in the loss of the entire plasmid. Plasmid free cells have a higher growth rate and can outgrow the plasmid containing cells. Since plasmid free cells do not produce the desired protein their domination results in a rapid decline in the efficiency of the culture. Plasmids with small DNA inserts are more stable. The production of a protein toxic to the host cell results in the death of the plasmid containing cells. Hence the plasmid free cells dominate the culture much faster. Dissolved oxygen concentration, growth temperature, presence of essential amino acids, dilution rates are some of the environmental factors affecting plasmid stability. The most common method of preventing plasmid free cells from dominating is using antibiotic selection. The plasmid harbors the gene for antibiotic resistance and when the cells are grown in antibiotic containing media, the loss of the plasmid results in cell death. However, in large-scale cultivations, a small portion of the cells can inactivate the antibiotic present in the medium and hence the plasmid free cells can propagate. Moreover, the cost and waste disposal considerations make antibiotic selection undesirable in large-scale productions. The stability of the plasmid can also be increased when a plasmid-encoded protein is essential for cellular metabolism. The plasmid stability problem can be solved definitely by the insertion of the target gene into the bacterial chromosome at the expense of having only one copy of the gene per cell which would limit the mRNA synthesis and hence protein production (Georgiou, 1988). Plasmids can also be maintained if they harbor the partitioning function, par, which ensures that the plasmid is efficiently segregated to the daughter cells at cell division (Old and Primrose, 1989).

2.1.4. Translational Regulation

Initiation of translation of *E. coli* mRNAs requires a ribosomal binding site followed by an initiation codon, which is most commonly AUG. Ribosome binding site is located downstream of the promoter and consists of the Shine-Dalgarno sequence complementary with the small subunit of the ribosomal RNA followed by an A+T rich translational spacer (Baneyx, 1999). The ribosomal binding site is important because it determines the efficiency of protein synthesis per mRNA molecule (Georgiou, 1988). The enrichment of the ribosome binding site with adenine and tyrosine residues increases efficiency of mRNA translation by reducing the potential for secondary structure formation at the 5' end of the transcript (Makrides, 1996). The secondary structure at the translation initiation region of mRNA is also important in the efficient gene expression. For good expression, the initiation and Shine-Dalgarno sites must be accessible. Thus the mRNA must be already folded in a specific secondary structure and then interact with the ribosome to initiate translation. The sequences immediately downstream of the start codon especially the triplet immediately preceding it are important as well. It is possible to increase protein expression by 20 fold by altering some nucleotides at the beginning of a gene (Georgiou, 1988). However, the rate-limiting reagent of protein synthesis is the mRNA because it is highly unstable due to endonuclease and 3' exonuclease attack. The rate of synthesis and the rate of degradation determine the level of mRNA. The average half-life of an mRNA molecule is only one or two minutes. The rate of synthesis may be enhanced using strong promoters, but the same effect can be achieved by preventing or reducing degradation (Old and Primrose, 1989). The stabilizing effect of the 5' hairpins was demonstrated in the case of the long-lived *ompA* mRNA (Baneyx, 1999).

Differences in codon usage between prokaryotes and eukaryotes can have a significant impact on heterologous protein production. The arginine codons AGA and AGG are rarely found in *E. coli* genes, whereas they are common in eukaryotes (Chen and Inouye, 1990). The presence of such codons in cloned genes affects protein accumulation levels, mRNA and plasmid stability and, in extreme cases, inhibits protein synthesis and cell growth (Sorensen *et al.*, 1989; Baneyx, 1999). Gene expression can be improved by replacement of the rare codons with more common ones, without modifying the protein product (Robinson *et al.*, 1984).

2.1.5. Cytoplasmic Expression

The cytoplasmic expression usually yields high amounts of the desired protein. However, the drawbacks of the cytoplasm usually eliminate this advantage. First of all, the cytoplasm is crowded with all the other proteins of the cell, making the purification of a recombinant protein from this pool a very tedious task. Additionally, recombinant proteins usually contain the amino acid methionine as their initiator codon, which cannot be removed in the cytoplasm unless the enzyme methionine amino peptidase is co-expressed (Makrides, 1996).

The cytoplasm is a reducing environment and does not facilitate disulfide bond formation; this usually hinders the expression of eukaryotic proteins in *E. coli* cytoplasm because proteins cannot assume their correct form. Bacterial cytoplasmic proteins contain few cysteine residues and most proteins that contain disulfide bonds are exported from the cytoplasm. Thioredoxin system composed of thioredoxin reductase and thioredoxin contribute to the reduction of disulfide bonds; therefore, thioredoxin reductase deficient *E. coli* strains may be employed to facilitate correct disulfide bond formation (Weickert *et al.*, 1996).

Overproduction of recombinant proteins in the cytoplasm of *E. coli* is often accompanied by their misfolding and segregation into insoluble aggregates known as inclusion bodies. A traditional approach to reduce protein aggregation is through fermentation engineering, most commonly by reducing the cultivation temperature. The more recent realization that *in vivo* protein folding is assisted by molecular chaperone systems such as DnaK-DnaJ-GrpE and GroEL-GroES, which promote the proper isomerization and cellular targeting of other polypeptides by transiently interacting with folding intermediates, and by foldases which accelerate rate-limiting steps along the folding pathway, has provided powerful new tools to combat the problem of inclusion body formation (Baneyx, 1999).

Protein folding and proteolytic degradation are intimately linked as catabolism is an efficient way to conserve cellular resources by recycling improperly folded or heavily damaged proteins into their constituent amino acids. In the cytoplasm of *E.coli*, most

degradation steps are carried out by proteases (Baneyx, 1999). Proteolysis occurs due to the presence of incomplete polypeptides, mutations, excessive synthesis of subunits from multimeric complexes and overproduction of recombinant proteins (Makrides, 1996). The stability of proteins depends on their amino acid sequence, especially in the amino terminal. Regions enriched in proline, glutamine, serine and threonine may destabilize proteins (Makrides, 1996). Flexibility and accessibility of the termini or the presence of chemically blocking amino terminal groups such as the acetyl group are also important factors that contribute to the stability of proteins. Moreover, exposure of protease cleavage sites on the surface of the protein makes it more susceptible to proteolytic degradation (Old and Primrose, 1989).

Several experimental approaches exist to prevent degradation by proteolysis. A possible strategy is to make use of host strains bearing mutations in protease genes such as *lon* gene coding for La protease, but since several proteases are usually involved in the degradation of a given protein substrate, multiple mutations in genes encoding proteases reduce cell growth rates and compromise strain fitness (Goldberg, 1992). Co-expression of the *pin* gene of bacteriophage T4 which was found to stabilize various labile eukaryotic proteins, optimization of the fermentation conditions, use of fusion partners or expression of the protein as inclusion bodies are some of the alternative approaches used to prevent proteolysis (Baneyx, 1999).

2.1.6. Periplasmic Expression

Periplasmic expression is preferred when the high yield does not compensate for the disadvantages of cytoplasmic expression. The periplasm contains only about four per cent of the total cell protein (Makrides, 1996) and hence purification of the target protein from the periplasmic pool is much simpler. In addition, proteolytic degradation is prevented to a great extent, increasing the yield (Talmadge and Gilbert, 1982). The oxidizing environment of the periplasm contains enzymes catalyzing the formation and rearrangement of disulfide bonds and hence facilitating the correct folding of proteins. Moreover, the signal peptide in the recombinant protein is removed during transportation to the periplasm, making the expression of the authentic native protein possible.

Secretory proteins are synthesized in the cytoplasm as precursors with an amino-terminal signal peptide that facilitates entry into the export pathway and translocation across the membrane. After translocation, the signal peptide is removed from the precursor protein and the mature moiety is released into the periplasm (Rosemond *et al.*, 1994). Typical signal sequences are 18-30 amino acids in length and consist of two or more basic residues at the amino terminus, a central hydrophobic core and a hydrophilic carboxyl-terminus. Many signal sequences derived from naturally occurring secretory proteins such as OmpA, OmpT, alkaline phosphatase, β-lactamase, protein A from *Staphylococcus aureus* and the human growth hormone support the efficient translocation of heterologous proteins across the inner membrane when fused to their amino termini (Baneyx, 1999).

Reduction in the expression of protein levels prevents the overloading of the translocation machinery, and co-production of proteins that participate in membrane transport processes improve the yield and efficiency of periplasmic expression. One of the drawbacks of periplasmic expression is that recombinant proteins may misfold or form inclusion bodies in this cellular compartment. However, co-expression of periplasmic chaperones such as Skp has been reported to improve the folding of a number of aggregation-prone single-chain antibody fragments (Bothmann and Plückthun, 1998; Hayhurst and Harris, 1999).

Protein disulfide isomerases that catalyze the oxidation of disulfide bonds and peptidyl-prolyl-cis-trans-isomerases that catalyze the isomerization of prolyl peptide bonds in outer membrane proteins may be overexpressed with the target protein to facilitate protein folding in the periplasm (Missiakas *et al.*, 1996). Also, co-expression of DsbA and heat shock factor σ^{32} at low temperature increased the yield of correctly folded T-cell receptor fragments (Hannig and Makrides, 1998).

2.1.7. Excretion of Recombinant Proteins to the Culture Medium

Excretion of the product to the culture medium is the best alternative from a biochemical engineering point of view. The culture medium contains a very small amount of contaminating proteins facilitating the purification and recovery of the product. Moreover, the absence of proteases in the medium stabilizes the protein (Hannig and

Makrides, 1998). The most advantageous feature of excretion into the medium is the possibility of employment of continuous processes in recombinant product synthesis. However, if the cells grow to a high cell density, they will need vigorous agitation and aeration. This will result in air liquid interfaces that can denature the protein (Old and Primrose, 1989).

The secretion of the recombinant proteins into the periplasm of *E. coli* using foreign or native sequences was found to be a prerequisite for their release into the culture medium (Gitter *et al.*, 1995). Excretion into the medium is possible by two basic methods. Either the protein is selectively secreted to the medium using pathways for truly selected proteins or the total cell proteins "leak", selectively or not, into the medium by using signal sequences, fusion partners, permeabilizing proteins or nutrients. Cloning, overexpression and high-level excretion of heterologous extracellular enzymes targeted into the periplasm of *E. coli* using *ompA* signal have recently been reported by several investigators (Lee *et al.*, 1998; Ko *et al.*, 1995). The excretion of the proteins targeted to the periplasm was facilitated by the controlled expression of the *kil* protein, using "leaky" host cells in which the gene encoding one of the outer membrane proteins has been disrupted by chemical permeabilization of the outer membrane or by the synergistic combination of chemical agents and controlled expression of the *kil* protein (van der Wal *et al.*, 1995; Miksch *et al.*, 1997; Miller *et al.*, 1998; Blanchin-Roland and Masson, 1989; Kato *et al.*, 1987; Kitai *et al.*, 1988; Nakamura *et al.*, 1990; Hsiung *et al.*, 1989). Manipulation of the cloned sequences, for example removal of the sequences coding for the first 13 amino acids of the recombinant protein, was found to enhance the secretion of the protein into the extracellular medium (Lee *et al.*, 1998). Also, *E. coli* strains bearing mutations in the *tolQRAB* gene cluster were found to release periplasmic proteins into the extracellular medium. The fragile phenotype of these strains were overcome by co-expressing the TolAIII protein and hence the proteins were efficiently released to the growth medium without forming inclusion bodies in the periplasmic space (Wan and Baneyx, 1998).

2.1.8. Fermentation Conditions

The main objectives of development strategies for recombinant protein fermentation are high volumetric productivity, high product concentration, high yield, high selectivity

and low cost of raw materials (Zabriskie and Arcuri, 1986). These goals can be reached by high rates of cell growth and product formation, high cell densities and increased product stability. Protein production can be increased significantly through the use of high-cell density culture systems in batch, fed batch and continuous systems (Georgiou, 1988). The composition of the growth medium, oxygen availability, pH, temperature and induction conditions are the parameters that influence recombinant protein overexpression (Hannig and Makrides, 1998).

The main problems arising from high cell density cultures are the solubility of solid and gaseous substrates in aqueous media, limitation of substrates with respect to growth, instability and volatility of substrates and products, accumulation of products or metabolic by-products to a growth-inhibitory level, degradation of products, high evolution rates of CO_2 and heat, high oxygen demand as well as the increasing viscosity of the medium in very dense cultures (Riesenberg and Guthke, 1999).

Acetic acid is the main metabolic by-product that accumulates during unrestricted aerobic growth with a carbon source in excess. Acetate accumulation is shown to be strain and growth medium specific; therefore, its inhibitory effects are prevented either by using low acetate producing strains such as *E. coli* BL21 (Shiloach *et al.*, 1996) or developing such strains by means of metabolic engineering (Aristidou *et al.*, 1994). Limited supply of the carbon source such as glucose and glycerol, or use of lactose and methyl α-glucoside instead of glucose, has been reported to prevent acetate accumulation as well (Chou *et al.*, 1994; Korz *et al.*, 1995; Aristidou *et al.*, 1999).

The dissolved oxygen concentration in the bioreactor is also important. When the rate of oxygen consumption exceeds the rate at which it is supplied, the culture becomes oxygen-limited. The adverse effects of oxygen limitation on growth rate and carbon yield can be prevented by increasing agitation, sparging pure oxygen or maintaining lower growth rates (Georgiou, 1988).

The composition of the growth medium must be carefully formulated and monitored, because it may have significant metabolic effects on both the cells and protein production.

The time and period of induction should also be optimized not to overburden the host cell metabolism (Makrides, 1996).

2.2. Gene Fusion Strategies

The use of gene fusions has greatly facilitated the expression and subsequent purification of heterologous proteins in *Escherichia coli.* Specific fusion peptides may confer advantages to the target protein during expression, such as increased solubility, protection from proteolysis, improved folding, increased yield and secretion. The presence of the additional amino acid sequences stabilizes the hybrid protein against degradation. Fusion of short peptide tags can be used to target the protein to different compartments of the cell or even to the culture medium, allowing accumulation of the protein in the most suitable environment. Furthermore, fusion tags can be removed by *in vitro* cleavage of the fusion protein, which makes the formation of the native amino terminus of the protein possible. The highly soluble fusion partner facilitates folding of the target protein. A problem of gene fusions is that the fusion proteins are sometimes cleaved *in vivo* at the junction between the fusion partner and the target protein, which creates problems and decreases yield of target protein if the fusion partner serves as an aid in purification (Ford *et al.,* 1991).

The protein of interest may be fused to itself to increase stability of the protein against proteolytic degradation. However, with this method, inclusion body formation may increase. The gene may be spliced directly after a signal sequence. With this strategy, if it is possible to remove the tag, a recombinant protein with a native amino terminus may be obtained. Moreover, the signal sequence can target the gene to different compartments of the cell; hence it may serve as a secretion tag (Ford *et al.,* 1991).

C-terminal fusions are commonly used constructs in which the gene of interest is fused to the 3' end of a gene, which codes for a highly expressed protein. This configuration is used frequently because transcriptional and translational signals are located at the 5' end of a gene and will function efficiently in the presence of a variety of different fusions to the 3' end of the gene. Moreover, many native proteins tolerate alterations in their N-terminal region (Carter, 1990).

N-terminal fusions can also be used to express genes. With N-terminal fusions, it is simple to perform N-terminal sequencing of the gene fusion and it is possible to design various expression assays. However, N-terminal fusions are disadvantageous because the transcription and translation signals must be engineered specific to the target gene. Additionally, the chemical cleavage of the fusion partner may result in nonnative N-terminal protein, which is undesirable especially in pharmaceutical recombinant proteins (Ford *et al.*, 1991).

Dual affinity fusions are especially used to express proteins that are highly susceptible to proteolysis. In this system, the gene of interest is fused to two fusion partners at each terminus, with specific affinity for two different ligands. This system requires two subsequent purification steps for complete recovery of the target protein. The solubility, stability, size, binding constants and subunit structure of both of the fusion partners must be considered in the construction of dual affinity fusions (Ford *et al.*, 1991).

Applications of fusion tail technology range from basic research to industrial production. On a laboratory scale, fusion recovery systems are powerful and elegant tools for one-step recovery and purification of recombinant proteins or identification of proteins encoded by cloned cDNAs. On an industrial scale, fusion technology can be used in the recovery and purification of both higher cost pharmaceuticals and lower- to medium-cost enzymes. Fusion tails that can be used to exhibit different biospecific and biochemical interactions are listed in Table 2.2. These include (a) entire enzymes with affinity for substrates or inhibitors, such as the β-galactosidase; (b) polypeptide binding proteins such as the staphylococcal protein A; (c) carbohydrate binding proteins or domains such as maltose binding protein and biotin; (d) antigenic epitopes with affinity to immobilized monoclonal antibodies; (e) charged amino acids for use in charge based recovery methods such as the poly (Arg); (f) poly (His) residues for binding to metal chelates and recovery by immobilized metal affinity chromatography; (g) other poly(amino acid)s with various binding specificities (Makrides, 1996; Ford *et al.*, 1991).

Table 2.2. Fusion partners and their applications.

Fusion partner	C-or N-terminus	Ligand / Matrix	Separation method
Enzymes			
β-galactosidase	N, C	TPEG-Seph	IS, IA
glutathione-S-transferase	N	Glutathione Seph	IS (GSH)
chloramphenicolacetyltransferase	N	Chloramphenicol-	IS
Polypeptide-binding protein			
staphylococcal protein A	N	Immunoglobin Seph	I-IgG
streptococcal protein G	C	Albumin	I- albumin
Carbohydrate-binding domains			
maltose binding protein	N	Amylose resin	IS
cellulose-binding protein	C	Cellulose	IS
Biotin-binding domain	C	Biotin	I-avidin
Antigenic epitopes			
RecA	C		IA
Flag	N	Antiflag monoclonal antibodies	IA
Charged amino acids			
Poly (Arg)	C	S-Seph	IEC
Glutamate	N		EEC
Poly (His)	N, C	Ni^{+2}-nitriloacetic acid	IMAC
Poly (Phe)	N	Phe-Superose	Phe- Superose
Poly (Csc)	N	Thiopropyl-Seph	PropylSeph -

The selection of the fusion tail depends on the desired final destination and authenticity of the protein, the stability of the protein and the purification method suitable for the recovery of the protein (Makrides, 1996).

The fusion partner may or may not be removed upon expression of the gene depending on the final desired authenticity of the protein and the biological interference of the fusion partner with the activity of the target protein. If the fusion partner interferes with the biological activity of the target protein, it is necessary to cleave the fusion protein and obtain the target in homogeneous form even in laboratory scale. If the target protein is to be used for therapeutic purposes, it is strictly necessary to obtain the target protein homogeneously in its authentic form not to cause immune response in the patient (Ford *et al.*, 1991; Uhlen and Moks, 1990).

It is possible to cleave the fusion protein enzymatically or chemically. The choice of method depends on the composition, sequence and characteristics of the particular protein. Chemical cleavage is efficient and the reagents are widely and cheaply available, making it possible to apply it on a large scale. However, it is nonspecific and harsh. Therefore, cleavage of the recombinant protein at additional cleavage sites occurs frequently. Moreover, the protein of interest may not always be able to withstand the harsh conditions that must be applied for chemical cleavage. Thirdly, a limited extent of amino acid chain modification may occur, leading to heterogeneity of the target protein. Enzymatic cleavage is desirable for relatively mild reaction conditions and, most importantly, for the high degree of specificity exhibited. Exopeptidases and endopeptidases may be used for cleavage of the fusion protein. However, enzymatic cleavage is affected by steric factors, therefore the cleavage site must be carefully engineered to be accessible by the enzyme. Among the useful enzymes are Factor Xa, thrombin, enterokinase, renin and collagenase. All of these enzymes have extended substrate recognition sequences (up to seven amino acids in the case of renin), which reduces the likelihood of unwanted cleavages elsewhere in the protein. Factor Xa and enterokinase are the most advantageous because they cleave on the carboxyl-terminal of their recognition sequences, allowing the release of the target protein with its authentic N-terminal (Uhlen and Moks, 1990).

2.3. pMAL Protein Fusion System

In the pMAL protein fusion system, the gene encoding the protein of interest is cloned downstream from the *malE* gene of *E.coli*, which encodes maltose binding protein (MBP), resulting in the expression of an MBP fusion protein. The method uses the strong

tac promoter and the *malE* translation initiation signals to give high-level expression of the cloned sequences and a one-step purification of the fusion protein using MBP's affinity for maltose (Figure 2.1).

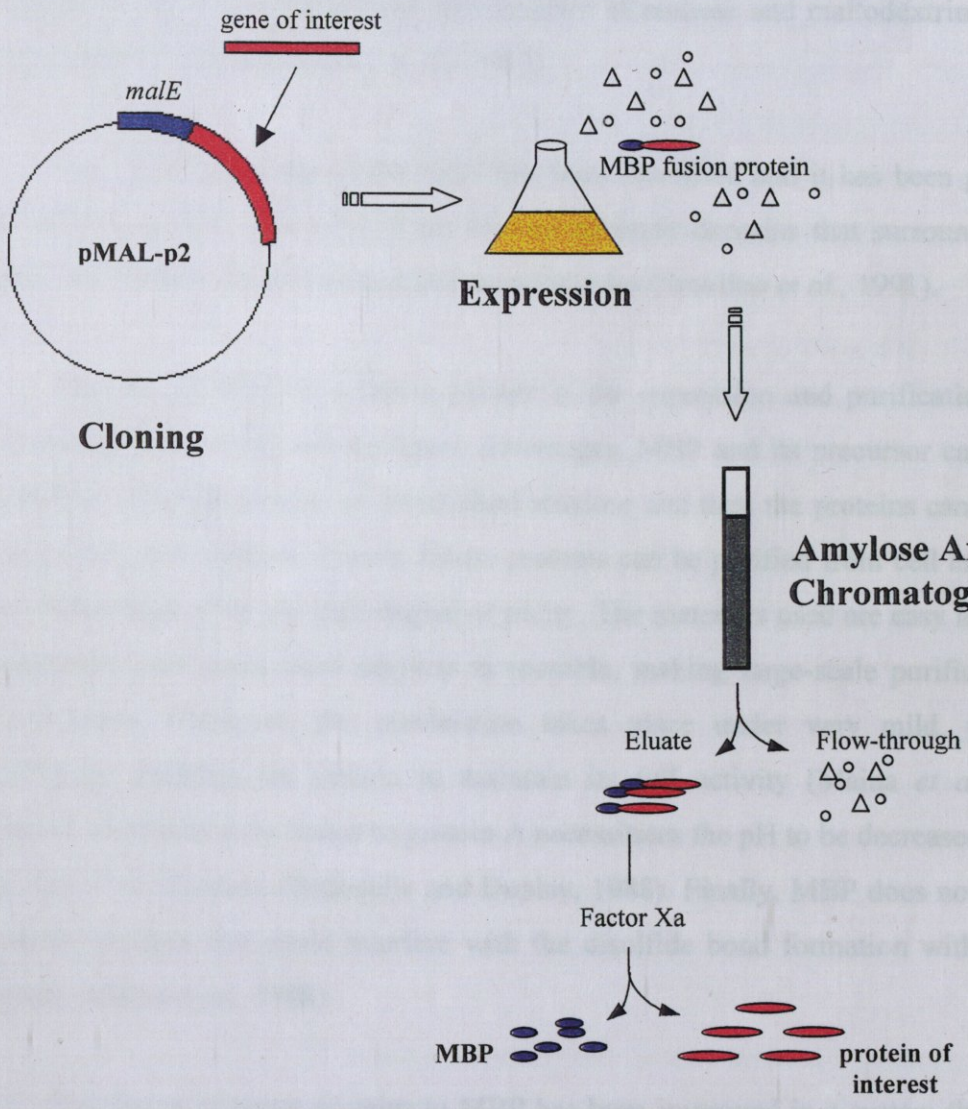


Figure 2.1. pMAL protein fusion and purification system

2.3.1. Maltose Binding Protein

Maltose Binding Protein (MBP) is the product of the *malE* gene of *Escherichia coli* K12. *MalE* gene encodes for the pre-protein of 396 amino acid residues, 26 of which constitute the signal peptide. The expression of MBP is controlled by *malEp*, a strong promoter that is activated by protein MalT in the presence of inducing maltose or

maltodextrins and repressed by glucose (Bedouelle and Duplay, 1988). It is a binding protein specific for maltose and maltodextrins. The binding site recognizes the αl-4 glycosidic bond linking the glucose moieties of maltose. MBP is exported across the cytoplasmic membrane and localized in the periplasm via its signal sequence, and it is essential for the energy dependent translocation of maltose and maltodextrins through the cytoplasmic membrane (Duplay *et al.*, 1984).

The crystal structure of the MBP has been identified and it has been proposed that the protein is made up of two discontinuous globular domains that surround a cleft that forms the binding site for maltose and maltodextrins (Spurlino *et al.*, 1991).

The use of MBP as a fusion partner in the expression and purification of foreign proteins in *Escherichia coli* has many advantages. MBP and its precursor can be purified by affinity chromatography on crosslinked amylose and then the proteins can be eluted by competition with maltose. Hence, fusion proteins can be purified from cell extracts in one step with a high yield and high degree of purity. The materials used are easy to prepare and inexpensive and crosslinked amylose is reusable, making large-scale purification simple and feasible. Moreover, the purification takes place under very mild, physiological conditions, enabling the protein to maintain its full activity (Maina *et al.*, 1988). In contrast, purification by fusion to protein A necessitates the pH to be decreased to 3.0, with the risk of inactivation (Bedouelle and Duplay, 1988). Finally, MBP does not contain any cysteine residues that could interfere with the disulfide bond formation within the target peptide (Maina *et al.*, 1988).

The fusion of target proteins to MBP has been improved in a system that allows the separation of the target domain by a site-specific proteolytic cleavage after purification. The fusion protein in this system contains the recognition sequence for the blood coagulation Factor Xa. Factor Xa recognizes and cleaves the protein at the tetrapeptide Ile-Glu-Gly-Arg. The presence of the recognition site before the target domain in a fusion would allow the cleavage at that site, releasing the target protein without any N-terminal residues.

2.3.2. pMAL-p2 Vector

The pMAL-p2 vector contains the *malE* gene that encodes for MBP together with its signal sequence under the control of the strong inducible *tac* promoter and the translation initiation signals of MBP. The vector is 6721 base pairs long and has an ampicillin resistance gene. There are approximately 20 copies of the vector per cell. The $lacI^q$ gene on the vector encodes the *lac* repressor, which keeps expression low until the *tac* promoter is induced with IPTG. Unique restriction endonuclease sites between *malE* and lacZα gene allow the insertion of coding sequences of interest, to generate fusions with foreign polypeptide (Figure 2.2). β-galactosidase α-fragment activity of the malE-lacZα gene fusion results in blue colony formation when α-complementing hosts such as TB1 or XL1 are transformed with the pMAL-p2 vector. Insertion of a DNA fragment into the polycloning site interrupts the malE-lacZα gene fusion, destroying α-fragment activity and hence allowing blue-white screening for inserts.

There is a spacer sequence coding for ten asparagines residues between the *malE* gene and the polylinker. This spacer insulates MBP from the target gene increasing affinity binding to amylose resin. The vector also has a sequence coding for the recognition site of Factor Xa which cleaves after the amino acids Ile-Glu-Gly-Arg allowing the removal of the MBP fusion tail after purification. The polylinker includes an *Xmn* I site superimposed on the Factor Xa site. When this site is used for cloning, no vector-derived residues remain on the target protein upon cleavage with Factor Xa.

The M13 origin of DNA replication in the vector allows the production of single stranded DNA for sequencing and oligonucleotide directed mutagenesis.

Several research groups have successfully induced *E. coli* to secrete foreign proteins into periplasmic space using MBP fusion system where the transport to the periplasmic space is directed by the MBP signal sequence. Bedouelle and Duplay (1988) reported the periplasmic expression of the *Staphylococcus aureus* nuclease A and the Klenow fragment of the *E. coli* DNA polymerase I enzymes as chimeric proteins with MBP as fusion tail.

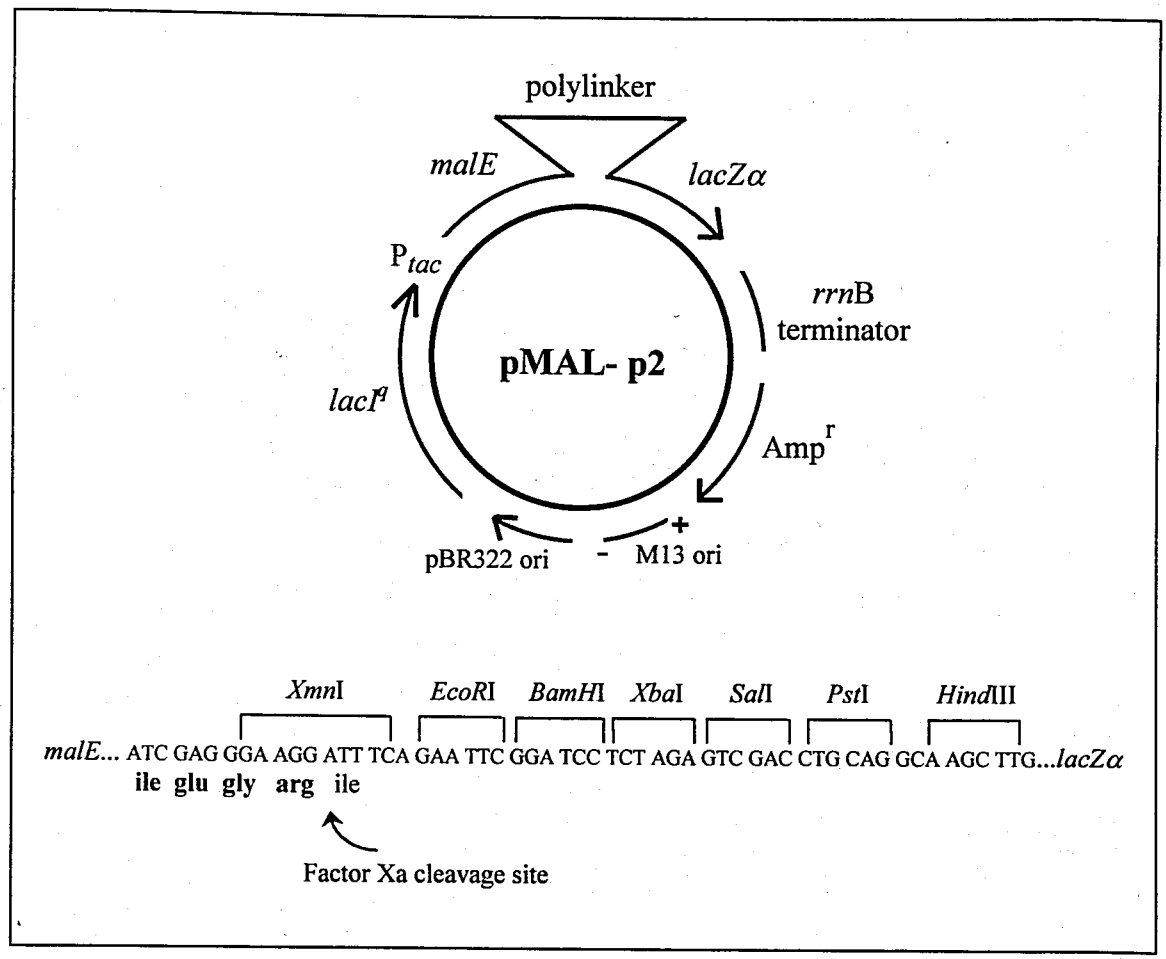


Figure 2.2. pMAL-p2 expression vector

Von Darl *et al.* (1994) used the pMAL-p2 vector for the periplasmic expression of the yolk ferritin which formed inclusion bodies when expressed in the cytoplasm, and the protein could be purified to homogeneity by affinity chromatography on cross-linked amylose. Also, the insoluble human translation initiation factor 4R could be purified from the periplasm of *E. coli* cells by using the the pMAL-p2 vector (Morino *et al.*, 1995). To study the functions of the neurospora-crassa metallothionein and phenylalanine-hydroxylase proteins, Pazirandeh *et al.* (1995) and Knappskog *et al.*, (1995) used the MBP fusion system, respectively. Betton *et al.* (1997) generated fully active artificial bifunctional hybrid proteins by inserting the β-lactamase enzyme into the permissive sites of the maltose binding protein. Kinetic analysis of the fusion proteins purified from the osmotic shock fluids by amylose affinity chromatography revealed that MBP stabilized the active site of the inserted β-lactamase enzyme. Miller et *al.* (1998) expressed the functional mature domain of pediocin AcH as an MBP fusion. They found out that the secreted fusion

protein was released into the culture medium when expressed in the periplasmic leaky host cell *E. coli* E609L in which the gene encoding the outer membrane protein, Braun's lipoprotein, has been disrupted.

2.4. pET Gene Expression System

The RNA polymerase of bacteriophage T7 has a high stringency for its own promoters, which do not occur naturally in *Escherichia coli*, and it is capable of efficiently producing complete transcripts from almost any DNA that is linked to a T7 promoter. Therefore, T7 RNA polymerase has been found very attractive to use for directing high-level expression of selected genes in *E. coli* (Tabor and Richardson, 1985; Studier and Moffat, 1986).

Unlike systems based on *E. coli* promoters (e.g., *lac, tac,* p_L), the pET system uses the bacteriophage T7 promoter to direct the expression of target genes. In the pET system, target genes are cloned in pET plasmids under control of the strong bacteriophage T7 transcription and translation signals. Expression of the cloned gene is induced by providing a source of T7 RNA polymerase in the host cell. Since *E. coli* RNA polymerase does not recognize the T7 promoter, there is virtually no transcription of the target gene in the absence of a source of T7 RNA polymerase, and the cloning step is thus effectively uncoupled from the expression step. Due to the high selectivity and activity of T7 RNA polymerase, almost all of the cell's resources are converted to target gene expression and the desired product can comprise more than 50 per cent of the total cell protein a few hours after induction. Another important benefit of the system is its ability to maintain target genes transcriptionally silent in the uninduced state (Mierendorf *et al.*, 1994).

Target genes are initially cloned using hosts that do not contain the T7 RNA polymerase gene, thus eliminating plasmid instability due to the production of proteins potentially toxic to the host cell. Recombinant plasmids are then transferred into expression hosts containing a copy of the T7 RNA polymerase gene under *lacUV5* control, and expression is induced by the addition of IPTG (Mierendorf et al., 1994).

Grodberg and Dunn (1988) have reported that the outer membrane protease of *E. coli* encoded by the *ompT* gene nicks T7 RNA polymerase after cell lysis. Therefore, the most widely used host in the pET System is the λDE3 lysogen of *E. coli* BL21, which has the advantage of being deficient in both *lon* and *ompT* proteases. In addition, this strain has been shown to posses an acetate self-controlling mechanism. Therefore, it was considered as an excellent host for IPTG induced recombinant genes since it was possible to grow it using simple high glucose batch operation, while leaving the problem of acetate accumulation to be handled by the bacterium itself (Shiloach *et al.*, 1996).

For more stringent control of the T7 RNA polymerase gene and to suppress the leaky expression of a highly toxic protein prior to induction, hosts carrying either pLysS or pLysE can be used. The pLys plasmids encode T7 lysozyme, which is a natural inhibitor of T7 RNA polymerase and thus reduces its ability to transcribe target genes in uninduced cells (Studier, 1991). pLysS hosts produce low amounts of T7 lysozyme, while pLysE hosts produce much more enzyme and, therefore, represent the most stringent control available in λDE3 lysogens.

Figure 2.3 illustrates in schematic form the host and vector elements available for control of T7 RNA polymerase levels and the subsequent transcription of a target gene in a pET vector.

2.4.1. The pET Vectors

The pET vectors were originally constructed by Studier and Moffat (1986). There are two general categories of the pET vectors: transcription and translation vectors. Transcription vectors are designed for expression of target genes that already carry their own prokaryotic ribosome binding site and ATG start codon. Translation vectors contain the highly efficient ribosome binding site from the phage T7 major capsid protein.

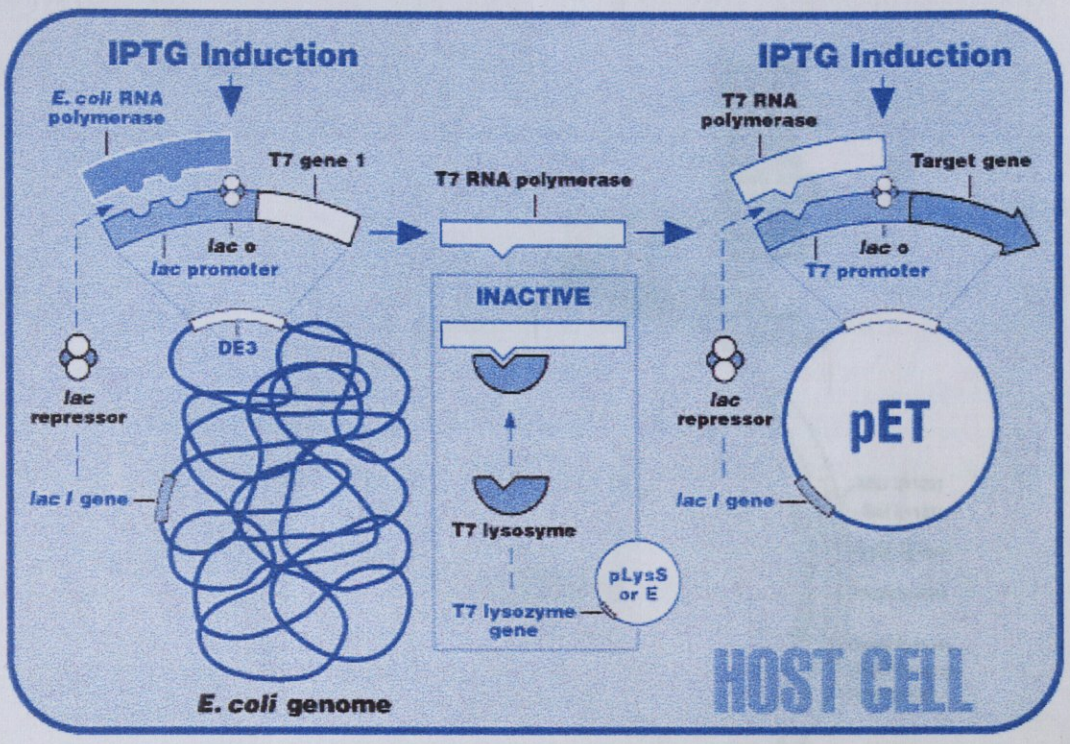


Figure 2.3. Control elements of the pET System

pET28a+ is a translation vector where gene expression is under control of T7*lac* promoter. The T7*lac* promoter contains a 25 bp *lac* operator sequence immediately downstream from the 17 bp promoter region. Binding of the *lac* repressor at this site effectively reduces transcription by T7 RNA polymerase, thus providing a second *lacI*-based mechanism (besides the repression at *lacUV5*) to suppress basal expression in λDE3 lysogens. pET28a+ also carries its own copy of *lacI* to ensure that enough repressor is made to titrate all available operator sites.

Figure 2.4 shows the circle map of the pET28a+ vector with the unique sites of many commercially available enzymes. Kanamycin resistance gene on the plasmid is in the opposite orientation from the T7*lac* promoter, so there should not be an increase in kanamycin gene product after induction due to read-through transcription from the T7*lac* promoter. The pET28a+ plasmid also contains f1 origin of replication that allows the production of single stranded plasmid DNA for mutagenesis and sequencing applications.

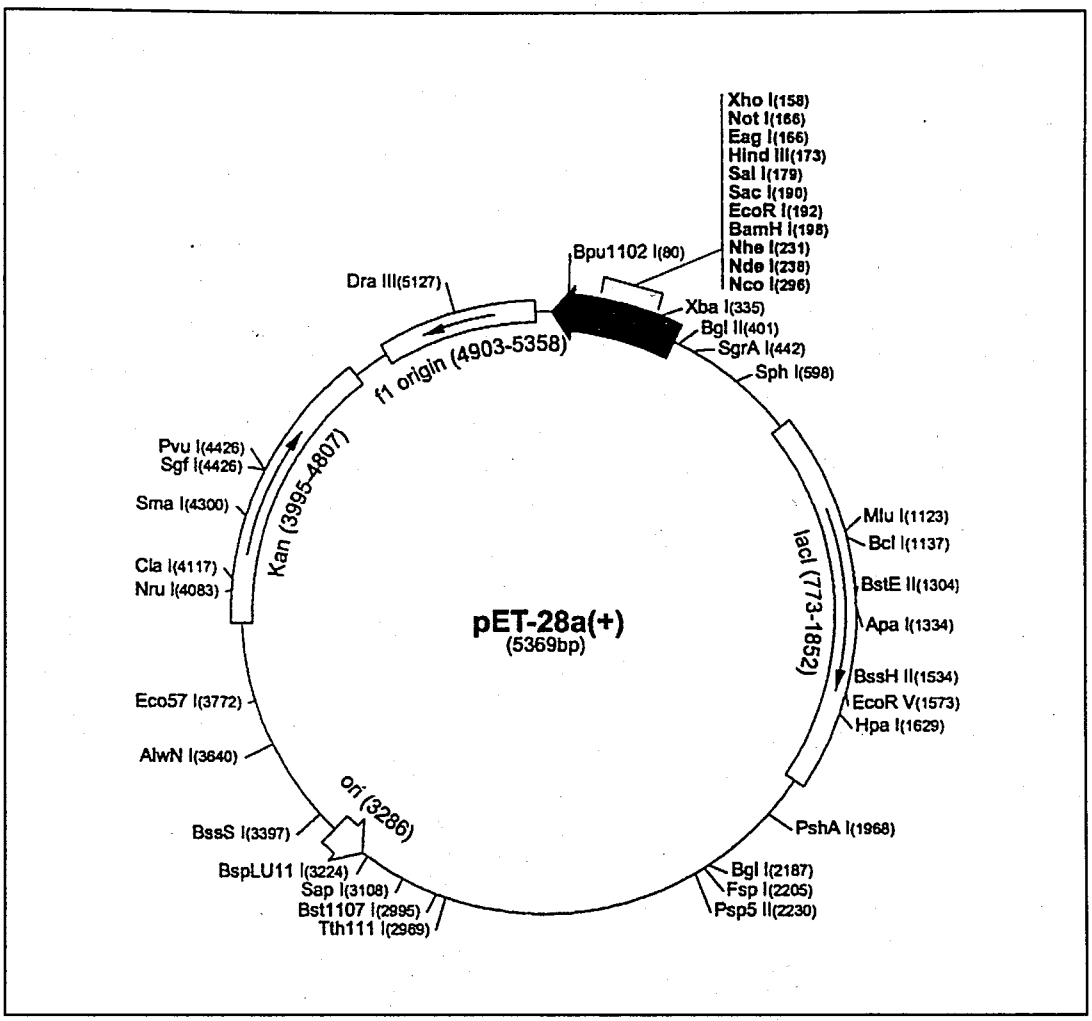


Figure 2.4. Circular map of pET28a+ vector

pET 28a+ vector also carries different sequences adjacent to the cloning sites that encode peptide tails, which perform various functions when fused to the target protein (Figure 2.5). N-terminal His-Tag/thrombin/T7-Tag configuration and the C-terminal His-Tag sequence facilitate the detection and one-step affinity purification of the target protein using nickel chelating resins. When the former configuration is used, the fusion tail can be removed from the target protein with thrombin protease enzyme after affinity chromatography (Hoffmann and Roeder, 1991).

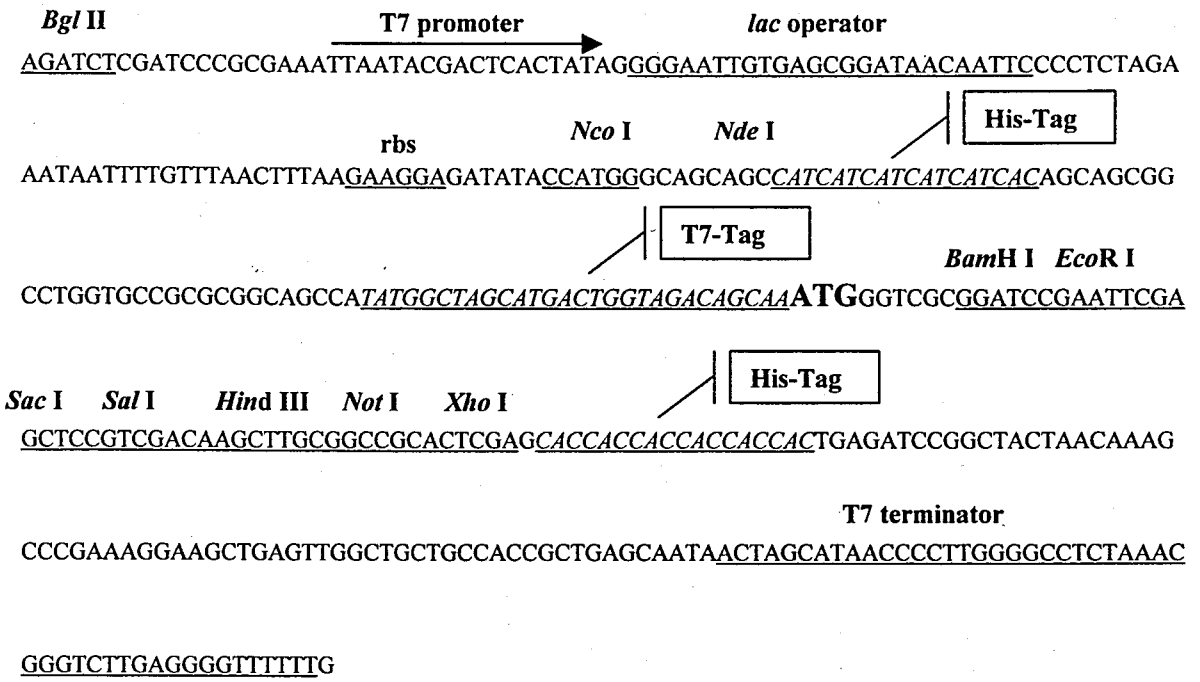


Figure 2.5. Cloning and expression region of pET28a+ vector

As shown in Figure 2.5 above, pET 28a+ vector contains translation stop codons in all three reading frames following the cloning and tag regions as well as a downstream T7 transcription terminator. The *Nde* I and *Nco* I sites within the multiple cloning region of the vector can be used for cloning into the AUG start codon at the 5' end of the insert coding sequence in order to express proteins that do not contain vector-encoded sequences. Similarly, proteins without vector-encoded C-terminal fusions can be obtained by including a translation stop codon in the insert.

