CHARACTERISATION OF AN EXPRESSION SYSTEM FOR COMMERCIAL PRODUCTION OF PROTEINS

by

LENE JØRGENSEN

Thesis submitted for the degree of Doctor of Philosophy

in

The University of Adelaide
Faculty of Engineering
and
Faculty of Science

May 1997

DECLARATION

This work contains no material which has been accepted for the award of any other degree or diploma in any university or any other tertiary institution, and to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

SIGNED: Lene Sørensen ..... DATE: 9-5-97 .....

ACKNOWLEDGMENTS

I would like to thank Dr. Anton P J Middelberg, Dr Brian O'Neill and Dr Connor Thomas for their supervision and encouragement. Their guidance in preparing this thesis was invaluable. I would also like to acknowledge Dr Renato Morona and Christopher Clark for many helpful discussions and comments during the molecular biology part.

I would also like to thank the CRC "Tissue Growth and Repair" for giving me the opportunity to conduct this work. I am grateful to Adelaide Brighton Cement for providing me with additional financial support through the William Schroder Scholarship in Chemical Engineering.

I am grateful to Department of Biochemistry for allowing me to use the Densitometer, French press and other facilities.

Finally, my gratitude goes to colleagues in both Department of Microbiology and Immunology and Department of Chemical Engineering for support and friendship.

SUMMARY

To enable the production of recombinant chloramphenicol acetyl transferase (CAT) in *Escherichia coli* JM101, a number of *promoter*::*cat* transcriptional fusions were constructed. These fusions contained either IPTG-inducible or stationary-phase inducible promoters in a bidirectional promoter probe vector. Recombinant protein expression in the engineered systems was quantified using validated assays developed in this thesis. Furthermore, mathematical models were employed to establish the relative efficiencies of transcription and translation.

The detection of bacterially-expressed *cat* mRNA by slot-blotting was found to be highly dependent on total RNA immobilised onto the solid support, as well as mRNA concentration. mRNA quantitation by comparison with a pure standard resulted in gross underestimation because of this possible steric hindrance. A new method to quantitate *cat* mRNA was therefore developed. The new protocol for *cat* mRNA detection included a three-dimensional standard calibration curve, constructed for each assay, and overcomes the confounding effect of contaminating RNA.

The French press was more efficient at disrupting cells and releasing proteins than sonication. French pressing disrupted all cells in suspension whereas a maximum of 80% of the cells were disrupted following sonication. The level of CAT release was highest when cells were totally disrupted. Additional treatment with the detergent Triton X-100 was necessary to maximise CAT recovery.

Promoters induced by IPTG are commonly used but have both cost and environmental penalties. Nevertheless, an IPTG-induced system was included in this work as a control, to compare the relative efficiency of a system with commercial potential (stationary phase promoter system). Maximal protein expression was achieved for 0.1 mM IPTG after induction at OD_{600} = 0.8 in both shake-flask and fermentation experiments. A concentration of 0.4 mM IPTG yielded maximal expression for induction at OD_{600} = 2.4. Maximum CAT protein expression was independent of oxygen concentration. However, CAT protein production was highly dependent on the growth phase of the culture at induction. Induction close to stationary phase produced lower levels of CAT compared to

induction in logarithmic phase. Also, inoculation with a stationary-phase culture gave better CAT protein yield than fermenters inoculated with a logarithmic phase culture.

CAT protein production under control of the *tac* promoter was clearly limited at the translational level. This was shown by constant CAT protein levels after induction for decreasing ribosomal (16S rRNA) levels. Furthermore, induction with IPTG concentration beyond optimal resulted in a concomitant increase in mRNA level but not CAT protein. Translational limitation was confirmed by a simple mathematical model to establish the relative efficiencies of transcription (16-81 %) and translation (0.2-1 %).

Separation of the growth and CAT production phase was achieved using the stationary phase *katE* gene promoter. Batch fermentation experiments in minimal media showed that the transition from logarithmic growth phase to stationary phase stimulated *katE* expression. However, despite published material to the contrary, no inducers of the *katE* gene promoter were identified in minimal media. Acetate and o-hydroxybenzoate did not stimulate promoter activity. O-hydroxybenzoate actually inhibited translational activities. Glucose addition did improve CAT protein levels, but this was probably due to increased translational activity.

Final CAT protein levels per cell for the IPTG inducible *tac* promoter system were about 250 times higher than the stationary-phase inducible *katE* gene promoter in batch fermentation experiments. However, CAT protein levels under control of the *katE* gene promoter was limited at the translational level in batch experiments. Protein yield was improved by continuously feeding glucose, thus establishing a growth-limited culture for CAT protein production. A 3-fold increase in CAT protein, and a 10-fold increase in *cat* mRNA levels, was achieved. Future work to define a critical growth rate below which the stringent control is induced, may allow the *katE* gene promoter to be employed for commercial purposes. Furthermore, more information regarding promoter regulation is needed.

PUBLICATIONS ARISING FROM THIS THESIS RESEARCH.

CONFERENCE PRESENTATIONS

1. **Jorgensen, L.**, Thomas, C. J., Middelberg, A. P. J. and O'Neill, B. K. "An improved method for the quantitation of chloramphenicol acetyl transferase messenger RNA using a Digoxigenin label". 10th International Biotechnology Symposium, Sydney, 1996. pp. 81-82.

2. **Jorgensen, L.**, O'Neill, B. K., Thomas, C. J., Morona, R. and Middelberg, A. P. J. "Comparative recovery of CAT from *Escherichia coli* after disruption by sonication and the French press". 4th Pacific Rim Biotechnology Conference, Melbourne, 1995. pp 268-269.

3. **Jorgensen, L.**, Thomas, C. J., Morona, R., Middelberg, A. P. J. and O'Neill, B. K. "Development of expression systems for the large scale production of IGF". 11th Australasian Biotechnology Conference, Perth, 1993. p. 173.

JOURNAL ARTICLES

1. **Jorgensen, L.**, Thomas, C. J., O'Neill, B. K. and Middelberg, A. P. J. (1997). Fermentation studies into the use of Stationary Phase-Inducible Promoters (*katE* and *katF*) for Recombinant Protein Expression in *E. coli*. Submitted to *Appl. Environ. Microbiol.*

2. **Jorgensen, L.**, Thomas, C. J., O'Neill, B. K. and Middelberg, A. P. J. (1997). Investigation of the relationship between protein, message and inducer concentrations in recombinant *E. coli* cells. *J. Microbiol. Biotechnol.* **1:** 21-24.

3. **Jorgensen, L.**, Middelberg, A. P. J. O'Neill, B. K. and Thomas, C. J. (1996). Quantitation of Chloramphenicol Acetyl Transferase (CAT) messenger RNA by filter hybridisation using a Digoxigenin label. *Biotech. Tech.* **10:** 83-88.

4. **Jorgensen, L.**, O'Neill, B. K., Thomas, C. J., Morona, R. and Middelberg, A. P. J. (1995). Release of CAT from Recombinant *Escherichia coli* by sonication and the French press. *Biotech. Tech.* **9:** 477-480.

LIST OF ABBREVIATIONS

Ab	antibody
Ap	ampicillin
ATP	adenosine 5'-triphosphate
bp	base pair
BSA	bovine serum albumin
CAT	chloramphenicol acetyl transferase
CIP	calf intestinal phosphatase
Cm	chloramphenicol
CsCL	caesium chloride
cst	carbon starvation
DEPC	diethyl pyrocarbonate
DIG	digoxigenin
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxy nucleotide tri-phosphate
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylene-diamine-tetra-acetic acid
ELISA	enzyme-linked immunosorbant assays
EtBr	ethidium bromide
fg	femto gram
G-CSF	granolocyte-colony stimulating factor
GM-CSF	macrophage/granolocyte-colony stimulating factor
×g	centrifugal force
h	hour
HBsAG	hepatitis B surface antigen
hGH	human growth hormone
HGT	high gelling temperature
HIV	human immunodeficiency virus
IPTG	isopropyl-b-D-thiogalactopyranoside
kb	kilobase
Klenow	Klenow fragment of *E. coli* DNA polymerase I
LGT	low gelling temperature
MCS	multiple cloning site
min	minutes
µg	micro grams
MOPS	3-[N-morpholino]propanesulphonic acid
MQ	millique
Mpa	mega Pascal (10^6 Pascal)
mRNA	messenger RNA
NA	nutrient Agar
NB	nutrient broth
NTP	nucleotide tri-phosphate
OD	optical density
oligo	oligodeoxynucleotides
PAGE	polyacrylamide gel electrophoresis

PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
pex	post exponential
POD	peroxidase
^{R}	resistant
RBS	ribosome binding site
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid
RNase	ribonuclease
SD	Shine-Dalgarno
SDS	sodium dodecyl-sulphate
TBE	tris-borate-EDTA buffer
TE	tris-EDTA buffer
tPA	tissue-type plasminogene activator
Tris	tris[hydroxymethyl]amino-methane
Triton X-100	α-[4-(1,1,3,3-tetramethylbutyl)phenyl]-ω-hydroxypoly(oxy-1,2-ethanediyl)
tRNA	transfer RNA
Tween	20 polyoxyy-ethylene-sorbitan monolaurate
UV	Ultraviolet
V/V	volume per volume
W/V	weight per volume
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Introduction

1.1 Recombinant DNA technology

In the past, many proteins of importance to medicine and industry have been either unavailable, or available only in limited supply because of limited sources. With the development of DNA technology over the past two decades, it is now possible to produce a wide range of proteins. The expression of recombinant proteins not only allows large quantities of proteins to be manufactured, but can also prevent product contamination by viruses and prions where products are traditionally isolated from natural sources.

The first recombinant protein to be produced commercially was human insulin. This protein can be synthesised by replacing a single amino acid in isolated porcine insulin (Novo Nordisk). However, an unlimited supply of insulin is obtainable by recombinant DNA techniques, eliminating the dependency on porcine pancreases. Furthermore, contamination with pancreatic peptides is avoided. Other recombinant proteins listed for therapeutic use include human growth hormone (hGH), Factor VIII, hepatitis B surface antigen (HBsAg), tissue-like plasminogene activator (tPA), erythropoietin, interferon, and granolocyte- and granolocyte/macrophage-colony stimulating factors (G-CSF and GM-CSF) (Dykes, 1993). hGH is used to treat hypopituitary dwarfism. Prior to the advent of genetic engineering, this protein was extracted from the pituarities of human cadavers. Not only was the supply of hGH limited but the extraction carried an additional disadvantage of possible contamination with prions accumulated in the brain. Factor VIII is a blood clotting agent used to treat haemophilia (Laurence and Bennett, 1992). This protein can be extracted from human blood but with the risk of contamination with human immunodeficiency virus (HIV). The advantages of producing the protein by recombinant means (as for hGH) include improved availability and safety. Hepatitis B causes chronic viral hepatitis and no effective cure exists for the disease. Instead, patients are treated with HBsAg vaccines. Risks associated with extracting this protein from human blood are as detailed previously. These few examples demonstrate the considerable medical impact of large-scale protein production methods, leading to a replacement of natural sources.

Examples of medically-important human proteins occuring at low levels include erythropoietin (a hormone used to treat renal failure-induced anaemia (Laurence and

Bennett, 1992)) and the interferons (effective against certain types of cancers, Kaposi's sarcoma in AIDS, genital warts (Laurence and Bennett, 1992) and Hepatitis (Berg *et al.*, 1996)). The sole method for production of sufficient levels of these proteins for clinical use is recombinant techniques. There are many other examples of recombinant proteins with potential therapeutic use. Another example is Insulin-like Growth Factor I (IGF(I)). It has reached stage 3 of the clinical trials protocol for the treatment of amyotrophic lateral sclerosis (ALS).

Successful production of recombinant proteins require careful optimisation of the expression process. This will be discussed later in this chapter. Important factors include maximising recombinant protein expression, minimising product degradation, and controlling the initiation of product expression. These factors, and their relation to the current work, will be discussed subsequently.

1.2 Hosts for recombinant expression systems

A wide range of hosts have been exploited for recombinant protein synthesis. Examples include prokaryotic cells, yeast, mammalian cells, insect cells, plant cells, and transgenic animals. The choice of host depends on the recombinant protein to be expressed but it may also be influenced by patent rights. *Escherichia coli*, *Saccharomyces cerevisae* and filamentous fungi dominate industrial-scale production. Examples include human insulin and hGH production in *E. coli* and HBsAg in *S. cerevisae*. Animal-cell cultures are used to produce monoclonal antibodies and human proteins when post-transcriptional modification (e.g. glycosylation) is essential. An example of a monoclonal antibody produced in animal-cell cultures is orthoclone OKT3 (anti-CD3), used in the treatment of acute renal allograft rejection (Webb, 1993).

There is an extensive body of knowledge concerning the use of *E. coli*. As an expression host for recombinant protein, this organism has a number of advantages. These include rapid generation of biomass, the availability of low-cost culture media and a large range of expression vectors. For example, *E. coli* is much less shear-sensitive than the extremely fragile mammalian cells. Hence, high-cell density cultures can be obtained

through vigorous mixing and oxygen transfer. For these reasons, *E. coli* is often the preferred host for the commercial production of proteins. Despite the many advantages of using *E. coli*, a few problems are associated with the production of recombinant proteins in this host. One disadvantage is low protein levels due to either inefficient translation or rapid degradation of small peptides. For example, direct expression of Met-IGF(I) in *E. coli* yielded no detectable protein (King *et al.*, 1992). However, when fused to a 48 amino acid leader peptide, up to 1 g of the fusion protein per L of fermentation broth could be recovered (King *et al.*, 1992). Other disadvantages of using *E. coli* include the production of endotoxins, a lack of post-transcriptional modifications such as glycosylation, and the presence of endogenous proteases which may degrade the expressed protein. Furthermore, high-level production of recombinant proteins often results in the formation of insoluble aggregates, known as inclusion bodies. The recovery of the protein from inclusion bodies adds substantial cost and time penalties to the production process. These are caused by the need for dissolution and refolding to a biologically-active form. One solution to inclusion body formation involves fusion of the gene of interest to the thioredoxin gene (*trxA*) (LaVallie *et al.*, 1993). Thioredoxin fusion proteins accumulate in adhesion zones[1] and can be recovered by osmotically shocking the cells.

E. coli remains an important host for commercial production of recombinant proteins. This thesis focuses on this organism for this reason. Another key factor is the choice of expression system, which is discussed in the next section.

1.3 *Escherichia coli* expression systems

Wild-type plasmids such as pSC101 and ColE1 served as two of the first vectors for cloning in *E. coli* (Balbas *et al.*, 1988). Both vectors have unique restriction sites, can autonomously replicate, and encode selection markers. However, these vectors do not have tight regulation of expression. High expression of recombinant protein can lead to a marked decrease in overall biomass yield (Andersson *et al.*, 1996). Also, it can have detrimental effects on both the host and recombinant protein. Hence, it is desirable to separate the growth and production phases by placing the structural gene under control of a

[1] Adhesion zones: osmotically sensitive sites located where the inner and outer membranes are contiguous.

regulatory promoter. A variety of bacterial expression systems offering strong transcriptional control have been designed. These systems encode promoters such as *lac* (Yanisch-Perron *et al.*, 1985), *trp* (Tacon *et al.*, 1980), *tac* (de Boer, 1983), *trc* (Amann *et al.*, 1983), P_L bacteriophage lambda under control of the thermosensitive cI repressor allele (Remaut *et al.*, 1981), *pho*A (Boidol *et al.*, 1982), *rec*A (Shirakawa *et al.*, 1984), and T7 (Studier *et al.*, 1990).

The *tac* and *trc* promoters are hybrid promoters produced by replacing the -35 region of the *lac* promoter with the -35 region of the *trp* promoter (de Boer, 1983 and Amann *et al.*, 1983). These promoters are both stronger than the parent promoters. A disadvantage of using these promoters is the need for induction by isopropyl-β-thiogalactopyranoside (IPTG). Chemical induction incurs both cost and environmental penalties for the process which are unattractive for large-scale production.

T7 RNA polymerase is highly selective for its own promoters, which do not occur naturally in *E. coli*. Furthermore, T7 RNA polymerase is a very active enzyme compared to the *E. coli* RNA polymerase. These properties open the probability of selectively expressing a target gene. A series of vectors constructed by inserting the T7 (Φ_{10}) promoter into the plasmid pBR322 (Bolivar *et al.*, 1977) are commercially available (Novagen, Inc., Madison, USA). After cloning the target gene under control of a T7 promoter, the plasmid is transferred into a suitable host. This host is a lysogen of bacteriophage DE3, which contains the T7 RNA polymerase gene under control of the *lac*UV5 promoter. Protein production is induced by adding IPTG. T7 RNA polymerase is synthesised which in turn transcribes the target DNA in the plasmid. For the reasons stated earlier, chemical induction may render this system unattractive for large-scale production.