The pET system has been widely used by various research groups for expression of different recombinant proteins in *E. coli*. Ernst *et al.* (1996) have cloned the heparinase I gene into four different pET vectors, expressed the heparinase I protein with an N-terminal histidine tag in *E. coli* BL21(DE3) lysogens and purified it to homogeneity by nickel-chelate chromatography. Among the four vectors, pET28a+ gave an expression yield of 19 heparinase I units per mg of protein in insoluble form, which was almost 20 times higher than that obtained with pET15b vector in soluble form. Ernst *et al.* (1997) have also made

a comparative analysis of the production costs of soluble and insoluble heparinase I expression and found that the purification process for soluble expression was 80 per cent more expensive than the insoluble one, mainly due to the low expression level, which resulted in large process volumes during fermentation and harvesting, leading to higher capital and raw materials costs.

Hanlon and Barford (1998) have used pET28a+ vector for the expression of the kinase associated phosphatase (KAP) as a His-tagged protein. Contrary to earlier unsuccessful attempts to purify this protein as a GST-fusion, by using pET system, they have obtained sufficient amounts of purified protein (3 mg of KAP per liter culture) to undertake crystallization studies. A similar result with the NS5B protein of Hepatitis C Virus (HCV) has been published by Yuan *et al.* (1997). They have first expressed this protein as a GST-fusion and then as His-tagged protein using pET vectors. Although the protein formed inclusion bodies in both cases, expression and recovery levels obtained in the latter system were much higher, making the pET system the preferred method for expression and purification of this protein.

Komai *et al.* (1997) have reported that co-expression of T7 lysozyme enzyme with the highly toxic HIV-1 protease resulted in ten times higher expression levels but also in a concurrent retardation of cellular growth. Poor biomass yields due to slower growth have also been observed by Kim *et al.* (1998) who have compared the expression of a bioadhesive precursor protein (BP) in BL21(DE3) strains with and without pLysS plasmid. The BP protein expression level was found to be lower in the strain harboring pLysS. Also in the periplasmic expression of the toxic Phospholipase D protein, pLysE and pLysS plasmids showed 22 per cent and 65 per cent lower product yields when compared with BL21(DE3) host, respectively (Mishima *et al.*, 1997). Therefore the use of pLys plasmid for controlled expression depend highly on the protein and its toxicity.

Due to the acetate controlling ability of BL21(DE3) strain, pET system has been used for the production of various proteins in high-cell-density cultures of *Escherichia coli*. Fed-batch bioreactor experiments with different feeding strategies have been reported by Shiloach *et al.* (1996) for *Pseudomonas* exotoxin A, Shin *et al.* (1998) for recombinant

human glucagon and human growth hormone, by Wong *et al.* (1998) for the bioadhesive protein, by Akesson *et al.* (1999) for xylanase and by Saraswat *et al.* (2000) for the human interleukin-2.

2.5. *Taq* I Restriction-Modification System

Restriction-modification systems are defense mechanisms of bacteria against invading foreign DNA. They are composed of a restriction endonuclease and a modification methylase. Restriction endonucleases cleave DNA at, or close to, their recognition sequences whereas modification methylases alter the sequences by methylation of internal adenine or cytosine residues such that the sequences are no longer susceptible to cleavage by the endonuclease. The function of restriction-modification systems is thought to be protective. The endonuclease protects the cell from infection by digesting unmethylated foreign DNA, and the methylase protects the cell from auto-digestion by modifying its DNA (Slatko *et al.*, 1987).

Brock and Freeze (1969) originally described the isolation of a new thermophilic bacterium, *Thermus aquaticus* YT1 in 1969. The optimum growth temperature of this microorganism is 70°C with a generation time of 50 minutes. *Taq* I restriction-modification system of *Thermus aquaticus* YT1, which was first described by Sato *et al.* (1977), consists of *Taq* I restriction endonuclease and *Taq* I methylase enzymes.

2.5.1. *Taq* I Restriction Endonuclease

Taq I is a type II restriction endonuclease that has been purified from *Thermus aquaticus* YT1 strain and is found to be very specific to the double-stranded palindromic 5' TCGA 3' target sequence (Sato *et al.*, 1977). *Taq* I restriction endonuclease cleaves both strands between the thymidine and cytosine residues leaving a 5' staggered end upon digestion. Its specificity is several million fold relative to sites that differ by only one base pair, a property common to the type II restriction endonucleases (Mayer and Barany, 1995). The enzyme cleaves bacteriophage lambda DNA at 121 sites, pBR322 at seven sites, sV4O at one site and bacteriophage ΦXl74 RF DNA at ten sites (Sato *et al.*, 1977). It has a temperature optimum of 65°C and a molecular weight of 31612 Daltons.

Kinetic analysis has shown that both high temperature and magnesium are required for specific binding and catalysis of *Taq* I endonuclease. In addition to its wide usage as a restriction endonuclease, it is also a model enzyme used to study DNA-protein interactions because of its high specificity. Moreover, the simplicity of its recognition site and its requirement for the cofactor Mg^{2+} makes the monitoring of its activity simple (Zebala, *et al.*, 1992a).

Taq I is a thermostable enzyme, maintaining its activity up to temperatures as high as 70°C. Moreover, *Taq* I restriction endonuclease is stable under conditions that would denature most of the proteins, it can cleave its canonical sequence in 50 per cent formamide or 7 M urea at 65°C (Barany, 1988a). The advantage is the inactivation of nonspecific endo- and exo-nucleases at this temperature.

The *Taq* I restriction endonuclease exhibits star activity under conditions of low salt, high pH, high enzyme concentrations, organic solvents such as dimethylformamide and ethylene glycol, 10 per cent DMSO and substitution of Mn^{2+} for Mg^{2+}. It was found that greater relaxation in specificity was found in 9LS buffer containing 20 per cent DMSO (9LS buffer: 10 mM Tris-HCl pH 9.0, 5 mM $MgCl_2$, 6 mM mercaptoethanol). Star activity is defined as the ability to cleave sites that differ by one base pair from the canonical site (Zebala *et al.*, 1992b; Barany., 1988b). It is reported that insertions at amino acid positions 47, 99, 101 and 108 result in less than 0.1 per cent wild type activity, suggesting that amino acids in those regions interact directly with the substrate (Mayer and Barany, 1995).

2.5.2. *Taq* I Methylase

Taq I methylase enzyme modifies the adenine residue on each strand of the palindromic 5' TCGA 3' target sequence by transferring a methyl group from its cofactor S-adenosyl-L-methionine. It consists of 421 amino acids and has a molecular weight of 47856 Daltons (Barany *et al.*, 1992).

Taq I methylase activity highly depends on pH, salt concentration and temperature. Methylase activity is retained in a narrow pH range with an optimum value of 7.4. The

enzyme activity decreases to about half at 0.1 M NaCl and is completely lost above 0.2 M NaCl concentrations. Being a thermostable enzyme, it maintains its activity up to temperatures as high as 70°C. Lowering the working temperature to 37°C results in 2.5-fold lower enzyme activities (Sato *et al.,* 1980).

From the three dimensional structure of *Taq* I methylase in complex with its cofactor S-adenosyl-L-methionine, the enzyme was reported to be divided into two domains. The N-terminal domain contained the cofactor inserted into a cavity, all conserved segments and the enzymatically functional parts, whereas the smaller C-terminal domain was less structured (Labahn *et al.*, 1994).

Site-directed mutagenesis at amino acid positions 108 and 196 resulted in dramatically reduced enzyme activities. Tyrosine at position 108 was found to be important for placing the target base adenine in an optimum position for methyl group transfer, and phenylalanine at position 196 was crucial for stabilizing the adenine residue at this position (Pues *et al.*, 1999).

2.5.3. Expression of *Taq* I Restriction-Modification System in *E. coli*

Taq I restriction-modification system was first cloned and sequenced in 1987 by Slatko *et al.* They prepared libraries by ligating restriction fragments of *Thermus aquaticus* YT1 genomic DNA into pBR322 vector and propogating the recombinant plasmids in *Escherichia coli.* It was found that the restriction endonuclease gene is 702 bp in length coding for the enzyme of a molecular weight of 27523 Daltons. The strain dependent viability of the recombinant cells was studied and it was observed that the restriction endonuclease is not lethal in the absence of modification methylase. This could be due to the fact that the cells are grown at 37 ^{0}C and hence the restriction endonuclease can only nick the DNA at its recognition sequence, which is repairable by *Escherichia coli* ligases (Slatko *et al.,* 1987).

The *Taq* I restriction endonuclease gene was cloned under the control of the alkaline phosphatase (phoA) promoter to isolate and characterize *Taq* I restriction endonuclease mutants and to study specificity of protein DNA interactions (Barany, 1987). The aim of

this study was to separate the methylase and endonuclease genes, to place the restriction endonuclease behind a strong promoter and to remove extraneous restriction sites thus allowing for subsequent two-codon insertion mutagenesis. When *Taq* I gene was introduced into a small pBR322 derivative, the recovery was 5000 Units from five ml of culture. The *Taq* I restriction endonuclease gene was then cloned under the phoA promoter yielding 1000 Units from five ml of culture. When the nucleotide sequences between phoA and *Taq* I endonuclease gene were deleted, the yield increased to 10000 Units from five ml of culture. The endonuclease and phoA were then cloned under T7 promoter yielding 30000 Units from five ml culture. When the endonuclease gene was finally reunited with the methylase gene, the yield increased to 45000 Units from five ml culture.

This study also revealed that insertion of two amino acids by two codon mutagenesis at position nine or ten had no effect on the activity while in frame deletions between amino acid ten and 22 almost completely abolished activity, indicating the importance of the N terminus.

The *Taq* I endonuclease gene was fused to the first four codons of the alkaline phosphatase signal sequence and overproduced under the phoA promoter with a yield of five per cent of the total cellular proteins. The sequence of the *Taq* I restriction endonuclease gene sequence revealed a 14 base pair hairpin structure. Replacement of the hairpin region by degenerate codons increased the yield. When the methylase gene was co-expressed, the increase in yield was remarkable, overproducing the restriction endonuclease to 30 per cent of the cellular protein. Three different *E. coli* strains RR1, HB 101 and MM294 were used in this study. It was observed that RR1 strain tolerated *Taq* I methylase containing plasmids because it lacked an endonuclease that cleaves DNA containing N^6-methyladenine (Barany, 1988c).

In 1992, Barany *et al.* (1992) redetermined the originally reported endonuclease and methylase gene sequences and found that the C terminus of the methylase gene overlapped with the N terminus of the restriction endonuclease gene by thirteen codons. It was suggested that the overlap region allows for formation of a hairpin structure, which allows termination of transcription. Failure to translate a full-length methylase might result in premature transcription termination at the hairpin loop, preventing the synthesis of the

restriction endonuclease, which is lethal, since the methylase is not expressed. However, proper translation of the methylase would occlude formation of the hairpin, allowing transcription to continue. Thus, the overlap with a hairpin region serves as a protector against degradation of DNA and hence loss of viability (Barany *et al.*, 1992).

According to the new sequence, *Taq* I restriction endonuclease and *Taq* I methylase enzymes are encoded by gene sequences of 792 base pairs and 1266 base pairs long, respectively (Figure 2.6 and 2.7). There are seven cognate sequences in the *Taq* I restriction endonuclease gene for the enzyme, while there are none in the methylase gene (Slatko *et al.*, 1987). This is thought to be important in the sequential expression of the two genes. The restriction endonuclease gene cannot be expressed until the DNA has become fully modified. When the restriction endonuclease is expressed, it cleaves its own DNA, interrupting its further transcription and hence preventing further endonuclease synthesis.

1 **ATG**GCTTCCA CACAAGCCCA GAAAGCGCTC GAAACTTTTG AGCGTTTTCT
51 CGCAAGCTTG GACCTCGAGT CCTACCAGCA AAAGTACCGC CCTATCAAAA
101 CGGTTGAACA AGACCTGCCT AGGGAGCTGA ACCCGCTTCC GGACCTGTAC
151 GAGCATTATT GGAAAGCGCT TGAGGATAAC CCTTCCTTCC TGGGCTTCGA
201 AGAGTTCTTT GACCACTGGT GGGAAAAGCG CCTACGGCCC TTGGACGAGT
251 TCATACGCAA ATACTTTTGG GGATGCTCCT ACGCGTTTGT TCGCTTGGGC
301 CTCGAGGCTA GGCTGTACCG AACAGCCGTT TCCATCTGGA CTCAGTTTCA
351 CTTCTGCTAC CGCTGGAACG CCTCCTGCGA GCTTCCTCTA GAAGCTGCCC
401 CAGAACTCGA CGCCCAAGGG ATAGACGCGC TGATTCATAC AAGCGGGTCC
451 TCAACAGGAA TCCAGATCAA AAAGGAAACT TACCGTTCCG AGGCCAAGAG
501 CGAGAACCGC TTTTTAAGGA AGCAAAGAGG CACCGCCCTC ATCGAGATTC
551 CCTACACCCT GCAGACACCA GAGGAGCTCG AAGAAAAAGC CAAACGGGCA
601 AGAGTGAACG GAGAAACCTA CCGTCTATGG GCCAAGGTTG CACACCATTT
651 GGACCGTCTA GAAAACGGAT TCGTCATTTT TCGGGAAAGT TATGTGAAAA
701 GCATTGAGCT TTTTCTCCAG AAAAACGCTC CTACCCTATC TGGGCTCATC
751 CGCTGGGACA GGGTGGCCCA GGAAGCCCTC ACCGCCCCGT **GA**GGTAGACA
801 CGAAGCACAA GCCCCACAGC AAAGAGCAGG CCCACGGCAC GG

Figure 2.6. The nucleotide sequence of the *Taq* I restriction endonuclease gene (Barany *et al.*, 1992).

1.**ATG**GGCCTGC CACCCCTTCT GTCCTTACCT TCCAACTCCG CCCCCAGGAG
51 CCTGGGCCGG GTGGAGACCC CCCCGGAGGT GGTGGACTTC ATGGTCTCCC
101 TGGCCGAGGC GCCCAGGGGG GGAAGGGTGC TGGAGCCCGC CTGCGCCCAT
151 GGGCCCTTCC TCCGGGCTTT CCGGGAGGCC CACGGGACGG GCTACCGCTT
201 CGTGGGGGTG GAGATAGACC CAAAAGCCCT GGACCTCCCC CCCTGGGCCG
251 AGGGCATCCT GGCGGACTTC CTCCTCTGGG AGCCGGGGGA GGCCTTTGAC
301 CTGATCCTGG GCAATCCGCC TTACGGCATC GTAGGAGAAG CCAGCAAATA
351 CCCCATTCAC GTCTTCAAAG CGGTCAAGGA CCTCTACAAG AAGGCCTTTT
401 CCACCTGGAA GGGCAAGTAC AACTTGTACG GGGCCTTTCT TGAAAAGGCC
451 GTTCGCCTTC TTAAGCCTGG TGGGGTCCTC GTCTTTGTAG TCCCGGCCAC
501 CTGGCTTGTC CTGGAGGATT TTGCCCTCCT TCGCGAGTTC CTTGCCCGGG
551 AAGGGAAAAC ATCTGTATAC TACCTTGGCG AGGTTTTCCC GCAAAAAAAG
601 GTTAGCGCTG TAGTGATTCG CTTCCAGAAG AGCGGAAAAG GCCTTTCACT
651 TTGGGATACC CAAGAAAGCG AAAGCGGGTT CACGCCCATC CTCTGGGCTG
701 AATATCCACA TTGGGAAGGA GAGATTATCC GCTTTGAAAC AGAGGAGACG
751 CGGAAGCTGG AAATATCGGG AATGCCACTG GGAGACCTCT TTCATATCCG
801 CTTCGCCGCA AGAAGCCCTG AATTCAAGAA ACATCCAGCA GTGAGAAAGG
851 AACCGGGGCC AGGTCTTGTG CCTGTGCTCA CAGGAAGAAA TTTAAAGCCG
901 GGGTGGGTAG ATTACGAGAA AAACCACTCC GGGCTTTGGA TGCCCAAGGA
951 AAGGGCCAAG GAGCTCAGGG ACTTCTATGC CACGCCCCAC CTGGTGGTAG
1001 CCCACACCAA GGGGACTAGA GTGGTGGCCG CTTGGGACGA AAGGGCCTAC
1051 CCCTGGCGGG AGGAGTTCCA CCTCCTGCCC AAGGAAGGTG TGAGACTGGA
1101 CCCCTCGTCC CTGGTGCAGT GGTTAAACTC CGAAGCGATG CAGAAGCACG
1151 TCAGGACGCT TTATCGCGAC TTCGTGCCCC ACCTGACGCT GAGGATGCTA
1201 GAAAGGCTTC CTGTAAGGAG GGAATATGGC TTCCACACAA GCCCAGAAAG
1251 CGCTCGAAAC TTT**TGA**

Figure 2.7. The nucleotide sequence of the *Taq* I methylase gene (Barany *et al.*, 1992).

3. MATERIALS

3.1. Bacterial Strains and Plasmids

E. coli TB1 [F$^-$, ara, Δ(lac-proAB), rpsL (Str ^{r}), ϕ80dlacZΔM15, hsdR (r_k^-, m_k^+)], *E. coli* ER2508 [lon:: Δ16Δ17, Δ(malB)zkb::tn5, Δ(lacZYA-argF)U169, Δ(mcrC-mrr)20, ara-14, galK2, rpsL20, xyl-5, mtl1, supE44, leuB6, fhuA2] and pMAL-p2 were supplied by the manufacturer New England BioLabs (Beverly, MA). *E. coli* ER2508 strain has a partial deletion of the *mal*E gene. Recombination deficient *E. coli* XL1 strain [F' : tn10 (tet ^{r}), Δ(lac-proAB), lacI^q, recA1, endA1, gyrA96 (Nal ^{r}), thi-1, hsdR17(r_k^-, m_k^+), supE44, relA1, Δ(lac), λ$^-$] was from our laboratory stock. *E. coli* BL21(DE3) [F$^-$, ompT, hsdS_B (r_B^-, m_B^-), gal, dcm, Δ(srl-recA)306 :: tn10 (tet ^{r})(DE3)] and pET28a+ were from Novagen (Madison, WI).

Recombinant plasmid pH185, where the expression of MBP-*Taq* I endonuclease fusion protein is under the control of the *tac* promoter, was constructed in our laboratory using pMAL-p2 as the cloning vector, as described elsewhere (Özdinler, 1996).

3.2. Chemicals

All chemicals and solutions used in this study were from MERCK (GERMANY) or SIGMA (USA) or DIFCO (USA) unless stated otherwise in the text.

The restriction enzymes were from PROMEGA (USA) and from New England Biolabs (USA). The enzyme *Pfu* DNA Polymerase was purchased from STRATAGENE (USA). DNA and protein size markers were purchased from PROMEGA (USA) and SIGMA (USA) respectively. The Centricon-30 and -100 spin dialysis membranes were purchased from AMICON (USA). The instant films (667) were from POLAROID (USA).

3.3. Oligonucleotide Primers

Oligonucleotide primers used for the amplification of *Taq* I endonuclease and *Taq* I methylase genes were designed using the sequence information given by Barany *et al.* (1992).

3.3.1. Primers for the Amplification of *Taq* I Endonuclease Gene

T7NcoI : 5' GGA GGG ACC ATG GCT TCC ACA CAA GCC CAG 3'
T7EcoRI : 5' TAT GAA TTC TCA CGG GGC GGT GAG GGC TTC CTG GGC CAC 3'

3.3.2. Primers for the Amplification of *Taq* I Methylase Gene

METF : 5' CGG TAG GAT CCG CAA ATG GGC CTG CCA CCC 3'
METR : 5' CTT GCG AGG ATC CGC TCA AAA GTT TCG AGC 3'

3.4. Bacterial Cell Media

Defined Medium	
Citric acid	1.7 g
KH_2PO_4	13.5 g
$(NH_4)_2HPO_4$	4 g
Casamino acids	0.5 g

dissolved in 933 ml deionized and distilled water. pH was adjusted to 7.0 with 5N NaOH. After autoclaving at 121°C for 15 minutes, 100mg filter sterilized thiamine, 1.2 g $MgSO_4$ $.7H_2O$ (0.6 g/ml stock solution sterilized separately), 10 ml trace metal solution and 50 ml glucose solution were added to one liter final volume.

Trace Metal Solution

$CoCl_2 .6H_2O$	2.5 g
$FeSO_4 .7H_2O$	10 g
$CaCl_2$	2 g
$ZnSO_4 .7H_2O$	2.25 g
$MnSO_4 .H_2O$	1.29 g
$CuCl_2 .2H_2O$	0.3 g
$(NH_4)_6Mo_7O_{24} .4H_2O$	0.1 g
H_3BO_3	0.45 g

per liter of deionized and distilled water. Sterilized by filtration and stored in a dark bottle.

LB Medium

Tryptone	10 g
Yeast Extract	5 g
NaCl	5 g
pH = 7.0	

per liter of deionized and distilled water.

LB-Agar Medium

Agar	15-20 g
pH = 7.0	

per liter of LB Medium.

Superbroth (SB) Medium

Tryptone	32 g
Yeast Extract	20 g
NaCl	5 g
pH = 7.0	

dissolved in 700 ml deionized and distilled water. pH was adjusted to 7.0 with 50 mM MOPS. After autoclaving at 121°C for 15 minutes, 100 ml sterile 20 per cent (v/w) Glucose solution and 200 ml sterile 5X M9 salts were added to one liter final volume.

5X M9 Salts

$Na_2HPO_4 7H_2O$	64 g
KH_2PO_4	15 g
NaCl	2.5 g
NH_4Cl	5 g

per liter of deionized and distilled water.

SOC Medium

Tryptone	0.4 g
Yeast Extract	0.1 g
NaCl	0.012 g
KCl	0.004 g
$MgCl_2$	0.04 g
$MgSO_4$	0.024 g
Glucose	0.08 g
pH = 7.0	

in 20 ml of deionized and distilled water.

TB Medium

Tryptone	12 g
Yeast Extract	24 g
Glycerol	55 mM
KH_2PO_4	17 mM
K_2HPO_4	72 mM

per liter of deionized and distilled water.

YT Medium

Tryptone	16 g
Yeast Extract	10 g
NaCl	5 g
pH = 7.0	

80-100 μg/ml of ampicillin and/or 50 μg/ml kanamycin were added into either the liquid medium or the agar medium after they cooled down to 50°C following sterilization.

3.5. Buffers and Solutions

3.5.1. Buffers and Solutions Used in the Plasmid DNA Isolation

STE Buffer

Tris-HCl (pH 8.0)	10 mM
NaCl	100 mM
EDTA (pH 8.0)	1 mM

Solution I

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

Solution II

NaOH	0.2 N
SDS	1%

Solution III

5 M Potassium Acetate	60 ml
Glacial Acetic Acid	11.5 ml
dH_2O	28.5 ml

TE Buffer	
Tris-HCl (pH 8.0)	10 mM
EDTA (pH 8.0)	1 mM

Phenol	50 mM TE saturated phenol
Chloroform / Isoamylalcohol	24:1 chloroform and isoamylalcohol
Ammonium Acetate	3 M NH_4OAc
RNase A	10 mg/ml in dH_2O
Ethanol	100 %

Cracking 2X Buffer	
5 M NaOH	2 ml
10% SDS	2.5 ml
Sucrose	10 g
dH_2O	to 50 ml.

3.5.2. Sonication Buffer

Tris-HCl (pH 8.0)	20 mM
β-mercaptoethanol	10 mM
EDTA (pH 8.0)	1 mM
PMSF	0.15 mM

3.5.3. Amylose Column Wash Buffer

Tris-HCl (pH 8.0)	20 mM
NaCl	200 mM
EDTA (pH 8.0)	1 mM
DTT	1 mM

3.5.4. Agarose Gel Electrophoresis Buffers

5XTBE (Tris-Borate) Buffer	
Trizma base	445 mM
Boric acid (pH 8.3)	445 mM
EDTA	10 mM

10X Loading Buffer	
Bromophenolblue	2.5 mg
SDS	1% in one ml glycerol.

3.5.5. Polyacrylamide Gel Electrophoresis Buffers

TEMED	N,N,N,N-tetramethylethylenediamine

Acrylamide-bisacrylamide mixture (30:0.8)	
Acrylamide	29.2 g
N'N'-bis-methylene-acrylamide	0.8 g
dH_2O	to 100 ml.

Electrophoresis Separating Gel (12.5%)	
1.5M Tris-HCl (pH 8.8)	2.5 ml
Acrylamide/Bis(30:0.8)	4.2 ml
10%(w/v) SDS	100 µl
10%(w/v) Ammonium persulfate	50 µl
TEMED	5 µl
dH_2O	to 10ml.

Electrophoresis Separating Gel (10%)	
1.5M Tris-HCl (pH 8.8)	2.5 ml
Acrylamide/Bis(30:0.8)	3.33 ml
10%(w/v) SDS	100 µl

10%(w/v) Ammonium persulfate	50 µl
TEMED	5 µl
dH_2O	to 10 ml.

Electrophoresis Stacking Gel (5%)

0.5M Tris-HCl (pH 6.8)	2.5 ml
Acrylamide/Bis(30:0.8)	1.7 ml
10%(w/v) SDS	100 µl
10%(w/v) Ammonium persulfate	50 µl
TEMED	10 µl
dH_2O	to 10 ml.

Sample Buffer

0.5M Tris-HCl (pH 6.8)	1 ml
Glycerol	0.8 ml
10%(w/v) SDS	1.6 ml
2-β-mercaptoethanol	0.4 ml
0.05%(w/v) Bromophenol blue	0.2 ml

5X Running Buffer

Tris-base	15 g
Glycine	72 g
SDS	5 g

complete to one liter with dH_2O.

Stain

Coomassie blue R-250	0.1%
Methanol	40%
Acetic acid	10%

Destain

Isopropanol	25%
Acetic acid	10%

3.5.6. Buffers Used in PCR

Promega Mg^{++} free Buffer (1X)

Tris-HCl pH 7.9	100 mM
KCl	500 mM
Triton X-100	1%

10X PCR Buffer for amplification with *Pfu* DNA Polymerase

Tris-HCl (pH 8.0)	200 mM
KCl	100 mM
$(NH_4)_2SO_4$	60 mM
$MgCl_2$	20 mM
Triton X-100	1%
Nuclease free BSA	100 µg/ml

3.5.7. Buffers and Solutions for the Isolation of DNA from Low-Melting Point Agarose

TE Buffer

Tris-HCl (pH 8.0)	10 mM
EDTA (pH 8.0)	1 mM

Phenol	50 mM TE saturated phenol
Chloroform / Isoamylalcohol	24:1 chloroform and isoamylalcohol
Ammonium Acetate	3 M NH_4OAc
Ethanol	100 %

3.6. DNA Size Markers Used in Electrophoresis

λ-DNA / Hind III DNA size marker	23130, 9416, 6557, 4361, 2322, 2027, 564, 125 bp
ϕX174 / HaeIII DNA size marker	1353, 1078, 872, 603, 310, 281, 271, 234, 194, 118, 72 bp

3.7. SDS Molecular Weight Markers

Albumin, bovine serum	66 kDa
Albumin, egg	45 kDa
Carbonic Anhydrase	29 kDa

3.8. Restriction Enzymes and Their Recognition Sites

*Bam*H I	5'...G / GATCC...3'
*Eco*R I	5'...G / AATTC...3'
*Hin*d III	5'...A / AGCTT...3'
Nco I	5'...C / CATGG...3'
Pst I	5'...CTGCA / G...3'
Sph I	5'...GCATG / C...3'
Xmn I	5'...GAANN / NNTTC...3'

3.9. Laboratory Equipment

Autoclave	Medexport, CIS, C.W.I.S. Model MAC-601(Eyela, Japan)
Balance	80A-200M (Precisa, Switzerland)
Camera	DS-34 Direct Screen Instant Camera (Polaroid, USA)
Centrifuges	SORVALL RC-5B Refrigirated Superspeed Centrifuge (DuPont, USA) Biofuge 28 RS (HERAEUS, Germany) Centrifuge 5415 C (EPPENDORF, Germany) Centrifuge NF 615 (NÜVE, Turkey)
Cold Room	VWR SCIENTIFIC, VCR 422DBA (USA)
Deepfreezers	-80°C, (HETO, Denmark) -20°C, (BOSCH, Germany) -20°C, (ARÇELIK, Turkey)
Electrophoresis	Horizon 58, Model200, Horizontal Gel Electrophoresis Apparatus (BRL, USA) Miniprotean II, (BIORAD, USA)
Electroporation Equipment	Gene Pulser II, Pulse Controller II, Capacitance Extender II, (BIORAD, USA)

Fermentor	Bioflo III Batch/Continuous Fermentor, (New Brunswick, England)
Fraction Collector	Model 2110, (BIORAD, USA)
Freeze Drier	VR-1, (HETO, Denmark) GD-1, (HETO, Denmark)
Gel Dryer	Vacuum Dryer, (BIOMETRA, Germany)
Ice Machine	AF-30, (Scotsman, USA)
Incubators	EN500, (NÜVE, Turkey) FN500, (NÜVE, Turkey) Lab-Line Ambi-High-Low Chamber, (Lab-Line Instruments, USA)
Laminar Flow	HBB 2460 LaminAir, (Holten, Denmark)
Orbital Shakers	GFL 3032, (GFL,Germany) INNOVA 4340 Illuminated refrigerated Incubator Shaker (New Brunswick Scientific,USA)
pH meter	HI 8521, (HANNA Instruments, Singapore)
Power Supply	Power/Pac 300, (BIORAD, USA)
Refrigerators	+4°C, (ARÇELİK, Turkey) +4°C, (SİMTEL, Turkey)

Sonifier	Branson Sonifier 450, (VWR Scientific, USA)
Spectrophotometer	Lambda 3 UV/VIS, (PERKIN ELMER CETUS, USA)
Thermo-cyclers	Thermal Reactor TR1, (HYBAID, England) Gene Amp PCR System 9600 (PERKIN-ELMER CETUS, USA)
Transilluminators	Reprostar II, (CAMAG, Switzerland) Foto/UV 15, (Fotodyne, USA)
UV Monitor	EconoUV Monitor, (BIORAD, USA)
Vortex	ELEKTROMAG, Turkey
Water Baths	BM102, (NÜVE, Turkey) DT Hetotherm, (HETO, Denmark) CB 8-30e DT_1, (HETO, Denmark) CB 8-30e AT_{110}, (HETO, Denmark)
Water Purification Systems	MILLI-Q UF Plus, (MILLIPORE, USA) MILLI-RQ Plus, (MILLIPORE, USA) Model 2004, (GFL, Germany)

4. METHODS

4.1. Sterilization

During the experiments, sterilized equipment was used. Pipetman tips, eppendorf tubes, centrifuge tubes, columns, solutions and culture media were all sterilized at 1.02 atm and 121°C for 20 minutes in an autoclave. Glassware was sterilized at 180°C for two hours in an oven.

For the sterilization of the fermentor, motor drive was removed from the top of the vessel and placed on the motor mount at the top of the cabinet. Air lines on the inlet filter side were disconnected and exhaust condenser water lines and jacket were removed. Sampler rubber bulb was removed and then glass wool was inserted into sampler port and sampler valve was closed. During sterilization, growth medium was put into the vessel to allow the probe tips to be moisturized during sterilization. Temperature and pressure were kept at 121°C and at 1.02 atm for 30 minutes for sterilization. Prior to autoclaving, the pH probe was calibrated, whereas the DO probe was calibrated after autoclaving.

4.2. Preparation of Preculture

The strains of *E. coli* were kept in glycerol solution at -70°C. Frozen glycerol cultures were used by streaking on LB agar plates for the preparation of master plates. During the experiments, subplates, produced from the master plates, were used in order to remove one isolated colony when it was required.

Preparation of plates, inoculation and all other experimental work that require sterile environment were carried out under sterile laminar flow.

10 ml of sterile liquid nutrient medium was inoculated with a single colony of bacteria from a slant culture. The inoculating loop was sterilized by heating in a flame and

cooled on the agar plate to avoid death of the living cells that were about to be transferred. The preculture was incubated in an orbital shaker at 37°C overnight.

4.3. Growth Conditions

In shake flask experiments, cells were grown in GFL or Innova orbital shaker at 37°C and with an agitation rate of 180 rpm. Culture volume was kept at one fifth of the flask volume for proper aeration. The pH was allowed to follow its natural course.

For the bioreactor fermentations, 2.5 L New Brunswick BioFlo III bench-top fermentor with a working volume of 1.5 L was employed. The temperature was monitored and kept constant at 37°C. The pH was measured by glass electrode (Mettler Toledo) and controlled at 7.0 through the addition of either 1N HCl or 1N NaOH for experiments with the fusion protein and 5N H_2SO_4 or 5N NaOH when working with the pET expression system. Dissolved oxygen (DO) was monitored using a polarographic oxygen electrode (Phoenix) and sustained above 30 per cent saturation. Filter sterilized air was supplied to the culture at a constant rate of one vvm. The agitation speed was set at 600 rpm throughout the run. Foam control was achieved by the addition of the 3%(v/v) silicone antifoaming agent. The size of the inoculum was 1%(v/v) to initiate growth in all experiments.

4.4. Isolation of Plasmid DNA

4.4.1. Large-Scale Plasmid DNA Isolation

Large-scale plasmid DNA isolation was carried out by alkaline lysis method (Sambrook *et al.*, 1989). 50 ml culture of the bacterial strain carrying the plasmid of interest was grown to an OD_{600} of approximately 0.6. 500 ml LB medium containing the appropriate antibiotic was inoculated with 15 ml of the late-log-phase culture and incubated for 16-20 hours at 37°C with vigorous shaking.

Bacterial cells were harvested by centrifugation at 6000 rpm for 15 minutes at 4°C, resuspended in 100 ml STE Buffer and collected by centrifugation at the previous conditions. The washed bacterial pellet was resuspended in 18 ml Solution I, 100 mg lysozyme and 40 ml freshly prepared Solution II with 20 minute incubation intervals at room temperature. Then 20 ml of ice-cold Solution III was added to the bacterial lysate and shaked vigorously until a flocculent white precipitate formed. After 15 minutes of incubation on ice, the lysate was centrifuged at 6000 rpm for 15 minutes at 4°C and the supernatant was filtered into a sterile centrifuge tube.

0.6 volume of isopropanol was added to the supernatant, mixed and incubated for 30 minutes at room temperature. Nucleic acids were recovered by centrifugation at 10000 rpm for 15 minutes at room temperature, the pellet was washed with 70 per cent ethanol and dissolved in 10 ml sterile TE Buffer.

DNase-free pancreatic RNase was added to this solution to a final 10 µg/ml concentration and incubated for one hour at room temperature. One volume of phenol:chloroform (1:1) was added to the solution and shaken vigorously to remove the proteins. After centrifugation at 6000 rpm for five minutes, the upper phase was mixed with one volume of chloroform:isoamylalcohol (24:1). Plasmid DNA was precipitated by the addition of 0.3 M ammonium acetate and 2.5 volumes of absolute ethanol.

The final mixture was kept at -70°C for one hour or at -20°C overnight, the plasmid DNA was recovered by centrifugation at 15000 rpm for 30 minutes at 4°C. The pellet was washed with 70 per cent ethanol, dried and dissolved in 500 µl sterile TE Buffer.

All centrifugations were carried out in SORVALL RC-5B refrigerated superspeed centrifuge (DuPont, USA) using SS-34 type of rotor.

4.4.2. Mini-Prep Method for Small-Scale Plasmid DNA Isolation

Plasmid DNA was isolated from overnight cultures of *Escherichia coli* picked from single colonies by the mini-prep method (Sambrook *et al.*, 1989). Ten ml LB medium

containing 100 µg/ml ampicillin was incubated with a single bacterial colony at 37°C overnight with vigorous shaking. 1.5 ml of the overnight culture was placed into a micro-centrifuge tube and centrifuged at 12000 g for one minute. The medium was removed and the cell pellet was resuspended in 100 µl of ice-cold Solution I. After five minutes of incubation at room temperature, 200 µl of fresh Solution II was added, mixed by inversion and incubated five minutes on ice. To the viscous lysate formed, 150 µl Solution III was added, vortexed gently and incubated for five minutes on ice.

The mixture was centrifuged at 12000 g for five minutes and the supernatant was transferred to a fresh tube, avoiding the white precipitate. RNase A was added to the supernatant to a final concentration of 20 µg/ml and incubated at 37°C for 20 minutes.

The solution was extracted once with TE saturated phenol:chloroform (1:1) and once with chloroform:isoamylalcohol (24:1) by vortexing for one minute and centrifuging at 12000 g for two minutes at each step.

Plasmid DNA was recovered from the aqueous phase by the addition of 0.3 M ammonium acetate and 2.5 volumes of absolute ethanol. After five minutes of incubation at -70°C, the plasmid DNA was precipitated by centrifugation at 14000 g for 20 minutes. The pellet was washed with 70 per cent ethanol, dried and dissolved in 30 µl sterile TE Buffer.

4.4.3. Determination of Quantity and Quality of DNA

Spectrophotometric measurement and agarose gel electrophoresis were used to determine the quantity and quality of DNA. For the spectrophotometric method, DNA was diluted with distilled water and the UV absorption of DNA was read at 260 nm using quartz cuvettes. Since 50 µg of double stranded DNA has an absorbance of 1.0 at 260 nm, the concentration of the DNA in the samples were calculated using the formula:

$$50\ \mu g/ml \times OD_{260} \times \text{dilution factor} = \text{Concentration in } \mu g/ml$$

The OD_{260}/OD_{280} ratio was also used to determine the purity of the DNA samples. A ratio of 1.8 was considered to be pure whereas higher and lower values suggested RNA and protein contamination, respectively.

For the latter method, gels were prepared by boiling the appropriate mass of agarose in 0.5 x TBE, cooling to 50°C, adding ethidium bromide to a concentration of 5 μg/ml and pouring this in the gel former. Different dilutions of DNA samples were loaded on the appropriate concentration of agarose gel, after mixing with 0.5 volume of DNA loading buffer and were run at 150 V. The known amounts of λ DNA were also run on the same gels. Ethidium bromide intercalates DNA, and in this state fluoresces when illuminated by UV light. DNA was visualized by illuminating the gel with short wavelength UV light on a transilluminator. The unknown amount was estimated by comparing the intensity of the fluorescence in each sample.

4.5. Cloning of the *Taq* I Methylase Gene into pBR322 Vector

The recombinant plasmid pMETaq was constructed by the amplification of the *Taq* I methylase gene from the *Thermus aquaticus* YT1 bacterial genome by specifically designed primers and by ligating the amplified insert gene into the *Bam*H I site of the pBR322 plasmid vector.

4.5.1. PCR Amplification of the *Taq* I Methylase Gene

Genomic DNA from *Thermus aquaticus* YT1 strain was used as a template for the PCR amplification of the 1266 base pair *Taq* I methylase gene using primers METF and METR. All PCR reactions were carried out in a reaction volume of 25 μl containing 12.5 pmole of each primers, 2.5 μl of *Pfu* DNA Polymerase 10X Buffer, 0.3 mM dNTP, 1.25 Units of *Pfu* DNA Polymerase and 0.2 μg of *Thermus aquaticus* genomic DNA. Amplification was performed in the Hybaid thermal cycler, with an initial denaturation step at 96°C for five minutes followed by 25 cycles of alternating incubations at 93°C for one minute, at 66°C for one minute , at 72°C for three minutes, ending with ten minutes of incubation at 72°C.

4.5.2. Isolation of PCR Products From Low-Melting Temperature Agarose Gels

The PCR amplification product was purified from excess primers, dNTPs, DNA polymerase and unspecific products using low melting point agarose gel electrophoresis (Sambrook *et al.*, 1989). PCR products were collected together, mixed with 1X Loading Buffer and loaded onto one per cent low melting point agarose gel which was prepared in 1X TBE Buffer. After electrophoresis, the 1398 bp fragment was cut out of the gel and was kept at 65°C for 15 minutes. When the gel melted, equal volume of TE buffer was added and the solution was kept at 65°C for an additional 20 minutes until it melted completely.

The mixture was then extracted twice with hot phenol, and once with chloroform. The desired DNA fragment was recovered from the aqueous phase by the addition of 0.1 volume of 3 M ammonium acetate and two volumes of ice-cold ethanol and centrifuged at 15000 rpm, using SS-34 type of rotor in Sorvall centrifuge, for 30 minutes. The pellet was dried and dissolved in sterile TE buffer.

4.5.3. Digestion of the Vector and Insert with *Bam*H I

One µg of the purified PCR fragment and five µg of pBR322 plasmid vector were both digested with 100 U *Bam*H I in a total volume of 50 µl as described by the manufacturer and incubated at 37°C overnight. Then the digestions were extracted once with phenol:chloroform and once with chloroform:isoamylalcohol and precipitated with ethanol. The dried pellets were dissolved in 50 µl sterile distilled water. The sticky ends of the vector were dephosphorylated with Calf Intestinal Alkaline Phosphatase (CIAP) in a total volume of 50 µl as described by the manufacturer and purified by spin dialysis using Centricon-100 filters. Using the Eppendorf 5415c bench top centrifuge, the filter was centrifuged two times at 5000 rpm for ten minutes. Then the filter was inverted and spinned at 5000 rpm for two minutes. Approximately 50 µl linear vector was recovered which was pure enough for the ligation reaction.

4.5.4. Ligation Reaction

In a microfuge tube, 100 ng pBR322 vector, 300 ng insert and distilled water were mixed in a total volume of 16 µl, incubated at 65°C for five minutes and cooled on ice. Two µl 10X Ligase buffer and 800 Units T4 DNA Ligase were added to this mixture, which was then incubated at 16°C overnight. After the completion of the reaction, two or three µl of samples were loaded on a 0.8 per cent agarose gel to check the efficiency of the ligation reaction before transformation.

4.5.5. Preparation of Competent Cells

100 ml of LB was inoculated with one ml of overnight *Escherichia coli* cell culture and was kept with vigorous shaking at 37°C until the OD_{600} reached 0.5. The cells were then chilled on ice for thirty minutes before being centrifuged at 4°C at 6000 rpm, using SS34 type of rotor, for six minutes.

The LB medium was discarded and the pelleted cells were washed twice with 100 ml of sterile distilled water, kept on ice for ten minutes before centrifugation at 6000 rpm for six minutes at 4°C using SS-34 type of rotor. The cells were then washed with two ml of ten per cent glycerol centrifuged and resuspended in 200 µl of ten per cent glycerol. The competent cells were then aliquoted in 40 µl volume of samples and were kept at -70°C for storage. All centrifugations were carried out in SORVALL RC-5B refrigerated superspeed centrifuge (DuPont, USA).

4.5.6. Transformation of *Escherichia coli* Cells by Electroporation

Electroporation was employed to transform the competent cells with the ligation mixture using a modified method described by Dower *et al.* (1988). Three tubes which contained 40 µl of competent cells were prepared. Two µl of the ligation mixture was added to the first tube, one µg of undigested, untreated pBR322 vector was added to the second tube and nothing was added to the last tube. Competent cells cannot survive in the presence of ampicillin unless they harbor a plasmid that confers antibiotic resistance. The

third tube was thus used as a control to check if the competent cells were free of any contamination. Competent cells were then placed into the electroporation cuvette on ice avoiding bubbling in the cuvette. The electroporation apparatus was adjusted to a voltage of 2.5 kV, resistance of 200 Ω, capacitance of 25 μF before applying the electric shock to the competent cells. Two ml of cold SOC medium was immediately added to the electroporation cuvette and the competent cells were then transferred to 37°C water bath for two hours of incubation with gentle shaking time to time.

Different dilutions of the transformed cells were spread on to LB agar plates containing ampicillin. The plates were kept at room temperature for about ten minutes to allow the spread cells to diffuse completely and were then transferred to 37°C oven for 16-20 hours of incubation.

4.5.7. Screening of Recombinant Colonies

4.5.7.1. Cracking Procedure for Rapid Estimation of Plasmid Size. Individual colonies from agar plates were picked using sterile pipet tips, and dissolved in 25 μl 10 mM EDTA, pH 8.0. 25 μl of freshly made cracking 2X buffer was added and the samples were incubated at 70°C for about ten minutes and cooled down to room temperature. 1.5 μl of 4 M KCl and 0.5 μl of 0.4 per cent bromophenolblue were added to the samples, vortexed and incubated ten minutes on ice. Then they were centrifuged at 14000 g for three minutes and 10-20 μl of the supernatant was immediately loaded on an agarose gel. Colonies containing plasmids of a known size were cracked together with the other samples and used as markers (Sambrook *et al.*, 1989).

4.5.7.2. PCR Analysis. The plasmids of the suspected colonies were isolated by mini-prep method as described in 4.4.2. Using the purified plasmid as template and the identical PCR conditions used in the amplification of the insert, the methylase gene contained within the recombinant plasmids were selectively amplified and the size of the insert was analyzed on one per cent agarose gels.

4.5.7.3. Restriction Enzyme Analysis. Orientation of the *Taq* I methylase gene was checked by digesting 100ng of plasmid DNA with 20 Units of *Eco*R I in a total volume of 30 µl as described by the manufacturer and incubated at 37°C for two hours. Digestions were analyzed on 1.2 per cent agarose gel and the plasmids giving the correct band pattern were identified.

4.6. Measurement of *Taq* I Methylase Activity

Presence of *Taq* I methylase activity was determined by digesting 100 ng of the recombinant plasmids with 10 Units of *Taq* I restriction endonuclease enzyme in a total volume of 20 µl at 65°C for one hour in 1X Multicore Buffer. Plasmids displaying resistance to endonuclease digestion were identified by electrophoresis on agarose gels.

4.7. Cloning of *Taq* I Methylase Gene into pET28a+ Vector

100 ng pMETaq and pET28a+ plasmids were digested with 10 Units of *Sph* I enzyme in a total volume of 100 µl as described by the manufacturer and incubated at 37°C overnight. Linearized pET28a+ was extracted once with phenol:chloroform and once with chloroform:isoamylalcohol and precipitated with ethanol. The band corresponding to the metylase gene was cut from low-melting temperature agarose gel and purified. Both DNA fragments were then digested with 50 Units of *Hin*d III enzyme in a total volume of 100 µl as described by the manufacturer and incubated at 37°C overnight. After purification of the vector and insert from low-melting temperature agarose gels, they were ligated with 800 Units of T4 DNA Ligase enzyme by incubating at 16°C overnight. The ligation mixture was used to transform competent *Escherichia coli* cells by electroporation.

Screening of the recombinant colonies, formed on kanamycin containing LB plates, was carried out by the cracking method and PCR analysis. Suspected recombinants were analyzed by incubating them with *Bam*H I and *Xmn* I restriction enzymes at 37°C for two hours. Expression of methylase gene was determined by digesting the plasmids with *Taq* I endonuclease in a total volume of 20 µl at 65°C for one hour in 1X Multicore Buffer.

4.8. Cloning of *Taq* I Restriction Endonuclease Gene into pET28a+ Vector

Taq I restriction endonuclease gene was amplified by PCR from *Thermus aquaticus* YT1 genomic DNA using primers T7NcoI and T7EcoRI. All PCR reactions were carried out in a reaction volume of 25 µl containing 12.5 pmole of each primers, 2.5 µl of *Pfu* DNA Polymerase 10X Buffer, 0.3 mM dNTP, 1.25 Units of *Pfu* DNA Polymerase and 0.2 µg of *Thermus aquaticus* genomic DNA. Amplification was performed in the GeneAmp PCR System 9600 thermal cycler, with an initial denaturation step at 96°C for five minutes followed by 25 cycles of alternating incubations at 96°C for 30 seconds, at 66°C for 30 seconds, at 72°C for two minutes, ending with ten minutes of incubation at 72°C. PCR products were collected together and purified from excess primers, dNTPs, DNA polymerase and unspecific products using low melting point agarose gel electrophoresis.

Ten µg of PCR product and pET28a+ plasmid were digested with 30 Units of *Nco* I enzyme in a total volume of 50 µl as described by the manufacturer and incubated at 37°C overnight. Endonuclease gene and linearized pET28a+ were extracted once with phenol:chloroform and once with chloroform:isoamylalcohol and precipitated with ethanol. Following the resuspension of the dried pellets with distilled water, both DNA fragments were digested with 100 Units of *Eco*R I enzyme in a total volume of 100 µl and incubated at 37°C overnight. After purification of the vector and insert from low-melting temperature agarose gels, they were ligated with 800 Units of T4 DNA Ligase enzyme by incubating at 16 °C overnight. The ligation mixture was used to transform the competent *Escherichia coli* cells by electroporation.

Screening of the recombinant colonies, formed on kanamycin containing LB plates, was carried out by the cracking method and PCR analysis and suspected recombinants were analyzed by incubating them with *Hin*d III restriction enzyme at 37°C for two hours.

4.9. Measurement of *Taq* I Restriciton Endonuclease Activity

Taq I restriction endonuclease activity was determined by digesting two µg pUC18 plasmid with appropriate dilutions of the samples at 65°C for one hour in 1X Multicore

Buffer. Digestions were analyzed by 1.2 per cent agarose gel electrophoresis. Parallel digestions were carried out with the commercial *Taq* I endonuclease to compare the activities at different dilutions. All conditions for digestion were kept constant both for the sample and the commercial *Taq* I restriction endonuclease. One Unit of *Taq* I endonuclease activity is defined as that required to digest one μg of lambda DNA to completion in one hour at 65°C.

4.10. Detection of the Expressed MBP-*Taq* I Gene Fusions

In the recombinant pH185 plasmid, the *Taq* I restriction endonuclease gene was cloned downstream from the *mal*E gene of *E. coli* which encodes for the periplasmic Maltose Binding Protein (Özdinler, 1996). *E. coli* cells harboring pH185 plasmid express MBP-*Taq* I fusion protein under control of the strong *tac* promoter therefore induction is done by the addition of the chemical inducer IPTG to the culture.

Recombinant *E. coli* cells were grown in liquid medium at 37°C and periplasmic expression of the fusion protein was initiated by adding IPTG to the growth medium at a certain phase of growth. Induced cells were harvested by centrifugation at 6000 rpm for 20 minutes at 4°C. Proteins from the extracellular medium were collected by $(NH_4)_2\ SO_4$ precipitation. $(NH_4)_2\ SO_4$ was added to the supernatant up to 85 per cent of saturation and the precipitate was collected by centrifugation at 15000 rpm for 30 minutes at 4°C and resuspended in sonication buffer. To prepare the periplasmic fraction, harvested cells were resuspended in 0.2 culture volume of 30 mM Tris-HCl containing 20 per cent sucrose. 1 mM EDTA was added to the suspension and incubated for 20 minutes at room temperature with gentle stirring. Then the cells were reharvested by centrifugation at 9000 rpm for 20 minutes at 4°C and shocked by resuspending and gentle stirring for 20 minutes in 0.03 culture volume of ice-cold 5 mM $MgSO_4$. Cytoplasmic cell extract was prepared by ultrasonication of the osmotically shocked cells by three or four rounds of 30 seconds of sonication on ice. Insoluble proteins and membranes were removed from the lysate by centrifugation at 9000 rpm for 30 minutes at 4°C. The presence of the fusion protein in these three compartments was shown by SDS-PAGE electrophoresis as described in 4.11.

All centrifugations were carried out in SORVALL RC-5B refrigerated superspeed centrifuge (DuPont, USA) using SS-34 type of rotor.

4.10.1. Chromatographic Purification of the Fusion Protein on Cross-Linked Amylose

A sterile ten ml syringe plugged with glass wool was used as the column. Ten ml amylose resin was poured and washed with three column volumes of wash buffer. The flow rate was adjusted to be approximately one drop per second. The crude cell extract, osmotic shock fluid or medium proteins were loaded onto the column and the flow-through fractions were collected into sterile glass tubes for further testing. The column was washed with 30 ml of column wash buffer. The MBP-fusion protein was eluted with wash buffer containing 10 mM maltose. Fractions of 500 µl were collected on ice and their protein concentration was determined by the Bradford method as described in 4.12.4. Ten µl samples of the flow-through fractions and eluate fractions were analyzed by SDS-PAGE. Enzyme activity of the fractions were determined as described in 4.9.

4.10.2. Cleavage with Factor Xa

The fractions containing the fusion protein were collected and the protein concentration was adjusted to one mg/ml. The protease digestion was accomplished by adding 0.2-1 µg Factor Xa to 50 µg of the fusion protein in a buffer containing 20 mM Tris-HCl, 100 mM NaCl and 2 mM $CaCl_2$. Reaction mixtures were incubated at 23 °C for 16 hours.

4.11. SDS-Polyacrylamide Gel Electrophoresis of Proteins

Sodium dodecyl sulfate-poyacrylamide gel electrophoresis (SDS-PAGE) analysis of proteins was carried out as described by Laemmli (1970) using Miniprotean II (BIORAD, USA) electrophoresis equipment.

The gel sandwich was assembled by placing the spacers between the two thin glass plates in the holder. The separating gel was prepared by mixing acrylamide-bisacrylamide, 1.5 M Tris-HCl and SDS at the specified percentage with TEMED and ten per cent APS added last and poured carefully into the gel sandwich about two cm from the top of the front plate. The top of the separating gel was covered with water to prevent contact with air and to keep the gel surface flat. The gel was allowed to polymerize and then the water was poured off. Five per cent stacking gel was prepared by mixing acrylamide-bisacrylamide, 0.5 M Tris-HCl and SDS with TEMED and ten per cent APS added last again and was poured on top of the separating gel. The comb was carefully inserted and the stacking gel was allowed to polymerize for about half an hour. After polymerization, the comb was removed and the gel was placed in the electrophoresis chamber. Electrophoresis buffer was added to the inner and outer reservoirs and any air bubbles trapped in the wells were removed by rinsing.

The protein sample was mixed with sample buffer in an eppendorf tube and the mixture was heated at 98°C for three minutes. The sample solution and molecular weight standards were loaded to the wells and the electrophoresis chamber was connected to the power supply previously set to 200 V. The electrophoresis was carried out at constant voltage setting until the dye front had migrated to the bottom of the gel. Then the chamber was disconnected from the power supply, gel plates were pried apart, the gel was removed and placed into staining solution for half an hour. The stain was then poured out and destain was added to get rid of unbound stain. After half an hour, destain was poured out and fixing solution was added to fix the proteins. The gels were then dried onto chromatography paper under vacuum for 30 minutes at 80^0C.

4.12. Analyses

4.12.1. Determination of the Growth Curve

Growth of the cells was followed by measuring the optical densities at 600nm and by determining the dry cell weight from known volumes of culture samples.

At the desired time intervals an aliqout of the cell suspension was taken and the optical density of the sample was measured at 600nm using cell-free broth as the blank. When necessary, samples were diluted with growth medium to keep the spectroscopic reading within reliable limits of 0.2-0.7.

For biomass estimation, known volumes of samples were centrifuged at 5000 rpm for 15 minutes using NF 615 NÜVE centrifuge. Pelleted cells were resuspended in one ml distilled water and transferred to 1.5 ml preweighed eppendorf tubes. Cell precipitate was collected by centrifugation at 14000 rpm for ten minutes using an Eppendorf 5415c bench-top centrifuge and then dried at 60°C for one day and reweighed.

For each *Escherichia coli* strain, a calibration chart was prepared to correlate the optical densities to dry weights in complex or defined media. Table 4.1 shows the correlation factor values, which are defined as biomass in grams dry cell weight per liter culture volume divided by OD at 600nm.

Table 4.1. Correlations of dry weights to optical densities for *E. coli* strains.

Escherichia coli strain	Medium	Correlation factor
BL21(DE3)	Complex	0.507
BL21(DE3)	Defined	0.577
ER2508	Complex	0.359
TB1	Complex	0.482
XL1	Complex	0.401
XL1	Defined	0.581

4.12.2. Determination of the Plasmid Stability

To determine the plasmid stability, serial dilutions of culture samples were spread onto LB medium agar plates and incubated overnight at 37°C. At proper dilutions, these

plates allowed the growth of both plasmid containing and plasmid free cells as single colonies. About 100 of these colonies were gently picked with the tips of sterile wooden sticks and transferred to selective LB plates containing ampicillin or kanamycin. After overnight incubation at 37°C, plasmid containing cells which formed colonies on selective plates were counted. Plasmid stability was calculated as the per cent ratio of the number of plasmid containing cells to the total cell number.

4.12.3. Determination of the Cell Viability

To determine the viable cell number, 200 µl of serially diluted culture samples were spread onto selective and non-selective LB medium agar plates and incubated overnight at 37°C. Plates having colony numbers in the range one to 100 were counted. The number of counted colonies were multiplied with the dilution factor to estimate the cell viability as colony forming units (CFU) per ml of culture.

4.12.4. Determination of the Protein Content

Bradford Dye-binding Method was used to estimate the protein content of the samples . Bovine serum albumin was used as standard (Bradford, 1976).

A series of tubes containing 0.1 to 10 µg of bovine serum albumin and increasing dilutions of the sample were prepared in a final volume of 1000 µl. Standard and sample dilutions were added to one ml of 1:4 diluted Bradford reagent. After five minutes of incubation at room temperature, the absorbance of each sample was read at 595 nm against a blank containing only the protein buffer and the diluted Bradford reagent. A standard curve showing protein amount in µg versus absorbance at 595 nm was obtained. The protein concentration in each sample was calculated using the calibration curve.

4.12.5. Determination of Glucose Concentration

The amount of residual glucose concentration in the medium was measured using the D-Glucose Kit of Boehriger Mannheim as described by the manufacturer. D-Glucose is

phosphorylated to glucose-6-phosphate (G-6-P) in the presence of the enzyme hexokinase (HK) and adenosine-5'-triphosphate (ATP) with the simultaneous formation of adenosine-5'-diphosphate (ADP).

$$\text{D-Glucose} + \text{ATP}^{+} \xrightarrow{\text{HK}} \text{G-6-P} + \text{ADP}$$

In the presence of the enzyme glucose-6-phosphate dehydrogenase (G6P-DH), G-6-P is oxidized by nicodinamide-adenine dinucleotide phosphate (NADP) to gluconate-6-phosphate with the formation of reduced nicotinamide-adenine dinucleotide phosphate (NADPH).

$$\text{G-6-P} + \text{NADP}^{+} \xrightarrow{\text{G6P-DH}} \text{gluconate-6-phosphate} + \text{NADPH} + \text{H}^{+}$$

The amount of NADPH formed in this reaction is stoichiometric with the amount of D-Glucose. The increase in NADPH is measured by means of its absorbency at 340 nm.

The test combination contains Solution I consisting of triethanolamine buffer, pH 7.6, NADP, ATP, magnesium sulfate and stabilizers, and Suspension II consisting of hexokinase and glucose-6-phosphate dehydrogenase as well as Solution III as D-Glucose standard.

Samples were placed into a water bath at 80^0C for 15 minutes to stop enzymatic reactions. They were centrifuged and the supematants (diluted accordingly) were used for the assay. 50 µl of sample, 500 µl of solution I and 950 µl of distilled water were pipetted in a cuvette. 500 µl of solution I and one ml of distilled water were pipetted in another cuvette as blank. After about three minutes, absorbencies (A_1) were read at 340 nm against air. Then, by the addition of 0.01 ml of suspension II the reaction was started. At the end of the reaction, after about 10-15 minutes, absorbencies (A_2) were read again at the same conditions both for the sample and for the blank.

The absorbency difference of the blank (A_2 - A_1) was subtracted from the absorbency difference of the samples (A_2 - A_1), thereby obtaining $\Delta A_{D\text{-}Glucose}$. The concentration of D-Glucose was calculated by the following equation:

$$C_{D\text{-}Glucose} = D \times (5.441 / \varepsilon) \times \Delta A_{D\text{-}Glucose} \text{ (g D-Glucose / L sample)}$$

where ε = adsorption coefficient of NADPH at 340 nm. = 6.3 (1 / mmol / cm)
D = Dilution factor of the samples

4.12.6. Determination of Isocitrate Dehydrogenase Activity

Isocitrate Dehydrogenase Kit of Sigma was used to determine the ICD activity in the medium and cytoplasm of cells. ICD catalyzes the following oxidative decarboxylation:

$$\text{L-Isocitrate} + \text{NADP} \xrightarrow[Mn^{++}]{\text{ICD}} \text{NADPH} + \text{2-Oxoglutarate} + CO_2$$

ICD activity is measured in terms of the increase in absorbance at 340 nm that occurs when NADP is reduced to NADPH. Activity is proportional to the concentration of NADPH formed.

The test combination contains isocitrate substrate consisting of DL-isocitric acid trisodium salt, 10 µmol/ml, pH 7.8, 0.67 mg/ml NADP and 0.15 mol/l $MnCl_2$.

125 µl of diluted sample, 300 µl isocitrate substrate and 15 µl $MnCl_2$.were pipetted in a cuvette. 125 µl of distilled water, 300 µl isocitrate substrate and 15 µl $MnCl_2$ were pipetted in another cuvette as blank. After about five minutes, absorbencies (A_1) were read at 340 nm against a blank. Then, 0.3 ml of NADP was added to start the reaction and absorbencies (A_2) were read after five minutes under the same conditions.

The concentration of ICD was calculated by the following equation:

$$\text{ICD (Sigma Units/ml)} = (A_2\text{-}A_1) \times 48000 \times D$$

where D is the dilution factor of the samples. One Sigma Unit of ICD activity is defined as the quantity of enzyme that will produce one nanomole of NADPH in one hour at 25°C.

4.12.7. Determination of Acetic Acid Concentration

Using the Acetic acid Kit of Boehriger Mannheim as described by the manufacturer, the amount of acetic acid concentration in the medium was measured. Acetic acid (acetate) is converted in the presence of the enzyme acetyl-CoA synthetase (ACS) with adenosine-5'-triphosphate (ATP) and coenzyme A (CoA) to acetyl-CoA.

$$\text{Acetate} + \text{ATP} + \text{CoA} \xrightarrow{\text{ACS}} \text{acetyl-CoA} + \text{AMP} + \text{pyrophosphate}$$

Acetyl-CoA reacts with oxaloacetate to citrate in the presence of citrate synthase (CS).

$$\text{Acetyl-CoA} + \text{oxaloacetate} + H_2O \xrightarrow{\text{CS}} \text{citrate} + \text{CoA}$$

Oxaloacetate required for this reaction is formed from L-malate and nicotinamide-adenine dinucleotide (NAD) in the presence of L-malate dehydrogenase (L-MDH) where NAD is reduced to NADH.

$$\text{L-malate} + NAD^+ \xrightarrow{\text{L-MDH}} \text{oxaloacetate} + \text{NADH} + H^+$$

The increase in light absorbance at 340 nm is proportional to NADH formed at the end of the reaction.

The test combination consists of Solution I: containing triethanolamine buffer, pH 8.4, L-Malic acid and magnesium chloride, Solution II: ATP, NAD and CoA, Solution III: L-MDH and CS, Solution IV: ACS.

50 µl of diluted sample, 500 µl of solution I, 100 µl of solution II and 950 µl of distilled water were pipetted in a cuvette. 500 µl of solution I, 100 µl of solution II and one

ml of distilled water were pipetted in another cuvette as blank. Absorbencies of the samples and blank (A_1, B_1) were read at 340 nm against air. Five µl of solution III was added to samples and blank and absorbencies (A_2, B_2) were read. Then, by the addition of 0.01 ml of solution IV the reaction was started. At the end of the reaction, after about 10-15 minutes, absorbencies (A_3, B_3) were read again at the same conditions both for the sample and for the blank.

The concentration of Acetic acid was calculated by the following equation:

$$\text{Acetic acid}(g/L) = D \times \left\{ \left[\frac{(A_3 - A_1) - (A_2 - A_1)^2}{(A_3 - A_1)} \right] - \left[\frac{(B_3 - B_1) - (B_2 - B_1)^2}{(B_3 - B_1)} \right] \right\} \times \frac{1.940}{6.3}$$

where D is the dilution factor of the samples.

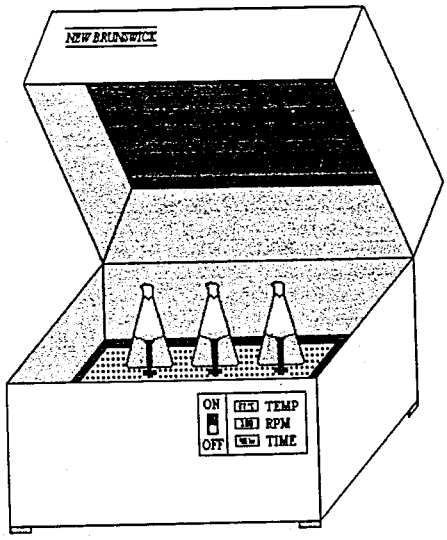
4.12.8. Quality Control Analyses

Overdigestion assay was performed by digesting pUC18 plasmid with increasing amounts of samples for 16 hours under appropriate conditions. The maximum number of units giving a clear, sharp and normal banding pattern was determined by agarose gel electrophoresis. In the cut-ligate-recut assay, pUC18 plasmid was digested with samples to completion, ligated with T4 DNA Ligase and recut with the same sample. Cut, ligated and recut DNAs were analyzed by agarose gel electrophoresis.

4.13. Experimental Set-Up

Fermentation experiments were performed either in shake flasks or in a 2.5 L New Brunswick BioFlo III batch/continuous fermentor. Figures 3.2 shows the experimetal set-up.

Orbital Shaker



BioFlo III batch/continuous fermentor

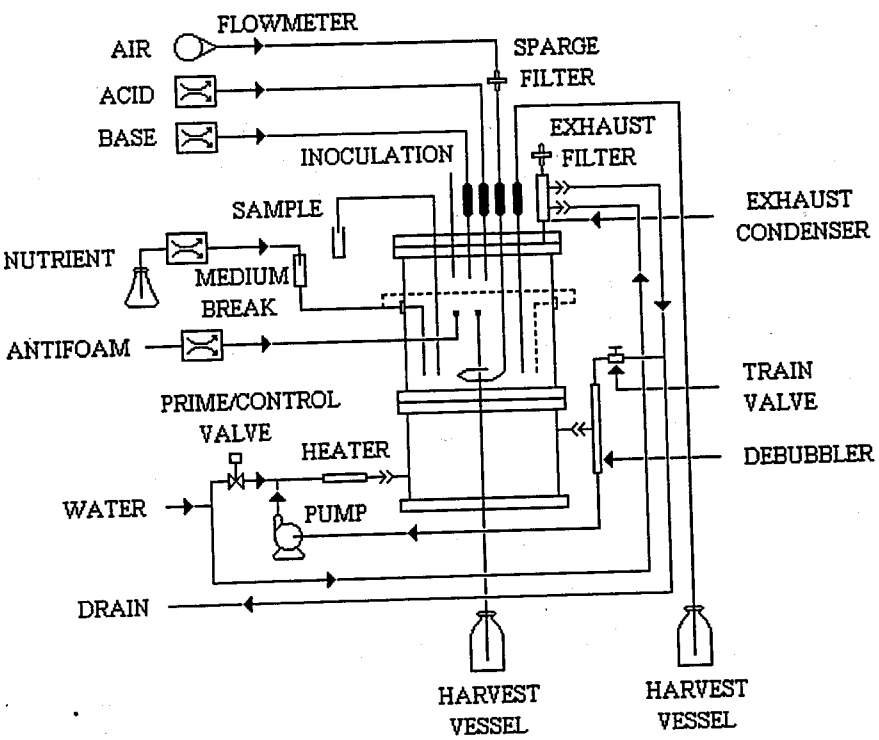


Figure 4.1. Experimental set-up.

4.14. Instrumentation and Control of the Fermentor

- **Description of Vessel.** The vessel parts of the fermentor consist of a flanged glass tube (thick walled) vessel (2.5 L) body that is detachable from the bottom-dished head and a stainless steel head plate. The dished head is jacketed for circulation of temperature controlled water. For the addition of antifoam, nutrients, and for the vessel overflow in continuous culture studies, four sterilizable polypropylene compression ports are provided in the glass wall. Ports are provided in the headplate for inoculation, base and acid addition, a thermowell for a resistance temperature detector, a temperature sensor, a sparger, a harvest port, a sampling port, an exhaust condenser, dissolved oxygen and pH electrodes. The drive bearing housing is allocated on the headplate.
- **Temperature Control.** The culture temperature is controlled by a microprocessor based PI (proportional and integral) controller. The range of temperature is in the range from 20°C to 60°C (± 0.1°C). The medium temperature is sensed by an RTD (resistance temperature detector) submerged in the thermowell.
- **Agitation System.** At the top of the bearing housing a removable agitation servo motor is located which is connected to the agitation shift with a multi jaw coupling. During sterilization of the vessel, motor can easily be disconnected and replaced after sterilization. Agitation speed range of the motor is, 12-1000 (±1) rpm.
- **Aeration.** Sterile air is controlled by the needle valve of the flowmeter. The medium can be aerated through the ring sparger which can provide 1.5 (working volume) of sterile air through 0.2 μm replaceable cartridge filter. The filter is sterilizable with the vessel. With the system, oxygen transfer rates of 350 mM O_2/L/hr may be obtained.
- **pH Control.** The pH of the medium is sensed by a glass electrode. Control is maintained by PID controller which operates two peristaltic pumps connected to acid and base addition ports. pH is controlled in the range of 2.00-12.00 (±0.01).
- **Dissolved Oxygen Control.** DO is sensed by a polarographic DO electrode and control is maintained by the PID controller that changes the speed of agitation. The range of control is in the range of 5-95% (± 1%).

  The polarographic dissolved oxygen electrode consists of polarized platinum and silver electrodes, with the electrolyte separated from the sample by teflon/silicone gas

permeable membrane. Oxygen diffuses across the electrode membrane and is reduced to hydroxyl ions at the platinum cathode according to the reaction;

$$O_2 + 2\,H_2O + 4e^- \longrightarrow 4\,OH^-$$

The electrons necessary for this process are produced by a reaction at the silver anode. Because the electrolyte contains chloride ions, this reaction occurs as;

$$4\,Ag + 4\,Cl^- \longrightarrow 4\,AgCl + 4\,e^-$$

At any given temperature, the current flow between cathode and anode is directly proportional to the level of oxygen outside the membrane.

- **Foam Control.** During batch fermentation, foam is controlled by the antifoam probe that is located in the headplate. The controller operates the antifoam addition pump that adds chemical defoamer through the port in the wall of the vessel. Sensitivity and time adjustments are made on potentiometers located on a small printed circuit board mounted on the underside of the electronic shelf, easily accessible when the rear door is open.
- **Medium Flow Control.** For the continuous culture fermentation, a nutrient feed peristaltic pump is provided. With the proper tubing, 10 mL/min maximum flow rate can be attained.
- **Exhaust System.** The exhaust gases pass into the exhaust condenser where moisture is removed and returned to the vessel. The remaining air passes into the 0.2 µm exhaust filter.
- **Sampling System.** The sampler has a rubber suction bulb to facilitate collection of representative samples without contamination. The hooded sampler is attached to a sampling tube that extends to the bottom of the vessel. A 25 mL screw cap container serves as a reservoir.
- **Flowmeter.** The design of the flowmeter is based on the variable area principle. It is a simple, precise means of indicating flow rates in fluid systems.

4.15. Calibration of the pH Probe

After the electrode was connected to the probe cable, the power was switched to the on position when the agitation switch was off. The selector switch was set to pH and the mode switch to ZERO. pH probe was immersed into an external pH 7.00 buffer solution. After waiting for a certain time interval, the display was set to read the pH value of the buffer solution by INC/DEC switch. Then the probe was immersed into a second external buffer solution which is several pH units above or below the pH selected in the previous step and the mode switch was turned to SPAN. Again by using INC/DEC switch, the display was set to the value of the second buffer solution. After autoclaving, the pH of a sterile solution at known pH was measured with the probe to check the calibration.

4.16. Calibration of the Dissolved Oxygen Probe

After the electrode was connected to probe cable, the power was switched to ON position when the agitation switch was OFF. The selector switch was set to DO and the mode switch to ZERO. DO probe cable was removed from the DO probe and after waiting for a certain time interval, the display was set to read DO value of zero by INC/DEC switch. Then the probe cable was reconnected to the DO probe. The agitation speed was set to 600 rpm and sterile air was sparged into the vessel vigorously. Mode switch was turned to SPAN. Again by using INC/DEC switch, the display was set to 100.

5. RESULTS AND DISCUSSION

5.1. Expression of *Taq* I Restriction-Modification System Using MBP Protein Fusion and Purification System

The gene encoding *Taq* I restriction endonuclease was PCR amplified by using specifically designed primers and cloned into pMAL-p2 vector downstream from the *mal*E gene of *E. coli* which encodes maltose binding protein (MBP) under the control of *tac* promoter. Bacterial secretion signal within the *mal*E gene targets the expressed MBP fusion proteins into the periplasmic space of *E. coli* cells. The recombinant plasmid pH185 containing an intact *Taq* I endonuclease gene was identified by restriction enzyme analysis and confirmed by sequencing (Özdinler, 1996).

In the first part, *E. coli* cells were transformed with pH185 plasmid and the expression and secretion of the MBP-*Taq* I endonuclease fusion protein under various fermentation conditions were investigated. For the second part, *E. coli* cells were co-transformed with pETMET plasmid in order to analyze the effect of methylase protection on extracellular secretion of MBP-*Taq* I fusion protein both in shake flasks and bioreactor conditions.

5.1.1. Effect of the Induction Period on Cellular Location of MBP-*Taq* I Fusion Protein

The recombinant *E. coli* XL1 cells harboring pH185 were grown in LB medium at 37°C in shake flasks and first induced with 1 mM IPTG at their late exponential phase of growth. Samples were taken at regular time intervals and *Taq* I endonuclease activity was measured in the cytoplasmic cell extracts, in the periplasmic fraction and in the medium as described in Methods 4.10. This analysis showed the presence of *Taq* I endonuclease activity in the cytoplasm and periplasm of the recombinant cells as well as in the medium after ten hours of induction. Figure 5.1 shows the effect of the induction period on the distribution of *Taq* I endonuclease activity. The enzyme activity detected in the cytoplasm

of the induced recombinant cells remained almost constant for 20 hours after induction. The extracellular enzyme activity as well as the total activity which is the sum of the activities recovered from the three compartments were found to reach a maximum after 18 hours of induction.

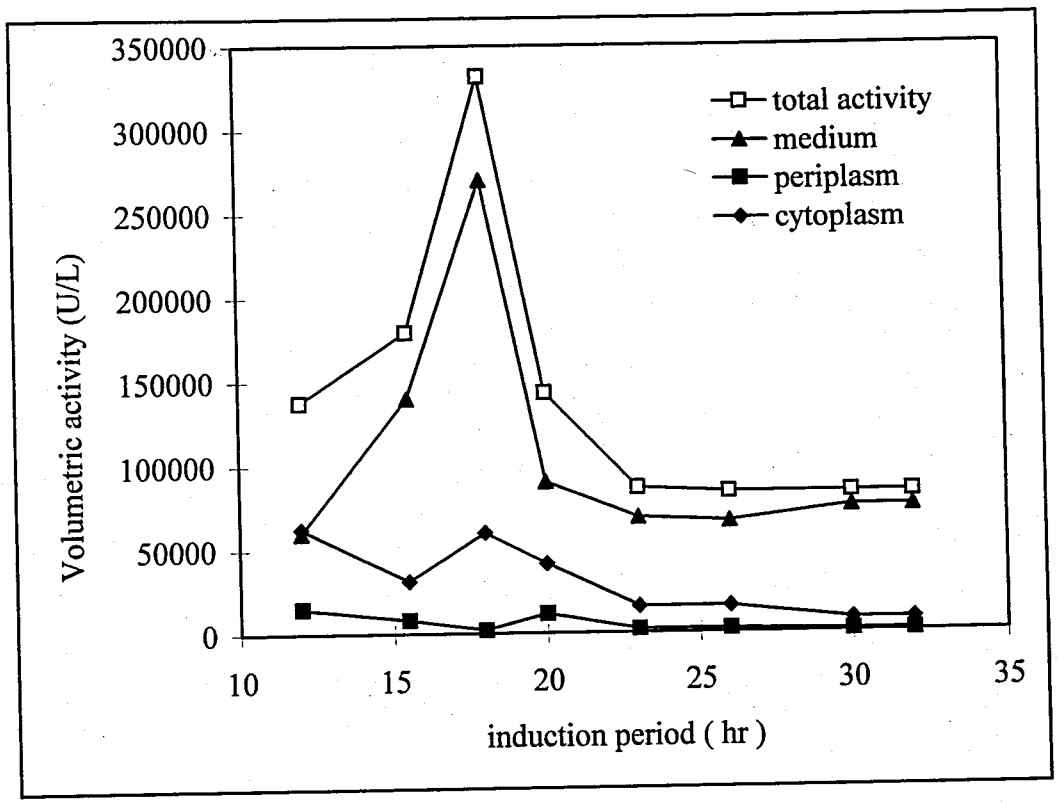


Figure 5.1. Effect of induction period on the distribution of *Taq* I endonuclease activity in *E. coli* XL1 cells containing pH185

An optimum induction period of 18 hours was found to be essential for the production and excretion of *Taq* I endonuclease activity by recombinant *E. coli* cells. Miksch *et al.* (1997) have also reported the positive effect of the prolonged cultivation time on the *kil* mediated release of a hybrid β-glucanase by *E. coli*. An induction period of 12-14 hours was found to be associated with the efficient excretion of the recombinant streptokinase directed to the periplasm of *E. coli* JM109 cells by *omp*A signal (Ko *et al.*, 1995, Lee *et al.*, 1997). Longer induction periods leading to the accumulation of the recombinant protein seem to make the cells increasingly leaky during the stationary phase and facilitate the release of proteins targeted into the periplasm.

In induced cultures, the maximum extracellular activity was reached after 18 hours followed by a sharp decrease between 18-20 hours, probably as the result of the degradation by proteases which are released into the medium by the activation of autolytic activities as suggested by Suominen *et al.* (1987). The use of *E. coli* strains with improved growth phenotypes and increased resistance to adverse conditions may possibly lead to a better productivity by delaying or preventing this activation of the autolytic activity (Weikert *et al.*, 1998).

5.1.2. Effect of Host Strain and Induction Time on MBP-*Taq* I Expression

The differences in the efficiency of various *E. coli* strains in the extracellular secretion of recombinant proteins were already reported (van der Wal *et al.*, 1995 , Duenas *et al.*, 1994). Therefore, the *E. coli* strains, ER2508 which is protease deficient and defective in the chromosomal *mal*E gene and TB1 which is the recommended host by the manufacturer for the cloning of the target gene, were also used as hosts. These cells were transformed with pH185, and the recombinant colonies were grown in shake flasks and induced at different phases of growth for 18 hours. The effect of the induction time on the production and excretion of the *Taq* I endonuclease activity is summarized in Table 5.1.

A maximum of 3.32 10^5 U/L total *Taq* I endonuclease activity was obtained when *E. coli* XL1 (pH185) cells were induced at their late-exponential growth phase, and 81 per cent of the *Taq* I activity was excreted into the growth medium. Induction at mid-exponential phase of the growth decreased the total *Taq* I activity to 1.57 10^5 U/L without affecting the percentage of the extracellular enzyme fraction. Induction at the beginning of the exponential phase resulted in a 90 per cent decrease in the total *Taq* I endonuclease production, however, more than 88 per cent of this activity was determined to be extracellular.

Table 5.1. Effect of host strain and induction time on the secretion of *Taq* I restriction endonuclease activity. Numbers in parentheses represent percentages against total activity.

	Taq I activity (U/ml) ± 5 %				
E. coli	Extracellular	Periplasmic	Intracellular	Total	(U/g DCW)
Induced at late-exponential phase					
XL1	270 (81)	2 (0.6)	60 (18)	332	164700
TB1	65 (58)	1.5 (1.4)	45 (41)	112	58500
ER2508	145 (91)	1.5 (1)	12 (8)	159	58900
Induced at mid-exponential phase					
XL1	125 (80)	1.5 (1)	30 (19)	157	94200
TB1	150 (94)	1.5 (1)	7.8 (5)	159	61800
ER2508	70 (93)	1 (1.3)	4 (5)	75	28100
Induced at pre-exponential phase					
XL1	22.5 (88)	1 (4)	2 (8)	25.5	11600
TB1	12.5 (86)	0.6 (4)	1.5 (10)	14.6	5400
ER2508	16 (89)	0.5 (3)	1.5 (8)	18	4900

The same behaviour was also observed with recombinant *E. coli* ER2508 (pH185). The highest total *Taq* I endonuclease activity obtained was 1.59 10^5 U/L when the cells were induced at the late exponential phase and approximately 92 per cent of this activity was found to be excreted. Although the fraction excreted into the medium remained the same (ca. 90 per cent), the total activity decreased twofold and ninefold when *E. coli* ER2508 (pH185) cells were induced at mid- and early exponential phases of the growth, respectively. The same highest amount of total activity (1.59 10^5 U/L) and excretion was recorded with recombinant *E. coli* TB1 (pH185) cells when the cells were induced at mid-exponential phase of growth. Shifting the induction time from mid-exponential to late exponential phase of the growth resulted in an approximately 45 per cent decrease in the extracellular *Taq* I activity in the latter case, and 40 per cent of the total *Taq* I endonuclease activity remained in the cytoplasm. In all cases, the total enzyme production expressed as units per gram dry cells was highest with the strain *E. coli* XL1 (pH185).

The production under the control of the *tac* promoter and excretion of the *Taq* I endonuclease activity directed to the periplasm by MBP signal of *E. coli* was found to be strain dependent also. Expression and cellular location of MBP-*Taq* I fusion protein in various strains was apparently affected by the phase of growth at which the inducer was added to the medium. Total and extracellular activities recovered from cultures induced at early phases of growth were found to be considerably low when compared with those induced at later phases of growth.

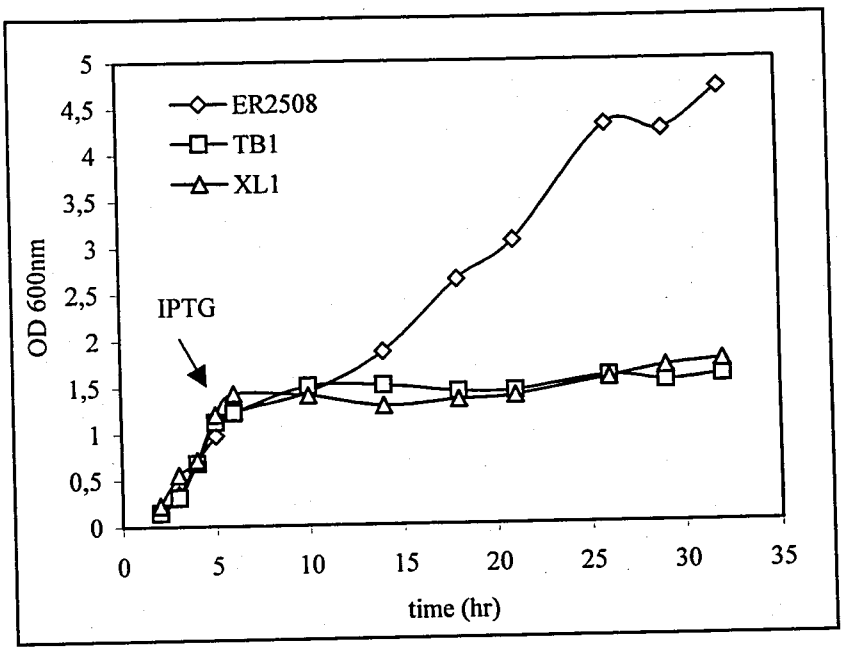


Figure 5.2. Growth of *E. coli* ER2508, TB1 and XL1 cells harboring pH185.

The growth behaviour of these three strains induced at their late-exponential phase with 1mM IPTG was also examined. As shown in Figure 5.2, the induction using 1mM IPTG forced the cells to enter stationary phase, the absorbance at 600nm remained constant, and no drop in culture turbidity was observed even after 26 hours of induction for *E. coli* TB1 and XL1 cells harboring the plasmid pH185. *E. coli* ER2508 (pH185) cells continued to grow after five hours of induction reaching absorbance values of 4.5-5 at 600nm.

The effect of protein production on cell viability of these three strains was also determined by counting colonies on selective LB agar plates as described in Methods

4.12.3. As shown in Figure 5.3, the number of viable cells decreased soon after the induction and remained relatively constant up to 20 hours after induction in the case of both *E. coli* ER2508 (pH185) and *E. coli* TB1 (pH185). The decrease in the number of viable cells in the case of *E. coli* XL1 (pH185), as a response to induction with 1mM IPTG, was lower and more gradual compared to the two other strains.

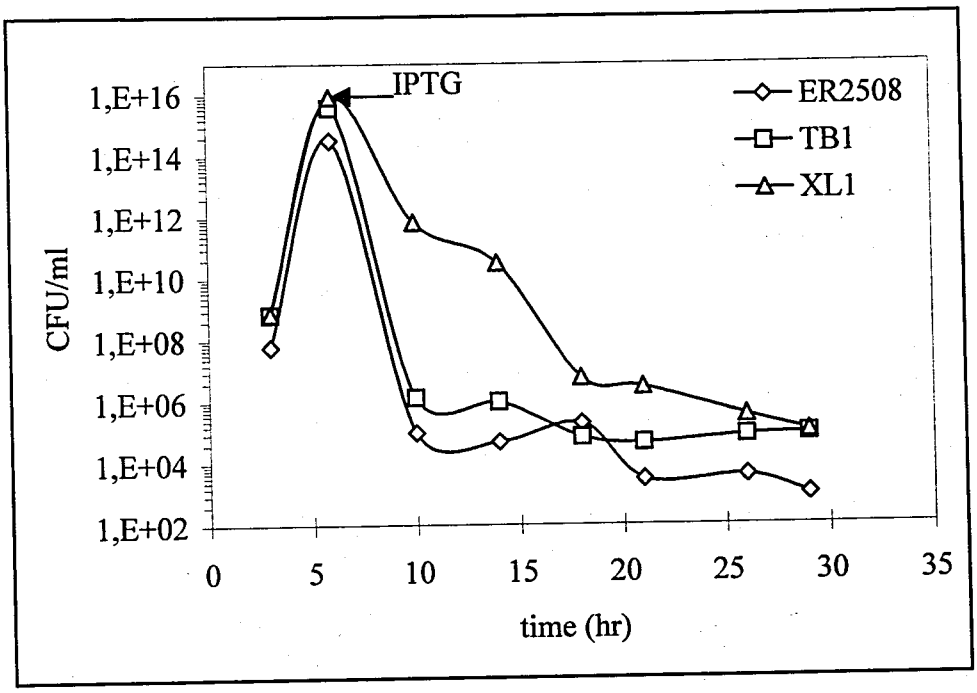


Figure 5.3. Cell viability of *E. coli* ER2508, TB1 and XL1 cells harboring pH185. Cells were induced at their late-exponential phase with 1mM IPTG.

5.1.3. Control of Cell Lysis

Isocitrate dehydrogenase was used as a cytoplasmic marker protein to ensure that the extracellular secretion did not come from cell lysis. Cytoplasmic and medium fractions were assayed in all cases, and almost all isocitrate dehydrogenase activity was found to be intracellular whereas negligible amounts could be detected in the medium. Bacterial lysis was also followed by monitoring the release of total cellular nucleic acids as described by Kloos *et al.* (1994) and no increase in the amount of DNA within the medium could be observed. These results indicated that cell lysis is not the major cause of the presence of the extracellular *Taq* I endonuclease activity. Furthermore, the absence of the extracellular enzyme activity after 18 hours of induction of the *E. coli* cells harbouring a recombinant

plasmid (pMAL-c2) which encodes cytoplasmic MBP-*Taq* I endonuclease fusion protein with *Taq* I endonuclease activity (Özkırımlı, 1998) provided additional evidence for the absence of a noteworthy lysis.

5.1.4. Purification of MBP-*Taq* I Fusion Protein by Amylose Affinity Chromatography

The MBP-*Taq* I fusion protein is a bifunctional protein displaying both *Taq* I endonuclease activity and the binding site for maltose. Therefore, it could easily be purified from cytoplasmic, periplasmic and extracellular fractions by a single-step affinity chromatography based on specific interactions of its MBP part with amylose, as described in Methods 4.10.1.

E. coli XL1(pH185) cells were grown in LB medium at 37°C and induced with 1mM IPTG at their late exponential phase of growth for 18 hours. According to the manufacturer, the binding capacity of the amylose column was three mg fusion protein per ml bed volume. Since expression levels under the control of the *tac* promoter do not exceed 30 per cent of the total cell protein, ten ml bed volume was chosen to be more than sufficient to reach high purification yields in this study. The amylose resin was poured into the column and washed with column wash buffer to baseline. The crude cell extract, osmotic shock fluid or medium proteins were prepared, and the protein concentrations were adjusted to remain below 2.5 mg/ml by diluting with column wash buffer in order to increase the efficiency of affinity binding. Samples were loaded onto the column and the flow-through fractions containing the unbound proteins were collected. After washing the column with column wash buffer, the MBP-fusions were eluted with wash buffer containing 10 mM maltose. Protein containing fractions were collected together to form the Flow-through and Eluate samples and their *Taq* I endonuclease activities were determined as described in 4.9. Results of the purification is summarized in Table 5.2.