A similar system was designed by Tabor and Richardson, (1985) but it avoided the need for chemical inducers. The system consists of two compatible plasmids, pGP1-2 and pT7-1. pGP1-2 carries the T7 RNA polymerase gene under control of the P_L promoter and the temperature sensitive repressor gene lambda cI857. pT7-1 carries the T7 promoter and a multiple cloning site. Elevated temperatures result in thermal destruction of cI857 repressor protein and hence transcription from the T7 promoter can occur. This system has successfully been used to produce recombinant proteins. As an example, $LongR^3$-IGF(I)

was produced in this system after alteration of the RBS and the spacer region between the RBS and ATG codons (Jørgensen, 1993, unpublished work). However, one disadvantage of induction by a temperature increase is the simultaneous induction of heat-shock proteins, including proteases which may reduce product yield and limit the usefulness of such a system.

The *pho*A promoter from the alkaline phosphatase gene (Boidol *et al.*, 1982) is induced by lowering the growth medium phosphate concentration. This promoter was investigated for its ability to express the finger domain of human tPA (Lubke *et al.*, 1995). The protein was expressed by transferring cells to a low-phosphate medium (below 0.05 mM). However, a large-volume cell inoculum (10% (v/v)) was needed to maintain a reasonable number of plasmid-bearing cells. For a large-scale fermentation such an option may be impractical. As an alternative, adjustment of the starting level of both glucose and phosphate in the fermenter, to ensure that phosphate becomes limiting when cells reach a desirable cell density, may permit the use of the *pho*A promoter for large-scale protein production. Further work may see successful application of this system at large scale.

*Lac*Z expression studies have shown that the *E. coli rec*A gene promoter is stronger than both *lac* and *trp* (Shirakawa *et al.*, 1984). The strength of the promoter was examined by assaying β-galactosidase activity expressed from a *cro-lac*Z fused gene placed downstream of the promoter. Upon induction by nalidixic acid in a *rec*A+ strain, high expression was observed for an extended period. After 5 h under inducing conditions, as much as 11% of the total cellular protein was cro-lacZ product (Shirakawa *et al.*, 1984). However, a significant cost penalty is incurred by induction of the *rec*A gene promoter using nalidixic acid. This route is unattractive for large-scale production.

The promoter systems described above have been designed to enable strong control over recombinant expression. This is generally important. However at commercial-scale production, other factors which might not be so important at bench- or small-scale also need to be considered. A suitable expression system for large-scale protein production is determined by the strength of its promoter, its ease of utilisation, and the economy associated with its use. Furthermore, it is often useful to evaluate more than one expression system, as performance may vary for different recombinant proteins.

Recently, expression systems specifically designed for fermentation conditions have been investigated. These vectors have employed an oxygen-responsive promoter (Dikshit *et al.*, 1990; Khosla *et al.*, 1990), a pH-regulated promoter system (Chou *et al.*, 1995), or a carbon-starvation inducible promoter (Tunner *et al.*, 1992).

The utility of an oxygen responsive promoter was investigated by cloning the natural promoter of the *Vitreoscilla* hemoglobin gene in front of the *lacZ* gene in *E. coli.* Under low oxygen conditions, up to ten percent of total cellular protein was β-galactosidase (Khosla *et al.*, 1990). A major problem in fed-batch fermentations is oxygen deficiency at high-cell density due to poor mass-transfer. Induction in response to low levels of oxygen therefore has an obvious advantage. However, a possible constraint on this system is the development of stress proteins including proteases.

In the second system, a CadA-β-galactosidase fusion protein was induced following the addition of acid (Chou *et al.*, 1995). The most striking aspect of this system was the ability to grade the expression levels with respect to pH, making it possible to directly control expression levels. However, this system is obviously not suitable for pH sensitive proteins (pH 5.5).

Nevertheless, the ultimate goal in recombinant gene-expression systems is to ensure maximum activity at high cell density. Continued growth diverts nutrient into biomass production rather than product formation. Consequently, protein production in a non-growing dense population is ideal. The use of starvation induced promoter systems renders this proposition feasible. Obviously, such systems also have disadvantages including a lack of energy and building blocks for protein synthesis during nutrient limitation and stress-response problems such as increased production of proteases. These problems associated with recombinant protein expression in stationary phase have been poorly characterised in the literature. These will be discussed further in section 1.3.1. The characterisation of stationary-phase promoters under fermentation conditions will be a focus of this thesis.

1.3.1 Recombinant protein production under control of stationary phase inducible promoters

E. coli bacteria possess the ability to sense and react to changes in their environment. Between thirty to fifty new proteins are induced in stationary phase following carbon starvation (Groat *et al.,* 1986). Genes induced by carbon starvation can be sub-divided into two groups, known as the *cst* (carbon starvation) genes and the *pex* (post exponential) genes. The *cst* genes are characterised by requirement for cAMP for their induction (Schultz *et al.*, 1988). In work by Tunner *et al.* (1992), the *lacZ* gene encoding the reporter protein β-galactosidase was placed under control of the *cst-1* gene promoter. Consequently, if the glucose concentration is high in the growth phase, then cellular levels of cAMP will be low, thereby giving low promoter activity. By contrast, at high cell density with correspondingly low glucose levels, cAMP levels and hence promoter activity increases. It has been shown that induction of the promoter follows glucose exhaustion in the batch fermentation mode (Tunner *et al.*, 1992).

Although it is possible to produce recombinant proteins in starved cells, protein production cannot continue indefinitely. Loss of rRNA and ribosomes imposes severe limitations on the translational machinery of the cell. Normally, when cells are actively growing RNA content is proportional to the growth rate and RNA levels exceed requirements. By contrast, twenty to thirty percent of the cellular RNA is lost during the first four hours of starvation (Dawes, 1976). Polysomes are broken down into individual ribosomes. The 70S monosomes are further degraded into 30S and 50S subunits by RNase I and RNase II (Davis *et al.,* 1986). 16S rRNA is degraded by polynucleotide phosphorylase. Increased protease activity during stationary phase occurs primarily following RNA breakdown (Dawes, 1976) to release amino acids, which are subsequently used in the synthesis of new proteins (Reeve, 1984). Bacteria can survive in this "dormant" state for extended periods of time. However, when cells are producing recombinant proteins in stationary phase, a feeding strategy must be established to ensure a sufficient supply of energy and hence adequate levels of rRNA, and to eliminate stress responses such as increased protease production. Recent work by Tunner *et al.* (1992) showed that the addition of acetate following glucose exhaustion in fed-batch mode

resulted in a three-fold increase in β-galactosidase levels, compared with batch fermentation in minimal media.

Cells use energy for functions other than growth. Apart from turnover of macromolecules as mentioned previously. These functions include cellular motility, maintenance of cellular osmolarity, and the maintenance of trans-membrane gradients. These constitute the maintenance energy demand. Most maintenance requirements are for energy use and can therefore be termed as ATP maintenance requirements. Maintenance represents a significant drain on metabolic sources, particularly at low growth rates (Vaarma and Palsson, 1993). Furthermore, when additional burdens are placed on the stationary-phase cells by introducing multiple copies of genes, even more energy is required. In the work of Tunner *et al.* (1992), this energy was in part provided by an acetate feeding strategy. ATP generation from acetate fuelling is 4.67 molecules of ATP per molecule of acetate. By contrast, glucose provides 18.7 molecules of ATP (Varma and Palsson, 1993). More energy is therefore supplied by feeding glucose compared to acetate, and one would expect the result to be higher levels of recombinant protein. Of course, this is balanced by the effect of intracellular ATP concentrations on promoter activity.

Any feeding strategy for recombinant protein production under control of the *cst* gene promoters will be difficult to implement in a controlled way, as these promoters are cAMP dependent. Another group of starvation-inducible genes, namely the *pex* genes, are either independent of, or are negatively regulated by, cAMP (Schultz *et al.,* 1988). The use of these promoters in fermentation experiments is compromised by promoter activity during the batch growth phase. However, these problems may be overcome by adding substances that inhibit promoter activity during the growth phase. Some evidence in the literature suggests that this is indeed possible (Mulvey *et al.,* 1990). In this thesis, protein expression under control of the stationary phase inducible *pex* gene promoters is characterised, because of their potential for use in fed-batch fermentations. An important consideration is the selection of a suitable reporter protein, discussed in the next section.

1.4 Reporter gene

The effectiveness of a bacterial expression system can be measured by linking the promoter sequence to a reporter gene. One commonly used reporter is *cat* which encodes the easily-detectable enzyme Chloramphenicol Acetyl Transferase (CAT). Promoter expression is monitored by measuring enzyme activity or by the total amount of CAT protein produced. Enzyme-linked immunosorbant assays (ELISA) have been developed for the quantitative estimation of total CAT. As enzyme activity is highly dependent on temperature and correct enzyme conformation, a total-CAT-protein measurement provides a more appropriate estimate of promoter strength. Another advantage using the CAT-ELISA assay is that fragments containing an intact antigenic epitope will be detected. By contrast, only intact molecules show enzymatic activity and protease activity becomes a greater concern if accurate quantitation of expression levels is required.

Meaningful determination of intracellular protein concentrations relies on a disruption method that completely releases the intracellular reporter protein without degradation. Samples must also be treated, stored and handled so that degradation is avoided. Detailed studies to resolve these important issues are conducted in this thesis (chapter 4), and overcame many of the problems associated with the semi-quantitative analyses often reported in the literature. Finally, CAT protein is soluble. However, *E. coli* host proteins, can form insoluble aggregates when overproduced under certain conditions (Georgiou *et al.*, 1986). The possibility of such formations must be excluded (section 5.1).

Expression of a reporter gene is only indicative and does not necessarily reflect expression of other structural genes. For example, the strength of the T7 promoter is believed to be superior to that of the *lac* promoter. However, compared to the *lac* promoter, the T7 promoter is far less efficient in driving the expression of protein-encoding genes such as *cat*, *neo* and *lacZ* (Lopez *et al.*, 1994). A possible reason for this observation is the de-synchronisation of the T7 RNA polymerase which travels faster than the ribosomes translating the reporter gene mRNA. This results in a lower polypeptide yield. Furthermore, transcriptional and translational efficiencies, and factors such as plasmid

stability, the mRNA, and cloned product depend on the particular protein produced, as discussed in section 1.5.

1.5 Maximising production of recombinant protein

The amount of recombinant protein expressed by a bacterial expression systems is determined by:

- Transcriptional efficiency (promoter strength and regulation of transcription);
- Translational efficiency (binding to RBS, secondary structure of mRNA, and distance between the RBS and ATG);
- Stability of mRNA (structural, nucleases);
- Stability of cloned product (proteases, structure of protein);
- Plasmid stability (structural/segregational);
- Gene-dosage effect (plasmid copy number, regulation of replication);
- Host cell (metabolic activity, host-vector interaction);
- Transcriptional termination signals and media;
- Growth conditions.

Transcription is initiated by the formation of a complex consisting of RNA polymerase-DNA (promoter sequence) and a sigma factor. The rate of transcription is determined by the strength of the promoter. For promoters recognised by the common sigma factor σ^{70}, this strength is proportional to the degree of homology to the consensus [TATAAT (-10) TTGACA (-35)]. For example, a combination of the -35 consensus region of the *trp* promoter with the -10 region of the *lac* promoter (de Boer *et al.*, 1983) resulted in the *tac* promoter that proved to be about eleven times stronger than the *lac* promoter.

Spacing between the -35 and -10 regions is another determinant of promoter strength. For example, the *trc* promoter was created by introducing 1 bp to the existing 16 bp spacing in the *tac* promoter (Amann *et al.*, 1983). When cloned upstream to the 4.5S RNA gene, the *trc* promoter showed 89% activity relative to *tac*, as determined by

quantification of the cloned RNA (Brosius *et al.*, 1985). By contrast, Mulligan *et al.* (1985) found the *tac* promoter to be half as active as the *trc* promoter *in vivo*.

The rate of transcription also depends on the number of gene copies. High copy-number systems typically produce more recombinant protein than low copy-number systems. However, this correlation often fails because high-copy-number systems are often unstable. Furthermore, it was recently reported that when copy number exceeds 200, a competition between plasmid derived mRNA and ribosomal mRNA results in a limited amount of ribosomal protein and thereby insufficient translation (Wood & Perretti, 1990).

Translation is initiated by formation of an "initiation complex". This complex consists of the 30S ribosomal subunit which, along with initiation factors, binds to the mRNA molecule and fMET-tRNA. Binding requires a RBS on the mRNA complementary to the anti-Shine-Dalgarno sequence, located at the 3' end of the 16S rRNA, in the 30S sub-unit. It is generally believed that a better complementarity between the two gives a more efficient RBS (Hui and de Boer, 1987; Jacob *et al.*, 1987). The initiation complex associates with the 50S ribosomal sub-unit to form the 70S initiation complex and chain elongation begins. Access to the initiation site is reduced if the RBS is masked by secondary structures. The extent of this structure may be regulated in response to growth conditions (Sørensen and Neuhard, 1991).

Another important parameter affecting efficiency of translation is the spacer sequence separating the RBS and the AUG start codon. Optimisation of this RBS/spacer sequence is dependent on the promoter used to drive gene expression (Jørgensen, 1993, unpublished results). The gene encoding the fusion-protein LR^3-IGF(I) has previously been optimised for the production of the protein under control of the *trc* promoter (King *et al.*, 1992). Substituting the *trc* promoter with the T7 promoter in a temperature induced expression system (Tabor and Richardson, 1985) did not lead to any detectable levels of LR^3-IGF(I) (Jørgensen, 1993, unpublished results). By altering the RBS and spacer region, LR^3-IGF(I) was successfully expressed from the T7 promoter, in amounts comparable to that obtained using a the *trc* promoter based expression system (Jørgensen, 1993, unpublished results).

Messenger RNA is degraded by endoribonucleolytic and exoribonucleolytic activities. Most natural mRNAs are protected against degradation by stable secondary structures (stem-loop) at their 3' end and decay is normally initiated by endonucleatic cleavage (Higgins *et al.*, 1993). No 5'-3' exonucleases have been identified in bacteria. As a consequence, functional half-life of mRNAs may be improved by altering the possibility for stem-loop formation or by substituting the existing RBS with a more efficient one that possesses a higher degree of homology to the 3' end of the 16s rRNA (Vind *et al.*, 1993).

Proteolysis of heterologous proteins can be a major problem. *E. coli* contains at least 24 identified endoproteases (Harcum and Bentley, 1993) which are produced constitutively or induced by cellular signals. Stress factors such as heat, oxidation, glucose and oxygen starvation increase the production of proteases. For example, a temperature up-shift will induce expression of the Lon protease and the proteolytic sub-unit of Clp protease (Clp_p) (Damarau & John, 1993). Furthermore, accumulation of abnormal or denatured protein generally results in increased activity of the Lon protease (Goff & Goldberg, 1985). The use of single-protease deficient strains has been examined for ability to produce proteolytically sensitive proteins. However, this approach has not been very successful due to problems of low yield and slow growth (Georgiou and Bancyx, 1996). Furthermore, most proteins are degraded by more than one protease. One would expect that increasing the number of gene deletions to produce multiple-protease deficient strains would result in non-viable cells. This expectation is based on the fact that proteolytic activity is necessary for the creation and recycling of cell building blocks. Surprisingly, a triple-protease deficient strain resulted in improved yield compared to not only its wild-type but also a similar single- and double-protease deficient strain (Georgiou and Bancyx, 1996).

Problems of plasmid segregational instability may be diminished by incorporation of a selection gene, eg. Ampicillin, by using suicidal mechanisms such as hok/sok (Gerdes, 1988), by incorporation of a partition locus, or by immobilisation of cells. Structural instability may be reduced by employing a *rec*A host. However, *rec*A strains generally have slow growth rates in minimal media. This make them less attractive for fermentation

purposes. For example, the *E. coli* DH5 strain has a specific growth rate of 0.4 h^{-1} when cultivated in minimal media (Jørgensen, (1992), unpublished results).

Finally, successful production of proteins in bacterial expression systems requires a strong transcription terminator to prevent transcriptional read-through (Brosius, 1984). The terminator hinders the RNA polymerase from transcribing beyond the gene's end and thus producing long mRNAs. Long mRNAs can interfere with the origin of replication leading to plasmid instability.

Recombinant protein expression can be maximised by optimisation of the above parameters. Indeed, much published work has characterised bacterial expression systems qualitatively. However, few attempts have been made to quantitiate the response and dynamics of expression systems under varied environmental conditions. Most importantly, the identification of transcriptional versus translational limitations will direct the focus of optimisation. A quantitative characterisation requires the use of accurate estimates, and extensive research to prove that estimates are in fact accurate is needed. The optimisation, validation and development of analytical methods to quantify relevant variables are all investigated in this thesis (Chapter 4), and used to characterise the *tac* (Chapter 5) and *pex* gene (Chapter 6) promoters.