Table 5.2. Purification of MBP-*Taq* I from *E. coli* XL1(pH185) cells by amylose affinity chromatography.

Sample	Total Volume (ml)	Protein conc. (mg/ml)	Total Protein (mg)	Activity (U/µl)	Total Activity (U/L culture)	Specific Activity (U/mg protein)	Recovery %	Purification fold
Medium	20	3.48	69.6	7.5	150000	2155	100	-
Flow-through	44	1.50	66.0	2.5	110000	1667	73	-
Eluate	3.2	0.18	0.58	2.5	8000	13793	5	6.4
Periplasm	10	0.61	6.1	0.62	6200	1030	100	-
Flow-through	20	0.27	5.4	0.16	3200	569	52	-
Eluate	0.8	0.29	0.23	1.25	1000	4340	16	4.21
Cytoplasm	40	1.98	79.2	0.62	24800	313	100	-
Flow-through	62	0.83	51.5	0.16	9920	193	40	-
Eluate	2.1	0.93	1.95	10	21000	10770	85	34.4

Table 5.3. Purification of MBP-*Taq* I from *E. coli* ER2508(pH185) cells by amylose affinity chromatography.

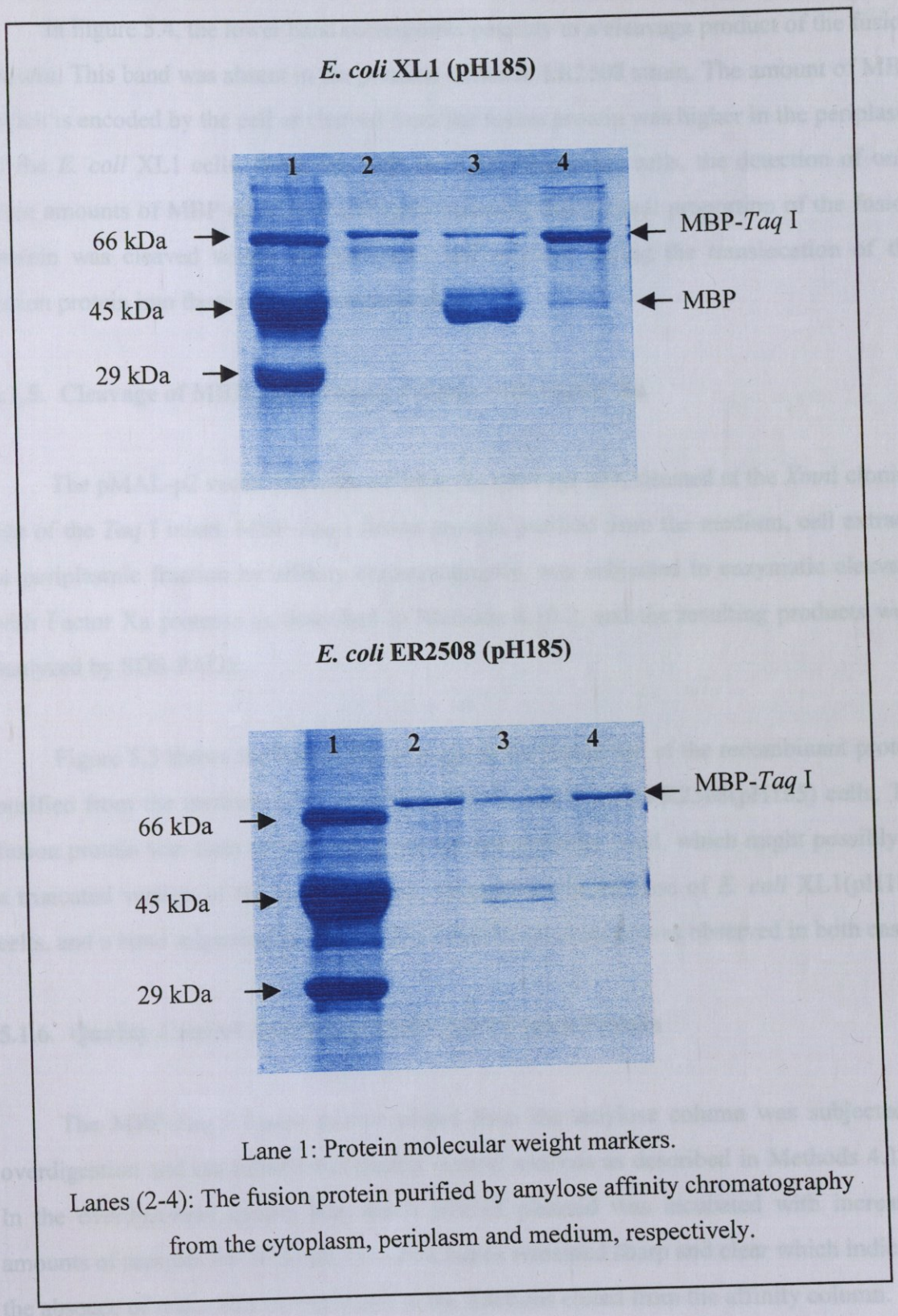
Sample	Total Volume (ml)	Protein conc. (mg/ml)	Total Protein (mg)	Activity (U/µl)	Total Activity (U/L culture)	Specific Activity (U/mg protein)	Recovery %	Purification fold
Medium	16	7.30	116.8	10	160000	1370	100	-
Flow-through	30	4.09	122.7	5	150000	1223	94	-
Eluate	6.6	0.02	0.13	0.16	1056	8123	1	5.9
Periplasm	8	1.72	13.8	0.18	1440	104	100	-
Flow-through	9	1.21	10.9	0.156	1404	129	98	-
Eluate	0.6	0.01	0.006	0.03	18	3000	1.25	28.8
Cytoplasm	36	8.95	322.2	0.63	22680	70	100	-
Flow-through	48	7.61	365.3	0.31	14880	41	66	-
Eluate	1.5	0.12	0.18	1.25	1875	10417	8.3	149

A total of 181000 units of *Taq* I endonuclease activity was recovered from the three cellular compartments, 83 per cent of which was excreted to the growth medium. While 85 per cent of the cytoplasmic enzyme activity was bound to the amylose column and recovered using maltose containing wash buffer, the active fraction bound to the column was 16 per cent and five per cent for periplasmic extracts and medium, respectively.

E. coli ER2508 strain has a partial deletion of the *mal*E gene, and hence, it cannot produce MBP from its bacterial genome. When this strain is used as the host organism, all the MBP is expressed by the recombinant plasmid as fusion tail. It also has a mutation in its *lon* gene which codes for the La protein, the primary protease degrading abnormally folded proteins in *E. coli*. The production and purification of MBP-*Taq* I fusion protein on cross-linked amylose was also examined in recombinant *E. coli* ER2508 cells harboring pH185 plasmid (Table 5.3).

A total of 184120 Units *Taq* I restriction endonuclease activity was measured and 87 per cent of this activity was found to be excreted to the extracellular medium. When the medium, osmotic shock fluid and cytoplasmic extracts prepared from these cells were loaded onto the amylose column, more than 90 per cent of the medium and periplasmic activities passed through and only one per cent could be recovered in eluate fractions. The specific activity of the cytoplasmic eluate was comparable with the XL1 strain but the recovery yield was 10 times lower with the ER2508 strain. Also, the protein concentration of the eluate fractions were 10 times lower when compared with the XL1 strain. These low recovery yields and protein concentrations indicated that ten ml bed volume was probably below the critical limits. Using an amylose column with a larger bed volume could enhance the purification yields by increasing the protein concentration of the eluate fractions and in turn improving the conformational stability and activity of the MBP fusions.

SDS-PAGE analysis of the cell extract, periplasmic fraction and the medium proteins eluted using column wash buffer containing 10 mM maltose showed the presence of the MBP-*Taq* I fusion protein with a molecular weight of approximately 74 300 Da, as expected from the fusion of the MBP (42 700 Da) and the monomeric *Taq* I endonuclease (31 600 Da) at the anticipated position (Figure 5.4).



Lane 1: Protein molecular weight markers.

Lanes (2-4): The fusion protein purified by amylose affinity chromatography from the cytoplasm, periplasm and medium, respectively.

Figure 5.4. SDS-PAGE analysis of MBP-*Taq* I fusion protein purified from *E. coli* XL1 (pH185) and *E. coli* ER2508 (pH185) cells.

In Figure 5.4, the lower band corresponds possibly to a cleavage product of the fusion protein. This band was absent in the protease deficient ER2508 strain. The amount of MBP which is encoded by the cell or cleaved from the fusion protein was higher in the periplasm of the *E. coli* XL1 cells. However, with the ER2508(pH185) cells, the detection of only trace amounts of MBP on SDS-PAGE gels indicated that a small proportion of the fusion protein was cleaved within the periplasm and medium during the translocation of the fusion protein into these cellular compartments.

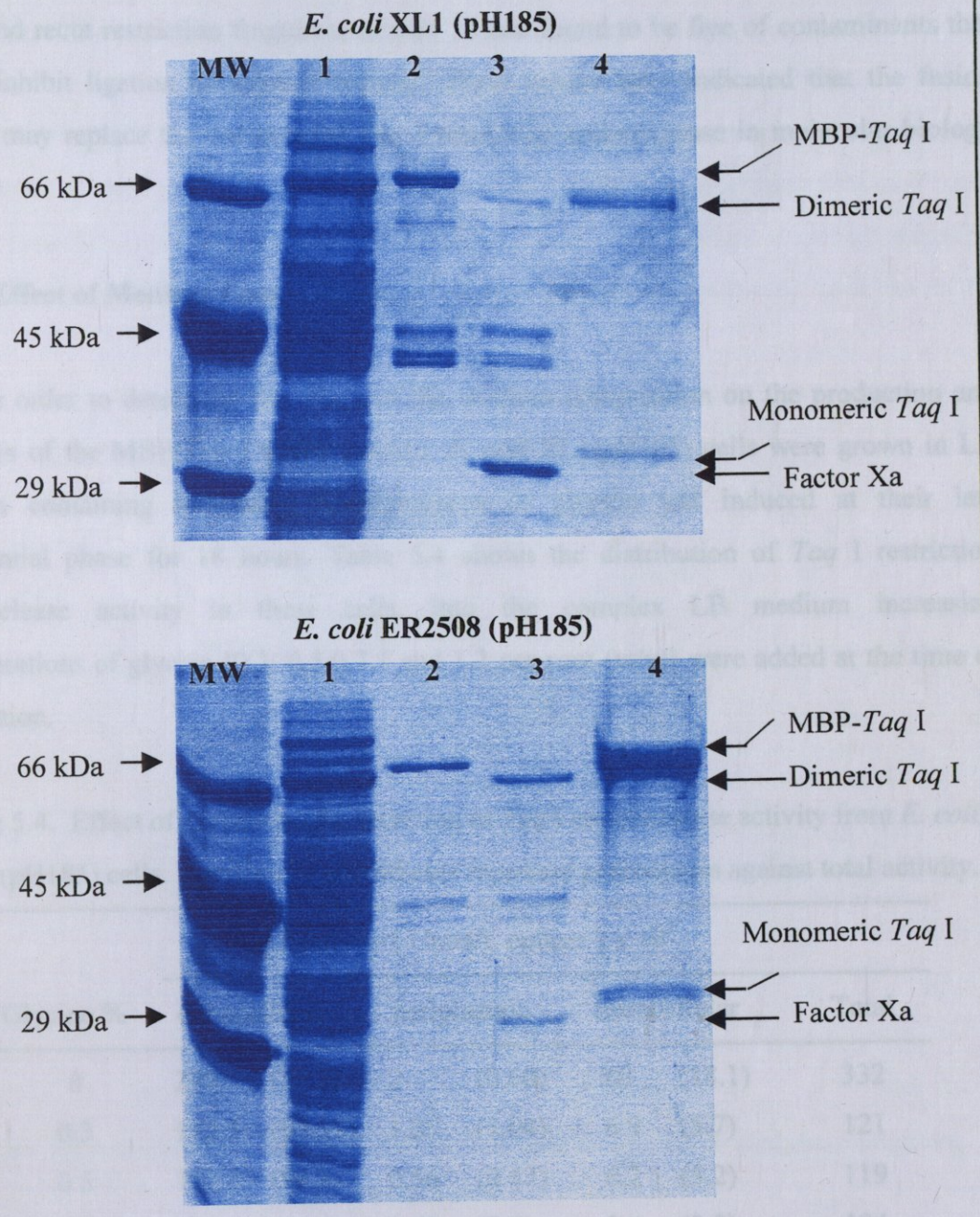
5.1.5. Cleavage of MBP-*Taq* I Fusion Protein with Factor Xa

The pMAL-p2 vector provides a Factor Xa cleavage site, situated at the *Xmn*I cloning site of the *Taq* I insert. MBP-*Taq* I fusion protein, purified from the medium, cell extracts or periplasmic fraction by affinity chromatography, was subjected to enzymatic cleavage with Factor Xa protease as described in Methods 4.10.2, and the resulting products were analyzed by SDS-PAGE.

Figure 5.5 shows the Factor Xa cleavage of the fusion tail of the recombinant protein purified from the medium of *E. coli* XL1(pH185) and *E. coli* ER2508(pH185) cells. The fusion protein was cleaved to completion, the lower fainter band, which might possibly be a truncated version of the fusion protein, disappeared in the case of *E. coli* XL1(pH185) cells, and a band migrating as dimeric *Taq* I restriction enzyme was observed in both cases.

5.1.6. Quality Control Analysis of MBP-*Taq* I Fusion Protein

The MBP-*Taq* I fusion protein eluted from the amylose column was subjected to overdigestion and cut-ligate-recut quality control analysis as described in Methods 4.12.8. In the overdigestion quality test where pUC18 plasmid was incubated with increasing amounts of samples for 16 hours, the DNA bands remained sharp and clear which indicated the absence of endo- and exonucleases in the fractions eluted from the affinity column.



MW: Protein molecular weight markers.

Lane 1: Medium proteins applied to the amylose column.

Lane 2: Medium proteins eluted from the amylose column.

Lane 3: Factor Xa cleavage of the column eluate.

Lane 4: Commercial *Taq* I (New England Biolabs).

Figure 5.5. Factor Xa cleavage of MBP-*Taq* I fusion protein.

The same samples containing the fusion protein were also tested for their ability to ligate and recut restriction fragments of pUC18 and found to be free of contaminants that would inhibit ligation or degrade termini. These results have indicated that the fusion protein may replace the commercial *Taq* I restriction endonuclease in molecular biology studies.

5.1.7. Effect of Medium Composition

In order to determine the effect of the medium composition on the production and secretion of the MBP-*Taq* I fusion protein, *E. coli* XL1(pH185) cells were grown in LB medium containing increasing concentrations of glycine and induced at their late exponential phase for 18 hours. Table 5.4 shows the distribution of *Taq* I restriction endonuclease activity in these cells. Into the complex LB medium increasing concentrations of glycine [0.3, 0.5.0.7.1 and 1.2 per cent (w/v)] were added at the time of inoculation.

Table 5.4. Effect of glycine on the secretion of *Taq* I endonuclease activity from *E. coli* XL1 (pH185) cells. Numbers in parentheses represent percentages against total activity.

	Taq I activity (Units/L culture) x 10^3						
Glycine %	extracellular		periplasmic		intracellular		Total
0	270	(81.3)	2	(0.60)	60	(18.1)	332
0.3	112.5	(93.27)	1.25	(1.04)	6.9	(5.7)	121
0.5	112.5	(94.3)	0.56	(0.47)	6.2	(5.2)	119
0.7	100	(96.1)	0.62	(0.6)	3.4	(3.3)	104
1	28	(86.5)	0.62	(1.92)	3.7	(11.54)	32
1.2	16	(79.7)	0.34	(1.71)	3.7	(18.6)	20

It has been observed that the total enzyme production decreases with the increasing concentrations of glycine. However, the percentage of the extracellular enzyme activity was found to increase from 81.3 per cent to 96.1 per cent in the presence of 0.7 per cent (w/v) glycine added to the medium at the time of inoculation. Higher concentrations of glycine in

the medium resulted in a significant decrease in both the total production and the percentage of the extracellular activity of *Taq* I endonuclease.

The growth behaviour of the recombinant cells in the presence of glycine was also investigated by measuring the culture turbidity at 600nm using a spectrophotometer (Figure 5.6). The resulting growth profiles showed that there was a considerable decrease in culture turbidity in the presence of one and 1.2 per cent of glycine soon after the addition of the inducer.

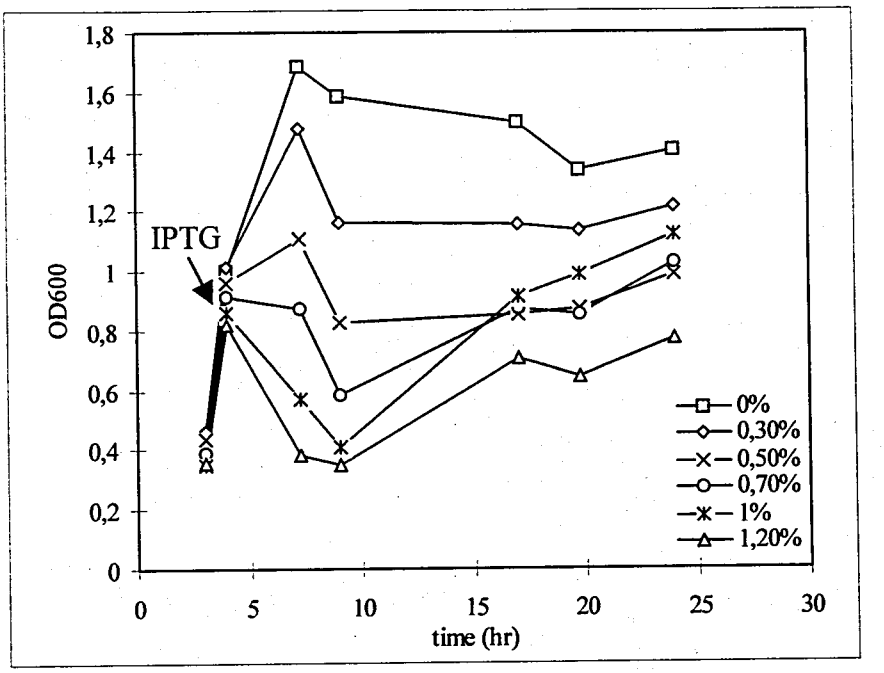


Figure 5.6. Growth of *E. coli* XL1 (pH185) cells in complex LB medium containing increasing concentrations of glycine.

Glycine has been shown to improve the rate and efficiency with which the bacterial system handles the recombinant protein and promote the release of the various proteins from the periplasmic space into the culture medium (Aristidou *et al.*, 1993; Yu and San, 1993). However, in this study, the presence of glycine did not increase the enzyme production as reported. Moreover, increasing concentrations of glycine had a detrimental effect on both cell growth and enzyme production.

The effect of the presence of the additional yeast extract, maltose, sucrose and sorbitol in the medium on the extracellular transportation of the fusion protein targeted to the periplasm was also investigated. Table 5.5 summarizes the effect of these compounds on the distribution of *Taq* I endonuclease activity.

Table 5.5. Effect of medium composition on the distribution of *Taq* I endonuclease activity in *E. coli* XL1 (pH185) cells.

	Taq I activity (Units/L culture)x10^3						
Culture Medium	extracellular		periplasmic		intracellular		Total
LB	270	(81.3)	2	(0.60)	60	(18.1)	332
LB + 2.5 g/L Yeast Extract	156.3	(98.4)	0.78	(0.49)	1.8	(1.11)	159
LB + 5 g/L Sucrose	70.3	(93.3)	0.78	(1.04)	4.3	(5.69)	75.4
LB + 5 g/L Sorbitol	19.5	(92.4)	0.7	(3.33)	0.9	(4.26)	21.1
LB + 5 g/L Maltose	31.3	(91.97)	0.39	(1.15)	2.3	(6.89)	34

The presence of the additional yeast extract in the medium resulted in an almost twofold decrease in the total enzyme activity. The presence of this compound did, however, enhance the extracellular transportation of the enzyme, and almost all enzyme activity (98.4 per cent) was found to be secreted to the extracellular medium. The presence of sucrose, maltose and sorbitol in the culture medium increased the percentage of the extracellular enzyme activity to a degree higher than 90 per cent leading to an approximately four, 10 and 16 fold decrease in the total enzyme production respectively.

Investigation of the growth behaviour of the recombinant *E. coli* XL1 (pH185) cells showed that the addition of maltose and sucrose to the medium did not alter the growth kinetics of the cells. However the cultures grown in media containing either yeast extract or sorbitol reached to higher densities than others (Figure 5.7).

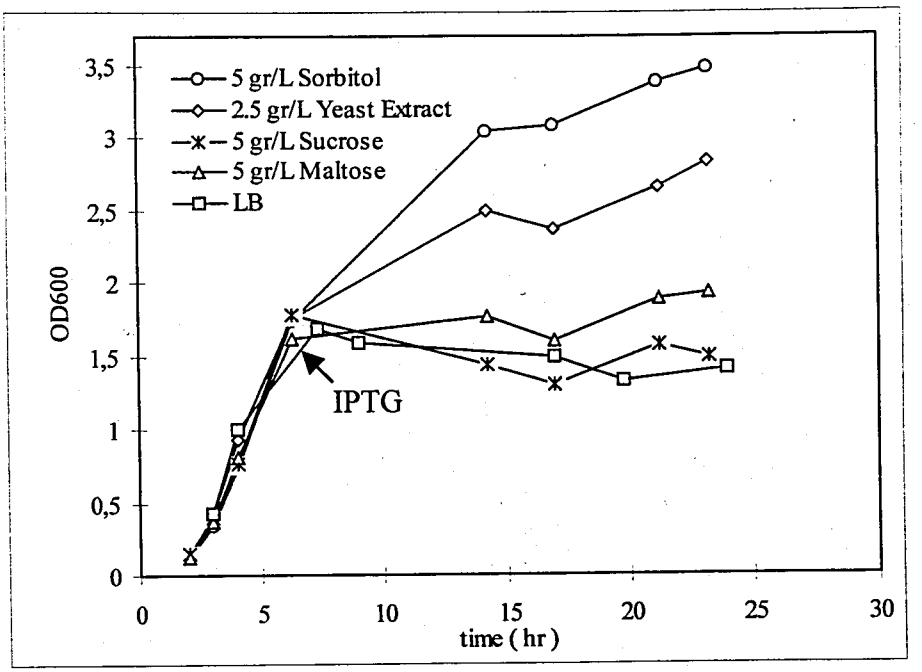


Figure 5.7. Growth profile of *E. coli* XL1 (pH185) cells in complex LB medium containing yeast extract, sucrose, sorbitol and maltose.

The presence of additional yeast extract, maltose, sucrose and sorbitol in the medium was also reported to enhance the extracellular transportation of proteins which are normally targeted to the periplasm (Harrison *et al.*, 1996; Zhang *et al.*, 1983; Scherrer *et al.*, 1994). However, in the present work, dramatic decreases in the level of the total enzyme production were observed in all cases, and the considerable improvement of the cell growth by the addition of sorbitol was not reflected in *Taq* I enzyme production, and very low levels of endonuclease activity could be recovered from the medium.

In all experiments, the percentage of colonies containing the recombinant pH185 plasmid was found to decrease to less than 10 after two or three hours of induction with IPTG. Although the optimum working temperature of *Taq* I endonuclease is between 65-70°C, it still retains 10 per cent of its activity at 37°C. Upon induction with IPTG, the MBP-*Taq* I fusion protein expressed by the pH185 plasmid was transported to the periplasmic space. Despite this, a certain amount of the fusion protein was found to be accumulated in the cytoplasm, most probably due to the saturation of the *sec* pathway. Moreover, the presence of enzyme activity in the affinity column eluates showed that the *Taq* I endonuclease enzyme retained its activity when expressed as a chimeric protein with

MBP as fusion tail. Therefore, cytoplasmic accumulation of this enzyme was found to be toxic to the host cell resulting in rapid drops in plasmid stability and cell viability.

5.1.8. Cloning of the *Taq* I Methylase Gene into pBR322 Vector

In order to protect the expression plasmid pH185 and host genomic DNA from the detrimental effects of *Taq* I endonuclease expression, the *Taq* I methylase gene was amplified from the *Thermus aquaticus* YT1 genome by PCR using specific primers and cloned under the promoter of the tetracycline resistance gene of the pBR322 vector.

The plasmid pBR322 is one of the most commonly used *E. coli* cloning vectors. In order to clone and express the *Taq* I restriction-modification system in *E. coli*, Slatko *et al.* (1987) constructed a *Thermus aquaticus* genomic DNA library using pBR322 vector. They found out that in colonies displaying *Taq* I methylase activity, transcription of the methylase gene derived from the tetracycline promoter of the vector. Using this information, pBR322 vector was used for the constitutive expression of the *Taq* I methylase gene under the control of the tetracycline resistance gene promoter. The cloning strategy, which was applied where the PCR amplified methylase gene was ligated into the unique *Bam*H I site within the tetracycline resistance gene of the pBR322 vector, is presented in Figure 5.8.

Using the 1266 base pair *Taq* I methylase gene sequence available in the literature (Barany *et al.*, 1992), the forward primer METF and reverse primer METR were designed such that they created the recognition sequence for the BamH I restriction endonuclease at both ends of the methylase gene. Genomic DNA from *Thermus aquaticus* YT1 strain was used as a template for the PCR amplification of the methylase gene using primers METF and METR as described in Methods 4.5.1.

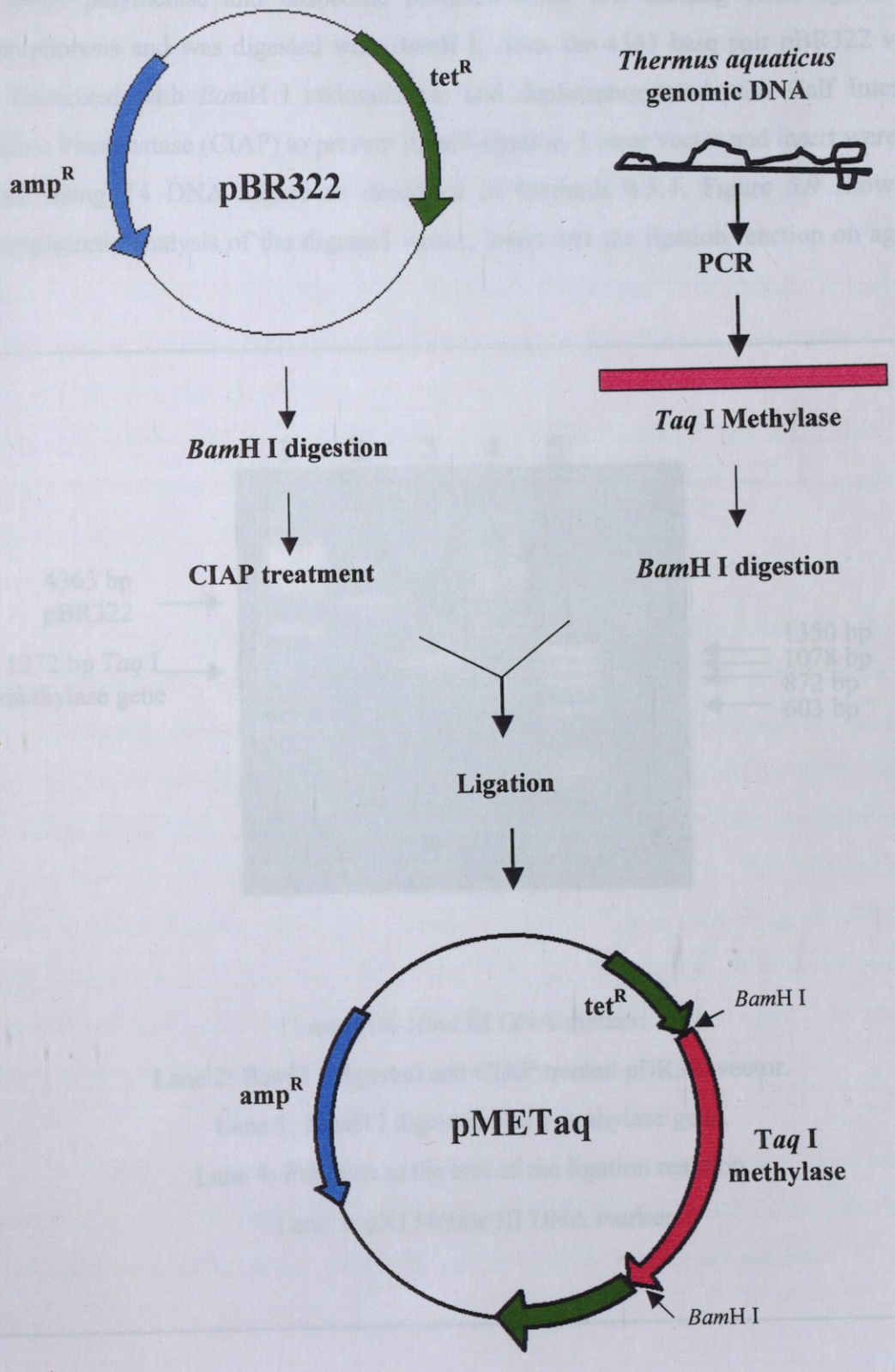
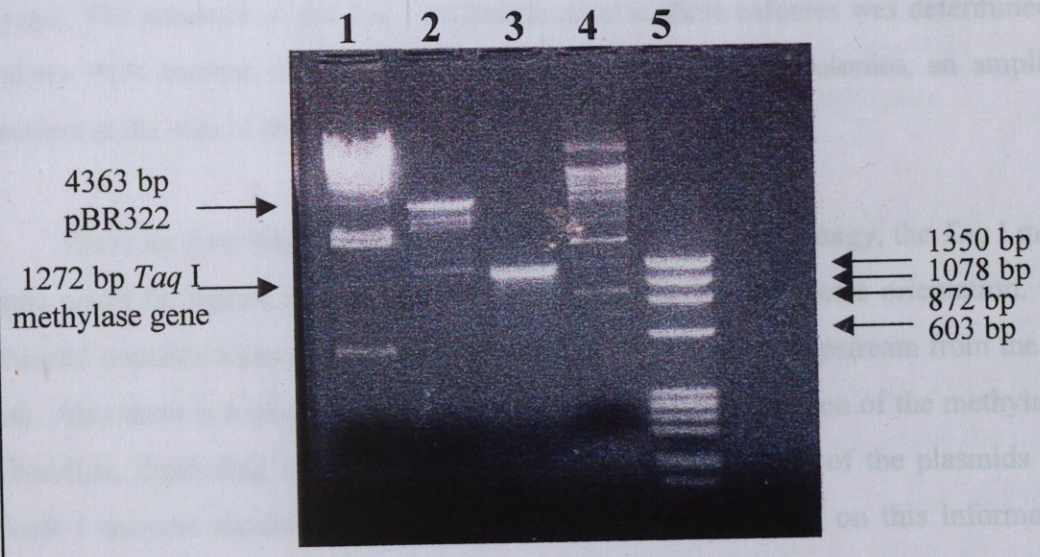


Figure 5.8. Cloning strategy of the *Taq*I methylase gene into pBR322 vector

The PCR product of 1296 base pairs long was purified from excess primers, dNTPs, *Pfu* DNA polymerase and unspecific products using low melting point agarose gel electrophoresis and was digested with *Bam*H I. Also, the 4363 base pair pBR322 vector was linearized with *Bam*H I endonuclease and dephosphorylated with Calf Intestinal Alkaline Phosphatase (CIAP) to prevent its self-ligation. Linear vector and insert were then ligated using T4 DNA Ligase as described in Methods 4.5.4. Figure 5.9 shows the electrophoretic analysis of the digested vector, insert and the ligation reaction on agarose gels.



Lane 1: λ-*Hind* III DNA marker.

Lane 2: *Bam*H I digested and CIAP treated pBR322vector.

Lane 3: *Bam*H I digested *Taq* I methylase gene.

Lane 4: Products at the end of the ligation reaction.

Lane 5: ϕX174/*Hae* III DNA marker

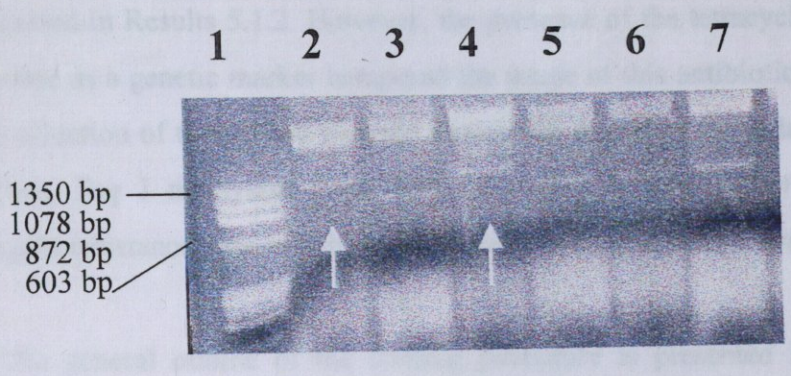
Figure 5.9. Electrophoretic analysis of *Bam*H I digested DNA fragments and the ligation reaction.

The observation of the presence of the bands larger than the pBR322 vector in the ligation mixture indicated that the ligation reaction was sucessfully completed. In lane 2, a band appeared at the size of the insert, which could be due to overloading the neighbouring well of the agarose gel.

The ligated product was introduced into competent *E. coli* ER2508 cells by electroporation, and after overnight incubation at 37°C, plasmid containing cells formed colonies on LB plates containing ampicillin. Screening of the recombinant colonies was carried out by the cracking method and PCR analysis. 88 colonies were analyzed by the cracking method and 28 colonies were identified as harboring plasmids larger than the vector. The presence of the *Taq* I methylase gene in these colonies was determined by the colony PCR method as described in Methods 4.5.7.2. In 18 colonies, an amplification product at the size of the *Taq* I methylase gene was observed.

Since no directional cloning was applied in this cloning strategy, the *Taq* I methylase gene could be ligated into the pBR322 vector in correct or reverse orientation. pBR322 plasmid contains a unique *Eco*R I site located at 377 base pairs upstream from the *Bam*H I site. Also there is a unique *Eco*R I site at the 819 base pair position of the methylase gene. Therefore, depending on the orientation of the insert, digestion of the plasmids with the *Eco*R I enzyme should generate different band patterns. Based on this information, we have estimated that the expected recombinants of 5635 base pairs long should produce DNA fragments of 1202 and 4433 base pairs for the correct orientation and DNA fragments of 828 and 4807 base pairs for the reverse orientation. Figure 5.10 shows the results of a typical restriction analysis experiment.

In Figure 5.10, recombinant plasmids in lanes 2 and 4 shown with the white arrow contained the *Taq*I methylase gene in the proper orientation, whereas in the other colonies DNA band sizes indicated that the insert DNA was in the reverse orientation.



Lane 1: ϕX174/*Hae* III DNA marker.

Lanes 2-7: *Eco*R I digested plasmids of 6 out of 18 selected recombinants.

Figure 5.10. Restriciton enzyme analysis of the recombinants.

The selected 18 colonies were also tested for their abiliy to express *Taq* I methylase activity in *E. coli*. Their isolated plasmids were digested with commercial *Taq* I restriction endonuclease at 65°C for one hour and analyzed by electrophoresis on agarose gels. Recombinant plasmids with the correct orientation of the insert showed *Taq* I methylase activity by resisting endonuclease digestion completely. In plasmids where the methylase gene was at the reverse orientation, partial and total digestion of the plasmid were observed. Hence the restriction analysis and methylase activity tests were in good accordance with each other. One of the properly oriented plasmids was selected for expression and cloning studies, and named as pMETaq.

5.1.9. Cloning of the *Taq* I Methylase Gene into pET28a+ Vector

Co-expression of the *Taq* I methylase gene in the pMAL protein fusion system required a selection marker other than ampicillin, since this antibiotic was used for the selection of pH185 plasmid in *E. coli* cells expressing MBP-*Taq* I fusion protein.

E. coli XL1 strain was chosen as the optimum host organism for the pMAL system, as explained in Results 5.1.2. However, the presence of the tetracycline resistance gene in its genome as a genetic marker hampered the usage of this antibiotic as a selective marker for the selection of the second plasmid containing the *Taq* I methylase gene in this strain. Therefore, *Taq* I methylase gene was sub-cloned into pET28a+ vector containing a kanamycin resistance gene which may be used as a second selection marker.

The general outline of the cloning procedure is presented in Figure 5.11. *Taq* I methylase gene together with the tetracycline promoter region was cut by digesting the pMETaq plasmid with *Hin*d III and *Sph* I restriction enzymes. In the pMETaq plasmid, *Hin*d III enzyme site is located at 57 base pairs upstream from the start of the tetracycline gene, and *Sph* I site is in the middle of this gene, 187 base pairs downstream the *Bam*H I site. On the other hand, the recognition site for *Hin*d III enzyme is located within the multiple cloning region of the pET28a+ vector and that for the *Sph* I enzyme is about 230 base pair upstream from the T7 promoter and terminator region of the same vector; hence, cutting this vector with these two enzymes would remove the T7 promoter region completely.

pMETaq and pET28a+ plasmids were double digested with *Sph* I and *Hin*d III enzymes as described in Methods 4.7 at 37°C overnight. In the pMETaq digestion, the 1805 base pair band consisting of the metylase gene and tetracycline promoter was cut from low-melting temperature agarose gel and purified. The 425 base pair band containing the T7 promoter region was removed from the rest of the pET28a+ vector by separating the bands on low-melting temperature agarose gel electrophoresis, cutting the 4944 base pair fragment and purifying. The purified vector and insert were ligated with T4 DNA Ligase enzyme by incubating at 16 °C overnight as described in Methods 4.7. The ligation mixture was used to transform the competent *E. coli* XL1 cells by electroporation.

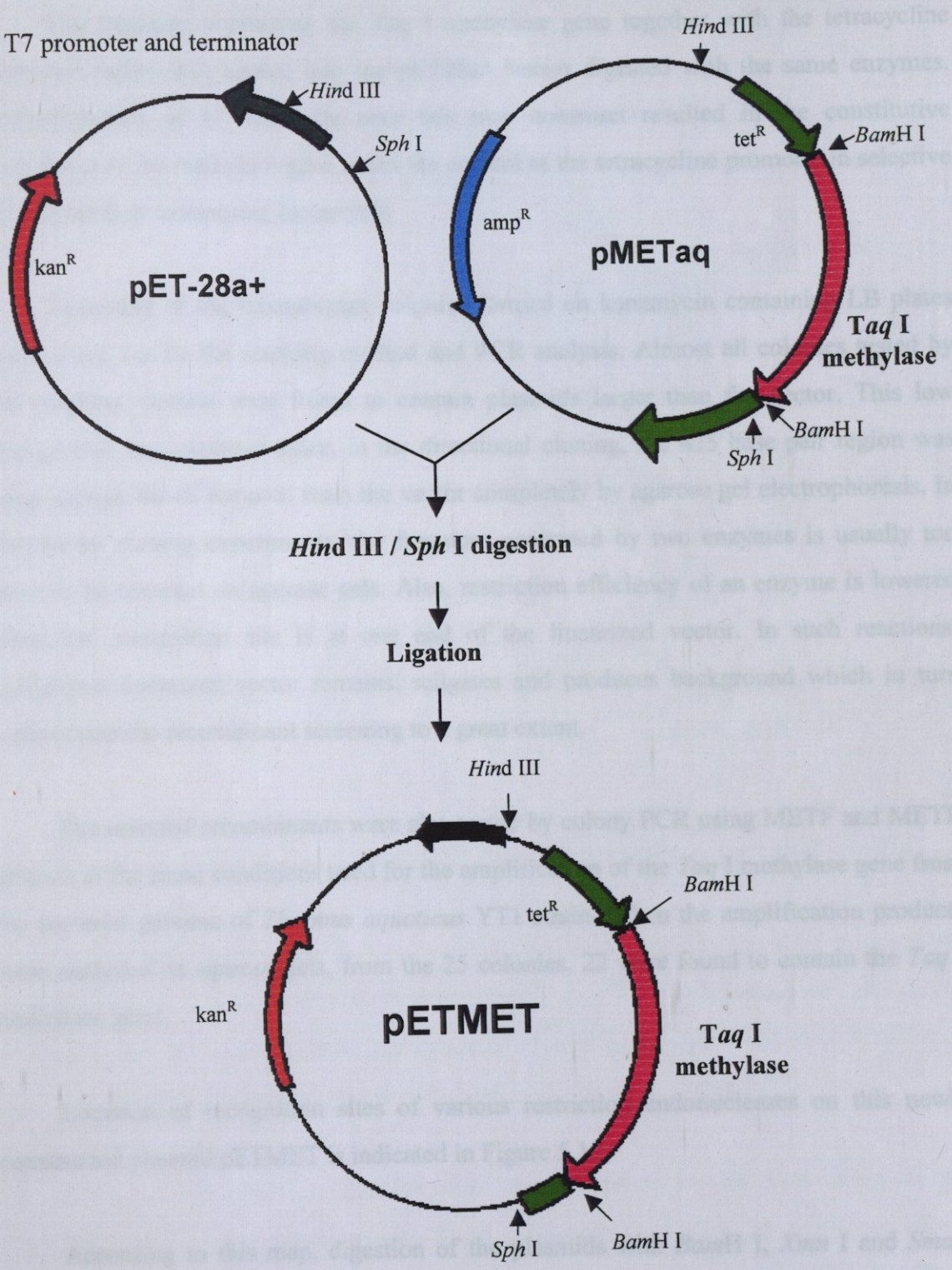


Figure 5.11. Strategy used in cloning of the *Taq* I methylase gene into pET28a+ vector

The fragment containing the *Taq* I methylase gene together with the tetracycline promoter region was ligated into the pET28a+ vector digested with the same enzymes. Transformation of *E. coli* cells with this new construct resulted in the constitutive expression of the methylase gene under the control of the tetracycline promoter in selective growth medium containing kanamycin.

Screening of the recombinant colonies formed on kanamycin containing LB plates was carried out by the cracking method and PCR analysis. Almost all colonies tested by the cracking method were found to contain plasmids larger than the vector. This low background was expected since, in the directional cloning, the 425 base pair region was large enough for its removal from the vector completely by agarose gel electrophoresis. In directional cloning experiments, the fragment generated by two enzymes is usually too short to be detected on agarose gels. Also, restriction efficiency of an enzyme is lowered when the recognition site is at one end of the linearized vector. In such reactions, undigested linearized vector remains, religates and produces background which in turn complicates the recombinant screening to a great extent.

The selected recombinants were also tested by colony PCR using METF and METR primers at the same conditions used for the amplification of the *Taq* I methylase gene from the bacterial genome of *Thermus aquaticus* YT1 strain. When the amplification products were analyzed on agarose gels, from the 25 colonies, 22 were found to contain the *Taq* I methylase gene.

Location of recognition sites of various restriction endonucleases on this newly constructed plasmid pETMET is indicated in Figure 5.12.

According to this map, digestion of the plasmids with *Bam*H I, *Xmn* I and *Sma* I enzymes would generate DNA fragments of sizes 1272 and 5477, 2439, 2277 and 2033, and 4613 and 2136 base pairs, respectively.

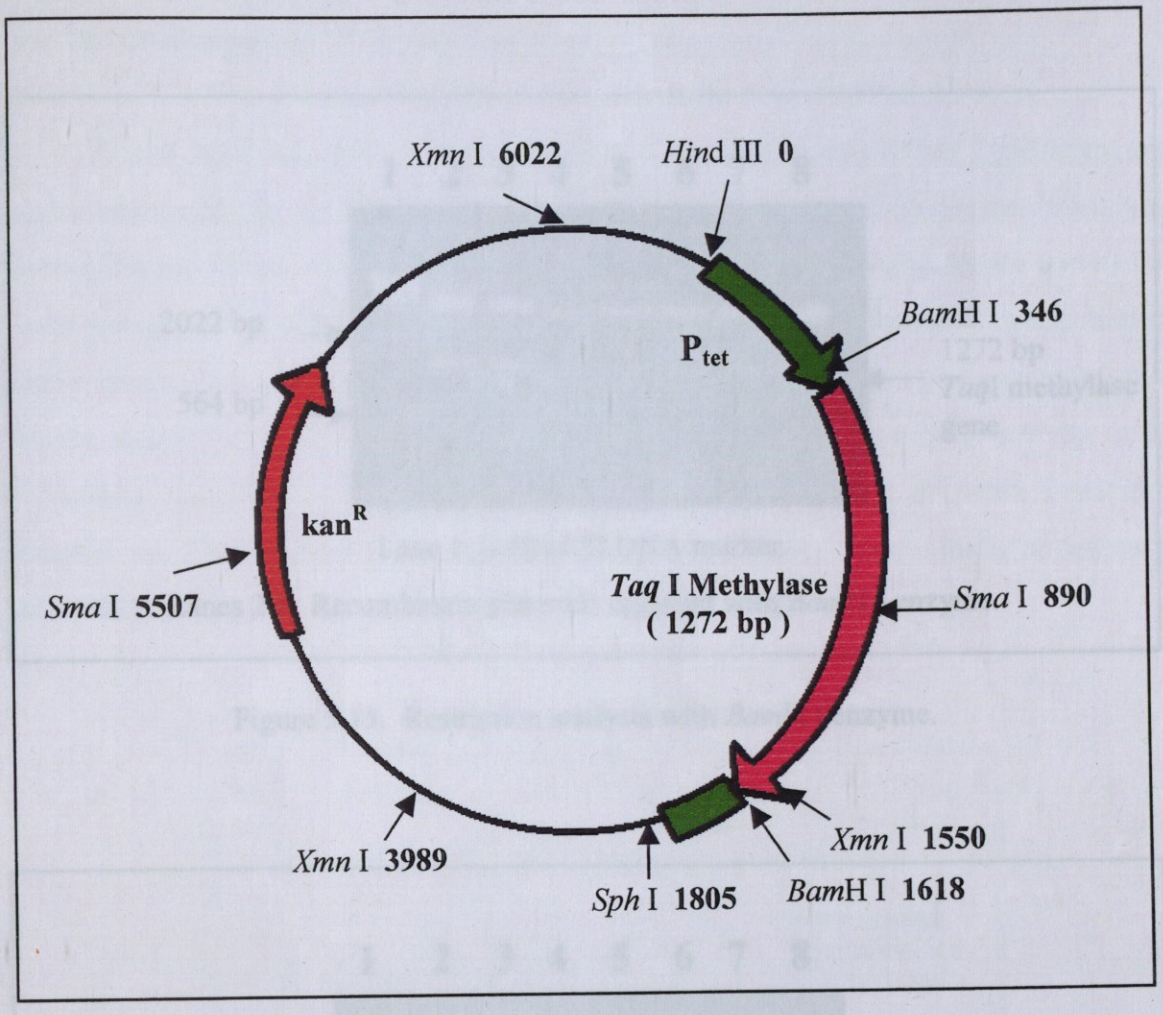
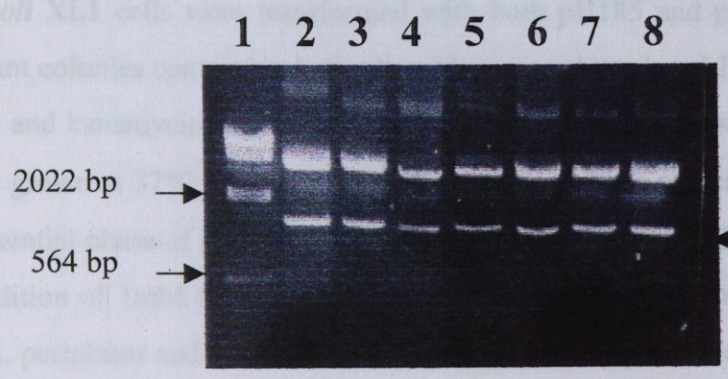


Figure 5.12. Restriction map of the 6749 base pair recombinant plasmid pETMET

Plasmids of 12 selected colonies were purified by the mini-prep method as described in Methods 4.4.2 and digested with *Bam*H I, *Xmn* I and *Sma* I enzymes. *Xmn* I and *Sma* I digestions gave the expected band patterns in all recombinants. Electrophoretic analysis of seven recombinant plasmids digested with *Bam*H I is shown in Figure 5.13.

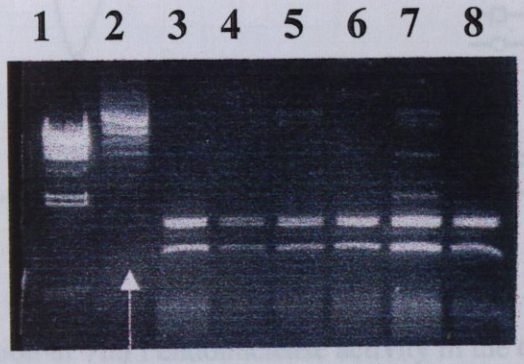
These 12 recombinant plasmids were also analyzed for the expression of the methylase gene by digesting them with *Taq* I endonuclease at 65°C for one hour. Figure 5.14 shows the agarose gel electrophoresis of seven recombinants, digested with *Taq* I endonuclease.



Lane 1: λ-*Hind* III DNA marker.

Lanes 2-8: Recombinant plasmids digested with *Bam*H I enzyme.

Figure 5.13. Restriction analysis with *Bam*H I enzyme.



Lane 1: λ-*Hind*III DNA marker.

Lanes2-8: Recombinant plasmids digested with *Taq* I endonuclease enzyme.

Figure 5.14. Expression of the methylase gene in the suspected recombinants.

One of the colonies, indicated with the white arrow, showed complete resistance to endonuclease digestion whereas others showed either partial or no resistance at all (Figure 5.14). This colony was chosen, named as pETMET and used for the subsequent expression studies.

5.1.10. Co-Expression of the *Taq* I Methylase Gene

E. coli XL1 cells were transformed with both pH185 and pETMET plasmids, and recombinant colonies containing both plasmids were selected on LB agar plates containing ampicillin and kanamycin. Selective LB medium was inoculated with a single colony and cells were grown at 37°C in an orbital shaker with a shaking velocity of 180 rpm. At their late-exponential phase of growth, expression of the MBP-*Taq* I fusion protein was initiated by the addition of 1mM IPTG to the growth medium. *Taq* I endonuclease activity in the cytoplasm, periplasm and extracellular medium of the samples taken at certain times were determined. The distribution of the enzyme activity in these three cellular locations is presented in Figure 5.15.

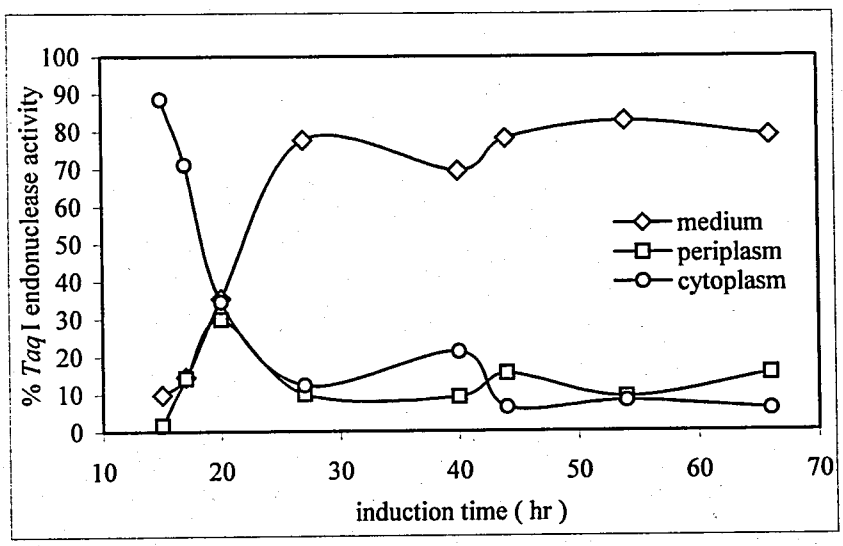


Figure 5.15. Distribution of *Taq* I endonuclease activity in the cytoplasm, periplasm and medium of *E. coli* XL1 (pH185, pETMET) cells.

Within the first 20 hours of induction, the percentage of cytoplasmic activity decreased, while that of the periplasm and extracellular medium increased. After 20 hours, release of enzyme activity from the periplasm to the medium was observed and a plateau was reached after 24 hours of induction. Time profile of the volumetric *Taq* I endonuclease activity recovered from the three cellular compartments is shown in Figure 5.16.

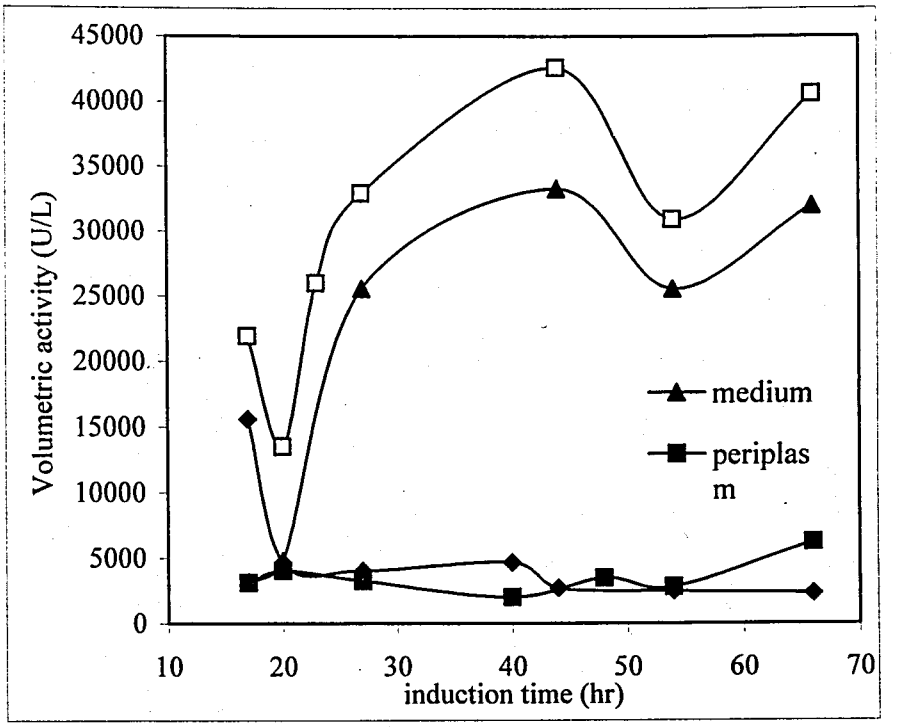


Figure 5.16. Time profile of volumetric *Taq* I endonuclease activity recovered from the cytoplasm, periplasm and medium of *E. coli* XL1 (pH185, pETMET) cells.

When the volumetric activities recovered from *E. coli* XL1 (pH185, pETMET) cells were compared with those of the *E. coli* XL1 (pH185) cells in Figure 5.1, about eight times lower total and extracellular volumetric *Taq* I endonuclease activities could be obtained from the methylase protected cells. However, in the single-plasmid system, the extracellular activity reached its maximum value after 18 hours of induction, and then decreased within the following two hours to one third of its maximum value. In the methylase protected cells, about 80 per cent of the total *Taq* I endonuclease activity was excreted to the growth medium after 24 hours of induction and maintained its value even after 66 hours of induction. In addition, maximum specific activity isolated from the osmotic shock fluid after 24 hours of induction was two times higher than that of the single vector system.

5.1.11. Effect of Medium Formulation

In order to compare the growth characteristics of *E. coli* XL1(pH185) and *E. coli* XL1(pH185, pETMET) cells, cultures were grown in selective LB medium at 37°C in an orbital shaker with a shaking velocity of 180 rpm. Samples were taken from these cultures at regular time intervals and the optical density and dry biomass concentrations were determined as described in Methods 4.12.1. Figure 5.17 shows the growth characteristics of these two cultures without induction.

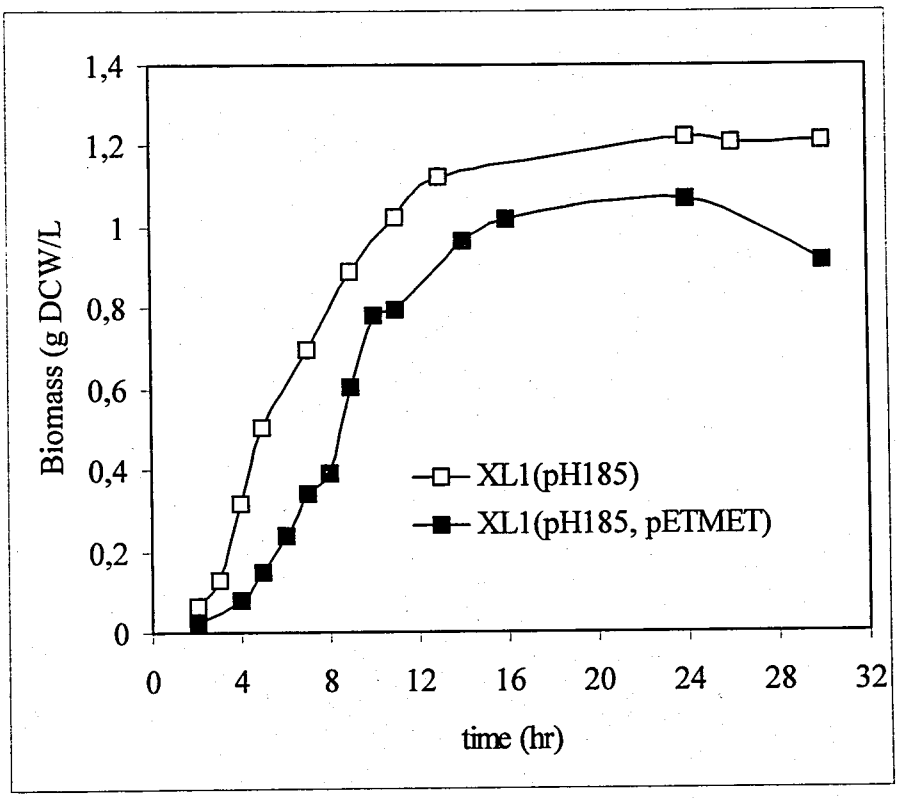


Figure 5.17. Growth characteristics of *E. coli* XL1(pH185) and *E. coli* XL1(pH185, pETMET) cells.

The *E. coli* XL1(pH185) and *E. coli* XL1(pH185, pETMET) cells were found to enter their exponential growth phase after about one and two hours of adaptation period, respectively. Elongation of the *lag* time in the cells harboring both plasmids was most probably due to the constitutive expression of the *Taq* I methylase gene. The maximum specific growth rates, μ_{max}, of the recombinant cultures were calculated from the slopes of lnx versus time plots by linear regression as 0.698±0.01 h^{-1} and 0.459±0.02 h^{-1} for *E. coli*

XL1(pH185) and *E. coli* XL1(pH185, pETMET) cells, respectively. Both cultures entered their stationary phase of growth after 14 hours. The maximum dry cell mass yields of the single- and binary-plasmid systems were 1.22 and 1.07 g DCW/L, respectively.

When the growth behaviour of *E. coli* XL1(pH185,pETMET) cells were compared with the single-plasmid system, 30 per cent decrease in maximum specific growth rate, lower biomass yields and longer adaptation times (t_{lag}) were observed. Since *Taq* I methylase enzyme is produced continuously during the growth of the cells, LB medium was likely to encounter nutrient limitations and thus might not reflect the true performance of the culture under prolonged induction conditions. Therefore, experiments were conducted in shake flask cultures using SB, TB and 2xYT media and the growth characteristics of the recombinant cells in these media were compared with that in LB medium (Figure 5.18).

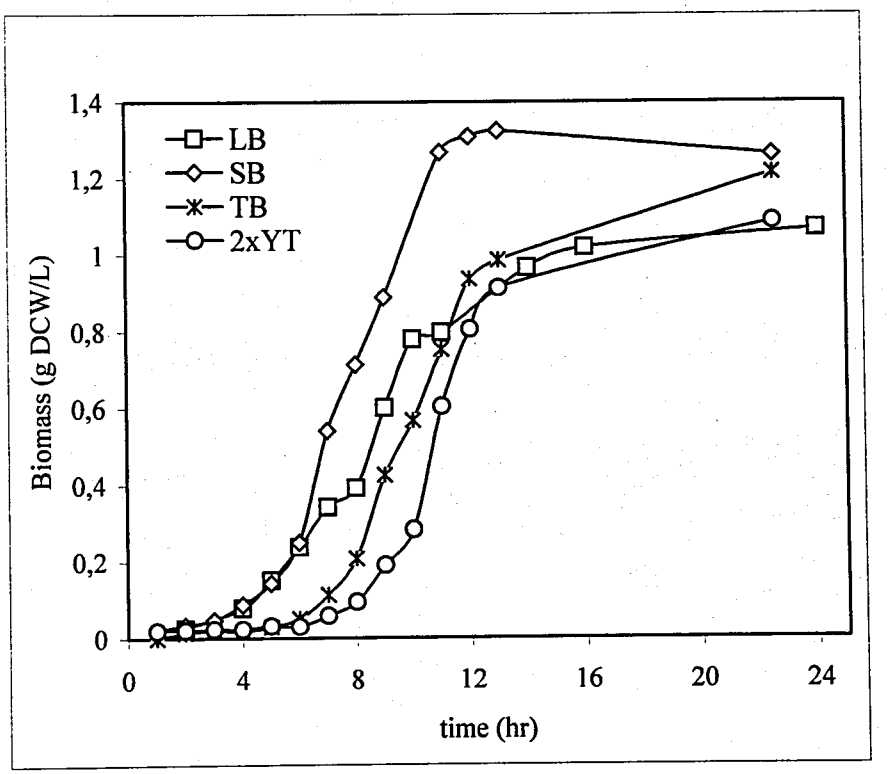


Figure 5.18. Effect of the medium composition on the growth characteristics of *E. coli* XL1(pH185, pETMET) cells.

Prolonged adaptation periods of five and six hours were observed in cells grown in TB and 2xYT media. The maximum specific growth rates during the exponential phase of the cells grown in SB, TB and 2xYT media were calculated as 0.591±0.005 h^{-1}, 0.699±0.003 h^{-1} and 0.577±0.008 h^{-1}, respectively. Although the μ_{max} values of cultures grown in TB and 2xYT media were 30-40 per cent higher than that of the LB medium, their adaptation times were almost doubled at the same conditions. All cultures reached the stationary phase of growth after 14 hours and SB medium gave the highest biomass yields of 1.32 g DCW/L.

The effect of medium type on enzyme production of *E. coli* XL1(pH185,pETMET) cells were also determined by analyzing the volumetric yields obtained from the three cellular compartments after 24 hours of induction (Figure 5.19).

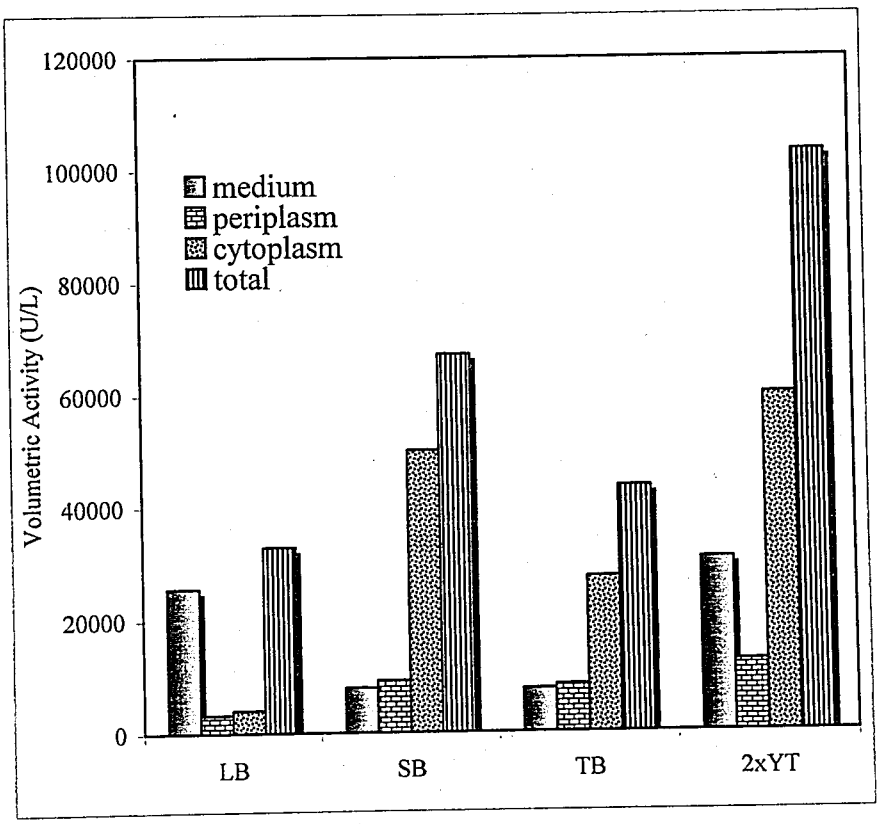


Figure 5.19. Effect of the medium formulation on the distribution of *Taq* I endonuclease activity in *E. coli* XL1(pH185,pETMET) cells after 24 hours of induction.

Highest cytoplasmic and periplasmic activities were obtained from SB and 2xYT media where 12 and 30 per cent of total volumetric activity were excreted to the growth medium, respectively. Total *Taq* I endonuclease activity recovered from cells grown in TB medium was 32 per cent higher than that of the LB medium, however, only 18 per cent of this activity was extracellular. On the other hand, almost 40 and 60 per cent of the cytoplasmic marker isocitrate dehydrogenase activity was extracellular in TB and 2xYT media respectively, which could be an indication of a possible cell lysis problem. The extracellular specific *Taq* I endonuclease activities were calculated as 472, 885, 97 and 326 U/mg for *E. coli* XL1(pH185,pETMET) cells grown in LB, SB, TB and 2xYT media, respectively.

Growing the recombinant cells in SB medium not only improved the growth characteristics but also total and extracellular *Taq* I endonuclease recovery yields without cell lysis. Therefore, LB medium was replaced with SB medium for the subsequent fermentor studies.

5.1.12. Production of MBP-*Taq* I Fusion Protein in Bioreactor

The culture pH, which is an important parameter for the stability of the excreted proteins during cultivation, may flactuate during the uncontrolled shake flask experiments. Therefore, in order to analyze the effect of methylase activity on extracellular secretion of MBP-*Taq* I fusion protein on a better basis, experiments were performed using SB medium supplemented with 20 g/L glucose in a 2.5 L New Brunswick BioFlo III batch/continuous fermentor with well-controlled environment of pH, temperature and dissolved oxygen concentration.

The fermentor was sterilized with SB medium where glucose and M9 salts were sterilized separately and added to the sterile fermentor vessel together with the filter-sterilized ampicillin and/or kanamycin to a final working volume of 1.5 liter. Calibration of the pH and dissolved oxygen probes were made as described in the Methods section. The temperature was kept constant at 37°C by heating or cooling the system with an external circulation of tap water through the coils. The pH was measured by glass electrode

and controlled at 7.0 through the addition of either 1N HCl or 1N NaOH. Dissolved oxygen (DO) was monitored using a polarographic oxygen electrode and sustained above 30 per cent saturation. The oxygen was supplied to the system with an air tube connected to the rotameter at the fermentor inlet. The air was filter-sterilized by a 0.2μm replaceable cartridge filter and supplied to the culture at a constant rate of 1vvm. The agitation speed was set at 600 rpm throughout the run. A silicone antifoaming agent was used at 1:10 dilution with sterile dH_2O.

The inoculum was prepared by transferring a single colony from selective LB plates into 50 ml SB medium containing the appropriate antibiotics and growing the cells in orbital shaker at 37°C and with an agitation rate of 180 rpm overnight. Fermentation was started by adding one per cent (v/v) inoculum to the fermentor and samples were taken throughout the experiments at certain time intervals. Expression of the MBP-*Taq* I fusion protein was initiated by inducing the cultures with 1mM IPTG just before they entered their stationary phase of growth. Time course of biomass formation, distribution of *Taq* I endonuclease activity, plasmid stability, extracellular glucose and acetic acid concentrations, dissolved oxygen and isocitrate dehydrogenase concentrations were followed in all cases.

5.1.13. Growth Characteristics

Time profiles for biomass formation, dissolved oxygen and glucose utilization of *E. coli* XL1(pH185) and *E. coli* XL1(pH185, pETMET) cells are summarized in Figure 5.20.

When the biomass profiles of the two systems were compared, unprotected cells entered the logarithmic growth phase after one hours of *lag* time, while two hours of adaptation period was observed in the *E. coli* XL1 cells harboring both plasmids. The maximum specific growth rates of the cultures were determined from the slopes of the lnx versus time plots in the exponential phase using linear regression as 1.451±0.05 h^{-1} and 1.344±0.07 h^{-1} for the unprotected and methylase protected cells, respectively. These values were almost equal to the maximum specific growth rate data reported by Minas and Bailey (1995) for various *E. coli* strains grown in a batch fermentor.

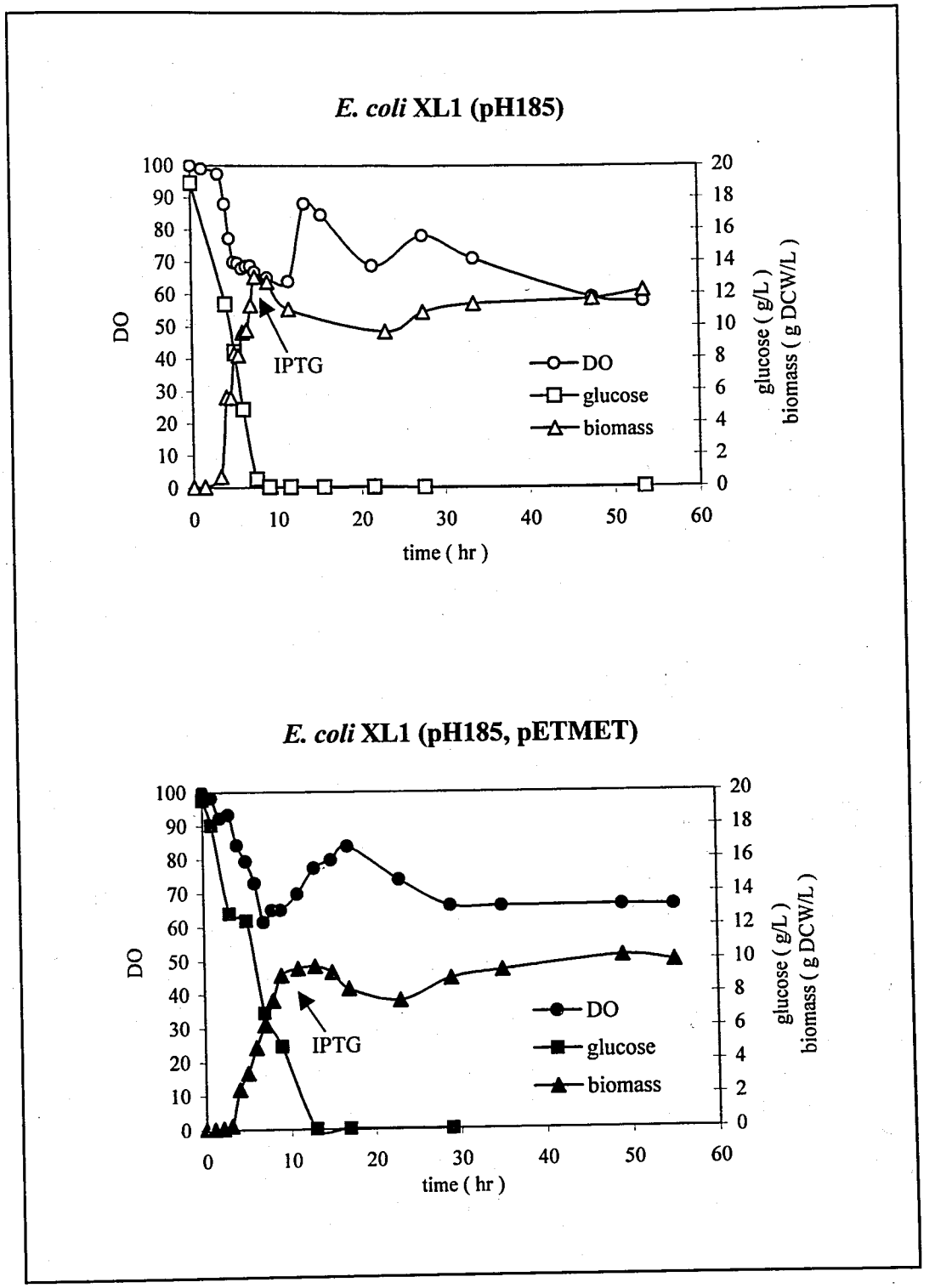


Figure 5.20. Time profiles for biomass concentration, glucose utilization and dissolved oxygen of *E. coli* XL1(pH185) and *E. coli* XL1(pH185, pETMET) cells.

The unprotected *E. coli* XL1(pH185) cells reached higher biomass concentrations throughout the course of fermentation. The maximum biomass concentrations reached by the induced cultures were 13.09 and 10.2 g DCW/L for the unprotected and methylase protected cells, respectively. Three times higher biomass yields could be obtained from *E. coli* XL1(pH185, pETMET) cells grown under controlled fermentor conditions when compared with the shake flask cultures.

Dissolved oxygen concentration decreased with increasing dry cell mass during the exponential growth phase in both cultures. When *E. coli* XL1(pH185) cells were induced with IPTG, a 20 per cent decrease in dry cell weight with a concomitant sudden increase in extracellular DO concentration was observed. This could be due to the fact that all of the glucose was utilized or that the onset of gene expression forced the cells to direct their metabolic activities from cellular growth towards recombinant protein production.

In *E. coli* XL1(pH185, pETMET) cells, on the other hand, induction with IPTG also resulted in a decrease in biomass concentration but the simultaneous increase in extracellular DO concentration was very smooth when compared with the unprotected cells. Since *Taq* I methylase was produced constitutively during the growth of these cells, the metabolic activities of these cells were accustomed to plasmid encoded recombinant protein synthesis and as a result, expression of MBP-*Taq* I fusion protein did not affect these cells as much as the unprotected cells.

20 g/L glucose added at the start of the fermentation was completely consumed after 9-11 hours of fermentation in both cultures with comparable glucose utilization trends. The biomass yield on glucose (Yx/s) for the single and double plasmid systems were 0.618 and 0.477 g of dry biomass per gram glucose, respectively. The maximum acetate accumulation for the unprotected and methylase protected systems were at nontoxic levels of 0.35 and 0.1 g/L, respectively.

5.1.14. Distribution of *Taq* I Endonuclease Activity

The cytoplasmic, periplasmic and medium proteins from the samples taken during the fermentation of the induced cultures were prepared as described in Methods 4.10 and their *Taq* I endonuclease activities were determined. Plasmid stability of the cultures was also determined by transferring colonies from plain LB plates to selective LB plates by the toothpicking method as described in 4.14.2. The plasmid stability and specific *Taq* I endonuclease activities recovered from the three cellular compartments of *E. coli* XL1(pH185) and *E. coli* XL1(pH185, pETMET) cells were shown in Figure 5.21.

The percentage of plasmid containing colonies dropped to less than 15 per cent within the first four hours of induction of *E. coli* XL1(pH185) cells. This rapid decrease in plasmid stability hindered the secretion of the MBP-*Taq* I fusion protein to the periplasmic space and hence resulted in its accumulation in the cytoplasm. However, in the binary plasmid system, where cellular DNA was protected from endonuclease attack by metylation activity, the plasmid remained highly stable (above 80 per cent) for ten hours of induction. Also the cytoplasmic accumulation of the fusion protein was prevented and its periplasmic secretion was highly improved such that more than three times higher specific *Taq* I activity could be isolated from the osmotic shock extracts. When the two systems were compared for their extracellular secretion efficiencies, in the single plasmid cells, *Taq* I activity could only be found after 40 hours of induction time whereas in the second system, extracellular secretion of enzyme activity started after 20 hours and specific enzyme production was 1.5 times higher than the unprotected cells.

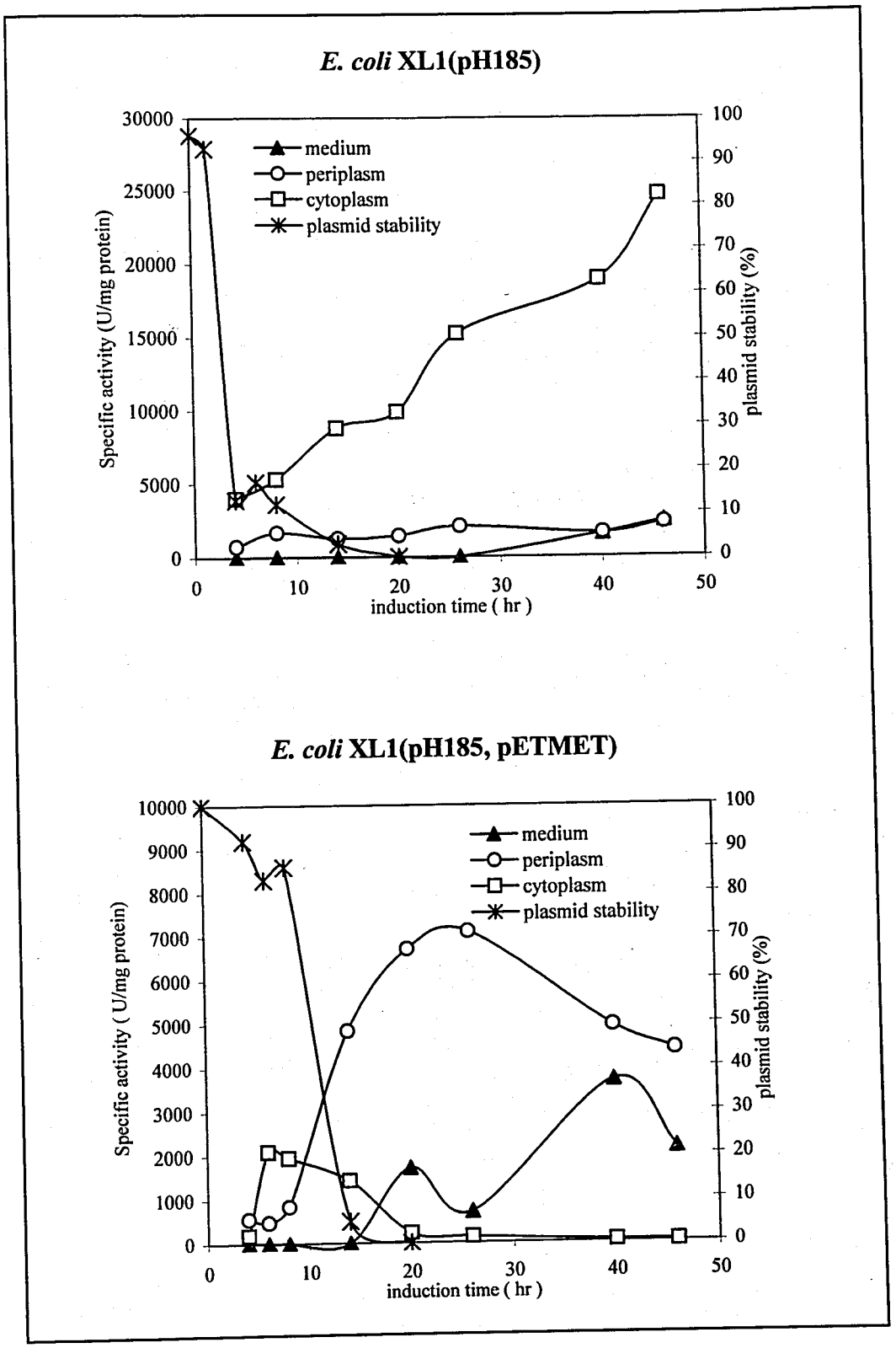


Figure 5.21. Plasmid stability and specific *Taq* I endonuclease activity recovered from the *E. coli* XL1(pH185) and *E. coli* XL1(pH185, pETMET) cells.

Total volumetric *Taq* I endonuclease activity recovered from the *E. coli* XL1(pH185) cells was $2.3x10^7$ U/L which was 70 times higher than that obtained from the shake flask cultures in LB medium (Figure 5.22). However, only 2.7 per cent of this activity was excreted to the growth medium after 46 hours of induction. Total volumetric *Taq* I endonuclease activity recovered from the *E. coli* XL1(pH185, pETMET) reached its maximum value of $1.9x10^6$ U/L after eight hours of induction then decreased and remained at values about $0.4x10^6$ U/L. Although ten times lower volumetric activities could be obtained from these cells, most of this activity was excreted to the growth medium (Figure 5.23).

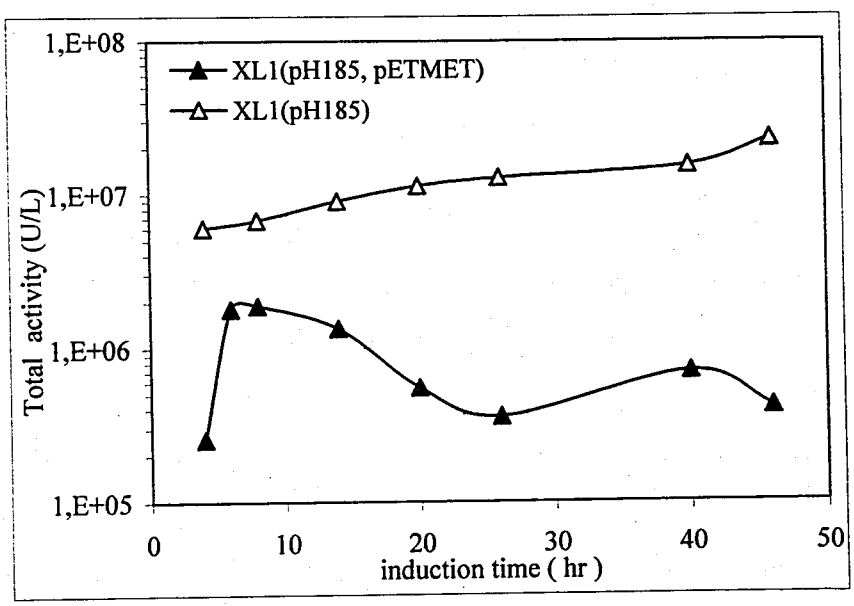


Figure 5.22. Time profile of total volumetric *Taq* I endonuclease activity yields recovered from *E. coli* XL1(pH185) and *E. coli* XL1(pH185, pETMET) cells.

In *E. coli* XL1(pH185) cells, most of the *Taq* I endonuclease activity remained trapped in the cytoplasm and periplasm of *E. coli* XL1 cells and only a small fraction (2.7 per cent) leaked into the culture medium. However, in the methylase protected cells, about 85 per cent of the total activity could be recovered from the growth medium after 40 hours of induction.

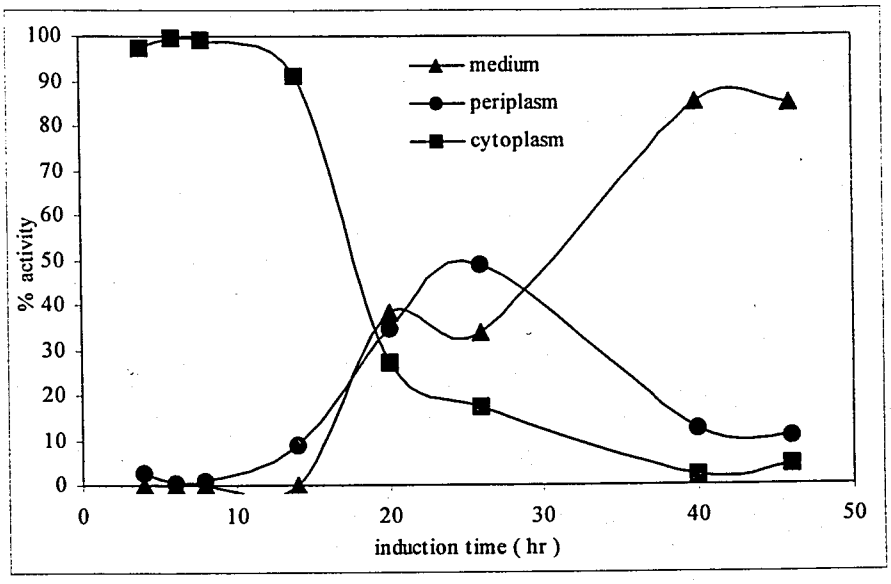


Figure 5.23. Distribution of volumetric *Taq* I endonuclease activity in the three compartments of *E. coli* XL1(pH185, pETMET) cells.

In all experiments, Isocitrate dehydrogenase was used as a cytoplasmic marker protein to ensure that the extracellular secretion did not come from loss of cell integrity. Cytoplasmic and medium fractions were assayed in all cases and almost all isocitrate dehydrogenase activity was found to be intracellular whereas only negligable amounts could be detected in the medium.

Co-expression of the methylase gene in *E. coli* XL1 cells producing MBP-*Taq* I fusion protein under the control of the *tac* promoter resulted in higher plasmid stabilities and excretion levels under controlled bioreactor conditions.

5.2. Expression of *Taq* I Restriction-Modification System Using pET System

pET system utilizes the bacteriophage T7 RNA polymerase to express cloned DNA under the control of the T7 promoter (Studier and Moffatt, 1986). Since T7 RNA polymerase is highly specific for its own promoter, this system is widely used for high-level expression of toxic proteins in *E. coli.*

In this study, two different expression systems were developed. In the first system, the coding sequence for *Taq* I endonuclease was amplified from the *Thermus aquaticus* bacterial genome by PCR and cloned into the pET28a+ expression plasmid under the control of T7 promoter. The recombinant construct, pETaqR, was introduced into the *E. coli* BL21(DE3) strain by electroporation. For the second system, *E. coli* BL21(DE3) cells were transformed with both pETaqR and pMETaq plasmids.

The two systems were investigated in many aspects, including medium type, fermentation characteristics, behaviour under bioreactor conditions, enzyme production, plasmid stability and cell viability.

5.2.1. Cloning of *Taq* I Restriction Endonuclease Gene into pET28a+ Vector

pET28a+ vector contains an *Nco* I site upstream from the fusion tag sequences for cloning into the AUG start codon at the 5' end of the insert coding sequence. Therefore when cloned at the *Nco* I site of the pET28a+ plasmid, the protein expressed by the insert gene does not contain vector-encoded sequences. The *Eco*R I site is located in the multiple cloning site of the vector upstream from the translation stop codons and T7 transcription terminator sequences. Consequently, as shown in Figure 5.24, *Nco* I and *Eco*R I sites has been chosen for directional cloning the *Taq* I endonuclease gene into pET28a+ vector.

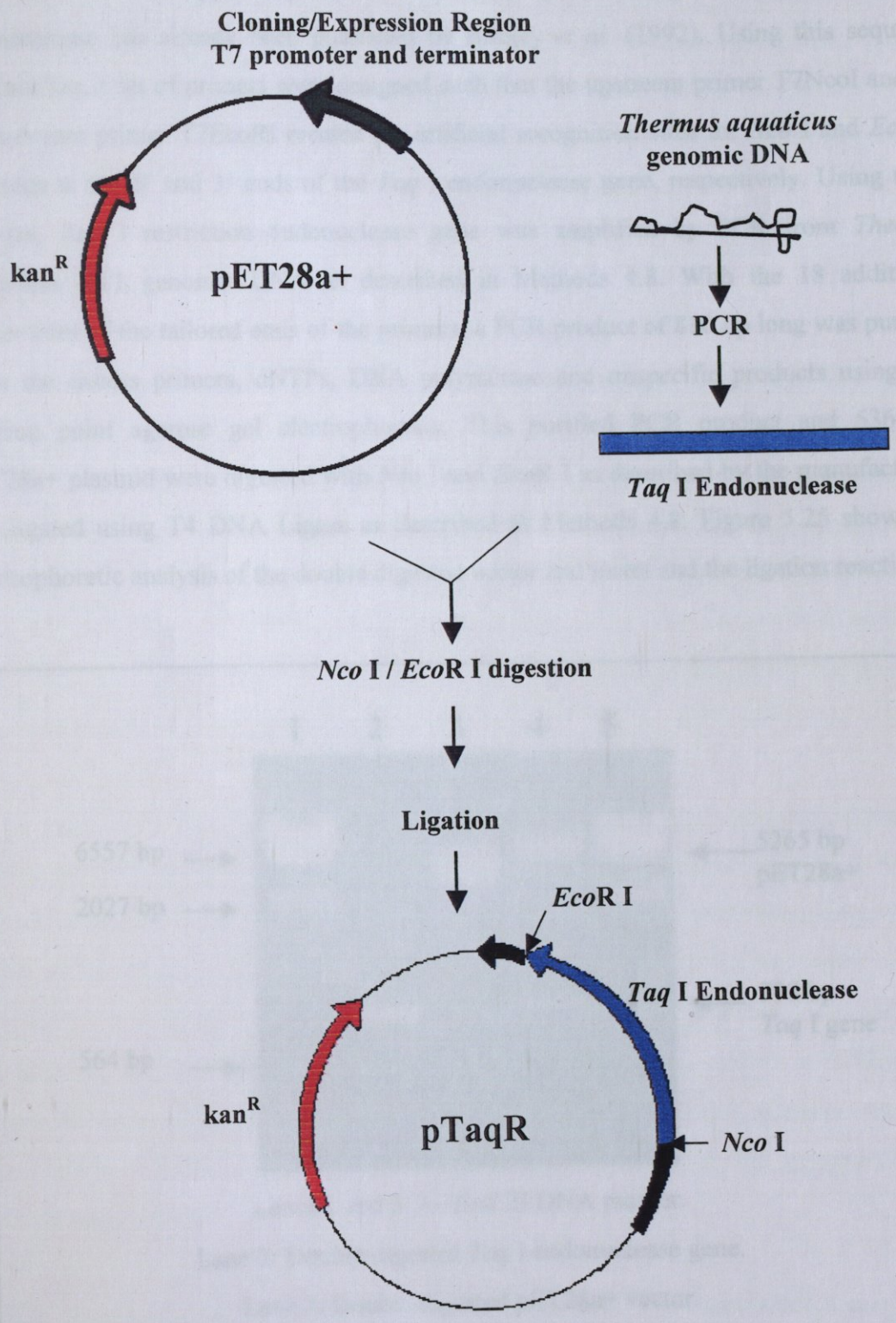
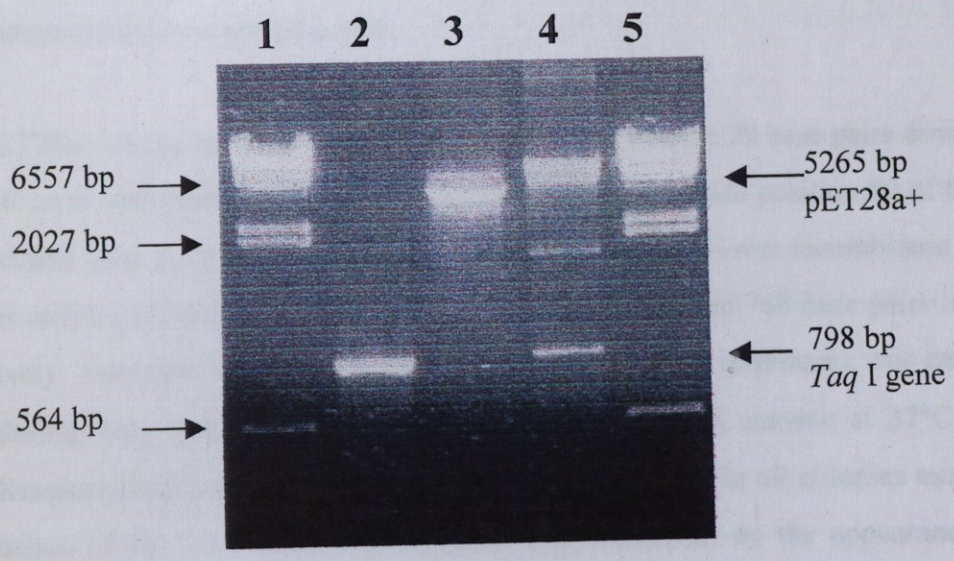


Figure 5.24. Directional cloning strategy of *Taq* I endonuclease gene into pET28a+ expression vector

The 792 base pair sequence of the entire gene coding for *Taq* I restriction endonuclease has already been published by Barany *et al.* (1992). Using this sequence information, a set of primers were designed such that the upstream primer T7NcoI and the downstream primer T7EcoRI created the artificial recognition sites for *Nco* I and *Eco*R I enzymes at the 5' and 3' ends of the *Taq* I endonuclease gene, respectively. Using these primers, *Taq* I restriction endonuclease gene was amplified by PCR from *Thermus aquaticus* YT1 genomic DNA as described in Methods 4.8. With the 18 additional nucleotides of the tailored ends of the primers, a PCR product of 810 bp long was purified from the excess primers, dNTPs, DNA polymerase and unspecific products using low melting point agarose gel electrophoresis. This purified PCR product and 5369 bp pET28a+ plasmid were digested with *Nco* I and *Eco*R I as described by the manufacturers and ligated using T4 DNA Ligase as described in Methods 4.8. Figure 5.25 shows the electrophoretic analysis of the double digested vector and insert and the ligation reaction.