1.6 Specific aims of this thesis

The prime aim was to characterise CAT protein expression under the control of stationary phase inducible *pex* promoters, and to quantitatively compare expression levels with those achieved under the control of IPTG-inducible promoters (eg. *tac*). The specific aims of this thesis include:

1. Construction of expression systems for CAT protein under control of the *katE* and *katF* stationary phase inducible gene promoters (members of the *pex* gene family):

    - Identification of potential inducers (glucose, acetate, o-hydroxybenzoate);
    - Characterisation of the stationary-phase systems through batch- and fed-batch-fermentation experiments;

- Description of CAT protein expression under control of the *katE* promoter in a simple mathematical model, to identify the limit to increased expression.

2. Construction of expression systems for CAT protein under control of IPTG-inducible promoters:

   - Identification of the best IPTG inducer concentration for a batch system;
   - Optimisation of the level of expression by examining oxygen availability, growth phase of the inoculum, and the starting OD in the fermenter;
   - Comparison of expression levels in shake flask and fermentation experiments, to establish whether promoter screening for large-scale work can be done using shake-flasks;
   - Determination of the rate-limiting steps in protein synthesis, by relating CAT protein to both *cat* mRNA and 16S rRNA;
   - Description of the IPTG induced expression systems in simple mathematical models to establish the relative efficiencies of transcription and translation.

3. Comparison of the above expression system to evaluate the efficiency of the stationary-phase induced system.

4. Finally, optimisation, validation and development of analytical methods to quantify relevant variables:

   - Detailed study of CAT, total soluble protein and RNA recovery;
   - Detailed study of CAT , total soluble protein and RNA integrity during storage, and RNA integrity during sampling and extraction;
   - Detailed study of the accurate quantitative detection of CAT protein, *cat* mRNA, 16S rRNA and plasmid DNA.

The work undertaken to achieve all of the above aims is summarised in Figure 1.1. Plasmids encoding IPTG- or stationary-phase-inducible promoters, as well as templates for the *in vitro* production of RNA standards is described in Chapter 3. The optimisation, validation and development of analytical assays are described in Chapter 4. The maximal

performance of an IPTG induced system is explored and characterised in a simple mathematical model in Chapter 5. Finally, the stationary phase induced systems are characterised and compared with the IPTG systems in Chapter 6.

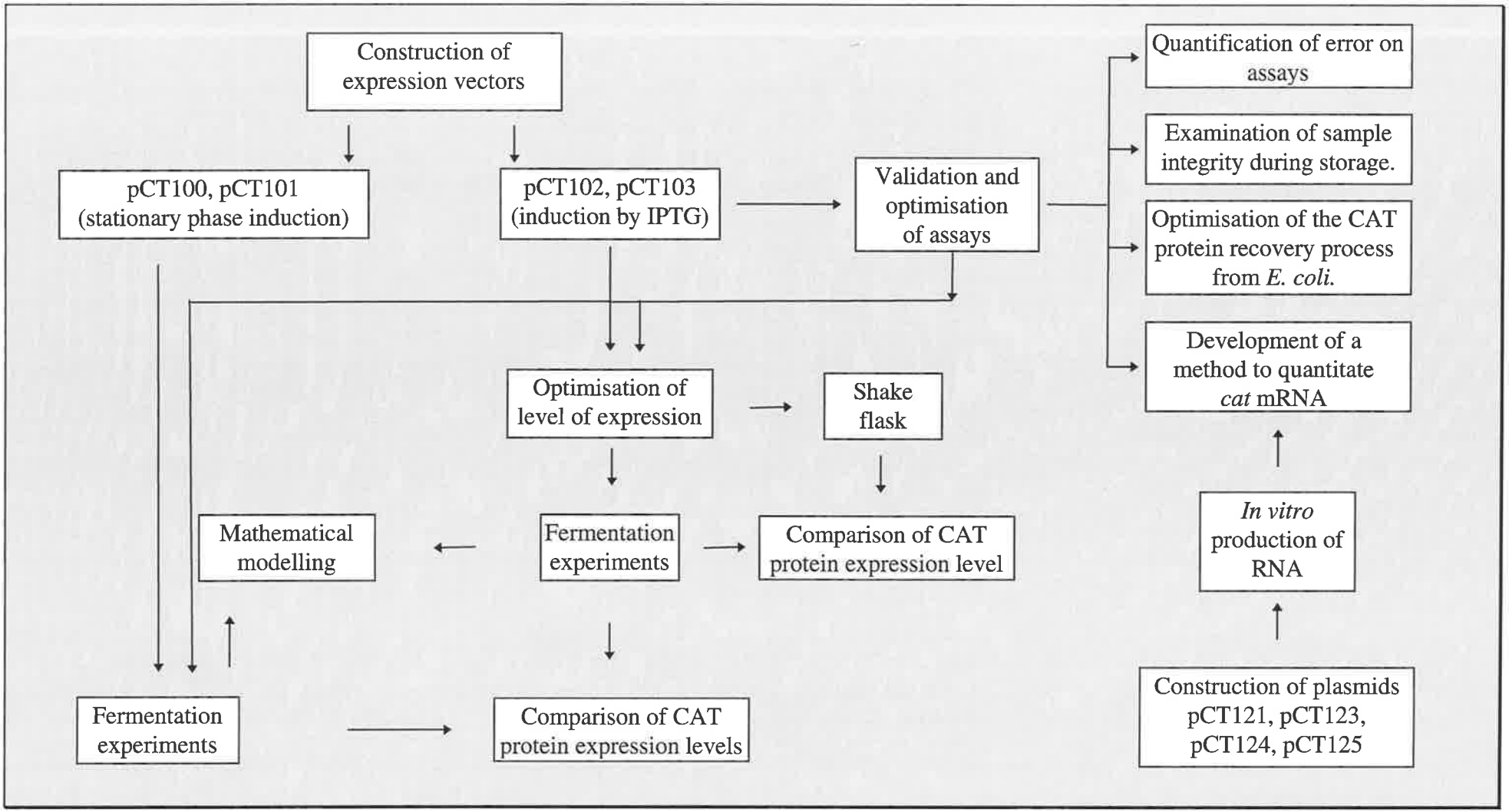

Figure 1.1. Outline of experimental work undertaken to achieve aims of this thesis.

Chapter 2

Materials and Methods

2.1 Chemicals and reagents

Chemicals were Analar grade. Unless otherwise specified chemicals were from Sigma Chemical Company (St Louis, Missouri, USA), BDH Laboratory supplies (Poole, Dorset, England), or Ajax Chemicals (Auburn, New South Wales, Australia). Phenol special grade, used for RNA work, was obtained from WAKO Pure Chemical Industries (Osaka, Japan). X-ray film, X-omat XK-1, was obtained from Kodak (Rochester, New York, USA). The Riboprobe system and RQ1 RNase-free DNase was obtained from Promega corp. (Madison, Wisconsin, USA). Tris base, IPTG, Digoxigenin-RNA labelling mix, Anti-Digoxigenin POD Fab fragments, BM Chemiluminescence Western blotting reagents, 16S/23S ribosomal RNA from *Escherichia coli* MRE 600, and blocking reagent for nucleic acid hybridisation and detection were obtained from Boehringer Mannheim (GmbH, Mannheim, Germany). X-gal was purchased from Progen Industries (Darra, Queensland, Australia). dNTP's were from Pharmacia (Uppsala, Sweden)

Milli Q water was water purified using a Milli Q purification system (Millipore Corp.) and used to prepare buffers and reagents for DNA and RNA manipulations. Additional reagents and buffers were prepared with deionised water.

All solutions for RNA work were treated for RNAse by adding DEPC to a final concentration of 0.1% (v/v), incubating overnight and autoclaving, whenever possible. Tris-containing solutions were made in DEPC-treated water and autoclaved.

2.2 Growth media

2.2.1 Rich media

Nutrient broth was Oxoid (Oxoid Ltd, London, England). Luria broth was prepared as described by Miller (1972) and Terrific broth was prepared as described by Maniatis (1982). McConkey agar base (Difco, Detroit, Michigan, USA) was prepared as described by the manufacturer. After autoclaving, the agar base was supplemented with a filter-sterilised solution of 20% (w/v) galactose to a final concentration of 1% (w/v).

Nutrient agar (NA) plates were prepared by adding Agar (Oxoid Ltd, London, England) (15 g L^{-1}) to nutrient broth before autoclaving.

2.2.2 Minimal media

2.2.2.1 Batch fermentation media

Salt solution (2.58 g L^{-1} NH_4Cl, 2.54 g L^{-1} KH_2PO_4, 4.16 g L^{-1} Na_2HPO_4, 1.94 g L^{-1} K_2SO_4) was prepared in the fermenter or shake flask and autoclaved. Trace metal solution was added after autoclaving to give a final concentration of 20 mg L^{-1} $FeSO_4.7H_2O$, 5.1 mg L^{-1} $MnSO_4.H_2O$, 8.6 mg L^{-1} $ZnSO_4.7H_2O$, 0.76 mg L^{-1} $CuSO_4.5H_2O$, 88.0 mg L^{-1} Na_3-citrate, 0.04 mL L^{-1} HCl (conc.). Glucose solution was autoclaved separately and added to give a final concentration of 3.33 or 10 g L^{-1} D-glucose and 0.67 g L^{-1} $MgSO_4$. Thiamine was dissolved in water, filter-sterilized (0.2 μm) and added to a final concentration of 40 μg mL^{-1}.

2.2.2.2 Fed-batch fermentation feeding solution

1.33 g L^{-1} KH_2PO_4, 18.67 g L^{-1} Na_2HPO_4, 11.27 g L^{-1} K_2SO_4, 2.07 g L^{-1} $MgSO_4.7H_2O$, 275 g L^{-1} glucose.

2.2.2.3 M9 minimal media.

2.58 g L^{-1} NH_4Cl, 2.54 KH_2PO_4, 4.16 g L^{-1} Na_2HPO_4, 1.94 g L^{-1} K_2SO_4. (Miller, 1972).

2.2.3 Antibiotics

Antibiotics were added to broth and solid media as required at the following concentrations: Ap 100 μg m L^{-1}, Cm 25 μg m L^{-1}. Ampicillin was added to the fermenters at a concentration of 50 μg m L^{-1}. NA plates containing antibiotics were prepared by adding the appropriate antibiotics from sterile stock solution to the medium before pouring the plates.

2.3 *E. coli* bacterial strains used in this study

Bacterial strains used in this thesis are listed in Table 2.1. All strains were stored as suspensions in glycerol-based medium (15% v/w glycerol, 1% Bacto-peptone) and maintained at –70°C. Cultures were also maintained *in vacuo* in sealed glass ampoules.

2.3.1 Selection of a JM101 *Ton*A mutant

T7 and T5 phage both use TonA as a receptor. TonA mutants resistant to T7 phage were isolated using the following procedure. Phage T5 was propagated in *E. coli* P400 as follows: Phage stock was diluted 10^{-2}, 10^{-4} and 10^{-6}. 100 µl of each dilution was mixed with 100 µl of *E. coli* P400 and incubated at 37°C for 10 min. Soft agar (NB:NA, 1:1) (3 mL) was added and the suspension poured onto a NA plate and incubated at 37°C overnight. One plaque was selected and transferred to NB (1 mL) containing 10 µl chloroform. 100 µl of this suspension was added to 100 µl of *E. coli* P400 as above. Phage were collected by collecting the soft agar overlay and suspending this in 10 mL NB containing 100 µl chloroform. The suspension was vortexed and the cellular debris removed by centrifugation.

A colony of *E. coli* JM101 resistant to T5 was selected as follows. Pure T5 was swabbed onto half of a NA plate. *E. coli* JM101 culture was then streaked onto both halves of this plate and incubated at 37°C until colonies were visible. A single resistant colony was isolated.

2.4 Plasmids used in this study

Plasmid clones and vectors used in this study are listed in Table 2.2.

Table 2.1 *E. coli* bacterial strains used in this study.

Strain	Source	Genotype/Phenotype
DH5	P Manning[1]	F- *endA1 recA1 hsd*R17($r_K^-m_K^+$) *deoR, thi1 supE44 gyrA96 relA1* λ-.
DH5α	P Manning[1]	F-,Φ80 *lacZ*ΔM15 Δ(*lacZYA-argF*)U169, *endA1 recA1 hsd*R17($r_K^-m_K^+$) *deoR thi1 supE44 gyrA96 relA1* λ-.
CB806	Schneider & Beck (1986)	F- Δ*lacZ lacY*+ *galK rpsL thi recA56 phoA8*.
P400	P Reeves[2]	*thr1 leu6 proA2 lacY1 supE44(?) galK non9 rpsL31 xyl5 mtl1 argE3 thi1 ara14* delete *His4*
JM101	New England Biolabs	F' *traD36 proAB*+ *lacI*q *lacZ*ΔM15 Δ(*pro-lac*) *supE mrcA thi* λ-).
JM101	This thesis	F' *traD36 proAB*+ *lacI*q *lacZ*ΔM15 Δ(*pro-lac*) *supE mrcA thi* λ-, *TonA*.
JM101	ATCC 33876	F' *traD36 proAB*+ *lacI*q *lacZ*ΔM15 Δ(*pro-lac*) *supE mrcA thi* λ-).
DH1	P Manning[1]	F- *supE44 recA1 endA1 gyrA96 (Nal*r*) thi1 hsdR17 ($r_K^-m_K^+$) recA1 spoT? rfbD1?*.

1. Department of Microbiology and Immunology, The University of Adelaide
2. Department of Microbiology, University of Sydney.

Table 2.2. Plasmids used in this study.

Plasmid	Genotype	Source	Reference
pAMkatE72	*katE::cat, bla*	PC Loewen[1]	Ossowski *et al* (1991)
pMF5	*katF::cat, bla*	PC Loewen[1]	Mulvey & Loewen (1989)
pPM1641	lac_p*::cat, galK, bla*	A Barker[2]	
pPM3024	*cat, galK, bla*	S Williams[2]	Williams & Manning (1991)
pEV601	*trc, bla*	B Egan[3]	
pPM1602	*cat, bla*	A Barker[2]	
pKK223-3	*tac, bla*	Pharmacia	Brosius & Holy (1984)
pGEM7ZF(+)	T7, SP6, *bla*	Promega Corp. Madison, Wisconsin. USA	
pCT100	*katEp::cat, galK, bla*	This thesis	
pCT101	*katFp::cat, galK, bla*	This thesis	
pCT102	*trc::cat, galK, bla*	This thesis	
pCT103	*tac::cat, galK, bla*	This thesis	
pCT121	T7, SP6, *cat, bla*	This thesis	
pCT123	T7, SP6, *16S rRNA, bla*	This thesis	
pCT124	T7, SP6, *bla*	This thesis	
pCT125	T7, SP6, *bla*	This thesis	

1. Department of Microbiology, University of Manitoba, Canada.
2. Department of Microbiology and Immunology, The University of Adelaide.
3. Department of Biochemistry, The University of Adelaide.

2.5 Oligonucleotides

Oligonucleotides used for PCR and sequencing are listed in Table 2.3.

2.6 Enzymes

Restriction enzymes were obtained from Boehringer Mannheim GmbH (Mannheim, Germany) or New England Biolabs (Beverley, MA, USA). T4 DNA ligase, Klenow fragment of DNA polymerase I, alkaline phosphatase (molecular biology grade) and lysozyme were from Boehringer Mannheim GmbH (Mannheim, Germany). *Taq* DNA polymerase (AmpliTaq) was obtained from Perkin Elmer Cetus Corp. (Norwolk CT,USA). DNase (RQ1) was from Promega Corp. (Madison, Wisconsin, USA).

2.7 Transformation

E. coli strains were transformed with plasmid DNA essentially according to the method described by Hanahan (1983).

2.7.1 Preparation of competent cells

E. coli strains were made competent for transformation with DNA as follows: An overnight culture in NB was subcultured 1:20 in 100 mL Φ_b broth (5g L^{-1} Bacto yeast extract, 20 g L^{-1}Bacto tryptone, 5 g $L^{-1}MgSO_4$) and cultured with shaking to an OD_{600} = 0.6. The cells were chilled for 5 min on ice, pelleted by centrifugation (8670×g, 5 min, 4° C), resuspended in 0.4 volume of ice-cold Tϕb_1 solution (30 mM Kac, 100 mM RbCl, 10 mM $CaCl_2.2H_2O$, 50 mM $MnCl_2.4H_2O$, 15% glycerol, adjusted to pH 5.8 with 0.2 M acetic acid), incubated on ice for 5 min. Cells were then pelleted by centrifugation (5640×g, 5 min, 4°C) and resuspended in 1/25 volume ice-cold Tϕb_2 solution (10 mM

Table 2.3. Oligonucleotides used in this study.

Number	Oligonucleotide Sequence
#362	5' CGGAGCGCAGCAGAGG 3'
#746	5' TAATACGACTCACTATAGGG 3'
#823	5' GCTAGTTATTGCTCAGCGG 3'
#719	5' GGAGGTAACATATGTTCCCAGC 3'
#720	5' ACAGCCAAGCTTGCATCATC 3'
#942	5' GTCTTCCAGAGAATT 3'
#B7420	5' TACATTGAGCAACTGACTGAA 3'
#866	5' GGAAGCTCATATGGAGAAAA 3'
#B7421	5' GTGCTGCAAGGCGATTAAGT 3'
#523	5' CCCGGATCTGATCATGGCTCAGATGGAA 3'
#534	5' CCCCTGCAGACTTCATGGAGTCGAGTTGC 3'
#M13 Forward	5' GTTTTCCCAGTCACGACG 3'
#BCAT	5' Biotin-ATTCTGCCGACATGGAAG 3'

MOPS, 75 mM $CaCl_2$, 10 mM RbCl, 15% glycerol, adjusted to pH 6.5 with 1 M KOH). The cells were allowed to stand at least 15 min before the addition of DNA or were snap frozen and stored at –70°C in 200 µL aliquots.

2.7.2. Transformation procedure

Competent cells were thawed on ice, mixed with DNA, and left on ice for at least 15 min. Cells were heat-shocked at 42°C for 90 seconds, placed on ice for 1 min and then 800 µL Φ_b broth (section 2.8.1) was added, followed by incubation at 37°C for 45 min. Cells were concentrated by centrifugation and resuspended in a small volume of Φ_b broth and plated onto selection plates.

2.8 DNA extraction procedures

2.8.1 Small scale plasmid DNA preparation

Small-scale plasmid preparation was performed using the method described by Maniatis (1982).

2.8.2 Caesium chloride gradient purification of plasmid

An overnight broth culture (250 mL) was harvested by centrifugation (8670×g, 10 min). The pellet was resuspended in 2.4 mL of solution 1 (50 mM glucose, 25 mM Tris-HCl pH 8, 10 mM EDTA) and 0.4 mL of 20 mg mL^{-1} lysozyme (Boehringer Mannheim GmbH, Mannheim, Germany) was added. Cells were incubated at room temperature for 10 min. Solution 2 (0.2 M NaOH, 1% SDS) (5.5 mL) was added and the suspension was incubated in ice for 5 min. Solution 3 (3 M KAc, 2M acetic acid, pH 4.8) (2.8 mL) was added and the suspension was incubated for 15 min on ice. The suspension was centrifuged (14900×g, 10 min), phenol:chloroform:isoamyl alcohol (25:24:1) extracted and separated by centrifugation in a bench-top centrifuge. DNA was precipitated with 0.6 volume of isopropanol, left for 10 min at room temperature and collected by centrifugation

(39200×g, 5 min). The pellet was washed in 70% (v/v) ethanol and dried. The plasmid DNA was further purified by re-suspension in 2.8 mL TE buffer (0.01 M Tris-HCl pH 8.0, 0.01 M $MgCl_2$) and centrifugation in a CsCl-EtBr gradient. CsCl (2.9 g) and 0.3 mL of EtBr (10 mg/mL) were added to the dissolved DNA. Additional crystals of CsCl were added to obtain a refractive index of 1.3860. The plasmid DNA was separated in 3.9 mL ultracentrifuge tubes in a Beckman Optima TLX Ultracentrifuge. The lower band consisting of covalently-closed circular plasmid DNA was withdrawn with a syringe. The EtBr was removed by extracting three times with isoamyl alcohol and dialysis overnight at 4°C against 1 L of TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8).

2.9 Analysis and manipulation of DNA

2.9.1 DNA quantitation

DNA was routinely determined by spectrophotometric assessment (1 OD_{260nm} unit = 50 μg/mL) or by computing densitometry (section 2.25).

2.9.2 Agarose gel electrophoresis of DNA

Agarose gel electrophoresis was performed as described by Maniatis (1982). Gels were prepared using either HGT agarose (Progen Industries Ltd., Darra, Australia), SeaKem GTG agarose (FMC Bioproducts, Rockland, ME, USA), LGT agarose (BioRad, Hercules, California, USA) or Biogel (Bio 101 Inc, CA, USA). Gels were normally run in TAE buffer (4.82 g L^{-1} Trizma base, 1.64 g L^{-1} NaAc, 0.336 g L^{-1} EDTA). Gels were stained in EtBr, destained in water, and DNA fragments were visualised by trans-illumination with short UV light (λ_{254}). Fragments to be isolated for cloning purposes were illuminated with medium UV light (λ_{305}) for no longer than 60 secs.

Gels for quantitation of plasmid DNA were run in TBE (10.8 g L^{-1} Tris-base, 5.5 g L^{-1} Boric acid, 0.93 g L^{-1} EDTA).

2.9.3 Isolation and purification of gene fragments

Plasmids containing the desired DNA fragment were digested with the appropriate restriction enzymes and fractionated on either LGT (BioRad, Hercules, California, USA), GTG (FMC Bioproducts, Rockland, ME, USA) or Biogel (Bio 101 Inc, CA, USA) electrophoresis gels. The desired fragment was cut out of the gel and purified using a Mermaid kit (Bio 101, Inc., CA, USA) (if smaller than 1 kb) or a Geneclean kit (Bio 101, Inc., CA, USA) (if larger than 1 kb).

2.9.4 Determination of restriction fragment size

The sizes of restriction enzyme fragments were estimated by comparing their relative mobility with that of *Eco*RI digested *Bacillus subtillus* bacteriophage SPP1 DNA (Bresatec, Adelaide, Australia).

2.9.5 Dephosphorylation of DNA using Alkaline Phosphatase

Calf intestinal phosphatase (CIP) was added to restricted vectors to remove 5'-phosphate groups and prevent self-ligation. Two units of the calf intestinal phosphatase was added directly to digested DNA together with dephosphorylation buffer and incubated for 30 min at 37°C. The reaction was terminated by addition of EDTA to a final concentration of 5 mM, followed by heating at 70°C for 1 hour. The volume was made up to 100 µL with MQ water and the mixture was extracted with Tris-saturated phenol:chloroform (1:1). DNA was precipitated with 2 volumes of 100% ethanol. DNA was recovered by centrifugation for 15 min in a microfuge. The pellet was washed in 70% (v/v) ethanol and dried. DNA was resuspended in MQ and stored at 4°C.

2.9.6 End-filling with Klenow fragment

Protruding ends were end-filled using Klenow fragment of *E.coli* DNA polymerase I. DNA in the restriction buffer was mixed with dNTP's to a final concentration of 50 μM, DTT to a final concentration of 1 mM, and 1 unit of Klenow. The mixture was incubated at 37°C for 30 min.

2.9.7 *In vitro* cloning

Ligation reactions were performed using insert and vector in a ratio of approximately 3:1. Cohesive ends ends were melted at 45°C for 5 min, then cooled on ice. Four units of T4 ligase was added together with T4 ligase buffer. The ligation mixture was incubated at room temperature for at least 2 hours, or overnight at 4°C, for "sticky end" ligations, or at 37°C for two hours followed by incubation at room temperature for at least 2 hours for "blunt end" ligations.

2.10 Synthesis of oligodeoxynucleotides

Oligonucleotides were synthesised using reagents purchased from Applied Biosystems or Ajax Chemicals (acetonitrile). Synthesis was performed on an Applied Biosystem 381 DNA synthesiser. Oligonucleotides were butanol extracted and stored at –20°C until use.

2.11 Polymerase Chain Reaction (PCR)

2.11.1 PCR amplification of *E. coli* DNA encoding 16S rRNA

PCR reactions were performed in a total volume of 50 μL. The reaction mixture consisted of PCR buffer (50 mM KCl, 10 mM Tris, pH 8.3, 2.5 mM $MgCl_2$), dNTP (200μM), primer #523 and #534 (0.2-1.0 μM), experimental DNA template (DH1

chromosomal DNA, 265 ng), AmpliTaq (1.25U). The mixture was overlayed with mineral oil prior to thermal cycling. The reaction mixture was incubated at 94°C for 3 min, followed by 5 cycles consisting of 1 min at 94°C (denaturation), 30 seconds at 52°C (annealing), 2 min at 72°C (extension) and then 20 cycles consisting of 1 min at 94°C, 30 seconds at 55°C, 2 min at 72°C and then a final extension at 72°C for 4 min. Product (5 μ L) was analysed on 2% GTG agarose electrophoresis gel. Contaminating mineral oil was removed from the reaction mixture by chloroform extraction. DNA was purified by further extraction with phenol:chloroform (1:1).

2.11.2 PCR amplification of *E. coli* DNA encoding CAT

PCR was carried out as in 2.17.1 in a 100 μL reaction volume using the primers #M13 Forward and BCAT. DNA was amplified from the template pCT121 for 30 cycles consisting of 30 sec at 94°C, 30 sec at 55°C, and 30 sec at 72°C.

2.12 DNA sequence analysis

2.12.1 Dye-Deoxy terminator sequencing

Sequencing of most constructs was conducted using PRISM Ready Reaction Dye-Deoxy terminator kit obtained from Applied Biosystems.

Template DNA was purified by PEG precipitation procedure as follows:
RNase (0.1 μg) was added to 20 μL DNA preparation, and incubated for 30 min. The solution was cooled on ice. NH_4Ac (7.5 M, 20 μl) and 40 μL isopropanol was added and the solution was further incubated for 15 min at room temperature. The mixture was then centrifuged in a microfuge at room temperature for 15 min at 12000×g, washed in first 70% ethanol and then 100% ethanol, and resuspended in 32 μL MQ water. NaCl (8.0 μl, 4 M) and 40 μL 13% (v/v) PEG-8000 was added and the solution was incubated at 4°C overnight. DNA was pelleted by centrifugation (12000×g, 15 min, 4°C) in a microfuge. The pellet was rinsed in 70% ethanol, dried and resuspended in 20 μL MQ water.

DNA was prepared for asymmetric PCR as described by the Prism manual. The mixture was overlayed with mineral oil prior to thermal cycling. Tubes were incubated in a thermal cycler preheated to 96°C, followed by 25 cycles consisting of 30 sec at 96°C (denaturation), 15 seconds at 50°C (annealing), and 4 min at 60°C (extension). After thermal cycling, the reaction volume was adjusted to 100 µL and extracted twice with phenol:water:chloroform (68:18:14) (v:v:v). DNA was precipitated with 300 µL of 100% ethanol and 30 µL of 3 M NaAc followed by incubation at –20°C over night. DNA was pelleted at 12000×g, 20 min, washed in 70% ethanol, and dried. Sequence analysis was carried out using a Applied Biosystems 373A DNA sequencer.

DNA sequence data were analysed using DNASIS version 6 and Clustal V multiple alignment computer software. The obtained sequence data and the actual sequence for the constructed plasmid were transported into DNASIS, in separate files. The two sets of data were together transported into the Clustal V program, aligned, and compared with respect to homology.

2.12.2 Dye-primer sequencing

pGEM7Zf(+) based subclones were sequenced using PRISM Ready Reaction Dye-Primer and Reverse Dye-Primer kit obtained from Applied Biosystems. Template DNA was purified on a Quiagen column and prepared for asymmetric PCR as described by the Prism manual. The mixture was overlayed with mineral oil prior to thermal cycling. Tubes were placed in a thermal cycler preheated to 95°C, followed by 15 cycles consisting of 30 sec at 95°C (denaturation), 30 sec at 55°C (annealing), and 60 sec at 70°C (extension). After thermal cycling the reaction mixture was recovered and DNA precipitated by adding 80 µL of 100% ethanol followed by incubation at –20°C for 30 min. DNA was pelleted at 39200×g, 30 min, washed in 70% ethanol and dried. Sequence analysis was carried out using a Applied Biosystems 373A DNA sequencer.

2.13 RNA analysis

2.13.1 RNA quantitation

Total RNA was quantified by spectrophotometric assessment (1 OD_{260nm} unit = 40 µg/mL). Samples with OD 260/280 ratios between 1.8 and 2.0 were accepted for use in hybridisation experiments. Samples with ratios below 1.8 were reextracted with phenol:chloroform (1:1). RNA on autoradiographs was quantified by scanning densitometry using a Computing Densitometer Model 300A (Molecular Dynamics, Sunnyvale, California, USA).

2.13.2 RNA extraction and purification

2.13.2.1 RNA extraction using the hot phenol method

Cells for RNA extraction were collected (9800×g, 5 min) and stored as a pellet at −70°C. Total cellular RNA was isolated using a modification of the hot phenol method described by Aiba *et al.* (1981). Cells were resuspended in 700 µL lysis buffer (20 mM NaAc, 1 mM EDTA, 0.5% SDS, pH 5.5) and transferred to a microfuge containing 500 µL phenol (65°C), equilibrated against 20 mM NaAc and 1 mM EDTA, pH 5.5. After vortexing and incubating at 65°C for 5 min the samples were centrifuged (12000×g, 2 min). The aqueous layer (500 µL) was removed and transferred into a fresh reaction tube containing 500 µL of hot phenol and 200 µL of lysis buffer. This extraction was repeated three times. RNA was precipitated by addition of 0.10 volume NaAc, pH 5.2 and 2.5 volumes of absolute ethanol followed by incubation at −70°C for at least 2 h. RNA was collected by centrifugation (39200×g, 20 min, 4°C). Samples were resuspended in 50 µL buffer (40 mM Tris-HCl, pH 7.9, 10 mM NaCl, 6 mM $MgCl_2$, 10 mM $CaCl_2$) and contaminating DNA was removed by addition of 1 µL DNase at 37°C for 30 min. Samples were purified with phenol:chloroform (1:1). RNA was re-precipitated as above and resuspended in 200 µL DEPC-treated MQ water. The integrity of extracted RNA and efficiency of DNA removal were examined by gel electrophoresis.

2.13.2.2 RNA extraction using the method of Gilman

Cells for RNA extraction (Gilman, 1989) were collected (9800×g, 5 min) and stored as a pellet at -70°C. The cell pellet was resuspended and incubated on ice for 15 min in 1 mL protoplasting buffer (15 mM Tris-HCl, pH 8.0, 0.45 M sucrose, 8 mM EDTA and 20 mM aurintricarboxylic acid) and 8 µL lysozyme (50 g L^{-1}). Protoplasts were collected (6600×g, 5 min) and the pellet was resuspended and incubated at 37°C for 5 min in 250 µL lysing buffer (10 mM Tris-HCl, pH 8.0, 10 mM NaCl, 1 mM Na_3-citrate and 1.5% SDS) with 7.5 µL DEPC. Saturated NaCl was added (125 mL), and the mixture was incubated on ice for 10 min. The insoluble fraction was removed by centrifugation (12000×g, 10 min, 4°C) and the supernatant was transferred to a new tube. Ethanol (2.5 volumes) was added, and the RNA precipitated at −70°C for a minimum of 2 h and collected by centrifugation (39200×g, 20 min, 4°C).

2.13.3 Slot-Blot hybridization analysis

Aliquots of RNA samples were adjusted to a volume of 50 µL with DEPC-H_2O and denatured in 150 µL RNA incubation solution (657 µL mL^{-1} formamide, 210 µL mL^{-1} formaldehyde (37% w/v), 133 µL mL^{-1} 10×MOPS (0.2M 3-[N-morpholino] propanesulfonic acid sodium salt), 90 mM sodium acetate, 10 mM EDTA, pH 7) at 65°C for 10 min. Samples were cooled on ice and 200 µL of 20×SSC (3.0 M NaCl, 0.3 M Na_3-citrate) was added. Samples were transferred to a positively charged Nylon membrane (Boehringer Mannheim GmbH, Mannheim, Germany) using a slot-blotting apparatus (Hoeffer, San Fransisco, California, USA). RNA was then fixed to the filter by baking at 80°C for 2 h *in vacuo*.

2.13.4 Hybridisation

Filters were placed in roller bottles and pre-hybridised in hybridisation solution (50% formamide, 5xSSC, 2% blocking reagent for nucleic acid hybridisation (Boehringer Mannheim GmbH, Mannheim, Germany), 0.1% N-lauroylsarcosine, 0.02% SDS).

Prehybridisation buffer was replaced by 10 mL fresh buffer containing the DIG-labelled RNA probe. Hybridisations were carried out for 14-16 h at 68°C.