Lanes 1 and 5: λ-*Hind* III DNA marker.

Lane 2: Double digested *Taq* I endonuclease gene.

Lane 3: Double digested pET28a+ vector.

Lane 4: ligation reaction.

Figure 5.25. Electrophoretic analysis of double digested DNA fragments and the ligation reaction.

Four major bands were detected in the ligation reaction, two of which correspond to the double digested vector and insert. The uppermost band with an approximate length of 6100 base pairs corresponded to the expected size of the recombinant. The second band from the bottom had an inferred length of 1600 base pair and was most likely generated by the ligation of two inserts.

5.2.2. Screening of Recombinant Colonies

The ligation reaction was used to transform the competent *E. coli* XL1 cells by electroporation and after overnight incubation at 37°C, plasmid containing cells formed colonies on LB plates containing kanamycin. Recombinant colonies were first screened by the cracking method. Analysis of 112 colonies by the cracking method resulted in the identification of 32 colonies on the basis of their plasmid size. The presence of the *Taq* I endonuclease gene in these colonies was determined by the colony PCR method as described in Methods 4.5.7.2. In 20 colonies, an amplification product at the size of the *Taq* I endonuclease gene was observed.

pET28a+ vector contains a *Hind* III restriction site located 20 base pairs downstream the EcoR I site and there is a unique *Hind* III site at the nucleotide position 50 of the *Taq* I endonuclease gene sequence. As a result, the digestion of the correct recombinant plasmid with this enzyme should create two DNA fragments of 5295 and 768 base pairs in length, respectively. Therefore, restriction analysis of the suspected recombinants was carried out by incubating their isolated plasmids with *Hind* III restriction enzyme at 37°C for two hours. Resulting fragments were analyzed on agarose gels and in all colonies except two, the presence of the *Taq* I endonuclease gene was confirmed by the appearance of the correct band pattern.

5.2.3. Expression of the *Taq* I Endonuclease Gene in *E. coli* BL21(DE3) Lysogens

A λDE3 lysogen of *E. coli* strain BL21 was used as host organism for the expression of the *Taq* I endonuclease gene. Bacteriophage λDE3 carries a DNA fragment containing the *lacI* gene, the *lacUV5* promoter, and the gene for T7 RNA polymerase (Studier and

Moffat, 1986). Once a DE3 lysogen is formed, the only promoter known to direct transcription of the T7 RNA polymerase gene is the *lacUV5* promoter, which is inducible by isopropyl-β-D-thiogalactopyranoside (IPTG). Addition of IPTG to the growing culture of the lysogen induces T7 RNA polymerase, which in turn transcribes the *Taq* I endonuclease gene from the recombinant pTaqR plasmid. Also, this *E. coli* BL21(DE3) strain is deficient in the La protease and lacks the ompT outer membrane protease that can degrade proteins during purification.

Among the 18 colonies selected from recombinant screening, three of them were used to check for the correct expression of the *Taq* I endonuclease gene. Competent *E. coli* BL21(DE3) cells were transformed with the plasmids of these colonies by electroporation. Recombinant lysogens were grown in liquid LB media containing kanamycin and induced at their exponential phase of growth with 1mM IPTG for two hours. Harvested cells were disrupted by sonication and *Taq* I endonuclease activity in the cell extracts were analyzed as described in Methods 4.9. All the three colonies were found to express the *Taq* I endonuclease gene at the same high level, therefore, one of them was chosen for further expression studies.

5.2.4. Optimization of Fermentation Conditions in Complex LB Medium

The effects of key fermentation variables such as growth characteristics, induction time and period, specific and volumetric enzyme productivities were systematically investigated to find optimal conditions maximizing the biomass production and the yield of *Taq* I endonuclease.

The culture temperature is an important parameter, which should be considered in optimizing recombinant gene expression. Shin *et al.* (1997) have investigated the effect of culture temperature on mini-proinsulin production and found that protein synthesis under the control of the T7 promoter was almost completely repressed at 30°C. Jeong and Lee (1999) have decreased the culture temperature from 37°C to 30°C to prevent the inclusion body formation of human leptin protein in recombinant *E. coli* BL21(DE3) lysogens. Although most of the leptin was recovered in soluble form, the total production yield

decreased by 30-40 per cent compared with that of the 37°C. Both research groups have concluded that the pET expression system was very temperature-sensitive and growth at sub-optimal temperatures resulted in lower product yields. Taking these published observations into consideration, 37°C was chosen to grow the recombinant cells throughout this study.

5.2.5. Determination of Growth Characteristics

E. coli BL21(DE3) cells were transformed with pTaqR and pMETaq plasmids by electroporation. Three different cultures consisting of the host *E. coli* BL21(DE3), *E. coli* BL21 cells containing pTaqR plasmid and *E. coli* BL21 cells containing both pTaqR and pMETaq plasmids were grown in conical flasks containing 200 ml complex LB medium in GFL 3032 orbital shaker at 37°C and at a rotational speed of 180 rpm. Samples were taken from these cultures at regular time intervals and the optical density and dry biomass concentrations were determined as described in Methods 4.12.1. Figure 5.26 shows the growth characteristics of these three cultures without induction.

Exponential growth of the recombinant cells containing both plasmids started after four hours of *lag* time whereas the host and single plasmid cells entered the logarithmic growth phase after one and two hours, respectively. This increase in adaptation time was consistent with the earlier observations mentioned in 5.1.11 for the MBP-fusion system and was most probably due to the constitutive expression of the *Taq* I methylase gene. The maximum specific growth rates, μ_{max}, of the recombinant cultures were calculated from the slopes of lnx versus time plots by linear regression as 1.166±0.001 h^{-1} for the host cells, 1.033±0.05 h^{-1} and 0.747±0.02 h^{-1} for the one and two plasmid systems, respectively. All the three cultures reached the stationary phase after eight hours of growth. The maximum dry cell mass yield of *E. coli* cells expressing the methylase gene was much lower (1.04 g/L) when compared with 1.50 and 1.34 g/L biomass yields obtained from the host and cells harboring pTaqR plasmid only.

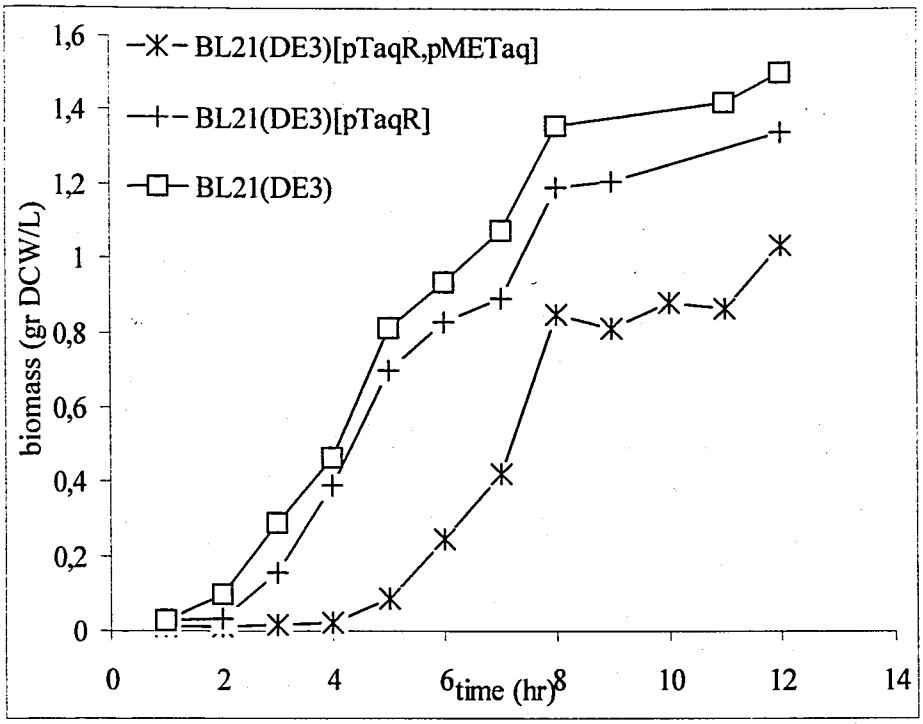


Figure 5.26. Growth characteristics of the host *E. coli* BL21(DE3), *E. coli* BL21(DE3)[pTaqR] and *E. coli* BL21(DE3)[pTaqR, pMETaq] cells.

Komai *et al.* (1997) reported that yeast extract in the LB medium affected the *lac* promoter for T7 RNA polymerase gene expression in *E. coli* BL21(DE3) and resulted in poor growth due to the leaky expression of the toxic HIV-1 protease gene. However, our results indicated that the uncontrolled expression of the *Taq* I endonuclease gene in uninduced *E. coli* BL21(DE3)[pTaqR] cells was not toxic to the host cell at detectable levels, and the growth is associated with slightly lower growth rate when compared to that of the host cells.

5.2.6. Optimization of the Induction Time and Period

The host *E. coli* BL21(DE3) contains a single copy of the T7 RNA polymerase gene in chromosome under the control of the *lac*UV5 promoter, therefore, the addition of IPTG induces the T7 RNA polymerase production, which in turn initiates high-level expression of the cloned *Taq* I restriction endonuclease gene. The recommended amount of the inducer IPTG required to titrate the repressor molecules by the manufacturer of the pET

system is 1 mM. Optimum IPTG concentration of 1 mM was also reported by Shin *et al.* (1997) and Jeong and Lee (1999). Therefore, in order to optimize the induction time, *E. coli*BL21(DE3)[pTaqR] and *E. coli*BL21(DE3)[pTaqR, pMetaq] cells were grown in selective LB medium and induced at different phases of their growth with 1mM IPTG. After two hours of induction, cells were harvested by centrifugation and *Taq* I endonuclease activity in crude cell extracts were determined as described in Methods 4.9.

Figure 5.27 shows the variation of the specific (-▲- , -Δ-) and volumetric (-◆- , -◊-) *Taq* I endonuclease activities of both cultures induced at different phases of their growth. The highest specific and volumetric *Taq* I endonuclease activities were obtained from *E. coli*BL21(DE3)[pTaqR] (open symbols) and *E. coli*BL21(DE3)[pTaqR, pMetaq] (closed symbols) cells when they were induced at their late- and mid-exponential phase of growth, respectively. The leftward shift of the peak in the two plasmid system was possibly due to the lowered nutritional quality of the growth medium since some of the essential metabolites were already utilized for the continuous expression of the *Taq* I methylase gene before the induction.

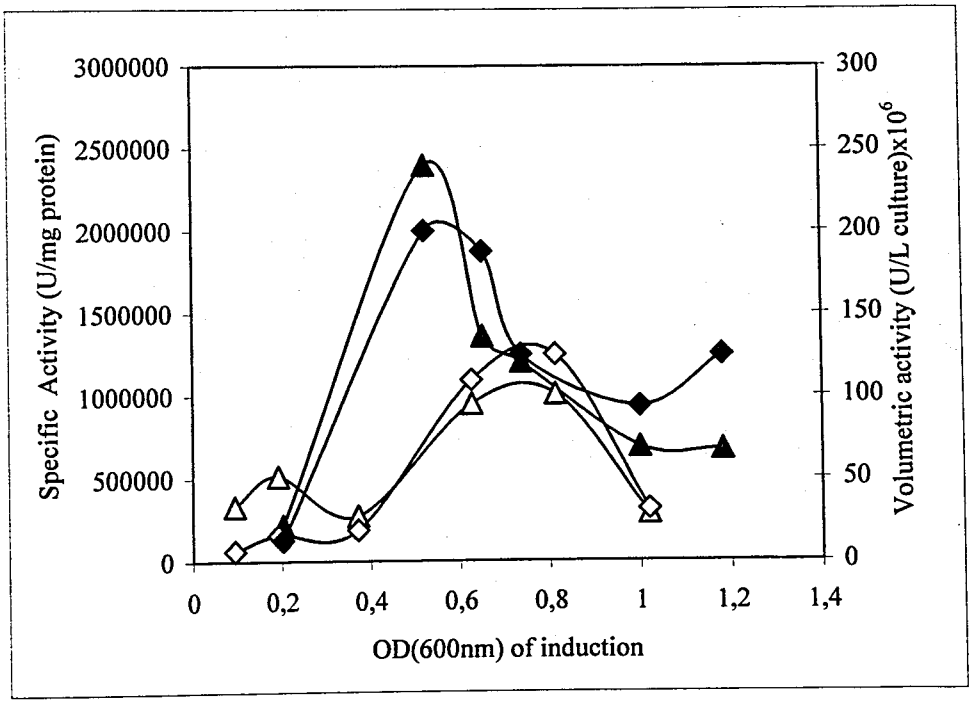


Figure 5.27. Optimization of induction time for the production of *Taq* I endonuclease.

The availability of nutrients especially during the phase when the cells are fully induced is crucial for recombinant product yield. Thus, since the protein production is growth associated, a high rate of production is achieved only when the cells are in their exponential phase (Gupta *et al.*, 1999).

In order to determine the optimum period of induction, *E. coli*BL21(DE3)[pTaqR] (open symbols) and *E. coli*BL21(DE3)[pTaqR, pMetaq] (closed symbols) cells were grown in selective LB medium and induced at late- and mid-exponential phases of their growth with 1mM IPTG. Samples were taken at regular time intervals, cells were harvested by centrifugation and biomass yields, specific (-▲- , -Δ-) and volumetric (-◆- , -◊-) *Taq* I endonuclease activities in crude cell extracts were determined (Figure 5.28).

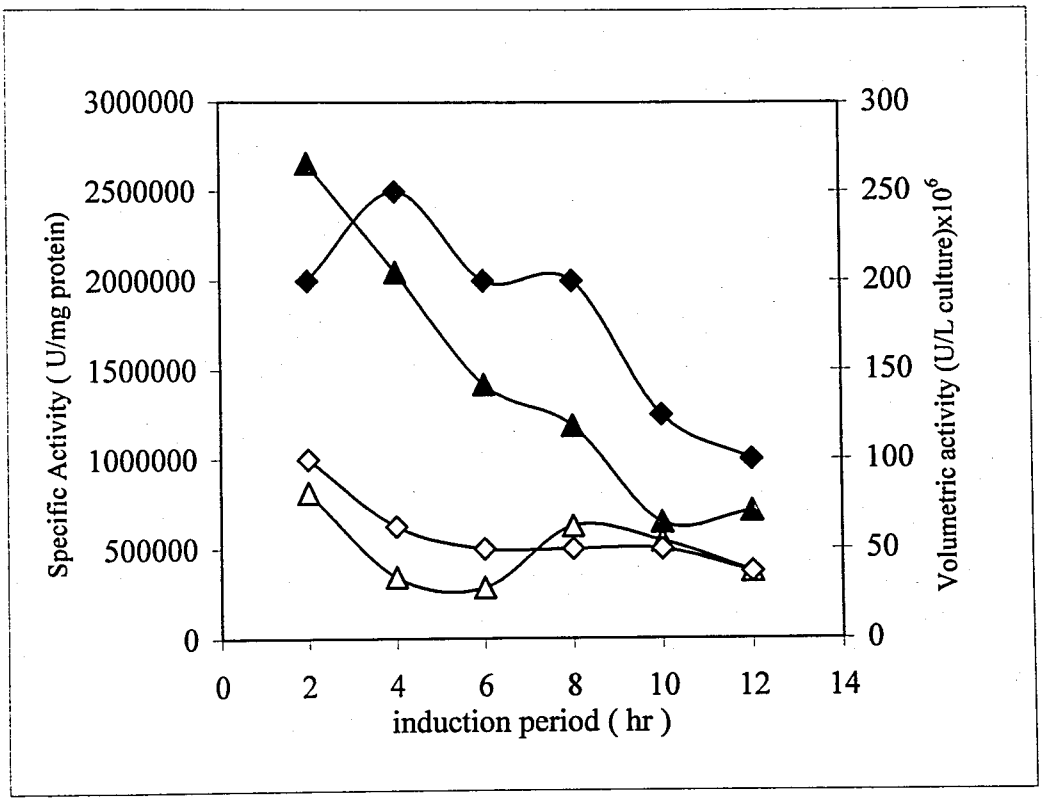


Figure 5.28. Optimization of induction period time for the production of *Taq* I restriction endonuclease.

Co-expression of *Taq* I methylase gene in *E. coli* BL21(DE3)[pTaqR, pMETaq] cells resulted in the production of at least two times or higher enzyme activities when compared to the unprotected cells. However, both specific and volumetric activities of the enzyme were found to decrease with the elongated induction periods.

Specific productivity is the time-average output of a process and defined as the product concentration per unit biomass divided by the total process time (Gaden, 2000). Since *Taq* I endonuclease is intracellularly accumulated in *E. coli* BL21(DE3) cells, productivity is proportional to both the biomass concentration and the specific productivity which is the amount of *Taq* I formed per unit dry cell mass per unit time. Figure 5.29 shows the biomass formation (-○- , -●-) and specific *Taq* I restriction endonuclease productivities (-□- , -■-) as a function of induction period for *E. coli* BL21(DE3)[pTaqR] (open symbols) and *E. coli* BL21(DE3)[pTaqR, pMETaq] (closed symbols) cultures.

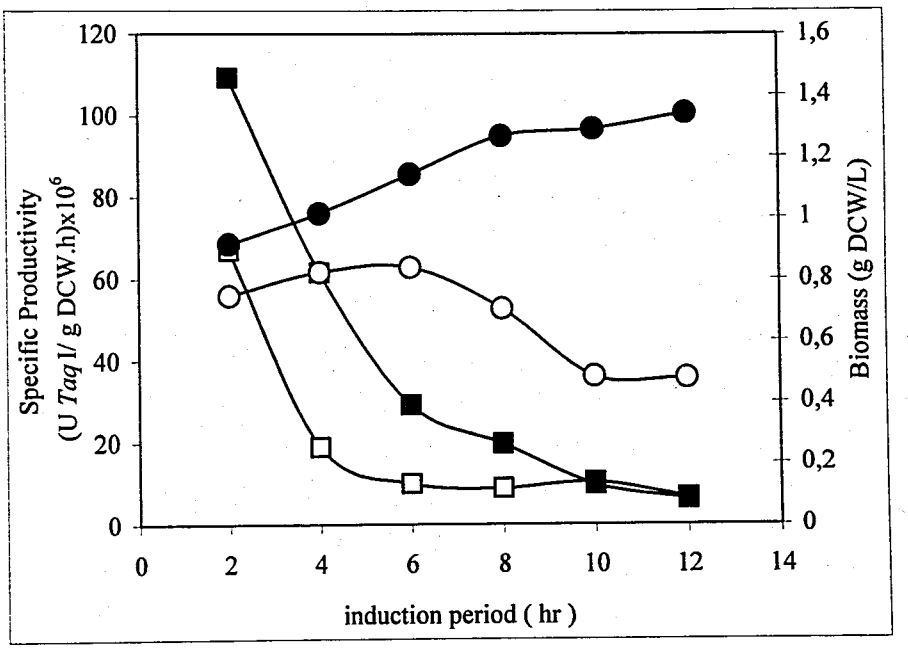


Figure 5.29. Productivity profiles of the two expression systems.

When the biomass profiles of the two systems were compared, induction periods longer than six hours were found to cause a decrease in the biomass accumulation of the *E. coli* BL21(DE3)[pTaqR] cells. The biomass concentration of the *E. coli* BL21(DE3)[pTaqR, pMETaq] culture maintained its increasing trend throughout the

whole induction period. A maximum of $109.3x10^6$ Units of *Taq* I restriction endonuclease were found to be produced per gram dry cell in one hour by recombinant cells containing both plasmids. However, 40 per cent lower *Taq* I endonuclease productivity was observed in unprotected cells. In both cultures, the specific productivity decreased with elongated induction periods which could be due to the overloaded host cell metabolism as suggested by Cserjan-Puschmann *et al.* (1999). After ten hours of induction, both cultures maintained equal levels of specific *Taq* I endonuclease productivity.

Delaying the induction time and prolonged induction periods had deleterious effects on the specific and volumetric *Taq* I endonuclease activity yields offsetting any gain achieved by the increase in cell biomass. This result was possibly due to nutrient starvation in the culture which lead to a lowering of the specific growth rate and hence a concomitant lowering of rates of *Taq* I endonuclease production.

In both expression systems, *Taq* I endonuclease was produced at very high amounts in complex LB medium when compared with the literature data. *E. coli* BL21(DE3)[pTaqR] cells produced $100x10^6$ U/L *Taq* I endonuclease activity when induced at their late-exponential phase of growth for two hours. This yield was about 17 times higher than $6x10^6$ U/L of *Taq* I activity reported by Barany *et al.* (1987) when *Taq* I endonuclease was expressed under the control of the alkaline phosphatase (phoA) and T7 promoters. This yield could be increased to only $9x10^6$ U/L when the *Taq* I methylase gene was also expressed but with a 20 amino acid deletion at its carboxyl terminus according to the corrected sequence published by Barany *et al.* (1992). However, in this work, co-expression of the methylase gene in the *E. coli* BL21(DE3)[pTaqR, pMETaq] cells resulted in a 2.5-fold increase in volumetric productivity yielding $250x10^6$ U/L *Taq* I endonuclease activity which was 28 times higher than the reported data.

One of the preliminary experiments conducted throughout this work was the cloning of the *Taq* I endonuclease gene into pUC18 expression vector and its expression under the control of the *lac* promoter. Effects of host strain, culture temperature, induction time and period, inducer concentration, plasmid stability and cell viability on *Taq* I endonuclease expression were analyzed in shake flask cultures. Maximum volumetric activity of 0.88

x10^6 U/L could be recovered at 37°C from *E. coli* TB1 cells when induced at their late-exponential phase with 0.1mM IPTG for 30 hours. By using pET expression system, about 100 times higher volumetric activities were obtained from the *E. coli* BL21(DE3)[pTaqR] cells.

5.2.7. Fermentation in Defined Medium

The fermentation medium has always been the critical component in the fermentation processes since by determining the chemical and nutritional environment it directly affects productivity and process economics (Zhang and Greasham, 1999). A chemically defined medium is composed of pure chemicals in precisely known proportions whereas a complex medium such as LB contains materials of undefined composition (Bailey and Ollis, 1986). When chemically defined medium is used in fermentation, the microorganism must synthesize all cellular components as well as products/by-products from simple, chemically defined substrates therefore for the selection of the optimum medium, all components that are necessary for growth should be identified very carefully (Zhang and Greasham, 1999).

Chemically defined medium has many advantages over complex media. Relative to complex media, equal or higher fermentation productivities with enhanced process consistencies can be reached with defined media. It has a low foaming tendency, high oxygen transfer rate and is less sensitive to sterilization conditions which enables improved scaling-up of commercial processes. Also the ability to characterize the medium offers the potential for maximal process control and monitoring (Zhang and Greasham, 1999).

Various defined medium formulations have been used for protein production using the T7 promoter system (Chavagnat *et al.*, 1996; Mishima *et al.*, 1997; Komai *et al.*, 1997; Teich *et al.*, 1998; Kim *et al.*, 1998; Shin *et al.*, 1998; Wong *et al.*, 1998; Akesson *et al.*, 1999; Cserjan-Puschmann *et al.*, 1999; Saraswat *et al.*, 2000). All these formulations are composed of a simple carbon source glucose or glycerol, nitrogen source in form of the ammonium ion, amino acids, vitamins, trace elements such as Fe, Ca, Mn, Zn, Cu, Co,

Mo, Al and B in form of inorganic salts and buffering components such as K_2HPO_4 or NaH_2PO_4 to maintain the pH at proper values.

In this work, a slightly modified form of R medium originally developed by Riesenberg *et al.* (1991) was used. In this new medium, glucose was used as carbon source since it can easily be metabolized to provide carbon and energy for biosynthesis. $(NH_4)_2HPO_4$ was used as nitrogen source since it has been reported that ammonia is the preferred nitrogen source for growth of *E. coli* on defined media with glucose as the carbon source (Yee and Blanch, 1993). Thiamine and hydrolyzed form of the casein protein (casaminoacids) were added as vitamin and amino acid sources respectively. In order to enhance the glucose uptake rate of the cells by extending the capacity of the tricarboxylic acid (TCA) cycle, the medium was also supplemented with citric acid.

5.2.8. Effect of Initial Glucose Concentration on Growth Properties

During balanced growth, only a single parameter μ (or the population doubling time) is required to characterize population growth kinetics. For this reason, the magnitude of the specific growth rate is widely used to describe the influence of the cells' environment on the cells' performance (Bailey and Ollis, 1986). The general goal in formulating a medium is to obtain good cellular growth and/or high rates of product synthesis. Contrary to intuitive expectations, this does not necessarily mean that all nutrients should be supplied in great excess since indiscriminate addition of nutrients may result in lower yields due to their interactions with various intracellular pathways and functions (Pilon *et al.*, 1996). Consequently, it is common practice to limit total growth by limiting the amount of one nutrient in the medium. This optimum amount of the essential medium constituent is found by varying its concentration while keeping the concentrations of all other medium components constant (Bailey and Ollis, 1986).

In order to analyze the effect of glucose concentration on the growth characteristics of the two recombinant systems, preliminary experiments were carried out in shake flask cultures containing defined media with increasing amounts of glucose as the sole carbon source. *E. coli* BL21(DE3)[pTaqR] and *E. coli* BL21(DE3)[pTaqR, pMETaq] cells were

grown in 200 ml selective defined medium containing 5, 10, 20 and 40 g/L glucose in GFL 3032 orbital shaker at 37°C and at a rotational speed of 180 rpm. Samples were taken from these cultures at regular time intervals to determine their growth characteristics in terms of optical density and biomass concentration as described in Methods 4.12.1. Residual glucose concentration in the medium was determined using an enzymatic assay kit as described in Methods 4.12.5.

The time-dependent variations in the dry cell weight and glucose concentrations of the two systems under uninduced conditions are shown in Figure 5.30 and Figure 5.31, respectively.

An approximately six hours of *lag* time was observed in *E. coli* BL21(DE3)[pTaqR, pMETaq] cells at all glucose concentrations. This adaptation period is two hours longer than that observed in the complex LB medium due to the limited nutrient sources in the defined medium. The maximum specific growth rates of the cultures were determined from the slopes of the lnx versus time plots in the exponential phase using linear regression (Table 5.6).

Table 5.6. Effect of glucose concentration on growth.

E. coli BL21(DE3) Strain	Initial Glucose Concentration (g/L)	Final Glucose Concentration (g/L)	Final Biomass Concentration (g DCW/L)	μ_{max} ±0.02 (h^{-1})
[pTaqR]	5	0	1.52	0.964
	10	0	1.95	0.739
	20	5.2	2.05	0.811
	40	23.5	1.98	0.769
[pTaqR, pMeTaq]	5	0	1.25	0.749
	10	0	1.53	0.693
	20	7.9	1.65	0.696
	40	28.3	1.61	0.570

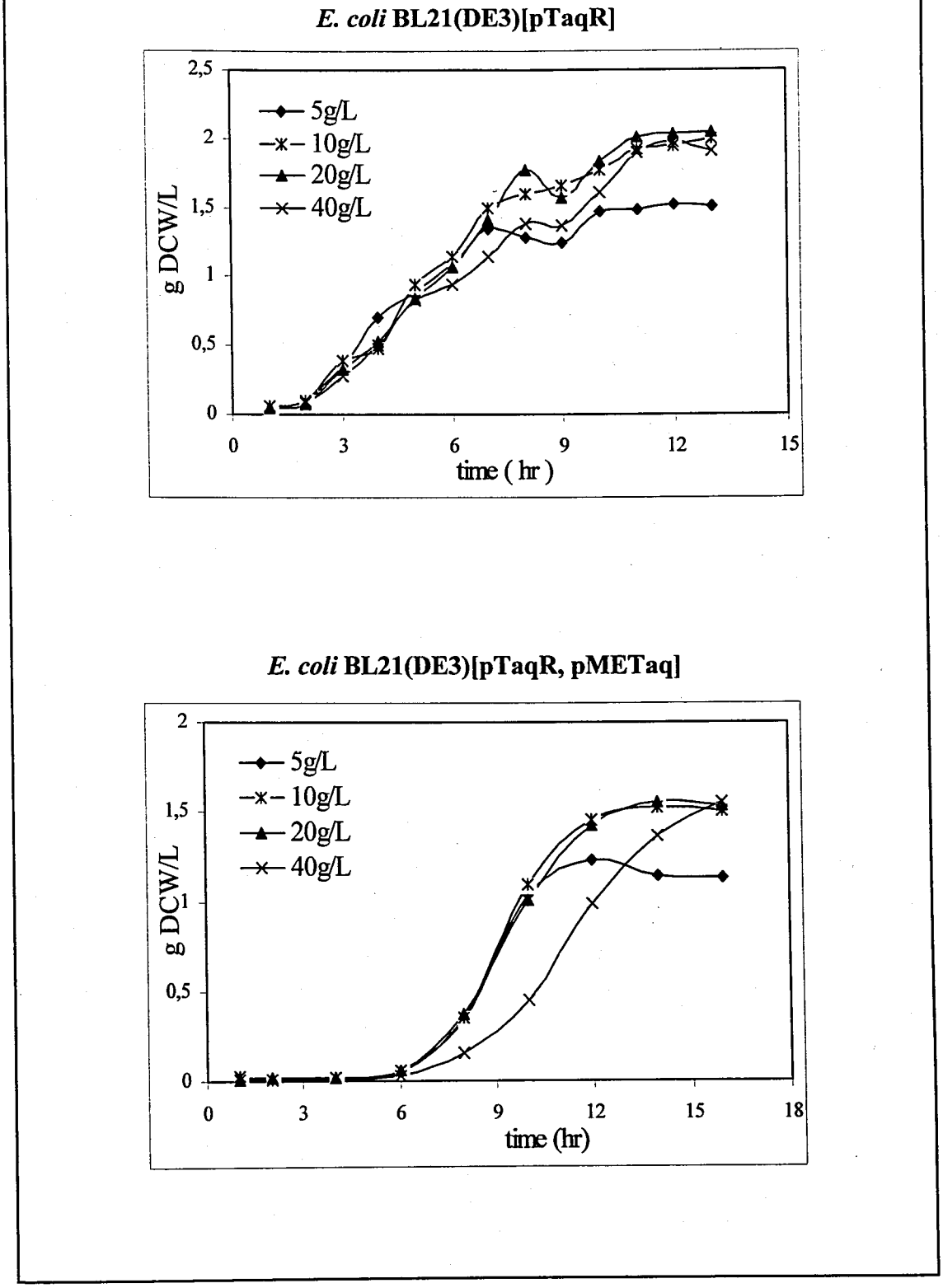


Figure 5.30. Time course of biomass formation of *E. coli* BL21(DE3)[pTaqR] and *E. coli* BL21(DE3)[pTaqR, pMETaq] cells on defined media containing glucose.

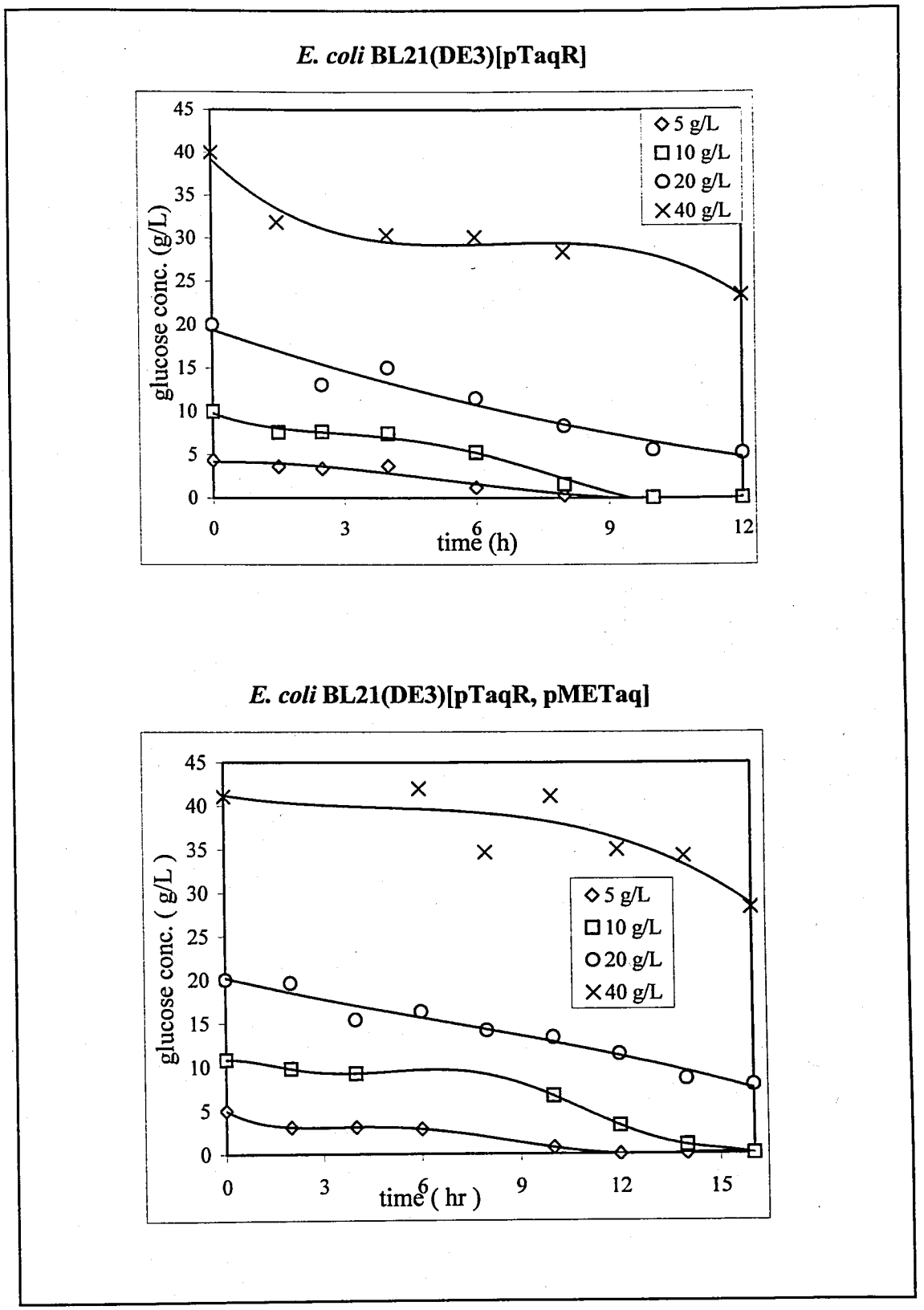


Figure 5.31. Time course of glucose utilization of *E. coli* BL21(DE3)[pTaqR] and *E. coli* BL21(DE3)[pTaqR, pMETaq] cells on defined media containing glucose.

The glucose was completely exhausted in the media containing 5 and 10 g/L initial glucose for both cultures. When the glucose concentration was increased to 20 g/L and 40 g/L, 26 and 59 per cent of the initial glucose remained unconsumed by the *E. coli* BL21(DE3)[pTaqR] cells. Similarly, with the *E. coli* BL21(DE3)[pTaqR, pMETaq] cells, 39 and 71 per cent of the 20 g/L and 40 g/L initial glucose concentrations remained unconsumed in the culture medium, respectively.

Highest μ_{max} values were obtained in the medium supplemented with 5g/L glucose for both strains. However, cells entered their stationary phase when all the glucose was depleted and hence the final biomass yields were lower than the other cases indicating that 5g/L glucose as carbon source was insufficient to support bacterial growth. The maximum biomass yields of the *E. coli* BL21(DE3)[pTaqR] cells in media containing 10g/L, 20g/L and 40 g/L glucose were very close to each other which suggests that the threshold value for these cells is between 10g/L and 20g/L above which the cells cannot grow and utilize glucose further. For the *E. coli* BL21(DE3)[pTaqR, pMETaq] cells, this threshold value seems to be 20g/L initial glucose since a slight increase in final dry cell mass was observed when compared with 10g/L initial glucose concentration. Retarded growth and a 17 per cent decrease in μ_{max} of the *E. coli* BL21(DE3)[pTaqR, pMETaq] cells in defined medium with 40g/L glucose concentration suggest that the excess substrate in the medium inhibited the growth rate of the cells as also reported by Holms (1986).

5.2.9. Effect of Initial Glucose Concentration on *Taq* I Endonuclease Activity

In order to determine the effect of initial glucose concentration on recombinant protein production, *E. coli* BL21(DE3)[pTaqR] and *E. coli* BL21(DE3)[pTaqR, pMETaq] cells were grown in 200 ml selective defined medium containing 5, 10, 20 and 40 g/L glucose in GFL 3032 orbital shaker at 37°C and at a rotational speed of 180 rpm and induced at late- and mid-exponential phases of their growth with 1mM IPTG. After three hours of induction, cells were harvested by centrifugation and *Taq* I endonuclease activity in the crude cell extracts was determined as described in Methods 4.9.

Figure 5.32 shows the specific *Taq* I endonuclease activities recovered from both cultures at different initial glucose concentrations.

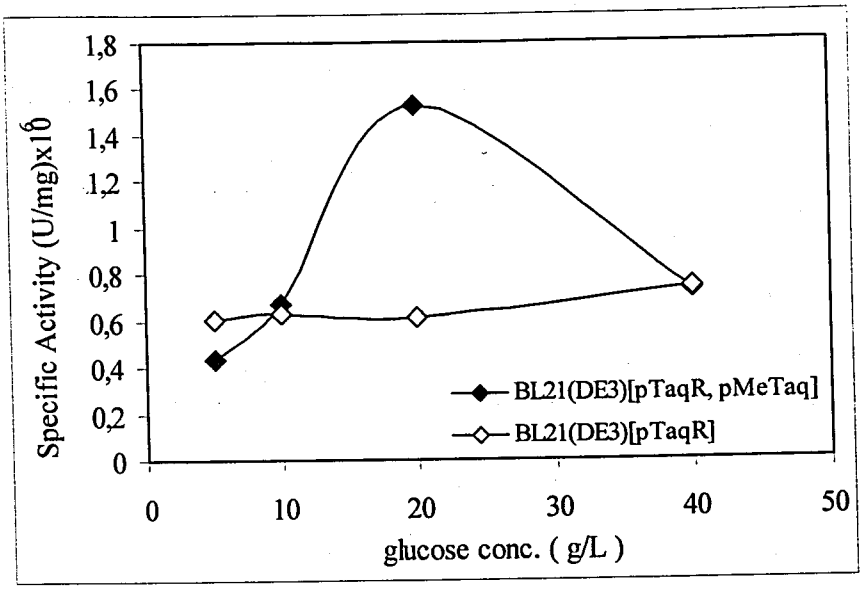


Figure 5.32. Specific *Taq* I endonuclease activity yields of *E. coli* BL21(DE3)[pTaqR] and *E. coli* BL21(DE3)[pTaqR, pMETaq] cells in defined media containing glucose.

Specific *Taq* I endonuclease activities recovered from *E. coli* BL21(DE3)[pTaqR] cells were almost constant and a maximum yield of 0.74×10^6 U/mg was observed at 40 g/L glucose concentration. One can deduce that initial glucose concentration of the defined medium did not have any considerable effect on final product yield for these cells. However, in the *E. coli* BL21(DE3)[pTaqR, pMETaq] cells, specific enzyme activity increased with increasing glucose concentration and maximum specific activity of 1.52×10^6 U/mg was observed at 20g/L glucose concentration which was two times higher than that of the unprotected cells under similar conditions. Increasing the initial glucose concentration to 40g/L resulted in a 50 per cent decrease in specific activity which indicated that the inhibitory effect of the excess substrate concentration was not only reflected in cell growth but also in recombinant protein production. Li *et al.* (1998) have also reported that glucose concentrations higher than 20g/L inhibited growth and glutathione biosynthesis in chemically defined media.

From these results, it may be concluded that for both single- and binary-plasmid systems, the defined medium containing 20g/L glucose as carbon source would be optimum to reach high yields both in the production of enzyme and biomass.

The complex LB medium and the defined medium were also compared in terms of the growth and *Taq* I endonuclease productivities of the two expression systems (Table 5.7).

Table 5.7. Effect of growth medium on growth and *Taq* I endonuclease production.

E. coli Strain	BL21(DE3)[pTaqR]		BL21(DE3)[pTaqR, pMeTaq]	
Medium	LB	Defined+20g/L glucose	LB	Defined+20g/L glucose
Biomass (g DCW/L)	1.34	2.05	1.04	1.65
μ_{max} (h^{-1})	1.033	0.811	0.747	0.696
Volumetric Activity (U/L) x10^6	100	100	250	125
Specific Productivity (U/g DCW.h) x10^6	67.06	26.5	109.3	43.4

When the biomass yields of the uninduced cultures were compared, growth in defined medium resulted in 47 and 41 per cent increases in dry cell mass per liter culture of single- and binary-plasmid systems, respectively. While the maximum specific growth rate, μ_{max}, of the unprotected cells in the complex LB medium was 21 per cent higher than that in the defined medium, only seven per cent decrease in μ_{max} was observed in methylase protected cells when grown in defined medium. However, it took *E. coli*BL21(DE3)[pTaqR, pMeTaq] cells six and four hours to adapt themselves to the defined and complex culture conditions, which was much longer than two and one hours of *lag* times observed for the unprotected cells in defined and complex media, respectively.

Rich media can sustain higher growth rates when compared with defined media since the growth rate of *E. coli* cells in defined medium supplemented with an organic carbon source depends on the concentration of the carbon substrate which serves as the source of intermediates, reducing equivalents and energy. On the other hand, cells grown in rich media do not depend solely on the sugar supplement for carbon skeletons, since many of the building blocks are provided by the rich media (Aristidou *et al.*, 1999). In accordance with this information, the lower specific growth rates and prolonged adaptation times in the defined medium could be attributed to the differences in nutrient limitations of the two medium formulations. Despite the lower growth rates, higher biomass yields could be obtained with the defined media, which is an important factor in maximizing the volumetric productivity of a recombinant protein.