2.13.5 RNA detection by chemiluminescence

Digoxigenin-labelled RNA probes were detected using the BM Chemiluminescence detection kit originally designed for Western blotting (Boehringer Mannheim GmbH, Mannheim, Germany). Filters were treated with blocking solution (5 % skim milk in buffer 1 (0.1 M Tris-HCl, 0.15 M NaCl, pH 7.5)) for 60 min with shaking at room temperature. The blocking solution was replaced with Buffer 1 containing Ab-POD conjugate (Boehringer Mannheim GmbH, Mannheim, Germany) in 1/5000 dilution and incubated with shaking for 1 h. Unbound conjugate was removed by washing 2 × 10 min with Buffer 1. Filter was equilibrated with PBS (1.21 g L^{-1} K_2HPO_4, 0.34 g L^{-1} KH_2PO_4, 8 g L^{-1} NaCl, pH 7.3) for 5 min. Blots were placed RNA side up in a transparent plastic bag and detected using a BM Chemiluminescence kit for Western blotting (Boehringer Mannheim). Filters were exposed to X-ray film which had been sensitised using a pre-flash unit (RPN 2051) (Amersham, UK).

2.14 Synthesis of the sense strand of *cat* mRNA

Sense strand *cat* mRNA was synthesised *in vitro* from clone pCT121 using the Riboprobe system from Promega Corp. (Madison, Wisconsin, USA) (Figure 2.1). Plasmid was prepared as template by linearising with *Mlu*I. DNA template was removed by treating with DNAse (37°C, 30 min). RNA was then purified by Phenol:Chloroform (1:1) extraction, and precipitated by addition of 0.10 volume 3 M NaAc and 3 volumes of absolute ethanol followed by incubation at −70°C for at least 2 h. RNA was collected by centrifugation (39200×g, 30 min, 4°C). Excess nucleotides were removed by resupending the RNA in 1M NH_4Ac and 3 volumes of absolute ethanol and the suspension incubated at −70°C for at least 2 h. Precipitated RNA was collected by centrifugation and the ammonium precipitation step repeated. Purified RNA was resuspended in sterile MQ

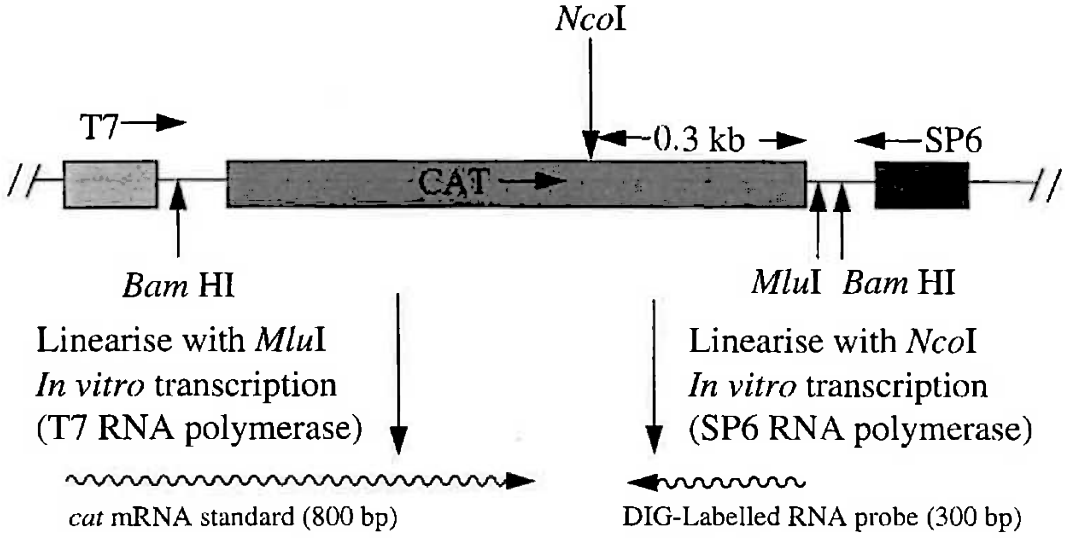

Figure 2.1. Construction of standard *cat* mRNA and DIG-labelled RNA probe using plasmid pCT121. The sense strand (standard *cat* mRNA) was synthesised *in vitro* from the T7 promoter, after linearisation of the template with *Mlu*I. This resulted in a full length transcript of the *cat* gene (0.8 Kb). The DIG-labelled anti-sense RNA probe was synthesised from the SP6 promoter after linearisation of the template with *Nco*I. The resulting probe was 0.3 Kb long.

water. The length of the standard sense strand was identical to that of the target *cat* mRNA (900 bp), as determined by gel electrophoresis.

2.15 Synthesis of DIG-labelled *cat* RNA probe

The anti sense probe was synthesised by *in vitro* transcription from plasmid pCT121 using the Riboprobe system from Promega Corp. (Madison, Wisconsin, USA) (Figure 2.1). Plasmid DNA was prepared as template by linearising with *Nco*I. RNA was synthesised as described in the kit except that nucleotides were substituted with RNA labelling mix (Boehringer Mannheim GmbH, Mannheim, Germany). RNA was precipitated by addition of 0.10 volume of 3M NaAc and 3 volumes 100% ethanol and incubated at -70°C for at least 2 h. RNA was recovered by precipitation (39200×g, 30 min, 4°C) and resuspended in sterile MQ water. The resulting DIG-labelled RNA probe, complementary to the CAT gene, was 300 bp in length. A 100 µL reaction produced enough labelled RNA for 20 probes.

2.16 Synthesis of the sense strand of 16S rRNA

Sense strand 16S rRNA was synthesised *in vitro* from clone pCT123 using the Riboprobe system from Promega Corp. (Madison, Wisconsin, USA) (Figure 2.2). Plasmid was prepared as template by linearising with endonuclease *Bam*HI. RNA was synthesised and treated as described in section 2.24. The length of the standard sense strand was confirmed at 1.4 kb as determined by gel electrophoresis.

2.17 Synthesis of DIG-labelled 16S rRNA probe

The antisense RNA probe was synthesised by *in vitro* transcription from plasmid pCT123 using the Riboprobe system (Figure 2.2). The plasmid was prepared as template by linearising it with *Xba*I. RNA was synthesised and treated as described in section 2.15. The resulting DIG-labelled RNA probe, complementary to the 16S rRNA gene, was 1.4 kb long. One 100 µL reaction provided enough labelled RNA for 20 probes.

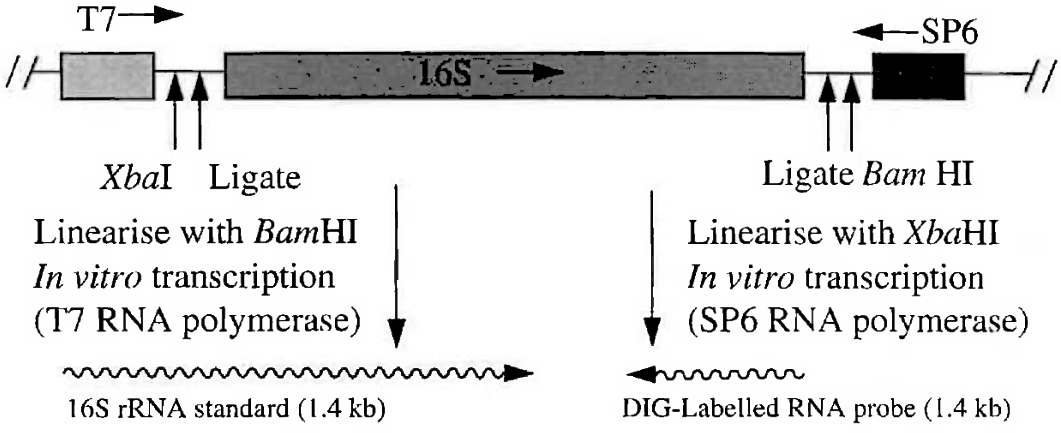

Figure 2.2. Construction of standard 16S rRNA and DIG-labelled RNA probe using the plasmid pCT123. The sense strand (standard 16S rRNA) was synthesised *in vitro* from the T7 promoter, after linearisation of the template with BamHI. This resulted in a full length transcript of the 16S gene (1.4 Kb). The DIG-labelled complementary RNA probe was synthesised from the SP6 promoter after linearisation of the template with *Xba*I. The resulting probe was 1.4 Kb long.

2.18 Preparation of Dynabeads

Dynabeads M-280 Streptavidin (Dynal, Skøyen, Norway) were prepared under RNAse free conditions as described in the manual (Dynal, 1995). A 100 bp or 720 bp fragment of CAT DNA, 5'-labelled with biotin, was synthesised by PCR (section 2.11.2). The 720 bp fragment encoded all of the CAT gene. The DNA fragments were attached to the Dynabeads and the secondary strand removed as described by the manual.

2.19 Cell counts

2.19.1 Total cell counts

Culture samples (0.5 mL) were mixed with 0.5 mL 0.4 % (v/v) formaldehyde/0.9 % (w/v) NaCl and diluted with Isoton II (Coulter Electronics Ltd, Hertfordshire, England). Cell counts were determined using a Coulter Counter (Model ZF, Coulter Electronics Ltd, Hertfordshire, England) fitted with a 30μm orifice. The Coulter Counter was calibrated with latex spheres (1.1 μm) (Coulter Electronics Ltd, Hertfordshire, England). Attenuation was 0.250, aperture current was 16, and threshold was 8. Isoton II was filtered (0.2 μm) prior to use.

2.19.2 Viable cell counts

Viable cell counts were obtained from counts on nutrient agar plates containing 100 μg/mL Ampicillin, and compared to counts on plates without selection.

2.20 Cell dry weight

Bacterial cells were collected by centrifugation (9800×g, 5 min), washed in MQ water, and dried at 105°C for 48-72 hours in pre-weighed reaction tubes. Tubes for this assay had been pre-dried as above.

2.21 Cell density

Culture absorbance at 600 nm was used as a measure of cell density. Absorbance was determined using a Unicam 8625 UV/VIS spectrophotometer.

2.22 Cell disruption

Cells used for chloramphenicol acetyl transferase (CAT) assays and total soluble protein were disrupted by two passes through a French press operated at 100 MPa. Disrupted samples were treated with detergent (Triton X-100, supplied with the CAT-ELISA kit) at room temperature for 30 min. This method maximises CAT protein recovery (see chapter 4). The effectiveness of cell disruption was quantified using an Applied Imaging disc centrifuge (Middelberg, 1992). A modified protocol that is considerably faster and simpler than reported previously was used, by mixing spin and buffer fluids prior to injection (Kleinig, 1997). The modified protocol is shown in Table 2.4. Feed cells and homogenate were treated with formaldehyde (0.02%) and stored on ice before disruption analysis (within 12 h).

2.23 CAT Protein

Bacterial cells from culture fluids were collected by centrifugation (9800×g, 5 min, 4°C), resuspended in 1 mL buffer (10 mM Tris, 1 mM EDTA, pH 7.8) and stored at –70°C. Quantitative determination of CAT was done with a CAT ELISA kit (Boehringer Mannheim GmbH, Mannheim, Germany). Samples were defrosted, diluted to the original culture sample volume and disrupted in a French press (100 MPa, 2 passes). Disrupted samples were treated with Triton X-100. This treatment gave maximal recovery of CAT (section 4.2.1). Cell debris was removed by centrifugation (12000×g, 10 min, 4°C). CAT protein in the supernatant was measured according to the CAT ELISA kit instructions.

Table 2.4. Operating conditions for the analytical disc centrifuge (Kleinig, 1997).

Spin Fluid (SF)	15 mL 10% w/w glycerol-water
Buffer fluid (BF)	1.0 mL water
Sample volume and Suspension	0.5 mL phosphate buffer[a]
Disc Speed	8000 rpm
Gain	6
Time = 0 s	inject SF and BF
Time = 90 s	inject sample
Time = 200 s	end

[a]Phosphate buffer was 20 mM Na2HPO4, 30 mM NaCl.

2.24 Total soluble protein

A Bicinchoninic acid-based protein assay kit (Pierce, Rockford, USA) was used to determine total soluble cell protein. Samples of culture were defrosted at 4°C, and disrupted in a French press (100 MPa, 2 passes). Cell debris was removed by centrifugation (12000×g, 10 min). The supernatants were collected and samples were analysed as described in the kit. BSA was used as a protein standard.

2.25 Plasmid copy number

A volume of cells, equivalent to a 1 mL culture of OD_{600}, were collected (12000×g, 1 min) and stored as a pellet at –20°C. DNA was extracted from cells by resuspending them in 100 µL of solution 1 (50 mM glucose, 25 mM Tris-HCl pH 8.0, 10 mM EDTA). Lysozyme (20 mg mL^{-1}, 20 µL), dissolved in solution 1, was added and the suspension was incubated at room temperature for 10 min. Solution 2 (0.2 M NaOH, 1% SDS) (240 µL) was added and cells were incubated on ice for 10 min. Solution 3 (3 M KAc, 2 M acetic acid, pH 4.8) (180 µL) was added and cells were further incubated on ice for 10 min. This suspension was centrifuged (12000×g, 1 min) to separate chromosomal DNA, protein, and high molecular weight RNA. The supernatant was then transferred to a fresh reaction tube and extracted once with an equal volume of phenol:chloroform (1:1). Plasmid DNA was precipitated by adding 2.5 volumes of absolute ethanol followed by incubation at –20°C for at least 2 hr. DNA was collected by centrifugation (39200×g, 30 min, 4°C). The pellet was washed with 70% ethanol and dried. The DNA pellet was resuspended in sterile MQ water in a volume equivalent to: (culture OD_{600} when collected) × (volume of cells precipitated (in mL)) × 100.

DNA (50 µL) was cut with *Eco*RV and DNA samples were run on 1% TBE (10.8 g L^{-1} Tris-base, 5.5 g L^{-1} Boric acid, 0.93 g L^{-1} EDTA) agarose gel. The gel was stained in ethidium bromide for 20 min, destained in water for 20 min, and photographed using a Polaroid 665 film. The negative was scanned using a Computing Densitometer Model

300A (Molecular Dynamics, Sunnyvale, USA). The amount of DNA was quantified by comparison with a CsCl purified standard (plasmid pBR322 linerised with *Hin*dIII).

2.26 Fermentations

2.26.1 Batch fermentations

Minimal medium (25 mL) supplemented with 40 µg/mL Thiamine, 100 µg/mL Ampicillin, 3.33 g/L glucose and 0.67 g/L $MgSO_4$ was inoculated with an a single colony of *E. coli* JM101, transformed with the relevant plasmid, and cultivated at 37°C. This culture was then subcultured (0.25 mL in 25 mL of fresh medium) and further incubated at 37°C. This culture was used to inoculate a 2 L fermenter (Applikon, Schiedam, Holland) containing 1.2 or 1.5 L of minimal media supplemented with 3.33 or 8.25 g/L glucose. Fermenters were operated at a temperature of 37°C and pH was maintained at 6.9 by automatic addition of 4 M NaOH. Dissolved oxygen was maintained above 30% of saturation at all times by adjusting stirrer speed. Foaming was controlled by addition of antifoam (Foamaster E 81 V, Henkell).

2.26.2 Fed-batch fermentation

A shake flask culture was prepared as described in section 2.26.1. This culture was used to inoculate four fermenters (2L Applikon, Schiedam, Holland) containing 1.5 L minimal media supplemented with 8.25 g/L glucose to a starting OD_{600} of 0.0005. Growth conditions were as in section 2.26.1. At OD_{600} = 2.5, 5 L of culture was transferred to the fermenter (Chemap-2000) containing 10 L of minimal media (glucose 8.25 g/L). Feeding was initiated at glucose exhaustion (OD_{600} = 10) at a rate of 1124 g/h. The measured glucose concentration (section 2.27) was 191.7 g/L after autoclaving. The fermenter was operated at 37°C and pH was maintained at 6.9 by automatic addition of NH_4OH (8 % v/v). Dissolved oxygen was maintained above 30% of saturation at all times by regulating the stirrer speed.

2.27 Glucose analysis

Glucose concentration was determined in sample supernatant using a YSI 2700 Select Glucose analyser (Yellow Springs Instrument Co. Inc., Yellow Springs, OHIO, USA).

2.28 Acetate analysis

Supernatants were stored at –20°C prior to determination of acetate concentration using a kit (Boehringer Mannheim GmbH, Mannheim, Germany, no 148 261).

Chapter 3

Preparation and nucleotide sequence of constructs

3.1 Strategy for construction of reporter plasmids with *cat*

The relative strengths of selected promoters were determined by assaying reporter gene products expressed from promoter-reporter gene transcriptional fusions. These fusions were prepared by subcloning DNA fragments containing the respective promoter into the bidirectional promoter probe vector, pPM3024 (Williams & Manning, 1991). The promoter activity and the direction of DNA transcription from the promoters were determined by transforming the constructs into a *gal*K strain and screening for Cm resistance or the ability to produce acid from galactose. The promoter strengths were determined by isolating the Ap^R colonies and measuring the total CAT produced by the bacteria. Sequence analysis of DNA was carried out to confirm that plasmids contained the relevant promoter in correct orientation.

3.2 Construction of *katE*::*cat* transcriptional fusion (pCT100)

A *katE*::*cat* transcriptional fusion was constructed by ligating a 1257 bp *Bam*HI, *Hin*dIII fragment containing the *katE* promoter to the promoter-less *cat* reporter gene in plasmid pPM3024. Details of the construction are shown in Figure 3.1. An Ap^R $galK^-$ colony was selected as a putative *katE*::*cat* transcriptional fusion and used for future analysis.

Plasmid DNA (pCT100) isolated from this strain was confirmed by restriction endonuclease digestion analysis. *Bam*HI, *Hind*III digestion of pCT100 gave a 1.25 and a 4.6 kb fragment consistent with correct insertion of the *katE* promoter fragment in the multiple cloning site of vector pPM3024.

Sequence analysis of pCT100 DNA was carried out by extension from the *gal*K and *cat* genes using the primers #362, #B7420 and #942 (Figure 3.2). Primer #362 reads out of the *galK* gene. Primer #B7420 reads backwards and out of the *cat* gene. Primer

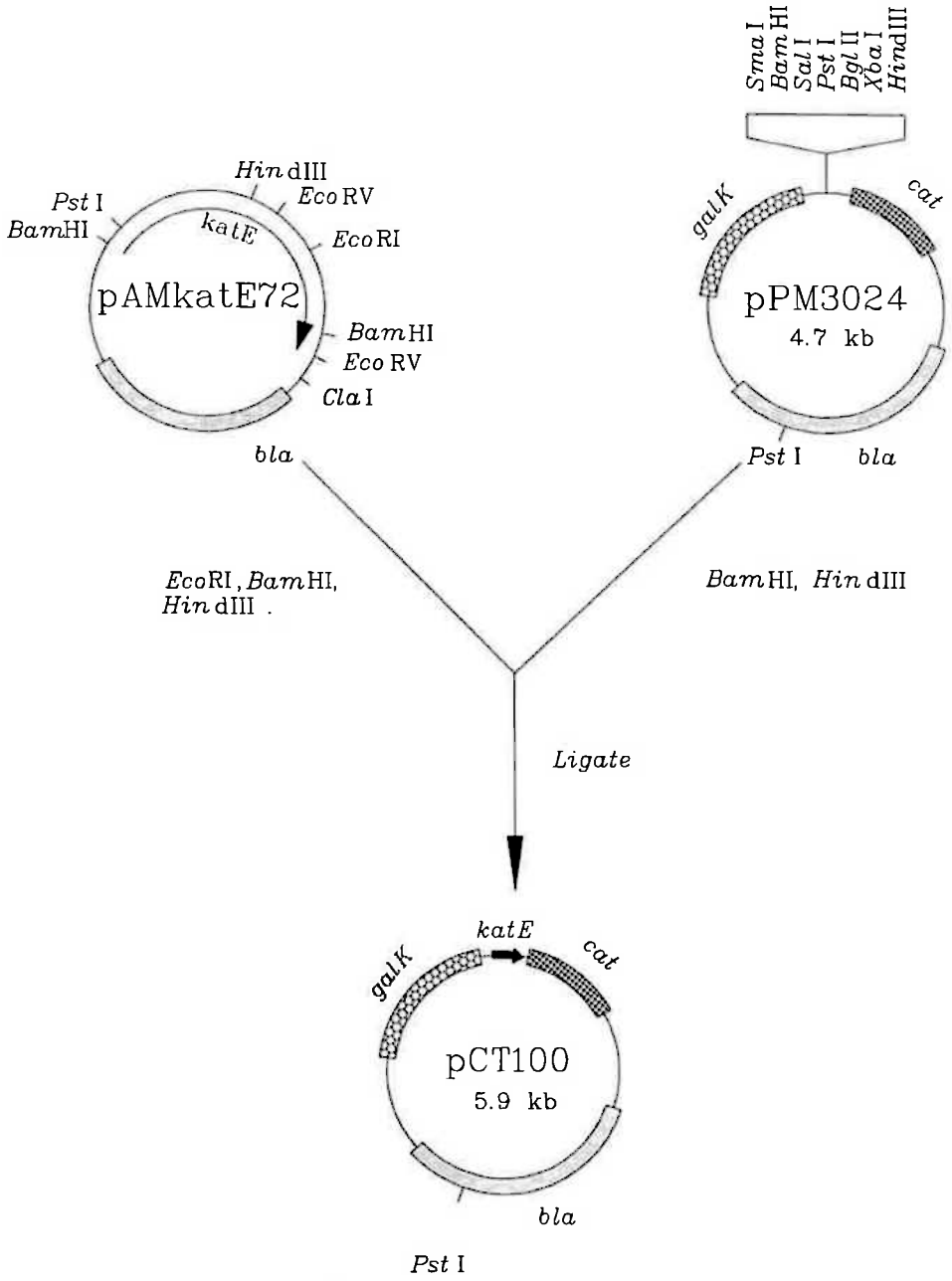

Figure 3.1. Construction of pCT100.
A *katE*::*cat* transcriptional fusion was constructed by ligating a 1257 bp *Bam*HI, *Hin*dIII fragment containing the *katE* promoter to the promoterless *cat* reporter gene in plasmid pPM3024. The *katE* fragment was isolated from plasmid pAMkatE72 following a *Bam*HI, *Hin*dIII, *Eco*RI digestion and cloned into *Bam*HI, *Hin*dIII digested plasmid pPM3024. This construct was transformed into the strain CB806 and plated onto Ap agar. Ap^R colonies were then plated on Cm/Ap plates. Colonies were found to be sensitive to Cm. To confirm the orientation of the insert in the vector, transformants were plated on McConkey/Ap agar containing galactose as a carbohydrate source. Clones containing a promoter driving the *galK* gene of pPM3024 produced bright red colonies whereas promoters driving the *cat* gene produced white colonies. Plasmid DNA isolated from a single white Ap^R colony was designated pCT100.

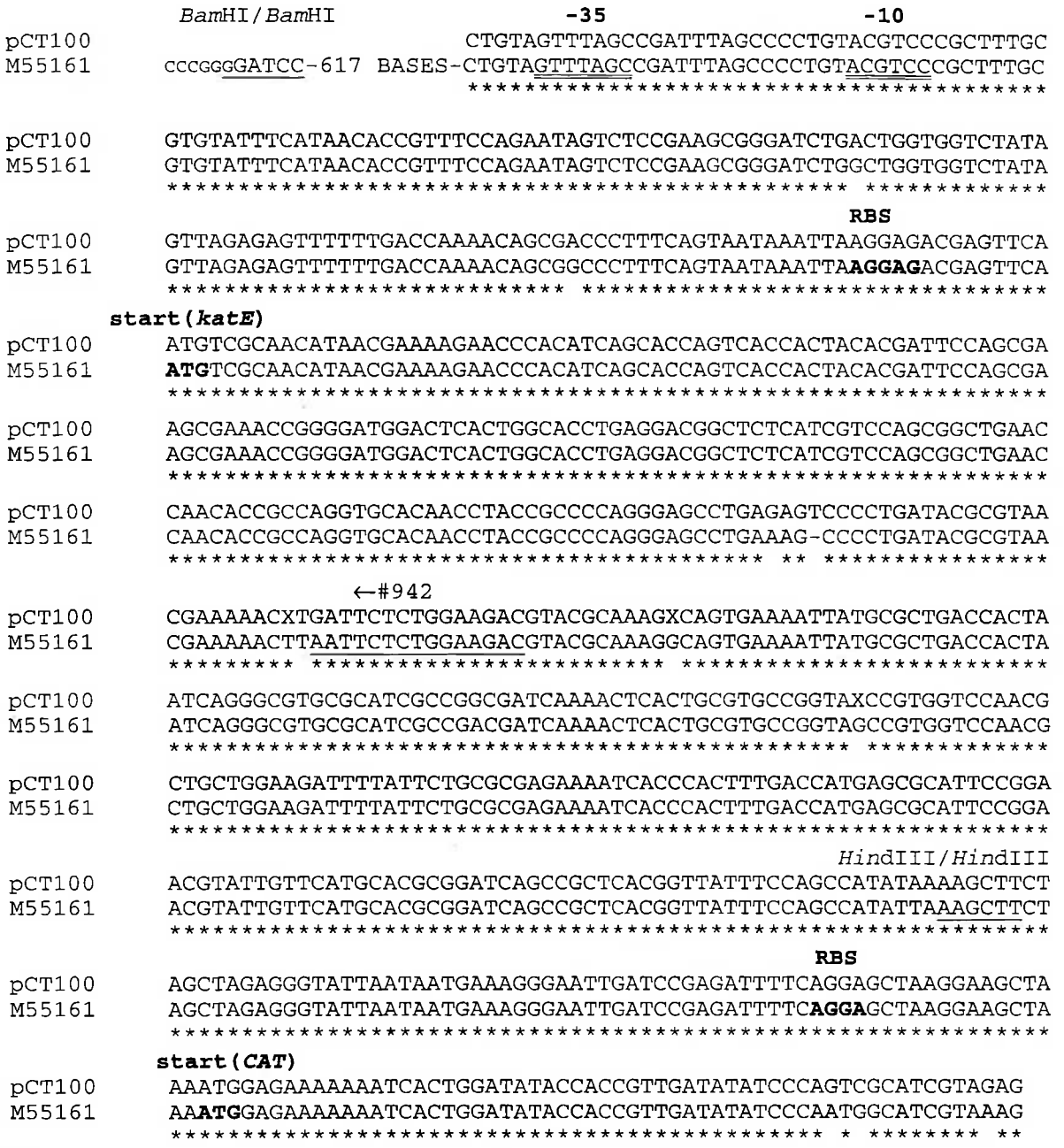

Figure 3.2. Sequence analysis of pCT100.

Alignment of actual sequence data derived from pCT100 compared with sequence information describing the constructed *katE*::*cat* transcriptional fusion, based on the sequence data for pPM3024 (Figure 3.9) and the *katE* gene (Genbank accession number M55161). Identities are indicated by an asterisk, gaps are shown as a hyphen. The *katE* promoter -10 and -35 sites are identified by double underline and ligation sites as an underline. The ribosome binding sites and start sites of the *katE* gene and the *cat* gene are also shown. Three sequencing primers were used in order to confirm the insert in pCT100. Primer #B7420 reads backwards and out of the *cat* gene. Primer #362 reads out of the *gal*K gene. Primer #942 was constructed to follow the sequence data obtained from the first reaction, using #B7420. #942 is identified by an underline and an arrow. Rows labelled pCT100 represent sequence data, whereas M55161 represents the fusion based on published data. Vector-derived nucleotides are presented in smaller fonts. X represents unidentified nucleotides.

#942 was constructed to follow the sequence data obtained from the first reaction, using #B7420. Sequence data obtained were aligned with sequence information based on that of pPM3024 (Figure 3.9) plus the 1257 bp fragment of the *katE* gene (Genbank accession number M55161) containing the putative promoter site as identified by von Ossowski *et al* (1991). Alignment of sequence data for pCT100 with that of assembled sequence information for a *katE::cat* fusion confirmed this plasmid contained the *katE* promoter in the correct orientation. Sequence data and assembled sequence information were in good agreement.

3.3 Construction of *katF*::*cat* transcriptional fusion (pCT101)

A *katF*::*cat* transcriptional fusion was constructed by ligating a 180 bp *Nco*I, *Pst*I fragment containing the *katF* promoter to the promoter-less *cat* reporter gene in plasmid pPM3024. Details of the construction are shown in Figure 3.3. A Cm/Ap^R galK$^-$ colony was selected as a putative *katF*::*cat* transcriptional fusion and used for future analysis.

Plasmid DNA (pCT101) isolated from this strain was confirmed by restriction endonuclease digestion analysis. *Hin*dIII digestion of pCT101 gave a 4.9 kb fragment consistent with correct insertion of *katF* promoter fragment in the multiple cloning site of the vector pPM3024 (Figure 3.9).

Sequence analysis of pCT101 DNA was carried out by extension from the *galK* gene using primer #362 (Figure 3.4). Sequence data obtained were compared with sequence information based on that of pPM3024 (Figure 3.9) plus the 180 bp *Nco*I/*Pst*I fragment of the *katF* gene (Genbank accession number X16400) containing the putative *katF* promoter site as identified by Mulvey and Loewen (1989). Alignment of sequence data for pCT101 with that of assembled sequence information for a *katF::cat* fusion confirmed this plasmid contained the *katE* promoter in the correct orientation. Sequence data and assembled sequence information were in good agreement.

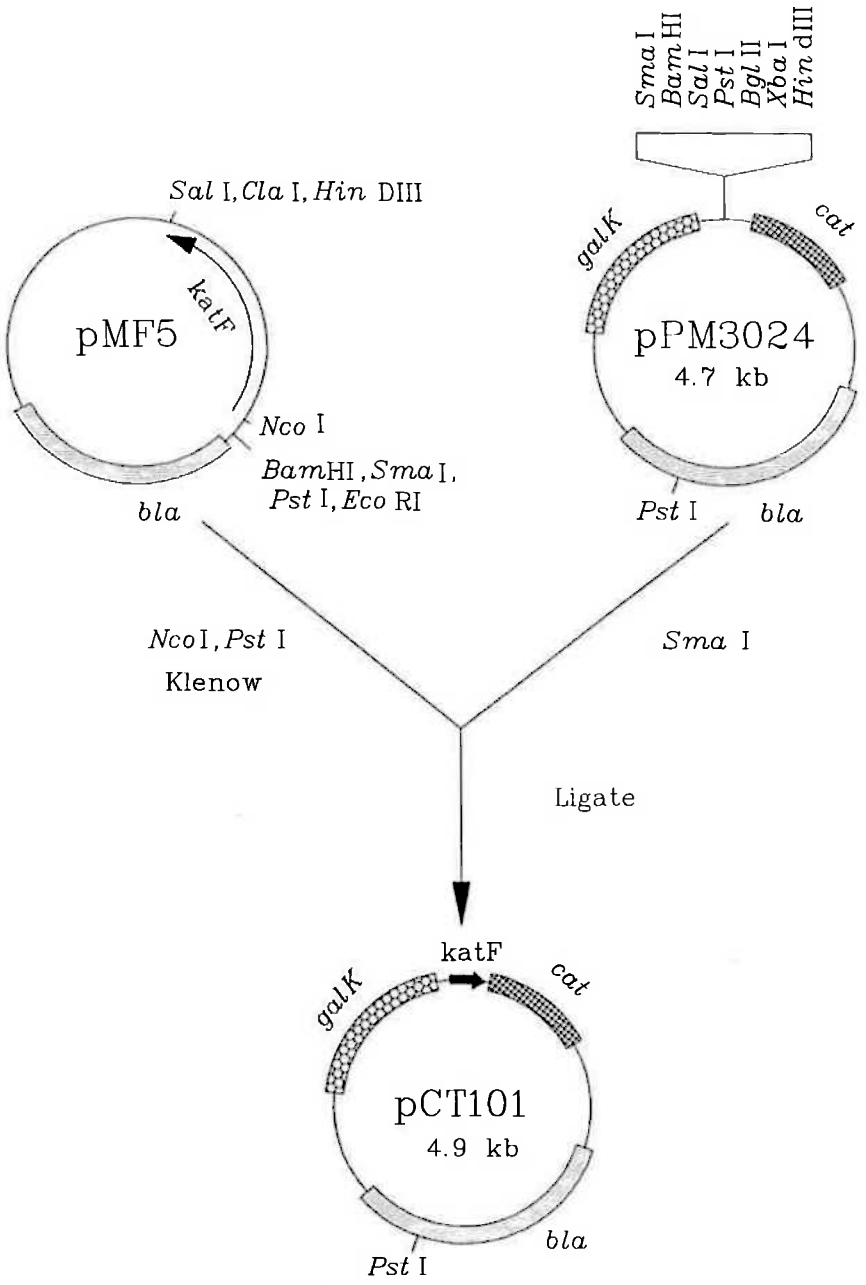

Figure 3.3. Construction of pCT101.
A *katF*::*cat* transcriptional fusion was constructed by ligating a 180 bp *Nco*I, *Pst*I fragment containing the *katF* promoter to the promoter-less *cat* reporter gene in plasmid pPM3024. The *katF* promoter was isolated from the plasmid pMF5 following a *Nco*I, *Pst*I digestion. After end-filling/chewing of protruding ends with Klenow fragment of DNA polymerase I the fragment was cloned into the *Sma*I site of the cloning vector pPM3024. The sub-cloned vector was transformed into the strain CB806 and plated onto Ap agar. Ap^R colonies were then plated onto Cm/Ap plates. Colonies were found to be resistant to Cm. To confirm the orientation of the insert in the vector, transformants were plated on McConkey/Ap agar containing galactose as a carbohydrate source. Clones containing a promoter driving the *galK* gene of pPM3024 produced bright red colonies whereas promoters driving the *cat* gene produced white colonies. Plasmid DNA isolated from a single white Ap^R colony was designated pCT101.