Equal volumetric activity yields were recovered from *E. coli* BL21(DE3)[pTaqR] cells grown in both synthetic and complex media. Changing the complex LB medium with the chemically defined medium brought about 50 per cent lower volumetric activity in *E. coli*BL21(DE3)[pTaqR, pMeTaq] cells. Onset from the rich medium conditions to defined conditions resulted in a 60 per cent decrease in specific *Taq* I endonuclease productivity in both systems. These lower productivity yields recovered from the cells grown in defined media were expected since LB medium was composed of highly rich nutrients such as tryptone and yeast extract. However, in defined medium, cells undergo severe nutrient limitations and lack sufficient cellular resources to quickly synthesize the recombinant protein after induction. Therefore three hours of induction might not be sufficient for these cells to reach high expression levels.

Data presented in Table 5.7 for the complex LB medium was based on cumulative results from several shake flask experiments conducted to find the optimum fermentation conditions. However in defined medium experiments, only the carbon-limited balanced growth, substrate utilization and *Taq* I endonuclease production at different initial glucose concentrations were analyzed. Further shake flask experiments in synthetic medium should also be conducted in order to find optimum induction conditions leading to higher protein productivities.

5.2.10. Bioreactor Experiments with Defined Medium

In order to analyze the two expression systems under controlled conditions, batch fermentation experiments were conducted in a 2.5 L BioFlo III bench-top fermentor. To achieve well-defined culture conditions, chemically defined medium was used as the growth medium. For the carbon-limited balanced growth of the cells, defined medium was supplemented with 20g/L glucose, which was found to be the optimum initial carbon source concentration from the preliminary shake flask experiments.

The fermentor was sterilized with a solution prepared from the heatstable compounds of the defined medium, namely citric acid, KH_2PO_4 , $(NH_4)_2HPO_4$ and casamino acids. Glucose and $MgSO_4$.$7H_2O$ were sterilized separately and added to the fermentor vessel together with the filter-sterilized trace element solution, thiamine, ampicillin and/or kanamycin to a final working volume of 1.5 liter. Calibration of the pH and dissolved oxygen probes were made as described in the Methods section. The temperature was monitored and controlled at 37°C. The pH was measured by glass electrode and controlled at 7.0 through the addition of either 5N H_2SO_4 or 5N NaOH. Dissolved oxygen (DO) was monitored using a polarographic oxygen electrode and sustained above 30 per cent saturation. Filter sterilized air was supplied to the culture at a constant rate of 1 vvm with an air tube connected to the rotameter at the fermentor inlet. The agitation speed was set at 600 rpm throughout the run. A silicone antifoaming agent was used at 1:10 dilution with sterile dH_2O.

Freshly transformed *E. coli* BL21(DE3) cells were used for the bioreactor experiments. The inoculum was prepared by transferring a single colony from selective LB plates into 100ml defined medium supplemented with 20g/L glucose and growing the cells in an orbital shaker at 37°C and with an agitation rate of 180 rpm overnight. Fermentation was started by adding one per cent (v/v) inoculum to the fermentor and samples were taken throughout the experiments at regular time intervals. Expression of the *Taq*I restriction endonuclease gene was initiated by inducing the cultures with 1mM IPTG within their exponential phase of growth. Time course of biomass formation, glucose utilization, acetic

acid formation, plasmid stability, enzyme activity and cell viability were followed in both systems.

5.2.11. Growth Characteristics

Growth of the cultures was followed by measuring the absorbance at 600nm and also by determining the dry cell mass concentration as described in 4.12.1. Time course of optical density and biomass formation of *E. coli* BL21(DE3)[pTaqR] and *E. coli* BL21(DE3)[pTaqR, pMETaq] cells are shown in Figure 5.33.

After one hour of adaptation period, *E. coli* BL21(DE3)[pTaqR] cells entered the exponential phase of their growth. As soon as the recombinant cells were induced with IPTG, growth almost ceased for two hours and then continued but at a slower rate. On the other hand, *E. coli* BL21(DE3)[pTaqR, pMETaq] cells entered the logarithmic growth phase after three hours of adaptation period. When compared with the unprotected cells, elongation of the *lag* time in shake flask cultures was also observed in the bioreactor conditions due to the constitutive expression of the methylase gene. For one hour after IPTG induction, growth of the methylase protected cells was not affected by the endonuclease gene expression, however, growth inhibition was also observed in these cells in the following three hours. Similar observation was also made by Gupta *et al.* (1999) that when recombinant organisms containing strong promoters, are induced, the specific growth rate declines sharply due to the overload of the host-cell metabolism.

The maximum specific growth rates of the cultures were determined from the slopes of the lnx versus time plots in the exponential phase using linear regression as 0.9361±0.06 h^{-1} and 0.4114±0.008 h^{-1} for the *E. coli* BL21(DE3)[pTaqR] and *E. coli* BL21(DE3)[pTaqR, pMETaq] cells, respectively. The maximum biomass concentrations reached by the induced cultures were almost equal being 6.52 and 6.58 g DCW/L for the unprotected and methylase protected cells, respectively. These values were three and four times higher than the maximum biomass yields obtained from the uninduced shake flask cultures.

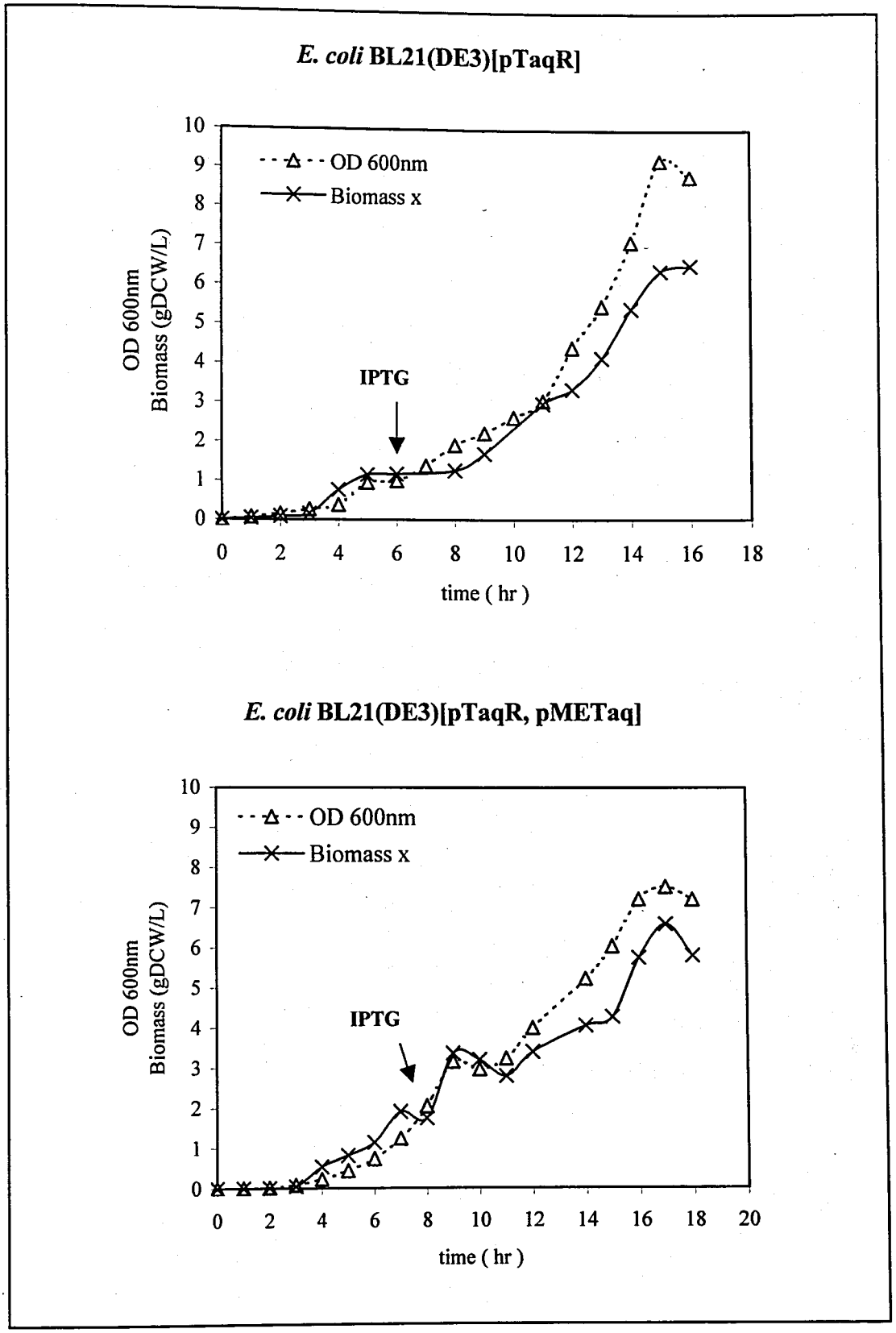


Figure 5.33. Time course of culture optical density and biomass formation of *E. coli* BL21(DE3)[pTaqR] and *E. coli* BL21(DE3)[pTaqR, pMETaq] cells on defined medium.

5.2.12. *Taq* I Endonuclease Activity

In order to determine the *Taq* I endonuclease activities expressed by the two cultures, samples were taken at regular time intervals. Specific (-▲- , -Δ-) and volumetric (-◆- , -◊-) enzyme activities in the crude cell extracts of *E. coli* BL21(DE3)[pTaqR] (open symbols) and *E. coli* BL21(DE3)[pTaqR, pMETaq] (closed symbols) cells were determined as described in the Methods section (Figure 5.34).

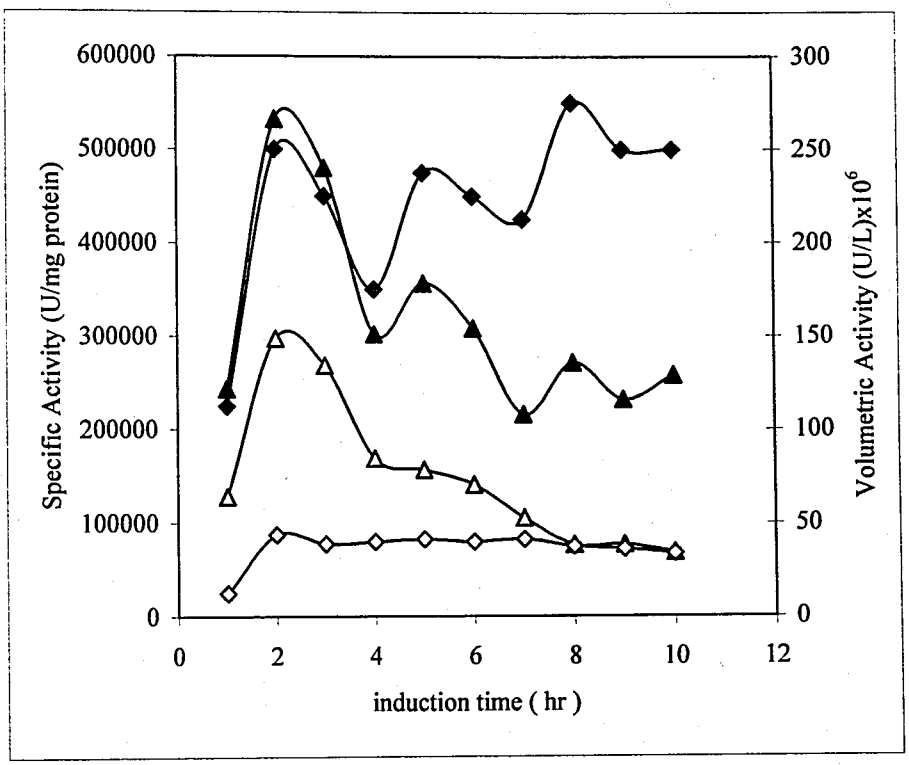


Figure 5.34. Time profiles of specific and volumetric *Taq* I restriction endonuclease activities of *E. coli* BL21(DE3)[pTaqR] and *E. coli* BL21(DE3)[pTaqR, pMETaq] cells.

Volumetric and specific *Taq* I endonuclease activities of the *E. coli* BL21(DE3)[pTaqR] cells reached their maximum value of 43.75×10^6 U/L and 3×10^5 U/mg within the first two hours of induction. During this period, cell growth was also ceased indicating that upon induction, almost all of the cells' resources were used for recombinant protein production, rather than cell growth. After this production period, while the volumetric *Taq* I restriction endonuclease activity of the cells remained constant for the rest of the fermentation, the specific *Taq* I restriction endonuclease activity decreased

rapidly and leveled off at a constant value of about $0.7x10^5$ U/mg. Although the biomass concentration increased three-fold in this time interval, they were probably not producing *Taq* I restriction endonuclease.

Three distinct peaks were observed in the time dependent volumetric and specific activity profiles of the *E. coli* BL21(DE3)[pTaqR, pMETaq] cells. The maximum specific *Taq* I restriction endonuclease activity of $5.3x10^5$ U/mg and a volumetric activity of $250x10^6$ U/L were recovered from the recombinant cells after the first two hours of induction which formed the first peak. Then, in the following two hours, both activities were decreased. Contrary to the unprotected cells, *E. coli* BL21(DE3)[pTaqR, pMETaq] cells started to produce *Taq* I restriction endonuclease forming the second peak. The maximum volumetric activity of $275x10^6$ U/L was obtained after eight hours of induction corresponding to the third peak. When these activity profiles were compared with the dry cell mass concentration profile of the culture, the positions of the peaks fitted well to the time intervals at which the growth of the cells were stopped temporarily. These results were in accordance with the fact that the cells' resources were used for recombinant protein production, rather than cell growth.

In order to analyze these observations on a better basis, the instantaneous rate of change of *Taq* I restriction endonuclease production by the two expression systems were determined. The specific *Taq* I restriction endonuclease production rates (qP) were calculated as Units of *Taq* I endonuclease produced per gram dry cell per unit time (Figure 5.35).

The recombinant protein production rate of the *E. coli* BL21(DE3)[pTaqR] cells reached its maximum value of $24.87x10^6$ Units *Taq* I/gDCW/h after two hours of induction and then decreased. After four hours of induction, rate of *Taq* I endonuclease production by these cells remained at values near zero. However, in the *E. coli* BL21(DE3)[pTaqR, pMETaq] cells, a quicker response to IPTG induction was observed such that the maximum *Taq* I endonuclease production rate of $33.5x10^6$ Units *Taq* I/gDCW/h was reached within the first hour of induction.

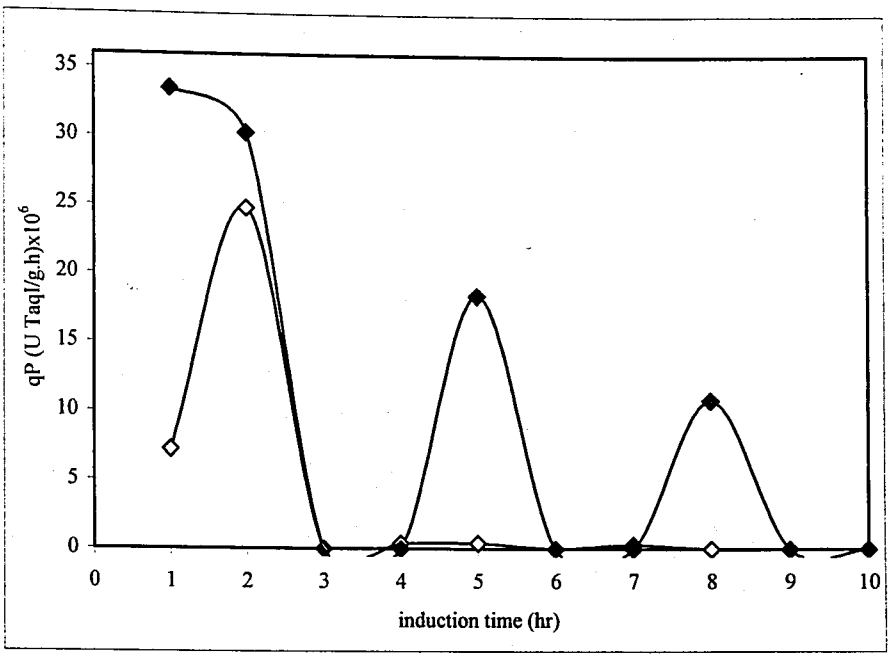


Figure 5.35. Specific *Taq* I production rates of *E. coli* BL21(DE3)[pTaqR] (open symbols) and *E. coli* BL21(DE3)[pTaqR, pMETaq] (closed symbols) cells.

This could be attributed to the fact that since the metabolic activities of these cells were already organized for plasmid encoded recombinant TaqI methylase synthesis, shorter periods were required for these cells to adapt themselves to the high-level *Taq* I endonuclease expression. Again three peaks with decreasing maxima were observed in the qP profile of these cells in accordance with the earlier observations. When compared with the unprotected cells, only a 35 per cent increase in the *Taq* I production rate was observed in the methylase protected cells. This was expected since in both systems, the *Taq* I endonuclease gene was expressed under the control of the T7lac promoter by the same host strain *E. coli* BL21(DE3). In these cells, constitutive expression of the *Taq* I methylase gene under the control of the tetracycline promoter could have protected the cells from the negative effects of cytoplasmic *Taq* I endonuclease accumulation resulting in higher production rates for longer periods.

The two systems were also compared in terms of their specific *Taq* I endonuclease productivities (Figure 5.36).

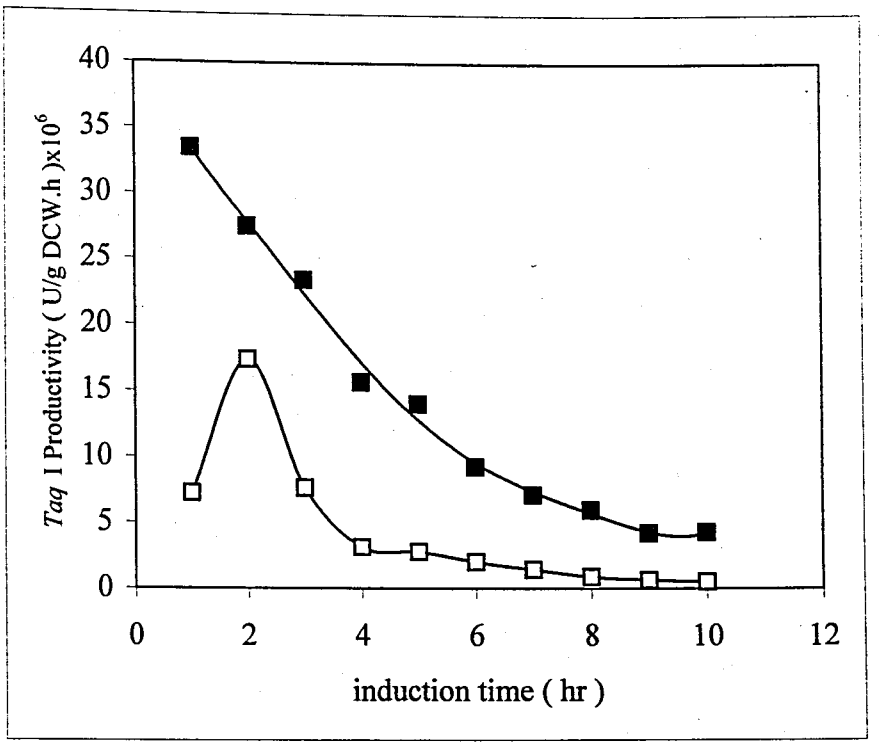


Figure 5.36. Specific productivity profiles of *E. coli* BL21(DE3)[pTaqR] (open symbols) and *E. coli* BL21(DE3)[pTaqR, pMETaq] (closed symbols) cells.

In accordance with the earlier observations, the specific *Taq* I restriction endonuclease productivity of the *E. coli* BL21(DE3)[pTaqR] cells reached its maximum of 17.36×10^6 Units *Taq* I/gDCW/h after two hours of induction and then decreased rapidly in the following two hours to one sixth of its maximum value.

The specific productivity of the *E. coli* BL21(DE3)[pTaqR, pMETaq] cells, on the other hand, reached its maximum value of 33.48×10^6 Units *Taq* I/gDCW/h within the first hour of induction which was about two times higher than that of the unprotected cells. Then, the specific productivity of the cells decreased gradually but at a much slower rate than the unprotected cells. After ten hours of induction, the specific productivity of the methylase protected cells was 4.31×10^6 Units *Taq* I/gDCW/h, 8.5 times higher than that of the unprotected cells.

Both in complex LB and chemically defined medium, two times higher specific *Taq* I productivities could be obtained from *E. coli* BL21(DE3)[pTaqR, pMETaq] cells. As also

observed in the shake flask cultures in complex medium, cells could not maintain their productivity for longer periods under controlled bioreactor conditions.

5.2.13. Cell Viability

A common objective of many fermentation processes is to maximize the final product concentration, which is the mathematical product of the specific production rate and total viable cell concentration. It follows that increases in either of these variables will increase final product concentration (Sauer *et al.*, 2000). Therefore, in order to analyze the cell viabilities in terms of colony forming units (CFU) of the two expression systems, culture samples taken during the course of fermentation were diluted with defined medium, spread on LB plates and after overnight incubation at 37°C, colonies formed on each plate were counted (Figure 5.37).

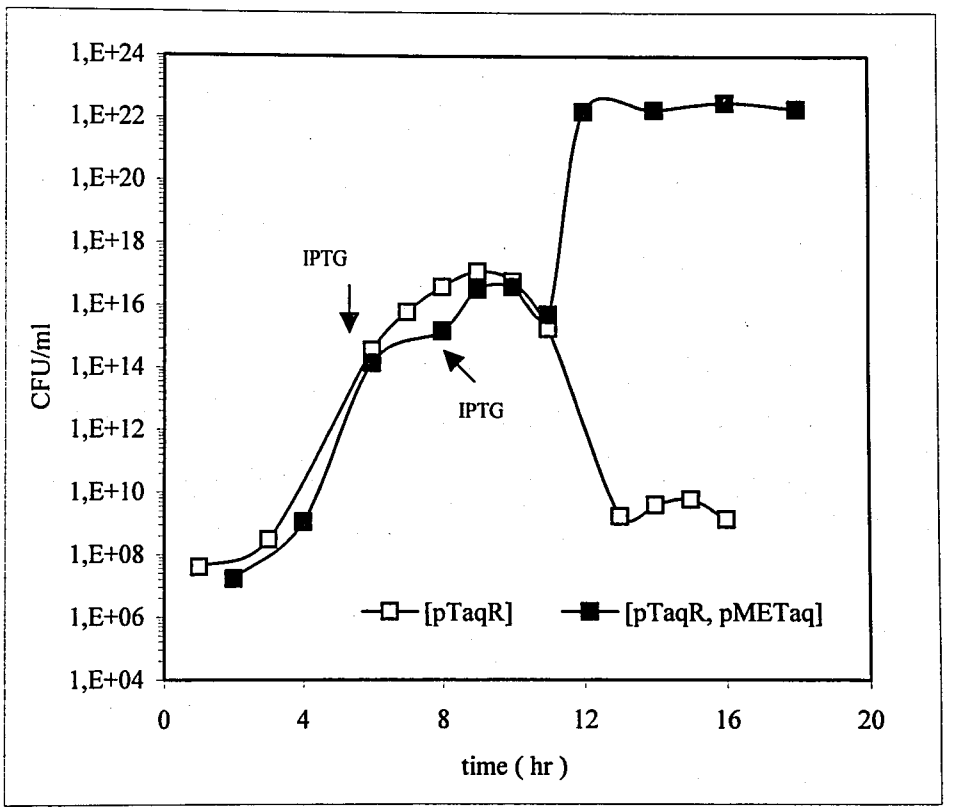


Figure 5.37. Cell viability profiles of *E. coli* BL21(DE3)[pTaqR] and *E. coli* BL21(DE3)[pTaqR, pMETaq] cells.

Cell viability of both cultures increased at the same rate during the pre-induction growth phase. After the addition of IPTG, no significant increase in cell viability was observed for a certain period hence the growth inhibition observed from the biomass concentration profiles was also reflected in the cells' colony forming ability. After this period, major differences occurred in cell viability of the two cultures.

After four hours of induction, a significant drop in viable cell concentration occurred in *E. coli* BL21(DE3)[pTaqR] culture which could be attributed to the morphological changes of the cells. Such a decrease was also reported by Cserjan-Puschmann *et al.* (1999) after seven hours of induction. Dry cell mass concentration, on the other hand, increased steadily during the same period. Also a disparity among the profiles of optical density and dry cell mass concentration was observed towards the end of the fermentation (Figure 5.33 a). All these observations were consistent with the results reported previously concerning the inability of host cells to divide and the concomitant elongation of the cells in the presence of excessive plasmid DNA synthesis and overexpression of foreign proteins (Flickinger and Rouse, 1993, Ryan *et al.*, 1996). As also suggested by Ryan *et al* (1996), due to the limitations of the OD measurements, one should always include auxiliary measurements such as viable cell count and dry cell weight, as part of the in-process monitoring in order to analyze the growth characteristics of the recombinant cells more properly.

On the other hand, the cell viability of the methylase protected cells increased and reached a steady-state value of about $3x10^{22}$ CFU/ml. Co-expression of the *Taq* I methylase gene promoted the viability and also the specific productivity of the recombinant cells.

5.2.14. Plasmid Stability

Stability of the pTaqR plasmid in the two expression systems were determined as described in the Methods section. As shown in Figure 5.38, this plasmid containing the *Taq* I endonuclease gene was highly stable during the first four hours of induction.

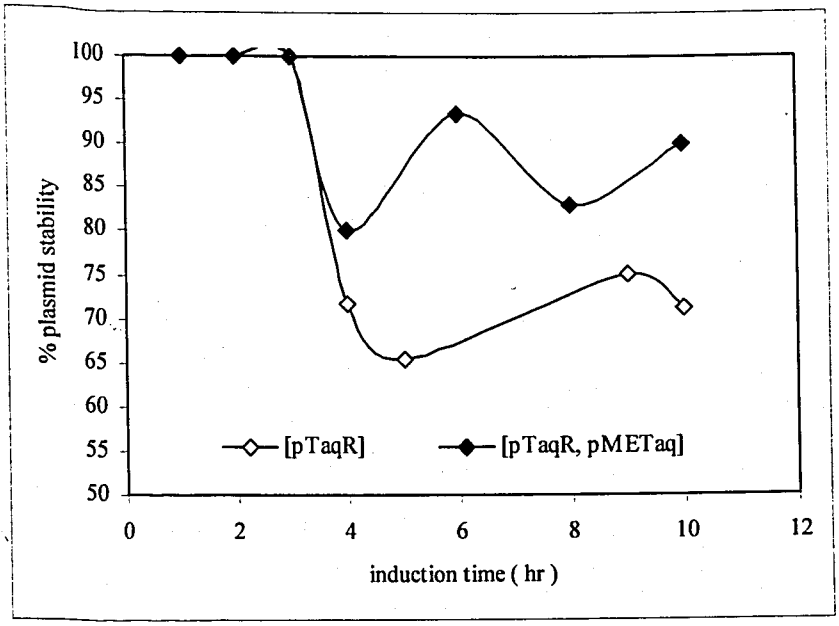


Figure 5.38. Plasmid stability of *E. coli* BL21(DE3)[pTaqR] and *E. coli* BL21(DE3)[pTaqR, pMETaq] cells.

After this period, its stability decreased to about 70 per cent in *E. coli* BL21(DE3)[pTaqR] cells for the rest of the fermentation. However, higher plasmid stabilities were observed in the methylase protected cells such that 85-90 per cent of the *E. coli* BL21(DE3)[pTaqR, pMETaq] colonies were found to retain this plasmid even after ten hours of induction. The presence of the second plasmid pMETaq containing the *Taq* I methylase gene was also checked by spreading diluted the *E. coli* BL21(DE3)[pTaqR, pMETaq] culture samples taken from the fermentor on LB plates containing both kanamycin and ampicillin. The number of colonies containing both plasmids were found to be 4.11×10^{11} CFU/ml at the end of the fermentation, which was 100 times higher than the viable cell count of the single-plasmid cells at the same time instant.

5.2.15. Glucose Utilization and Acetic Acid Excretion

Synthetic processes require the metabolic energy released by oxidation of primary carbon sources and sugar utilization is generally taken as an indication of the rate of energy release to the system. While proteins and fats are similarly degraded, with accompanying energy release, carbohydrate sources are ordinarily the major energy suppliers. At the same

time these materials are frequently the substrates from which specific products are formed (Gaden, 2000).

In this work, glucose was added to the defined medium as the primary carbon and energy source. However, one of the major problems preventing the achievement of high yield and high volumetric productivity during fermentation of *Escherichia coli* is the conversion of a significant fraction of the commonly used carbon source glucose, into acidic by-products, especially acetic acid. The accumulation of acetic acid above a certain level inhibits cell growth and therefore limits the maximum attainable biomass densities. Furthermore, acetic acid has been found to impede the cellular machinery responsible for the expression of recombinant proteins (Aristidou, 1999). Therefore, to analyze the time dependent glucose consumption and acetic acid accumulation profiles of the two cultures, extracellular concentration of these two chemicals were determined enzymatically by using commercially available kits as described in Methods section.

Figure 5.39 shows the glucose (-■- , -□-) and acetic acid (-▲- , -Δ-) concentrations of *E. coli* BL21(DE3)[pTaqR] (open symbols) and *E. coli* BL21(DE3)[pTaqR, pMETaq] (closed symbols) cells. Utilization of glucose by the *E. coli* BL21(DE3)[pTaqR] cells started at the sixth hour of fermentation, when *Taq* I endonuclease production was initiated by the addition of IPTG. After 15 hours, all of the glucose in the medium was depleted. Glucose was consumed by the *E. coli* BL21(DE3)[pTaqR, pMETaq] cells from the start of the growth most probably as an energy supply for the constitutive expression of the *Taq* I methylase gene. When *Taq* I endonuclease production was induced in these cultures after eight hours, glucose was utilized at a higher rate and after 17 hours, all of the glucose was used up.

During the period of glucose consumption, acetic acid was excreted reaching a level of about 2.59 g/L and 1.15 g/L for the *E. coli* BL21(DE3)[pTaqR] and *E. coli* BL21(DE3)[pTaqR, pMETaq] cultures, respectively. In unprotected cells, the maximum acetic acid concentration was two times higher than the methylase protected cells, however, these values were far below the growth-inhibitory concentration of 5 g/L reported by Lee (1996). *E. coli* BL21(DE3) cells were found to posses an acetate self-control

mechanism which enables the cells to maintain low acetate concentrations in its growth medium, regardless of the external glucose concentration as reported by Shiloach *et al.* (1996).

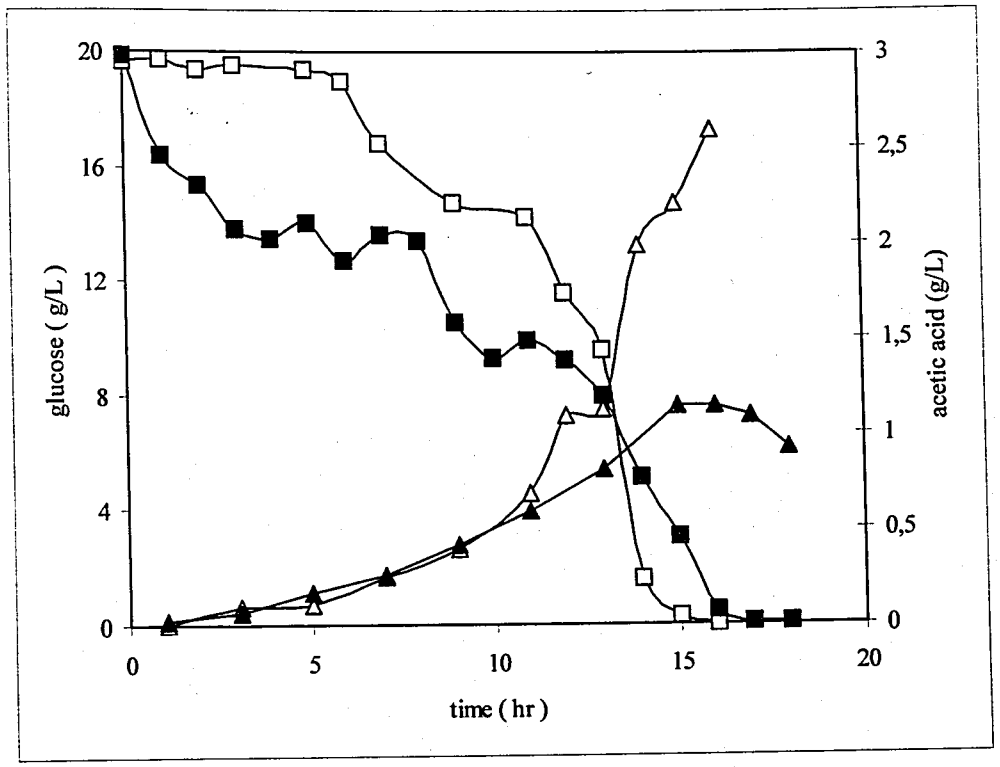


Figure 5.39. Glucose and acetic acid concentrations of *E. coli* BL21(DE3)[pTaqR] and *E. coli* BL21(DE3)[pTaqR, pMETaq] cells.

Glucose, acting as a catabolic repressor, is known to totally repress the glyoxylate shunt, whose unique inducible enzymes isocitrate lyase and malate synthase allow the cell to utilize acetic acid as carbon source (Clark, 1981). The inability of the cells to utilize acetic acid in the presence of glucose was apparent in both systems. When the glucose was totally depleted, *E. coli* BL21(DE3)[pTaqR, pMETaq] cells used acetic acid as the carbon source resulting in a decrease in acetic acid concentration towards the end of the fermentation.

5.2.16. The Overall Yield Factors

It has been frequently observed that the total amount of the cell mass formed by cell growth is proportional to the mass of substrate (typically carbon sources, energy source and oxygen) utilized. The yield factors are sometimes used to characterize system performance. The overall yield factors for biomass formation per substrate consumption, Yx/s (g cells formed per glucose consumed), the product formation per biomass production, Yp/x (Units of *Taq* I endonuclease formed per g cells formed) and product formation per substrate consumption, Yp/s (Units of *Taq* I endonuclease formed per glucose consumed) are given in Table 5.8 for both systems.

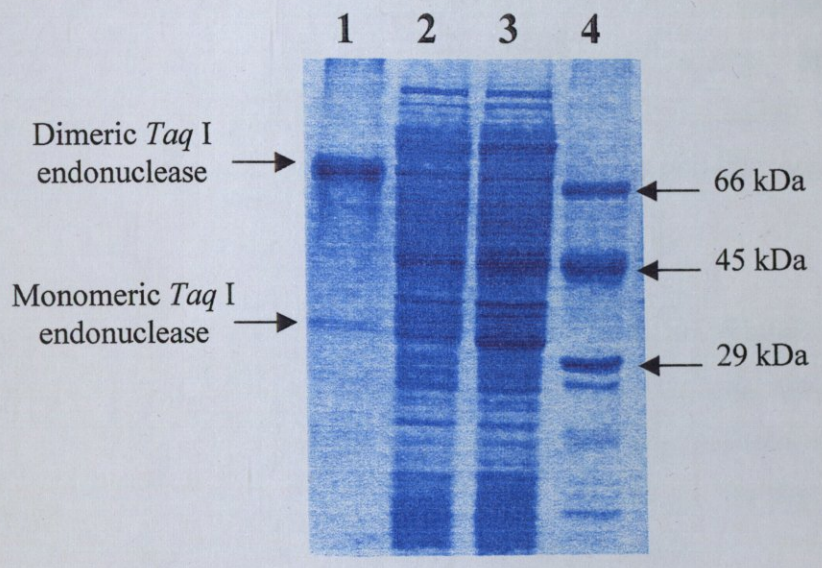
Table 5.8. Overall yield coefficients of the two expression systems.

	E. coli BL21(DE3) [pTaqR]	*E. coli* BL21(DE3) [pTaqR, pMETaq]
Yx/s (g cells/g glucose)	0.326	0.332
Yp/x (U *Taq* I/g cells)x10^6	34.72	47.83
Yp/s (U *Taq* I/g glucose) x10^6	8.79	14.21

The ultimate goal in using the pET expression system was to recover *Taq* I endonuclease in high yield from recombinant cells. The biomass yields on substrate glucose (Yx/s) were very close since both systems utilized all of the 20g/L initial glucose at different rates and reached equal dry cell mass concentrations. Yp/x of the methylase protected cells was 40 per cent higher since elevated *Taq* I endonuclease production rates could be maintained by these cells over prolonged induction times. When the Yp/s values are compared, per gram glucose consumed, 1.6 times higher *Taq* I endonuclease was produced by *E. coli* BL21(DE3)[pTaqR, pMETaq] cells than by the unprotected single-plasmid cells.

5.2.17. SDS-PAGE Analysis

The molecular masses of the full length *Taq* I restriction endonuclease and *Taq* I methylase enzymes has been reported to be 31600 Da and 47900 Da, respectively (Barany, 1992). Samples taken from the fermentor were also analyzed by SDS-PAGE on 12 per cent polyacrylamide gels to confirm the presence of these two enzymes in the crude cellular extracts (Figure 5.40).



Lane 1: Commercial *Taq*I endonuclease.

Lane 2: *E. coli* BL21(DE3)[pTaqR] cell extract (10 µg).

Lane 3: *E. coli* BL21(DE3)[pTaqR, pMETaq] cell extract (10 µg).

Lane 4: Molecular weight markers.

Figure 5.40. SDS-PAGE analysis of the samples taken from the fermentor.

The band corresponding to the monomeric *Taq* I endonuclease was apparent in both samples but at a higher intensity in the crude extract of *E. coli* BL21(DE3)[pTaqR, pMETaq] cells after eight hours of induction due to its elevated levels of expression compared to the unprotected cells (Lane 3). Also in the crude extract of these cells, the band corresponding to the *Taq* I methylase enzyme was at its expected position.

6. CONCLUSIONS AND RECOMMENDATIONS

6.1. Conclusions

Taq I restriction endonuclease is a thermostable enzyme with high specificity for its palindromic tetranucleotide recognition sequence 5'TCGA3'. Its extensive use in many molecular biology and biotechnology applications results in the consumption of high amounts of this enzyme. The low expression levels *Taq* I restriction endonuclease makes its purification from its native host *Thermus aquaticus* very complicated and uneconomical. Therefore, high-level expression of this enzyme in *Escherichia coli* is considered to be inevitable to achieve high productivity via cost-effective downstream processing.

The major aim of this study was to clone and express the *Taq* I restriction-modification system in *E. coli*. Two expression systems were developed and the growth, fermentation and *Taq* I endonuclease production characteristics of recombinant *E. coli* cultures were analyzed under various growth conditions. On the basis of the presented results, several conclusions may be drawn for each system.

6.1.1. pMAL Protein Fusion System

The recombinant plasmid construct pH185 contained the gene encoding *Taq* I restriction endonuclease that was genetically fused to the *mal*E gene of *E. coli* which encodes maltose binding protein (MBP) (Özdinler, 1996). In the recombinant *E. coli* cells, harboring pH185 plasmid, MBP-*Taq* I fusion protein was found to be synthesized in the cytoplasm, successfully transported to the periplasm by the MBP signal sequences and released into the medium.

An optimum induction period of 18 hours was found to be essential for the production and excretion of *Taq* I endonuclease activity by recombinant *E. coli* cells. Longer induction periods leading to the accumulation of the recombinant protein seem to

make the cells increasingly leaky during the stationary phase and facilitate the release of proteins targeted into the periplasm.

The production and excretion of the *Taq* I endonuclease activity directed to the periplasm by MBP signal of *E. coli* was found to depend both on the host strain and on the phase of growth at which the inducer was added to the medium. Total and extracellular activities recovered from cultures induced at early phases of growth was found to be considerably low when compared with those induced at later phases of growth. When compared with the *E. coli* TB1 and the protease deficient ER2508 strains, highest total *Taq* I endonuclease activity ($1.65x10^5$ U/g DCW) was recovered from *E. coli* XL1 (pH185) cells induced at their late-exponential phase of growth for 18 hours and 81 per cent of this activity was excreted to the growth medium. Also the harmful effect of protein production on the viability of the *E. coli* XL1 (pH185) cells was inferior to the other two strains.

The absence of cellular nucleic acids and the intracellular marker isocitrate dehydrogenase in the medium indicated that the release of the *Taq* I endonuclease activity was not associated with the cell lysis.

Using the specific interaction of MBP tail with amylose, MBP-*Taq* I fusion protein was purified from cytoplasmic, periplasmic and extracellular fractions of the *E. coli* XL1 (pH185) and *E. coli* ER2508 (pH185) cells by a single-step affinity chromatography. The MBP tail of the chimeric protein was removed by enzymatic digestion with Factor Xa protease resulting in the production of the native *Taq* I restriction endonuclease. Furthermore, quality control tests indicated that the purified fusion protein, as well as its digestion product, could replace the molecular biology grade *Taq* I endonuclease available in the market.

Low levels of active fraction, which displayed an affinity to amylose, indicated that MBP part of the chimeric protein was cleaved during the translocation through inner- and outer cytoplasmic membrane of *E. coli* cells. Accumulation of the MBP within the periplasmic space indicated that proteases residing in the membrane or periplasmic compartment were possibly responsible for the cleavage of the MBP portion of the secreted chimeric protein. Therefore, the one-step purification of the enzyme by affinity

chromatography from the medium may not be carried out with high yield. However it should be noted that extracellular production of the enzyme presents an advantage for downstream processing.

The presence of glycine up to 0.7 per cent in the medium resulted in an increase in the fraction *Taq* I endonuclease released into the culture medium. In fact, glycine has been shown to improve the rate and efficiency with which the bacterial system handles the recombinant protein and promote the release of the various proteins from the periplasmic space into the culture medium (Aristidou *et al.*, 1993; Yu and San, 1993). However, the presence of the glycine, in the present study, did not increase the enzyme production as reported. Increasing concentrations of glycine had a detrimental effect on both cell growth and enzyme production.

The effect of additional yeast extract, maltose, sucrose and sorbitol in the medium on the excretion of MBP-*Taq* I fusion protein was also investigated. The percentage of the extracellular fraction reached up to 98.4 per cent by the addition of 2.5 g/L yeast extract. Dramatic decreases in the level of the total enzyme production was observed in all cases and considerable improvement of the cell growth by the addition of sorbitol was not reflected in *Taq* I enzyme production and very low levels of endonuclease activity could be recovered from the medium.

In order to protect the expression plasmid pH185 and host genomic DNA from the detrimental effects of *Taq* I endonuclease expression, the *Taq* I methylase gene was amplified from the *Thermus aquaticus* YT1 genome by PCR using specific primers and cloned under the promoter of the tetracycline resistance gene of the pBR322 vector. To have a selection marker other than ampicillin, pETMET plasmid was constructed by subcloning the *Taq* I methylase gene together with the tetracycline promoter region into pET28a+ vector, which contained a kanamycin resistance gene.

Co-expression of the *Taq* I methylase gene in *E. coli* XL1(pH185, pETMET) cells resulted in about eight times lower total and extracellular volumetric *Taq* I endonuclease activities. However, similar excretion levels were obtained and the extracellular enzyme activity could be maintained for much longer induction periods when compared with the

unprotected cells. In addition, maximum specific activity isolated from the osmotic shock fluid after 24 hours of induction was two times higher than that of the single vector system.

When the growth behaviour of *E. coli* XL-1(pH185,pETMET) cells in LB medium was compared with the single-plasmid system, 30 per cent decrease in maximum specific growth rate, lower biomass yields and longer adaptation times (t_{lag}) were observed most probably due to the nutrient limitations of the LB medium. Although the μ_{max} values of cultures grown in TB and 2xYT media were 30-40 per cent higher than that was found in the LB medium, their adaptation times were almost doubled at the same conditions and SB medium resulted in the highest biomass yields.

Growing the recombinant cells in SB medium not only improved the growth characteristics, but also total and extracellular *Taq* I endonuclease recovery yields were increased without any cell lysis. Therefore, SB medium was used for the bioreactor studies.

When the growth characteristics of the two systems were compared under controlled fermentor conditions, the unprotected *E. coli* XL1(pH185) cells reached higher biomass concentrations throughout the course of fermentation with shorter adaptation period, when compared with the methylase protected cells. The maximum biomass concentrations reached by the induced cultures were 13.09 and 10.2 g DCW/L for the unprotected and methylase protected cells, respectively. Three times higher biomass yields could be obtained from *E. coli* XL1(pH185, pETMET) cells grown under controlled fermentor conditions when compared with the shake flask cultures.

Onset of MBP-*Taq* I fusion protein expression resulted in a decrease in biomass concentrations and oxygen consumption levels in both cultures, which indicated that the onset of gene expression forced the cells to direct their metabolic activities from cellular growth towards recombinant protein production. Since *Taq* I methylase was produced constitutively during the growth of *E. coli* XL1(pH185, pETMET) cells, their metabolic activities were accustomed to plasmid encoded recombinant protein synthesis. Therefore, expression of MBP-*Taq* I fusion protein did not affect these cells as much as the unprotected cells.

Both cultures consumed the 20 g/L initial glucose added at the start of the fermentation after 9-11 hours of fermentation in both cultures with comparable utilization trends and the extracellular acetic acid concentrations were at nontoxic levels.