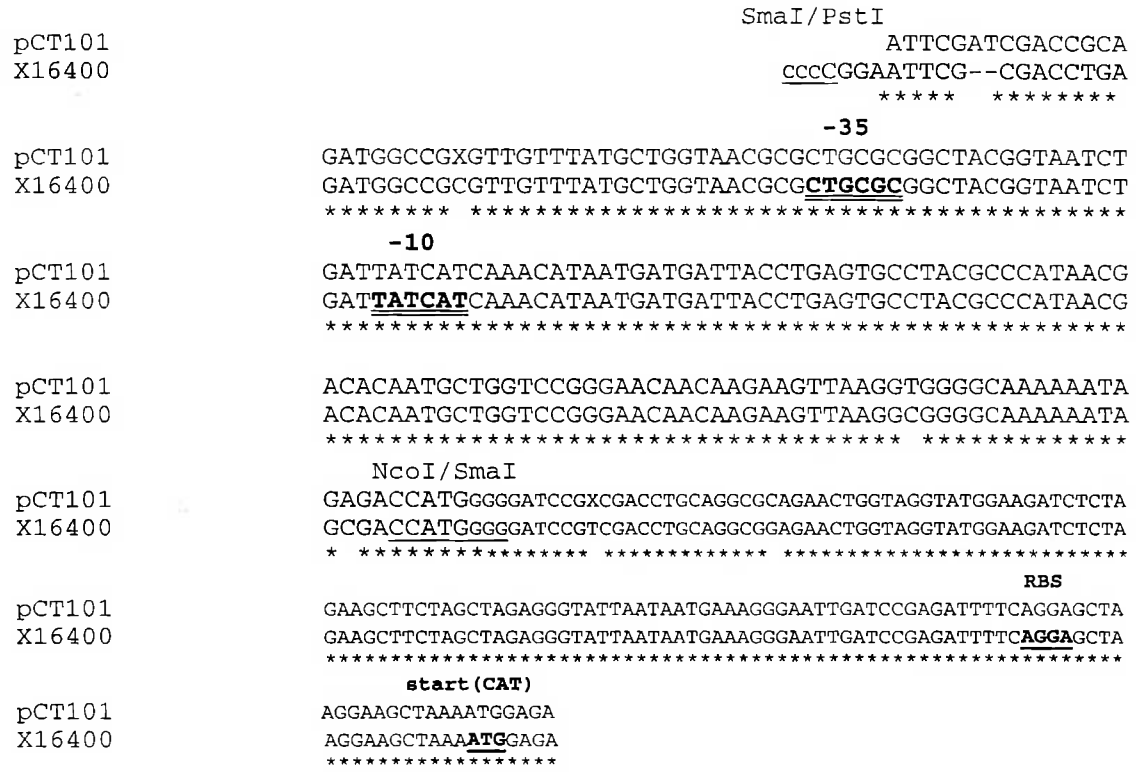

Figure 3.4. Sequence analysis of pCT101.
Alignment of actual sequence data derived from pCT101 compared with sequence information describing the constructed *katF*::*cat* transcriptional fusion, based on the sequence data for pPM3024 (Figure 3.9) and the *katF* gene (Genbank accession number X16400). Identities are indicated by an asterisk, gaps are shown as a hyphen. The *katF* promoter sites -10 and -35 are identified by bold double underline, ligation sites as an underline. The ribosome binding site and start site of the *cat* gene are also shown in bold with single underline. Rows labelled pCT101 represent sequence data whereas X16400 represents the Genbank derived sequence. Vector derived nucleotides are presented in smaller fonts.

3.4 Construction of *trc*::*cat* transcriptional fusion (pCT102)

A *trc*::*cat* transcriptional fusion was constructed by ligating a 285 bp *Eco*RI, *Hin*dIII fragment containing the *trc* promoter to the promoterless *cat* reporter gene in plasmid pPM3024. The *trc* promoter was isolated from the plasmid pEV601. Details of the construction are shown in Figure 3.5. A Cm/Ap^R $galK^-$ colony was selected as a putative *trc*::*cat* transcriptional fusion and used for future analysis.

Plasmid DNA (pCT102) isolated from this strain was confirmed by restriction endonuclease digestion analysis. *Nde*I digestion of pCT102 gave a 1877 bp and a 3096 bp fragment consistent with correct orientation of the *trc* promoter fragment in the multiple cloning site of the vector pPM3024.

Sequence analysis of pCT102 DNA was carried out by extension from the *galK* gene using primer #362 (Figure 3.6). Sequence data obtained were compared with sequence information based on that of pPM3024 (Figure 3.9) plus the 289 bp fragment containing the *trc* promoter (Genbank accession number U02439) (Amann & Brosius, 1985). Alignment of sequence data for pCT102 with that of assembled sequence information for a *trc*::*cat* fusion confirmed this plasmid contained the *trc* promoter in the correct orientation. Sequence data and assembled sequence information were in reasonable agreement.

3.5 Construction of *tac*::*cat* transcriptional fusion (pCT103)

A *tac*::*cat* transcriptional fusion was constructed by ligating a 269 bp *Bam*HI fragment containing the *tac* promoter to the promoterless *cat* reporter gene in the plasmid pPM3024. The *tac* promoter fragment was isolated from pKK223-3 (Pharmacia LKB Biotechnology, Sweden). Details of the construction are shown in Figure 3.7.

A Cm/Ap^R $galK^-$ colony was selected as a putative *tac*::*cat* transcriptional fusion

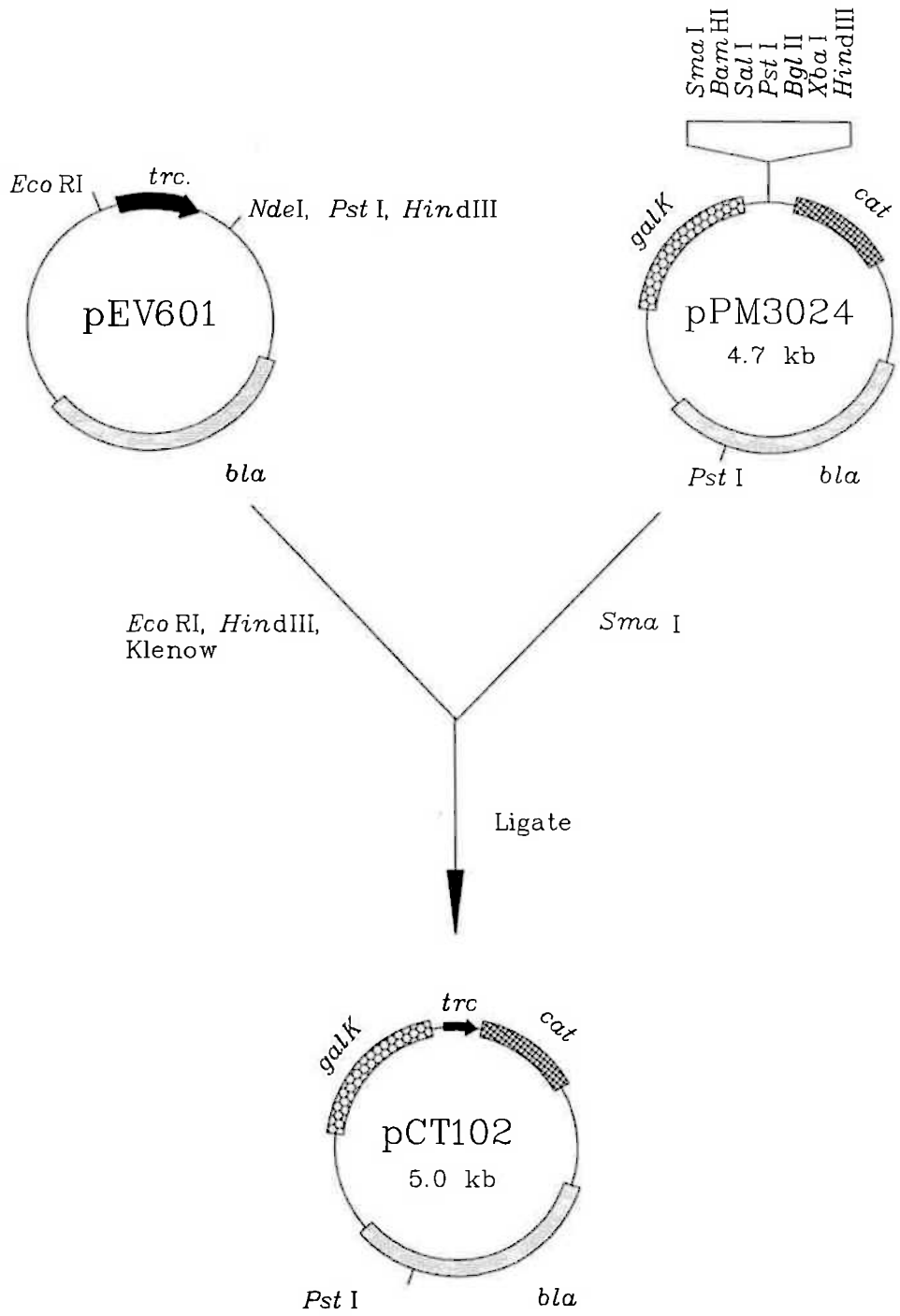

Figure 3.5. Construction of pCT102.
A *trc*::*cat* transcriptional fusion was constructed by ligating a 285 bp *Eco*RI, *Hin*dIII fragment containing the *trc* promoter to the promoterless *cat* reporter gene in plasmid pPM3024. The *trc* promoter was isolated from plasmid pEV601 after *Eco*RI, *Hin*dIII digestion. Protruding ends were end-filled using Klenow fragment of DNA polymerase I and the fragment cloned into the *Sma*I site of the cloning vector pPM3024. The vector was transformed into the strain CB806 and plated onto Ap/Cm agar. Colonies resistant to both antibiotics were screened on McConkey/Ap plates, containing galactose as a carbohydrate source. *GalK* positive clones produce bright red colonies on the McConkey/Ap medium whereas *cat* positive colonies are white. Plasmid DNA isolated from a Ap^R single white colony was designated pCT102.

Figure 3.6. Sequence analysis of pCT102.
Alignment of actual sequence data derived from pCT102 compared with sequence information describing the constructed *trc*::*cat* transcriptional fusion, based on the sequence data for pPM3024 (Figure 3.9) and the *trc* promoter (Genbank accession number U02439). Identities are indicated by an asterisk, gaps are shown as a hyphen. The *trc* promoter is identified by double underline, ligation sites as an underline. The ribosome binding site and ATG start site of the *cat* gene are also shown in bold. Rows labelled pCT102 represent sequence data whereas U02439 represents published sequence data. Vector derived nucleotides are presented in smaller fonts.

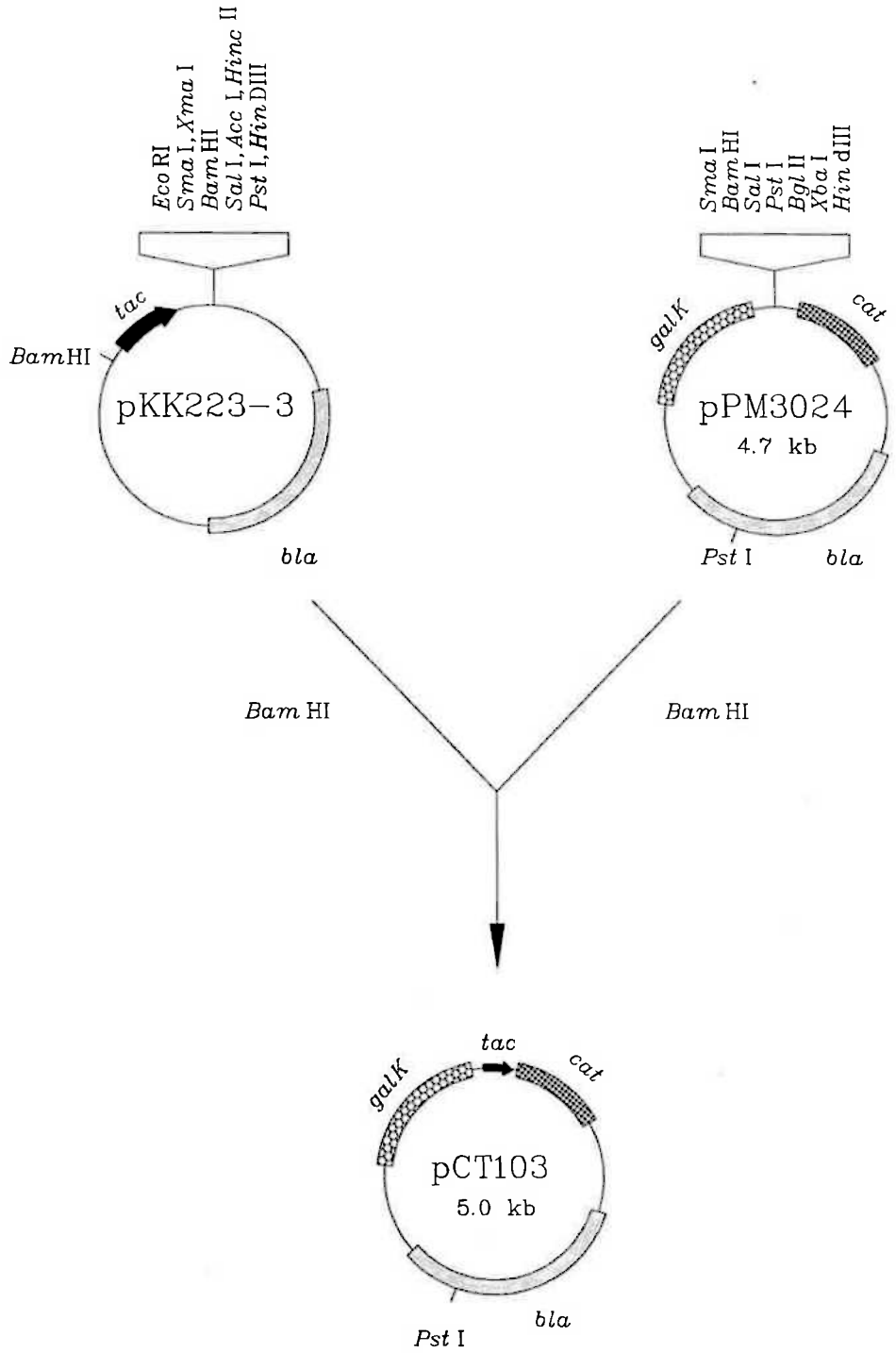

Figure 3.7. Construction of pCT103.
A *tac*::*cat* transcriptional fusion was constructed by ligating a 269 bp fragment containing the *tac* promoter to the promoterless *cat* reporter gene in the plasmid pPM3024. The *tac* promoter was isolated from the plasmid pKK223-3 after *Bam*HI digestion and cloned into the *Bam*HI site of the multiple cloning site of the reporter vector pPM3024. The sub-cloned vector was transformed into the strain CB806 and plated onto Ap/Cm agar. Colonies resistant to both antibiotics were further transferred to McConkey/Ap plates containing galactose as a carbohydrate source. Plasmid DNA isolated from a single Ap^R white colony was designated pCT103.

and used for future analysis. Plasmid DNA (pCT103) isolated from this strain was confirmed by restriction endonuclease digestion analysis. *Eco*RI digestion of pCT103 gave a 594 bp and a 3360 bp fragment consistent with correct orientation of the *trc* promoter fragment in the multiple cloning site of the vector pPM3024 (Figure 3.9).

Sequence analysis of pCT103 DNA was carried out by extension from the *galK* gene using primer #362 (Figure 3.8). Sequence data obtained were aligned with sequence information based on that of pPM3024 (Figure 3.9) plus the 269 bp fragment containing the *tac* promoter (Genbank accession number K01728) (Amann *et al.*, 1983). Alignment of sequence data for pCT103 with that of assembled sequence information for a *tac*::*cat* fusion confirmed this plasmid contained the *trc* promoter in the correct orientation. Sequence data and assembled sequence information were in reasonable agreement.

3.6 Construction of plasmids for *in vitro* production of DIG-labelled *cat* RNA probe and positive control

A modified RNA hybridisation technique was used to quantitate *cat* mRNA transcripts. RNA blots were detected using a complimentary DIG-labelled RNA probe and Chemiluminescence detection. The intensity of a produced signal was then related to a standard calibration curve. DIG-labelled RNA probe and the positive control was produced *in vitro* from plasmid pCT121.