The percentage of plasmid containing colonies in the *E. coli* XL1(pH185) culture decreased soon after the induction of cells which hindered the secretion of the MBP-*Taq* I fusion protein to the periplasmic space and hence resulted in its accumulation in the cytoplasm. However, in the binary plasmid system, where cellular DNA was protected from endonuclease attack by metylation activity, the plasmid remained highly stable for longer periods. Also the cytoplasmic accumulation of the fusion protein was prevented and its periplasmic secretion was highly improved such that more than three times higher specific *Taq* I endonuclease activity could be isolated from the osmotic shock extracts.

In *E. coli* XL1(pH185) cells, most of the *Taq* I endonuclease activity remained trapped in the cytoplasm and periplasm of *E. coli* XL1 cells and only a small fraction (2.7 per cent) leaked into the culture medium. However, although ten times lower volumetric activities could be obtained from the methylase protected cells, about 85 per cent of this activity was excreted to the growth medium after 40 hours of induction without cell lysis. 0.6×10^6 U/L *Taq* I endonuclease activity could be recovered from the extracellular medium of *E. coli* XL1(pH185, pETMET) cells under controlled bioreactor conditions, which was 60 and two times higher than the extracellular activity obtained from the shake flask cultures of protected cells in SB medium and unprotected cells in LB medium, respectively.

6.1.2. pET Expression System

The *Taq* I restriction endonuclease gene was amplified from the *Thermus aquaticus* bacterial genome by PCR and cloned into the pET28a+ expression plasmid under the control of the strong bacteriophage T7 RNA polymerase promoter. High levels of *Taq* I endonuclease activity was recovered from *E. coli* BL21(DE3) cells harboring the recombinant construct pTaqR. *E. coli* BL21(DE3) cells transformed with both pETaqR and pMETaq plasmids were used in order to investigate the effect of protective methylation on enzyme productivity both in shake flasks and under controlled bioreactor conditions.

When the growth characteristics of the host, unprotected and methylase protected cells were analyzed in shake flask cultures containing complex LB medium, elongated adaptation periods due to the constitutive expression of the *Taq* I methylase gene was observed. The biomass yield of *E. coli*BL21(DE3)[pTaqR, pMetaq] cells was much lower (1.04 g/L) when compared with 1.50 and 1.34 g/L biomass yields obtained from the host and cells harboring pTaqR plasmid only. The uncontrolled expression of the *Taq* I endonuclease gene in uninduced *E. coli* BL21(DE3)[pTaqR] cells was not toxic to the host cell at detectable levels, and the growth was associated with slightly lower growth rate when compared to that of the host cells.

The highest specific and volumetric *Taq* I endonuclease activities were obtained from *E. coli* BL21(DE3)[pTaqR] and *E. coli* BL21(DE3)[pTaqR, pMetaq] cells when they were induced at their late- and mid-exponential phase of growth, respectively. Co-expression of *Taq* I methylase gene in *E. coli* BL21(DE3)[pTaqR, pMETaq] cells resulted in the production of at least two times or higher enzyme activities when compared to the unprotected cells. However, specific and volumetric activities of the enzyme and the specific productivity were found to decrease with the elongated induction periods in both cultures, most probably due to the overloaded host cell metabolism.

Delaying the induction time and prolonged induction periods had deleterious effects on the specific and volumetric *Taq* I endonuclease activity yields offsetting any gain achieved by the increase in cell biomass. This result was possibly due to nutrient starvation in the culture which lead to a lowering of the specific growth rate and hence a concomitant lowering of rates of *Taq* I endonuclease production.

Co-expression of the methylase gene in the *E. coli* BL21(DE3)[pTaqR, pMETaq] cells resulted in a 2.5-fold increase in volumetric productivity yielding $250x10^6$ U/L *Taq* I endonuclease activity which was 28 times higher than the reported data (Barany, 1988). When compared with the *lac* promoter, T7 promoter resulted in 100 times higher volumetric activities.

The growth characteristics and enzyme productivity of both cultures were also analyzed in chemically defined medium containing increasing amounts of glucose as the

sole carbon source. The adaptation period of the *E. coli* BL21(DE3)[pTaqR, pMETaq] cells was two hours longer than that observed in the complex LB medium due to the limited nutrient sources in the defined medium. Highest μ_{max} values were obtained in the medium supplemented with 5g/L glucose for both strains. Retarded growth and decreased μ_{max} values were observed in both cultures growing in defined medium supplemented with 40g/L glucose which suggested that the excess substrate in the medium inhibited the growth of these cells.

In both cultures, the glucose was completely exhausted in the media containing 5 and 10 g/L initial glucose whereas unconsumed glucose was detected in cultures growing in defined media supplemented with 20 g/L and 40 g/L glucose. Onset from the rich medium conditions to defined conditions resulted in a 60 per cent decrease in specific *Taq* I endonuclease productivity in both systems.

For both single- and binary-plasmid systems, the defined medium containing 20g/L glucose as carbon source was found to be optimum to reach high yields both in the production of enzyme and biomass.

To achieve well-defined culture conditions and carbon-limited balanced growth of the cells, chemically defined medium supplemented with 20g/L glucose was used as the growth medium for the bioreactor studies.

When the growth characteristics of the two systems were compared, elongation of the *lag* time in shake flask cultures of *E. coli* BL21(DE3)[pTaqR, pMETaq] cells was also observed in the bioreactor conditions due to endonuclease gene expression. The maximum specific growth rate of the *E. coli* BL21(DE3)[pTaqR] cells was two times higher than that of the methylase protected cells but the maximum biomass yields of both cultures were almost equal and three to four times higher than obtained from the uninduced shake flask cultures. Growth inhibition due to perturbed balance of protein synthesis after induction was also observed in both cultures.

Volumetric and specific *Taq* I endonuclease activities of the *E. coli* BL21(DE3)[pTaqR] cells reached their maximum value of 43.75×10^6 U/L and 3×10^5 U/mg

within the first two hours of induction. After this production period, while the volumetric *Taq* I restriction endonuclease activity of the cells remained constant for the rest of the fermentation, the specific *Taq* I restriction endonuclease activity decreased rapidly.

The maximum specific *Taq* I restriction endonuclease activity of $5.3x10^5$ U/mg and a volumetric activity of $250x10^6$ U/L were recovered from the *E. coli* BL21(DE3)[pTaqR, pMETaq] cells after the first two hours of induction.

The maximum recombinant protein production rates of the *E. coli* BL21(DE3)[pTaqR] and *E. coli* BL21(DE3)[pTaqR, pMETaq] cells were $24.87x10^6$ and $33.5x10^6$ Units *Taq* I/gDCW/h, respectively. Higher production rates for longer periods were observed for the methylase protected cells. Both in complex LB and chemically defined medium, two times higher specific *Taq* I productivities could be obtained from *E. coli* BL21(DE3)[pTaqR, pMETaq] cells. As also observed in the shake flask cultures in complex medium, cells could not maintain their productivity for longer periods under controlled bioreactor conditions.

In accordance with the earlier observations, the specific *Taq* I restriction endonuclease productivity of the *E. coli* BL21(DE3)[pTaqR] cells reached its maximum of $17.36x10^6$ Units *Taq* I/gDCW/h after two hours of induction and then decreased rapidly in the following two hours to one sixth of its maximum value.

The specific productivity of the *E. coli* BL21(DE3)[pTaqR, pMETaq] cells was about two times higher than that of the unprotected cells and despite its decreasing trend, 8.5 times higher specific productivity levels were maintained even after 10 hours of induction.

Number of viable cells in both cultures increased at the same rate during the pre-induction growth phase. After the addition of IPTG, no significant increase in cell viability was observed for a certain period hence the growth inhibition observed from the biomass concentration profiles was also reflected in the cells' colony forming ability. After this period, a significant drop in viable cell concentration occurred in *E. coli* BL21(DE3)[pTaqR] culture which could be attributed to the morphological changes of the

cells. On the other hand, the cell viability of the methylase protected cells increased and reached a steady-state value of about $3x10^{22}$ CFU/ml. Co-expression of the *Taq* I methylase gene not only promoted the specific productivity but also the viability of the recombinant cells.

The pTaqR plasmid containing the *Taq* I endonuclease gene was highly stable during the first four hours of induction in both cultures. After this period, its stability decreased to about 70% in *E. coli* BL21(DE3)[pTaqR] cells for the rest of the fermentation. However, higher plasmid stabilities were observed in the methylase protected cells such that 85-90 per cent of the *E. coli* BL21(DE3)[pTaqR, pMETaq] colonies were found to retain this plasmid even after ten hours of induction. The presence of the second plasmid pMETaq containing the *Taq* I methylase gene was also checked and the number of colonies containing both plasmids were found to be 100 times higher than the viable cell count of the single-plasmid cells at the end of the fermentation.

Utilization of glucose by the *E. coli* BL21(DE3)[pTaqR] cells started with the onset of *Taq* I endonuclease production and after 15 hours, all of the glucose in the medium was depleted. However, glucose was consumed by the *E. coli* BL21(DE3)[pTaqR, pMETaq] cells from the start of the growth as an energy supply for the constitutive expression of the *Taq* I methylase gene. When *Taq* I endonuclease production was induced in these cultures after eight hours, glucose was utilized at a higher rate and after 17 hours, all of the glucose was used up.

In unprotected cells, the maximum acetic acid concentration was two times higher than the methylase protected cells, however, these values were far below the growth-inhibitory concentration. When the glucose was totally depleted, *E. coli* BL21(DE3)[pTaqR, pMETaq] cells used acetic acid as the carbon source resulting in a decrease in acetic acid concentration towards the end of the fermentation.

The biomass yields on substrate glucose (Yx/s) were very close since both systems utilized all of the 20g/L initial glucose at different rates and reached equal dry cell mass concentrations. Yp/x of the methylase protected cells was 40 per cent higher since elevated *Taq* I endonuclease production rates could be maintained by these cells over prolonged

induction times. When the Yp/s values are compared, per gram glucose consumed, 1.6 times higher *Taq* I endonuclease was produced by *E. coli* BL21(DE3)[pTaqR, pMETaq] cells than by the unprotected single-plasmid cells.

SDS-PAGE analysis of the crude cellular extracts revealed that the band corresponding to the monomeric *Taq* I endonuclease was apparent in both samples but at a higher intensity in the crude extract of *E. coli* BL21(DE3)[pTaqR, pMETaq] cells due to its elevated levels of expression compared to the unprotected cells. Also in the crude extract of these cells, the band corresponding to the *Taq* I methylase enzyme was at its expected position.

6.2. Recommendations

In the light of this study, several recommendations can be drawn for the future work related to the high level expression of the *Taq* I restriction-modification enzymes in pMAL and pET expression systems.

6.2.1. pMAL Protein Fusion System

The activity of the *Taq* I methylase enzyme expressed by the *E. coli* XL1(pH185, pETMET) cells should be determined quantitatively by developing an assay procedure using data available in literature.

The *Taq* I methylase gene can be cloned and expressed under a controllable promoter that may be regulated by a variety of signals such as pH, temperature, dissolved oxygen, nutritional starvation or by the addition of chemicals other than IPTG such as L-arabinose or galactose.

The conditions used for the purification of the fusion protein by amylose affinity chromatography may be optimized to improve the recovery yields.

Co-expression of the methylase gene in *E. coli* XL1 cells producing MBP-*Taq* I fusion protein resulted in higher plasmid stabilities and excretion levels under controlled

bioreactor conditions therefore, *E. coli* XL1(pH185, pETMET) cells can be used for the subsequent fermentor experiments.

Different medium formulations containing variable amounts of glucose, glycerol, maltose, yeast extract, glycine or sorbitol can be applied and their effects on the expression of MBP-*Taq* I fusion protein can be investigated.

The fermentor can be operated at the fed-batch mode by using various feeding strategies to maintain a certain glucose level in the system and hence to improve fusion protein production.

Chemically defined medium, where proteins and contaminating cellular components are not added in contrast to complex medium can be used to simplify the downstream processing of the fusion protein and also to improve the final product quality.

A new experimental set-up for the continuous protein production can be developed by connecting a cell-recycle system at the harvest port of the fermentor operating at the continuous mode. In this set-up, the quality of the growth medium can be maintained by continuously removing inhibitory or toxic compounds from the system and since the cells are recycled, they can be utilized more efficiently. In addition, an amylose affinity column may also be attached at the outlet of the recycle system such that the cell-free chemically defined medium containing the expressed MBP-*Taq* I fusion protein can pass directly through it. After replacing the column with a regenerated one during fermentation, the bound fusion proteins can be eluted with a buffer containing 10 mM maltose.

Different unstructured and structured mathematical models can be proposed to analyze the growth and MBP-TaqI excretion kinetics of the recombinant *E. coli* XL1(pH185, pETMET) cells.

Optimization of other fermentation variables such as culture temperature and pH, induction time and period, inducer concentration and aeration may provide further information about the commercial feasibility of this process.

A mathematical model describing the extracellular transport of the MBP-*Taq* I fusion protein through the sec pathway can be developed and validated against fermentation results obtained from recombinant *E. coli* cells expressing MBP-*Taq* DNA polymerase and MBP-Glucose isomerase fusion proteins.

6.2.2. pET Expression System

The two systems can also be compared on the basis of their ability to express the *Taq* I endonuclease gene on successive generations. The expression stability of the *Taq* I methylase gene needs also to be elucidated.

A promoter that is induced by other means than addition of IPTG can be used for the controlled expression of the *Taq* I methylase gene.

Glycerol can be used as the sole carbon source instead of glucose and its effect on *Taq* I endonuclease production and acetic acid accumulation can be investigated and compared with that of glucose.

Further shake flask experiments in synthetic medium should also be conducted in order to find optimum induction conditions leading to higher protein productivities.

The copy number of the constructed plasmids, which gives direct information about the gene dosage, need to be determined and monitored by a quantitative method.

The expressed proteins can be purified by a fast and sensitive method such as HPLC or capillary electrophoresis to reach high product qualities for their consequent use in molecular biology and biotechnology applications.

Bioreactor experiments with a complex medium can be carried out to compare the commercial feasibility and reproducibility of the process with that of the defined medium.

The formulation of the defined medium can further be improved by using a statistical technique such as response surface methodology (RSM), which is widely used for medium optimization purposes (Shin *et al.*, 1997).

In order to achieve high cell densities, the fermentor can be operated at the fed-batch mode and the effects of key fermentation variables such as growth and production medium compositions, specific growth rate, induction conditions and volumetric feed rate of the postinduction media can systematically be investigated to find optimal conditions maximizing the biomass production and the yield of *Taq* I restriction endonuclease.

The pre-induction and post-induction specific cell growth rates of the culture can be controlled by using an electronic control module and an associated software program to examine the correlation between the specific growth rate and the efficiency of recombinant product formation since it directly relates to the fed-batch control strategy.

Other fermentation variables such as dissolved oxygen concentration, DNA and RNA concentrations should be determined and this information can be used for the mathematical description of the process by applying various unstructured and structured models available in the literature.

Another plasmid construct containing both the *Taq* I endonuclease and *Taq* I methylase genes was also developed in our laboratory. Optimum fermentation conditions of the *E. coli* BL21(DE3) cells harboring this new construct should be determined and compared with the results presented in this study for the two expression systems.

REFERENCES

Akesson, M., E. N. Karlsson, P. Hagander, J. P. Axelsson and A. Tocaj, 1999, "On-Line Detection of Acetate Formation in *Escherichia coli* Cultures Using Dissolved Oxygen Responses to Feed Transients", *Biotechnology and Bioengineering*, Vol. 64, No. 5, pp. 590-598.

Aristidou, A. A., P. Yu and K. San, 1993, "Effects of Glycine Supplement on Protein Production and Release in Recombinant *Escherichia coli*", *Biotechnology Letters*, Vol. 15, No. 4, pp. 331-336.

Aristidou, A. A., K. San and N. G. Bennet, 1994, "Modification of Central Metabolic Pathway in *Escherichia coli* to Reduce Acetate Accumulation by Heterologous Expression of the *Bacillus subtilis* Acetolactate Synthase Gene", *Biotechnology and Bioengineering*, Vol. 44, pp. 944-951.

Aristidou, A. A., K. San and N. G. Bennet, 1999, "Improvement of Biomass Yield and Recombinant Gene Expression in *Escherichia coli* by Using Fructose as the Primary Carbon Source", *Biotechnology Progress*, Vol. 15, pp. 140-145.

Bailey, J. E. and D. F. Ollis (editors), 1986, *Biochemical Engineering Fundamentals*, McGraw-Hill, Singapore.

Baneyx, F., 1999, "Recombinant Protein Expression in *Escherichia coli*", *Current Opinion in Biotechnology*, Vol. 10, No. 5, pp. 411-421.

Barany, F., 1987, "A Genetic System for Isolation and Characterization of *Taq* I Restriction Endonuclease Mutants", *Gene*, Vol. 56, pp. 13-27.

Barany, F., 1988, "The *Taq* I Star Reaction: Strand Preferences Reveal Hydrogen-Bond Donor and Acceptor Sites in Canonical Sequence Recognition", *Gene*, Vol. 65, pp. 149-165.

Barany, F., 1988, "How *Taq* I Recognizes its Cognate Sequence", *Gene*, Vol. 74, pp. 63-65.

Barany, F., 1988, "Overproduction, Purification and Crystallization of *Taq* I Restriction Endonuclease", *Gene*, Vol. 65, pp. 167-177.

Barany, F., B. Slatko, M. Dazitz, D. Cowburn, I. Schildkraut and G. G. Wilson, 1992, "The Corrected Nucleotide Sequences of the *Taq* I Restriction and Modification Enzymes Reveal a Thirteen-Codon Overlap" *Gene*, Vol. 112, pp. 91-95.

Bedouelle, H. and P. Duplay, 1988, "Production in *Escherichia coli* and One-step Purification of Bifunctional Hybrid Proteins Which Bind Maltose", *European Journal of Biochemistry*, pp. 541-549.

Betton, J. M., J. P. Jacob, M. Hofnung and J. K. Broome-Smith, 1997, "Creating a Bifunctional Protein by Insertion of β-lactamase into the Maltodextrin-binding Protein", Nature Biotechnology, Vol. 15, pp. 1276-1279.

Blanchin-Roland, S. and J. M. Masson, 1989, "Protein Secretion Controlled by a Synthetic Gene in *Escherichia coli*", *Protein Engineering*, Vol. 2, pp. 473-480.

Blondel, A. and H. Bedouelle, 1990, "Export and Purification of a Cytoplasmic Dimeric Protein By Fusion to the Maltose-Binding Protein of *Escherichia coli*", *European Journal of Biochemistry*, Vol. 193, pp. 325-330.

Bollag, D. M., M. D. Rozycki and S. J. Edelstein (editors), 1996, *Protein Methods*, Second Edition, John Wiley and Sons Inc., New York.

Bothmann, H. and A. Plückthun, 1998, "Selection for a Periplasmic Factor Improving Phage Display and Functional Periplasmic Expression", *Nature Biotechnology*, Vol. 16, pp. 376-380.

Bradford, M., 1976, "A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-dye Binding", *Analytical Biochemistry*, Vol. 72, pp. 248-254.

Brock, T. D. and H. Freeze, 1969, "*Thermus aquaticus* gen.n. and sp.n., a Nonsporulating Extreme Thermophile," *Journal of Bacteriology*, Vol. 98, No. 1, pp. 289-297.

Carter, P., 1990, "Site-specific Proteolysis of Fusion Proteins", in M. R. Ladisch, R. C. Willson, C. C. Painton and S. E. Builder (editors), *Protein Purification: From Molecular Mechanisms to Large-scale Processes*, pp. 181-193, American Chemical Society Symposium Series No. 427, American Chemical Society, Washington, D.C.

Chavagnat, F., C. Duez, M. Guinand, P. Potin, T. Barbeyron, B. Henrissat, J. Wallach and J. M. Ghuysen, 1996, "Cloning, Sequencing and Overexpression in *Escherichia coli* of the Alginate-Lyase-Encoding *aly* Gene of *Pseudomonas alginovora*: Identification of Three Classes of Alginate Lyases", *Biochemical Journal*, Vol. 319, pp. 575-583.

Chen, G. T. and M. Inouye, 1990, "Suppression of the Negative Effect of Minor Arginine Codons on Gene Expression; Preferential Usage of Minor Codons Within the First 25 Codons of the *Escherichia coli* Genes", *Nucleic Acids Research*, Vol. 18, No. 6, pp. 1465-1473.

Chou, C. H., N. G. Bennet and K. San, 1994, "Effect of Modulated Glucose Uptake on High-Level Recombinant Protein Production in a Dense *Escherichia coli* Culture", *Biotechnology Progress*, Vol. 10, pp. 644-647.

Clark, D., 1981, "Regulation of Fatty Acid Degradation in *Escherichia coli*: Analysis by Operon Fusion", *Journal of Bacteriology*, Vol. 148, No. 2, pp. 521-526.

Cserjan-Puschmann, M., W. Kramer, E. Duerrschmid, G. Striedner and K. Bayer, 1999, "Metabolic Approaches for the Optimization of Recombinant Fermentation Processes", *Applied Microbiology and Biotechnology*, Vol. 53, pp. 43-50.

Dower, W. J., J. F. Miller and C. W. Ragsdale, 1988, "High Efficiency Transformation of *E. coli* by High Voltage Electroporation", *Nucleic Acids Research*, Vol. 16, pp. 6127-6133.

Duenas, M., J. Vazquez, M. Ayala, E. Söderlind, M. Ohlin, L. Perez, C. A. K. Borrebaeck and J. V. Gavilondo, 1994, "Intra- and Extracellular Expression of an scFv Antibody Fragment in *E. coli*: Effect of Bacterial Strains and Pathway Engineering Using GroES/L Chaperonins", *BioTechniques*, Vol. 16, No. 3, pp. 476-483.

Duplay, P., H. Bedouelle, A. Fowler, I. Zabin, W. Saurin and M. Hofnung, 1984, "Sequences of the *malE* Gene and of its Product, The Maltose-binding Protein of *Escherichia coli* K12", *The Journal of Biological Chemistry,* Vol 259, No. 16, pp. 10606-10613.

Ernst, S., G. Venkataraman, S. Winkler, R. Godavarti, R. Langer, C. L. Cooney and R. Sasisekharan, 1996, "Expression in *Escherichia coli*, Purification and Characterization of Heparinase I from *Flavobacterium heparinum*", *Biochemical Journal*, Vol. 315, pp. 589-597.

Ernst, S., O. A. Garro, G. Venkataraman, S. Winkler, R. Godavarti, R. Langer, C. L. Cooney and R. Sasisekharan, 1997, "Process Simulation for Recombinant Protein Production: Cost Estimation and Sensitivity Analysis for Heparinase I Expressed in *Escherichia coli*", *Biotechnology and Bioengineering*, Vol. 53, No. 6, pp. 574-582.

Flickinger, M. C. and M. P. Rouse, 1993, "Sustaining Protein Synthesis in the Absence of Rapid Cell Division: An Investigation of Plasmid Encoded Protein Expression in *Escherichia coli* During Very Slow Growth", *Biotechnology Progress*, Vol. 9, pp. 555-572.

Ford, C. F., I. Suominen and C. E. Glatz, 1991, "Fusion Tails for the Recovery and Purification of Recombinant Proteins", *Protein Expression and Purification*, Vol. 2, pp. 95-107.

Gaden, E. L., 2000, "Fermentation Process Kinetics", *Biotechnology and Bioengineering*, Vol. 67, No. 6, pp. 629-635.

Georgiou, G., 1988, "Optimizing the Production of Recombinant Proteins in Microorganisms", *AIChE Journal*, Vol. 34, No. 8, pp. 1233-1248.

Gitter, B., R. Diefenbach, H. Keweloh and D. Riesenberg, 1995, "Influence of Stringent and Relaxed Response on Excretion of Recombinant Proteins and Fatty Acid Composition in *Escherichia coli*", *Applied Microbiology and Biotechnology*, Vol. 43, pp. 89-92.

Goldberg, A. L., 1992, "The Mechanism and Functions of ATP-dependent Proteases in Bacterial and Animal Cells", *European Journal of Biochemistry*, Vol. 203, pp. 9-23.

Grodberg, J. and J. J. Dunn, 1988, "*OmpT* Encodes the *Escherichia coli* Outer Membrane Protease that Cleaves T7 RNA Polymerase During Purification", *Journal of Bacteriology*, Vol. 170, No. 3, pp. 1245-1253.

Gupta, J. C., M. Jaisani, G. Pandey and K. J. Mukherjee, 1999, "Enhancing Recombinant Protein Yields in *Escherichia coli* Using the T7 System Under the Control of Heat Inducible λP_L Promoter", *Journal of Biotechnology*, Vol. 68, pp. 125-134.

Hanlon, N. and D. Barford, 1998, "Purification and Crystallization of the CDK-associated Protein Phosphatase KAP Expressed in *Escherichia coli*", *Protein Science*, Vol. 7, pp. 508-511.

Hannig, G. and S. C. Makrides, 1998, "Strategies for Optimizing Heterologous Protein Expression in *Escherichia coli*", *Trends in Biotechnology*, Vol. 16, pp. 54-60.

Harrison, J. S., E. Keshavarz-Moore, P. Dunnill, M. J. Berry, A. Fellinger and L. Frenken, 1996, "Factors Affecting the Fermentative Production of a Lysozyme-Binding Antibody Fragment in *Escherichia coli*", *Biotechnology and Bioengineering*, Vol. 53, No. 6, pp. 611-622.

Hayhurst, A. and W. J. Harris, 1999, "*Escherichia coli* Skp Chaperone Co-Expression Improves Solubility and Phage Display of Single Chain Antibody Fragments", *Protein Expression and Purification*, Vol. 15, pp. 336-343.

Holms, W. H., 1986, "The Central Metabolic Pathways of *Escherichia coli*: Relationship Between Flux and Control at a Branch Point, Efficiency of Conversion to Biomass and Excretion of Acetate", *Current Topics in Cellular Regulation*, Vol. 28, pp. 69-105.

Hoffmann, A. and R. G. Roeder, 1991, "Purification of His-tagged Proteins in Non-denaturing Conditions Suggests a Convenient Method for Protein Interaction Studies", *Nucleic Acids Research*, Vol. 19, No. 22, pp. 6337-6338.

Hsiung, H. M., A. Cantrell, J. Luirink, B. Oudega, A. J. Veros and G. W. Becker, 1989, "Use of Bacteriocin Release Protein in *E. coli* for Excretion of Human Growth Hormone into the Culture Medium", *Biotechnology*, Vol. 7, pp. 267-271.

Jeong, K. J. and S. Y. Lee, 1999, "Secretory Production of Human Leptin in *Escherichia coli*", *Biotechnology and Bioengineering*, Vol. 67, No. 4, pp. 398-407.

Kato, C., T. Kobayashi, T. Kudo, T. Furusato, Y. Murakimi, T. Tanaka, H. Baba, T. Oishi, E. Ohtsuka, M. Ikehara, T. Yanagida, H. Kato, S. Moriyama and K. Horikoshi, 1987, "Construction of an Excretion Vector and Extracellular Production of Human Growth Hormone from *Escherichia coli*", *Gene*, Vol. 54, pp. 197-202.

Kim, Y. C., S. Kwon, S. Y. Lee and H. N. Chang, 1998, "Improved Production of a Bioadhesive Precursor Protein by Fed-batch Cultivation of a Recombinant *Escherichia coli* with a pLysS Vector", *Biotechnology Letters*, Vol. 20, No. 8, pp. 799-803.

Kitai, K., T. Kudo, S. Nakamura, T. Masegi, Y. Ichikawa and K. Horikoshi, 1988, "Extracellular Production of Human Immunoglobulin G Fc Region (hIgG-Fc) by *Escherichia coli*", *Applied Microbiology and Biotechnology*, Vol. 28, pp. 52-56.

Kloos, D. U., M. Stratz, A. Güttler, R. J. Steffan and K. N. Timmis, 1994, "Inducible Cell Lysis System for the Study of Natural Transformation and Environmental Fate of DNA Released by Cell Death", *Journal of Bacteriology*, Vol. 176, No. 23, pp. 7352-7361.

Knappskog, P., H. Eiken, A. Martinez, T. Flatmark and J. Apold, 1995, "The PKU Mutation S349P Causes Complete Loss of Catalytic Activity in the Recombinant Phenylalanine-Hydroxylase Enzyme", *Human Genetics*, Vol. 95, Iss. 2, pp. 171-173.

Ko, J. H., D. K. Park, C. I. Kim, S. H. Lee and S. M. Byun, 1995, "High-Level Expression and Secretion of Streptokinase in *Escherichia coli*", *Biotechnology Letters*, Vol. 17, No. 10, pp. 1019-1024.

Komai, T., Y. Ishikawa, R. Yagi, H. Suzuki-Sunagawa, T. Nishigaki and H. Handa, 1997, "Development of HIV-1 Protease Expression Methods Using the T7 Phage Promoter System", *Applied Microbiology and Biotechnology*, Vol. 47, pp. 241-245.

Korz, D. J., U. Rinas, K. Hellmuth, E. A. Sanders and W. D. Deckwer, 1995, "Simple Fed-Batch Technique for High Cell Density Cultivation of *Escherichia coli*", *Journal of Biotechnology*, Vol. 39, pp. 59-65.

Labahn, J., J. Granzin, G. Schluckbier, D. P. Robinson, W. E. Jack, I. Schildkraut and W. Saenger, 1994, "Three-Dimensional Structure of the Adenine-Specific DNA Methyltransferase M.*Taq* I in Complex with the Cofactor S-adenosylmethionine", *Proceedings of the National Academy of Sciences USA*, Vol. 91, pp.10957-10961.

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

Lee, S. Y., 1996, "High-Cell Density Culture of *Escherichia coli*", *Trends in Biotechnology*, Vol. 14, pp. 98-105.

Lee, S. J., I. C. Kim, D. M. Kim, K. H. Bae and S. M. Byun, 1997, "Enhanced Production and Secretion of Streptokinase into Extracellular Medium in *Escherichia coli* by Removal of 13 N-terminal Amino Acids", *Biotechnology Letters*, Vol. 19, No. 2, pp. 151-154.

Lee, S. J., I. C. Kim, D. M. Kim, K. H. Bae and S. M. Byun, 1998, "High Level Secretion of Recombinant Staphylokinase into Periplasm of *Escherichia coli*", *Biotechnology Letters*, Vol. 20, No. 2, pp. 113-116.

Li, Y., J. Chen, Y. Mao, S. Lun and Y. Koo, 1998, "Effect of Additives and Fed-Batch Culture Strategies on the Production of Glutathione by Recombinant *Escherichia coli*", *Process Biochemistry*, Vol. 33, No. 7, pp. 709-714.

Maina, V. C., P. D. Riggs, G. A. Grandea, B. E. Slatko, S. L. Moran, J. A. Tagliamonte, A. L. Mc Reynolds and C. Guan, 1988, "An *Escherichia coli* Vector to Express and Purify Foreign Proteins by Fusion to and Separation from Maltose-binding Protein", *Gene*, Vol. 74, pp. 365-373.

Makrides, S. C., 1996, "Strategies for Achieving High-Level Expression of Genes in *Escherichia coli*", *Microbiological Reviews*, Vol. 60, No. 3, pp. 512-538.

Mayer, A. M. and F. Barany, 1995, "Photoaffinity Cross-Linking of *Taq* I Restriction Endonuclease Using Aryl Azide Linked to the Phosphate Backbone", *Gene*, Vol. 153, pp. 1-8.

Mierendorf, R., K. Yeager and R. Novy, 1994, "The pET System: Your Choice for Expression", *Innovations, Newsletter of Novagen*, Vol. 1, No. 1, pp. 4-8.

Miksch, G., R. Neitzel, E. Fiedler, K. Friehs and E. Flaschel, 1997, "Extracellular Production of a Hybrid B-glucanase from *Bacillus* by *Escherichia coli* Under Different Cultivation Conditions in Shaking Cultures and Bioreactors", *Applied Microbiology and Biotechnology*, Vol. 47, pp. 120-126.

Miller, K. W., R. Schamber, Y. Chen and B. Ray, 1998, "Production of Active Chimeric Pediocin AcH in *Escherichia coli* in the Absence of Processing and Secretion Genes from the *Pediococcus* pap Operon", *Applied and Environmental Microbiology*, Vol. 64, No. 1, pp: 14-20.

Minas, M. and J. E. Bailey, 1995, "Co-Overexpression of *prlF* Increases Cell Viability and Enzyme Yields in Recombinant *Escherichia coli* Expressing *Bacillus stearothermophilus* α-Amylase", *Biotechnology Progress*, Vol. 11, pp. 403-411.

Mishima, N., K. Mizumoto, Y. Iwasaki, H. Nakano and T. Yamane, 1997, "Insertion of Stabilizing Loci in Vectors of T7 RNA Polymerase-Mediated *Escherichia coli* Expression Systems: A Case Study on the Plasmids Involving Foreign Phospholipase D Gene", *Biotechnology Progress*, Vol. 13, pp. 864-868.

Missiakas, D., J. M. Betton and S. Raina, 1996, "New Components of Protein Folding in Extracytoplasmic Compartments of *Escherichia coli* SurA, FkpA and Skp/OmpH", *Molecular Microbiology*, Vol. 21, pp. 871-884.

Morino, S., Y. Terako, T. Ishida, H. Ueda and S. Uesugi, 1995, "Soluble Expression of a Synthetic Gene for Human Translation Initiation Factor 4R in *Escherichia coli*", *Biological and Pharmaceutical Bulletin*, Vol. 18, Iss. 2, pp. 372-376.

Nakamura, S., T. Masegi, K. Kitai, Y. Ichikawa, T. Kudo, R. Aono and K. Horikoshi, 1990, "Extracellular Production of Human Tumor Necrosis Factor by *Escherichia coli* Using a Chemically-Synthesized Gene", *Agricultural and Biological Chemistry*, Vol. 54, pp. 3241-3250.

Old, R. W. and S. B. Primrose (editors), 1989, *Principles of Gene Manipulation*, Blackwell Scientific Publications, Oxford.

Özdinler, H., 1996, *Cloning of Taq I Endonuclease Gene into E. coli and Isolation of the Enzyme by Using pMAL-p2 Vector System*, M.S. Thesis, Boğaziçi University.

Özkırımlı, E., 1998, *Cloning and Expression of Taq I Endonuclease Gene Using pMAL-p2 Vector System in Different Strains of Escherichia coli*, M.S. Thesis, Boğaziçi University.

Pazirandeh M., L. Chrisey, M. Mauro, J. Campbell and B. Gaber, 1995, "Expression of the Neurospora- Crassa Metallothionein Gene in *Escherichia coli* and Its Effect on Heavy-Metal Uptake", *Applied Microbiology and Biotechnology*, Vol. 43, Iss. 6, pp. 1112-1117.

Pilon, A. L., P. Yost, T. E. Chase, G. L. Lohnas and W. Bentley, 1996, "High Level Expression and Efficient Recovery of Ubiquitin Fusion Proteins from *Escherichia coli*", *Biotechnology Progress*, Vol. 12, pp. 331-337.

Pues, H., N. Bleimling, B. Holz, J. Wölcke and E. Weinhold, 1999, "Functional Roles of the Conserved Aromatic Amino Acid Residues at Position 108 (Motif IV) and Position 196 (Motif VIII) in Base Flipping and Catalysis by the N6-Adenine DNA Methyltransferase from *Thermus aquaticus*", Biochemistry, Vol. 38, pp. 1426-1434.

Riesenberg, D., V. Schulz, W. A. Knorre, H. D. Pohl, D. Korz, E. A. Sanders, A. Rob and W. D. Deckwer, 1991, "High Cell Density Cultivation of *Escherichia coli* at Controlled Specific Growth Rate", *Journal of Biotechnology*, Vol. 20., pp. 17-28.

Riesenberg, D. and R. Guthke, 1999, "High-Cell-Density Cultivation of Microorganisms", *Applied Microbiology and Biotechnology*, Vol. 51, pp. 422-430.

Robinson, M., R. Lilley, S. Little, J. S. Emtage, G. Yarranton, P. Stephens, A. Millican, M. Eaton and G. Humphreys, 1984, "Codon Usage Can Affect Efficiency of Translation of Genes in *Escherichia coli*", *Nucleic Acids Research*, Vol. 12, No. 17, pp. 6663-6671.

Rosemond, M. J. C., S. M. Strobel, P. H. Ray and P. J. Bassford, 1994, "Ability of MBP or RBP Signal Peptides to Influence Folding and In Vitro Translocation of Wild-Type and Hybrid Precursors", *FEBS Letters*, Vol. 349, pp. 281-285.

Ryan, W., P. Collier, L. Loredo, J. Pope and R. Sachdev, 1996, "Growth Kinetics of *Escherichia coli* and Expression of a Recombinant Protein and its Isoforms under Heat Shock Conditions", *Biotechnology Progress*, Vol. 12, pp. 596-601.

Sambrook, J., E. F. Fritsch and T. Maniatis (editors), 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Saraswat, V., J. Lee, D. Y. Kim and Y. H. Park, 2000, "Synthesis of Recombinant Interleukin-2 via Controlled Feed of Lactose-Complex Media in Fed-Batch Cultures of *Escherichia coli* BL21(DE3)[pT7-G3IL2]", *Biotechnology Letters*, Vol. 22, pp. 261-265.

Sato, S., C. A. Hutchison and J. I. Harris, 1977, "A Thermostable Sequence-Specific Endonuclease from *Thermus aquaticus*", *Proceedings of the National Academy of Sciences USA*, Vol. 74, No. 2, pp.542-546.

Sato, S., K. Nakazawa and T. Shinomiya, 1980, "A DNA Methylase from *Thermus thermophilus* HB8", *Journal of Biochemistry*, Vol. 88, pp. 737-747.

Sauer, P. W., J. E. Burky, M. C. Wesson, H. D. Sternard and L. Qu, 2000, "A High-Yielding, Generic Fed-Batch Cell Culture Process for Production of Recombinant Antibodies", *Biotechnology and Bioengineering*, Vol. 67, No. 5, 585-597.

Scherrer, S., N. Robas, H. Zouheiry, G. Branlant and C. Branlant, 1994, "Periplasmic Aggregation Limits the Proteolytic Maturation of the *Escherichia coli* Penicillin G Amidase Precursor Polypeptide", *Applied Microbiology and Biotechnology*, Vol. 42, pp. 85-91.

Shiloach, J., J. Kaufman, A. S. Guillard, and R. Fass, 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*, Vol. 49, pp. 421-428.

Shin, C. S., M. S. Hong, C. S. Bae and J. Lee, 1997, "Enhanced Production of Human Mini-Proinsulin in Fed-Batch Cultures at High Cell Density of *Escherichia coli* BL21(DE3)[pET-3aT2M2]", *Biotechnology Progress*, Vol. 13, pp. 249-257.

Shin, C. S., M. S. Hong, D. Y. Kim, H. C. Shin and J. Lee, 1998, "Growth-Associated Synthesis of Recombinant Human Glucagon and Human Growth Hormone in High-Cell-Density Cultures of *Escherichia coli*", *Applied Microbiology and Biotechnology*, Vol. 49, pp. 364-370.

Slatko, B. E., J. S. Benner, T. J. Quinton, L. S. Moran, T. G. Simcox, E. M. Van Cott and G. G. Wilson, 1987, "Cloning, Sequencing and Expression of the *Taq* I Restriction-Modification System", *Nucleic Acids Research*, Vol. 15, pp. 9781-9796.

Sorensen, M. A., C. G. Kurland and S. Pedersen, 1989, "Codon Usage Determines the Translation Rate in *Escherichia coli*", *Journal of Molecular Biology*, Vol. 207, pp. 365-377.

Spurlino, J. C., G. Y. Lu and F. A. Quiocho, 1991, "The 2.3 Angstrom Resolution Structure of the Maltose- or Maltodextrin-binding Protein, a Primary Receptor of Bacterial Active Transport and Chemotaxis", *Journal of Biological Chemistry*, Vol. 266, pp. 5202-5219.

Studier, F. W. and A. B. Moffat, 1986, "Use of Bacteriophage T7 RNA Polymerase to Direct Selective High-level Expression of Cloned Genes," *Journal of Molecular Biology*, Vol. 189, pp. 113-130.

Studier, F. W., 1991, "Use of Bacteriophage T7 Lysozyme to Improve an Inducible T7 Expression System", *Journal of Molecular Biology*, Vol. 219, pp. 37-44.

Suominen, I., M. Karp, M. Lahde, A. Kopio, T. Glumoff, P. Meyer and P. Mantsala, 1987, "Extracellular Production of Cloned α-Amylase by *Escherichia coli*", *Gene*, Vol. 61, pp. 165-176.

Tabor, S. and C. C. Richardson, 1985, "A Bacteriophage T7 Polymerase/promoter System for Controlled Exclusive Expression of Specific Genes", *Proceedings in National Academy of Sciences USA*, Vol. 82, pp. 1074-1078.

Talmadge, K. and W. Gilbert, 1982, "Cellular Location Affects Protein Stability in *Escherichia coli*", *Proceedings of the National Academy of Sciences USA*, Vol. 79, pp. 1830-1833.

Teich, A., H. Y. Lin, L. Andersson, S. Meyer and P. Neubauer, 1998, "Amplification of ColE1 related Plasmids in Recombinant Cultures of *Escherichia coli* after IPTG Induction", *Journal of Biotechnology*, Vol. 64, pp. 197-210.

Toksoy, E., P. H. Özdinler, Z. İ. Önsan and B. K. Kırdar, 1999, "High Level Secretion of *Taq* I Restriction Endonuclease by Recombinant *Escherichia coli*", *Biotechnology Techniques*, Vol. 13, pp. 803-808.

Uhlen, M. and T. Moks, 1990, "Gene Fusions for Purpose of Expression: An Introduction", *Methods in Enzymology*, Vol. 185, pp. 129-143.

van der Wal, F. J., C. M. ten Hagen-Jongman, B. Oudega and J. Luirink, 1995, "Optimization of Bacteriocin-Release-Protein-Induced Protein Release by *Escherichia coli*: Extracellular Production of the Periplasmic Molecular Chaperone FaeE", *Applied Microbiology and Biotechnology*, Vol. 44, pp. 459-465.

Von Darl, M., P. M. Harrison and W. Bottke, 1994, "Expression in *Escherichia coli* of a Secreted Invertebrate Ferritin", *European Journal of Biochemistry*, Vol. 222, pp. 367-376.

Wan, E. W. and F. Baneyx, 1998, "TolAIII Co-Overexpression Facilitates the Recovery of Periplasmic Recombinant Proteins into the Growth Medium of *Escherichia coli*", *Protein Expression and Purification*, Vol. 14, pp. 13-22.

Weickert, M., D. Doherty, E. Best and P. Olins, 1996, "Optimization of Heterologous Protein Production in *Escherichia coli*", *Current Opinion in Biotechnology*, Vol. 7, pp. 494-499.

Weickert, C., U. Sauer and J. E. Bailey, 1998, "An *Escherichia coli* Host Strain Useful for Efficient Overproduction of Secreted Recombinant Protein", *Biotechnology and Bioengineering*, Vol. 59, No. 3, pp. 386-391.

Wong, H. H., Y. C. Kim, S. Y. Lee and H. N. Chang, 1998, "Effect of Post-Induction Nutrient Feeding Strategies on the Production of Bioadhesive Protein in *Escherichia coli*", *Biotechnology and Bioengineering*, Vol. 60, No. 3, pp. 271-276.

Yee, L. and H. W. Blanch, 1993, "Defined Media Optimization for Growth of Recombinant *Escherichia coli* X90", *Biotechnology and Bioengineering*, Vol. 41, No. 2, pp. 221-230.

Yu, P. and K. San, 1993, "Continuous Production of Cell-Free Recombinant Proteins Using *Escherichia coli*", *Biotechnology Progress*, Vol. 9, pp. 587-593.

Yuan, Z. H., U. Kumar, H. C. Thomas, Y. M. Wen and J. Monjardino, 1997, "Expression, Purification, and Partial Characterization of HCV RNA Polymerase", *Biochemical and Biophysical Research Communications*, Vol. 232, pp. 231-235.

Zabriskie, D. W. and E. J. Arcuri, 1986, "Factors Influencing Productivity of Fermentations Employing Recombinant Microorganisms", *Enzyme Microbiology and Technology*, Vol. 8, pp. 706-726.

Zebala, J., J. Choi and F. Barany, 1992, "Characterization of Steady State Single-Turnover, and Binding Kinetics of the *Taq* I Restriction Endonuclease", *Journal of Biological Chemistry*, Vol. 267, pp. 8097-8105.

Zebala, J., J. Choi, G. Trainor and F. Barany, 1992, "DNA Recognition of Base Analogue and Chemically Modified Substrates by the *Taq* I Restriction Endonuclease", *Journal of Biological Chemistry*, Vol. 267, pp. 8106-8116.

Zhang, Q., N. Tsukagoshi, S. Miyashiro and S. Udaka, 1983, "Increased Production of α-Amylase by *Bacillus amyloliquefaciens* in the Presence of Glycine", *Applied and Environmental Microbiology*, Vol. 46, No. 1, pp. 293-295.

Zhang, J. and H. Greasham, 1999, "Chemically Defined Media for Commercial Fermentations", *Applied Microbiology and Biotechnology*, Vol. 51, pp. 407-421.