3.6.1 Construction of T7/SP6::*cat* transcriptional fusion (pCT121)

A T7/SP6::*cat* transcriptional fusion was constructed by ligating a 800 bp *Bam*HI fragment containing the *cat* gene into the plasmid pGEM7Zf(+). Details of the construction are shown in Figure 3.10.

Sequence analysis of pCT121 DNA was carried out by extension from the SP6 and T7 priming sites using PRISM Ready Reaction Dye Primer kit (Figure 3.11). Alignment of sequence data with that of assembled sequence information confirmed

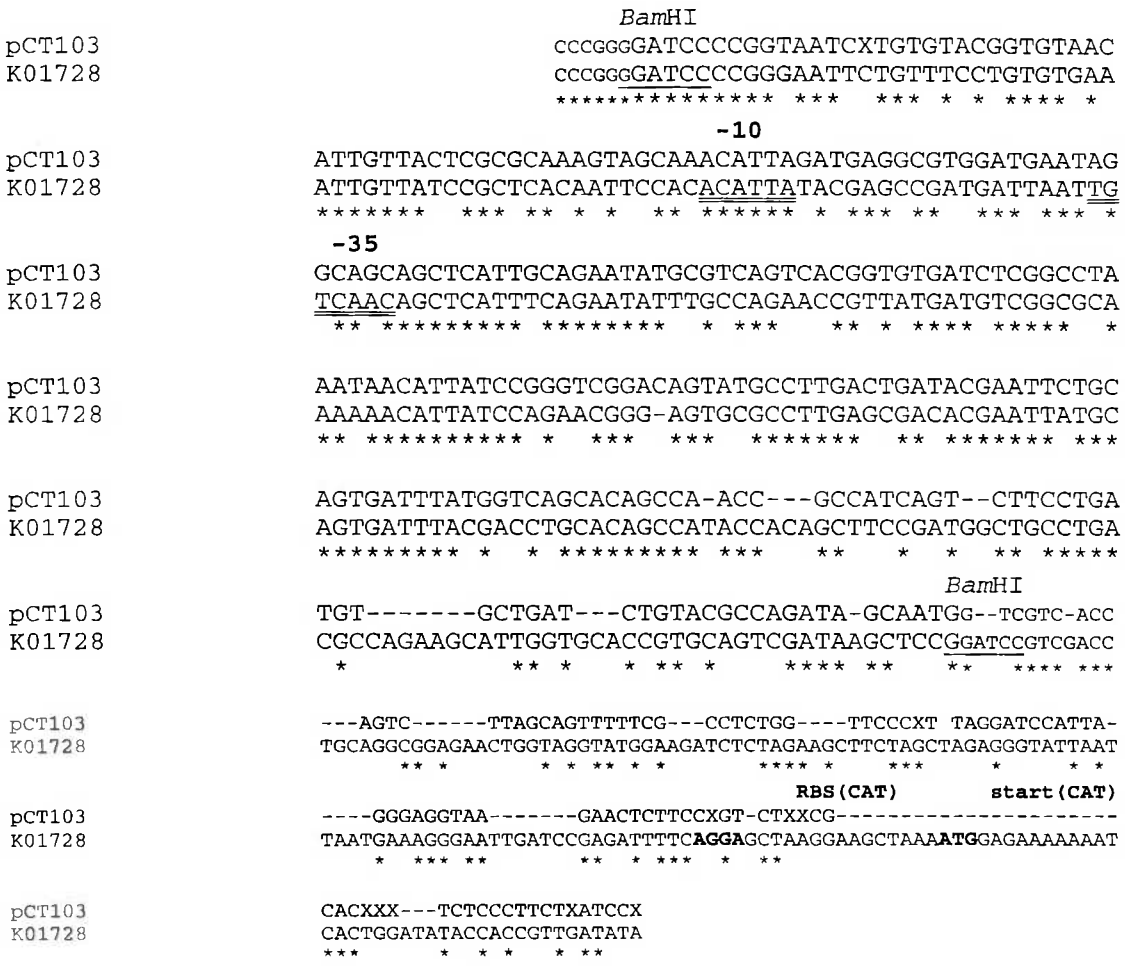

Figure 3.8. Sequence analysis of pCT103.
Alignment of actual sequence data derived from pCT103 with sequence information describing the constructed *tac*::*cat* transcriptional fusion, based on the sequence data for pPM3024 and the *tac* promoter (Genbank accession number K01728). Identities are indicated by an asterisk, gaps are shown as a hyphen. The *trc* promoter -10 and -35 sites are identified by double underline, ligation sites as an underline. The ribosome binding site and start site of the *cat* gene are also shown. Rows labelled pCT103 represent sequence data whereas K01728 represents published sequence data. Vector-derived nucleotides are presented in smaller fonts.

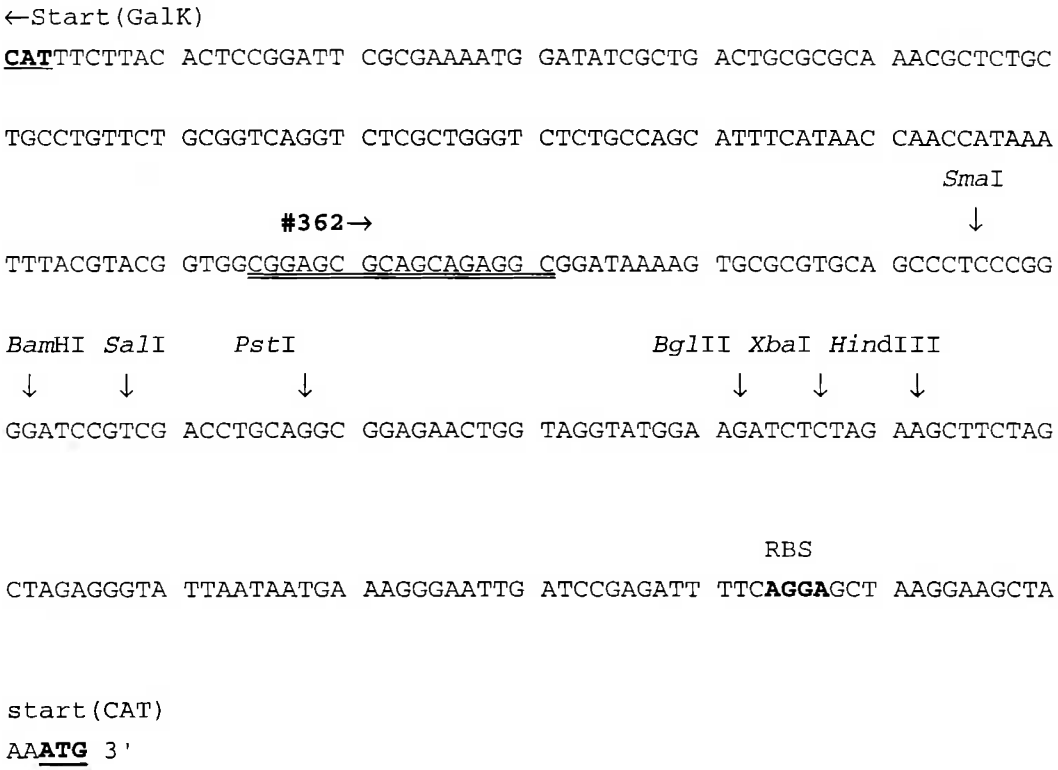

Figure 3.9. Sequence data for pPM3024 reporter vector extending from the *galK* gene over the MCS to the *cat* gene.
The ATG start site for the *galK* gene and the *cat* gene are both underlined and in bold. Restriction sites in the multiple cloning site (MCS) are indicated by arrows. The *cat* gene ribosome binding site is shown in bold. Sequence identical to that of primer #362 is double underlined. Sequence data were created from information published by Schneider & Beck (1986) and McKenny *et al.* (1981).

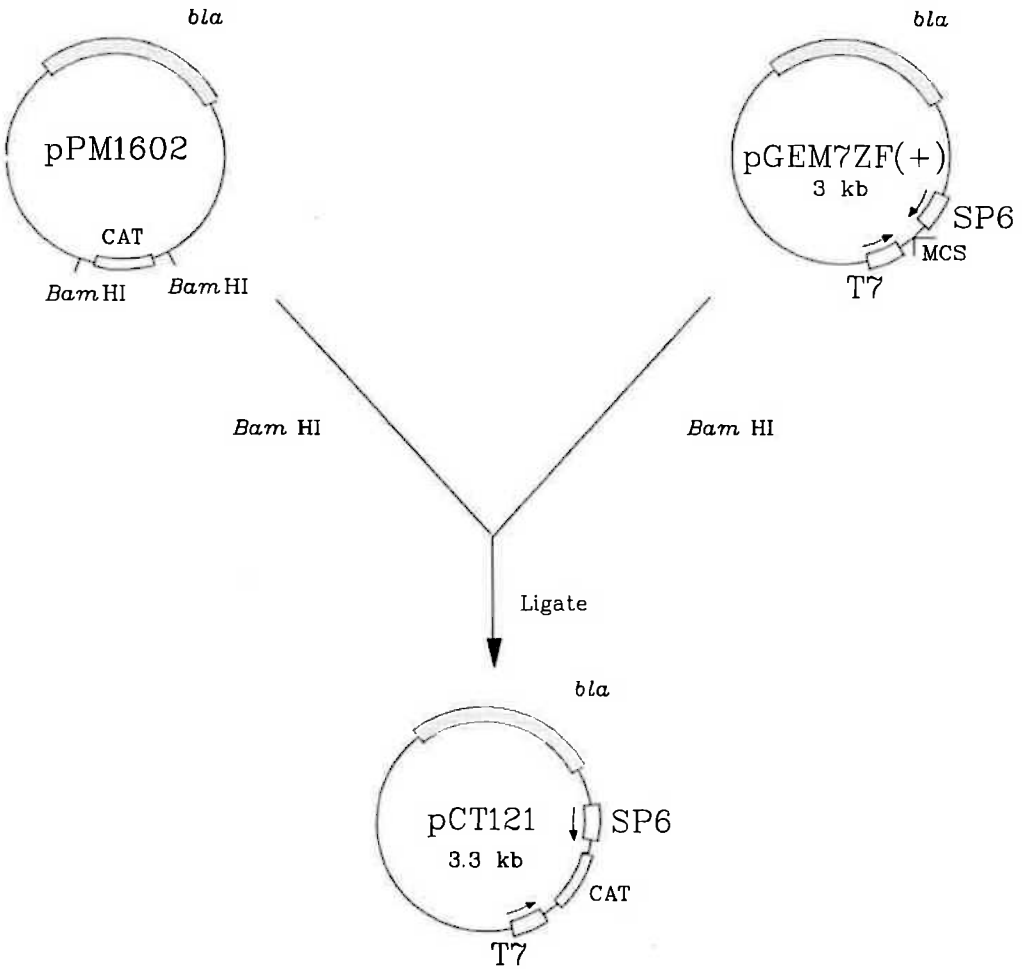

Figure 3.10. Construction of pCT121.
A T7/SP6::*cat* transcriptional fusion was constructed by ligating a 800 bp *Bam*HI fragment into the vector pGEM7Zf(+). The *cat* gene was isolated from the plasmid pPM1602 after *Bam*HI digestion. The fragment was cloned into the *Bam*HI site of the cloning vector pGEM7Zf(+). This cloned vector was transformed into the strain DH5α and plated onto X-gal/Ap agar. Positive colonies were identified as white β-galactosidase -ve, colonies. Plasmid DNA isolated from a single Ap^R white colony was designated pCT121.

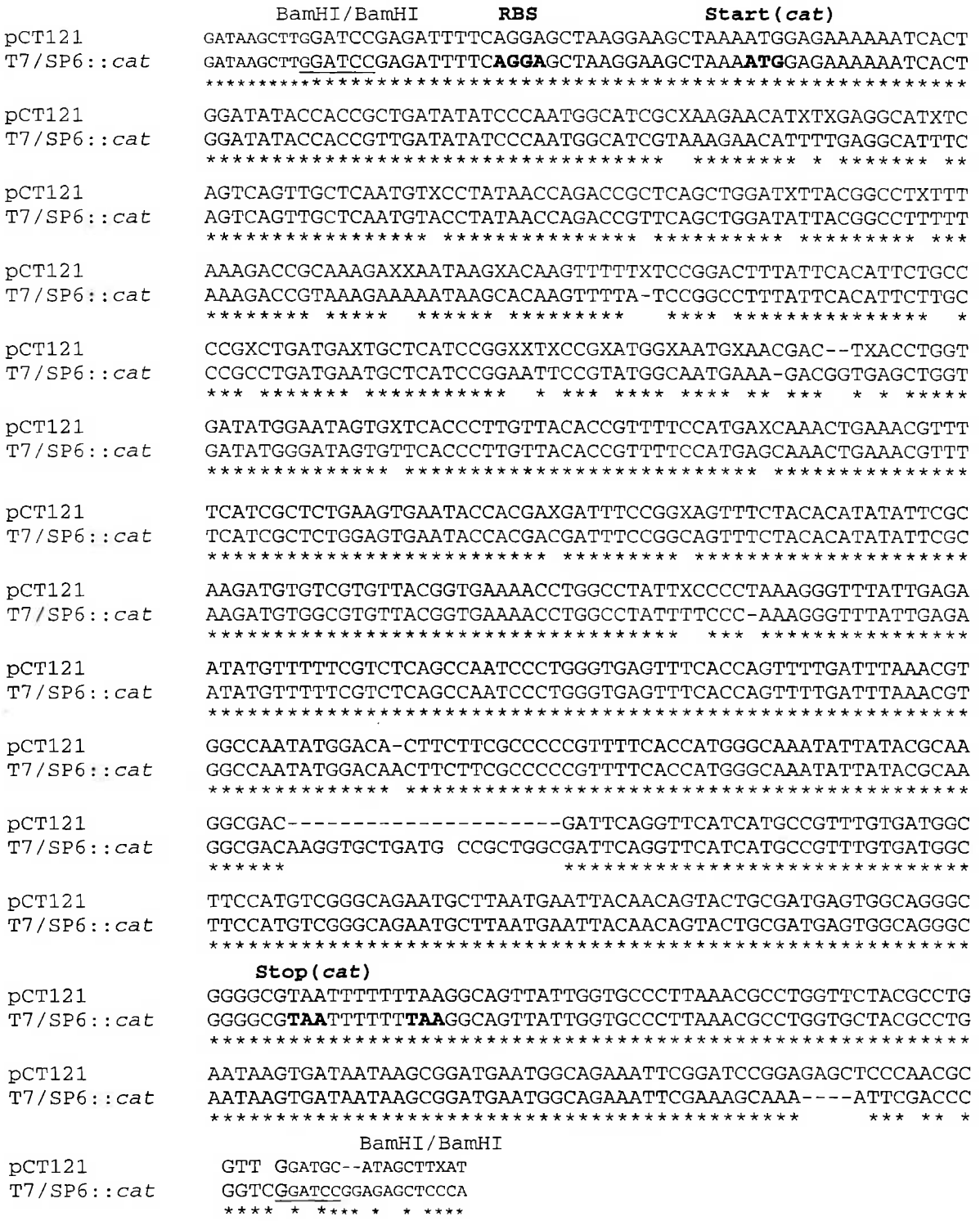

Figure 3. 11. Sequence analysis of pCT121.
Alignment of actual sequence data derived from pCT121 compared with sequence information describing the constructed T7/SP6 transcriptional fusion, based on the sequence data for pGEM7Zf(+) and pPM3024 (Figure 3.9). Identities are indicated by an asterisk, gaps are shown as a hyphen. Ligation sites are identified as an underline. The RBS, start- and stop site for *cat* is shown in bold. Rows labelled pCT121 represent sequenced data whereas T7/SP6::*cat* represents published sequence data. Vector derived nucleotides are presented in smaller fonts.

that pCT121 contained an 800 bp insert of *cat*. Sequence data and assembled sequence information were in good agreement.

3.7 Construction of plasmid for *in vitro* production of DIG-labelled 16S rRNA probe and positive control

The amount of Ribosomal RNA in bacterial cells was estimated by determining cellular concentrations of 16S rRNA. Total RNA was isolated and transferred to a membrane. The rRNA was detected after hybridisation with a complimentary DIG-labelled RNA probe using a Chemiluminescence detection system. The intensity of a produced signal was then related to that of a pure standard. DIG-labelled RNA probe and the positive control 16S rRNA were produced *in vitro* from plasmid pCT123.

3.7.1 Construction of T7/SP6::16S rRNA fusion (pCT123)

A T7/SP6::16S rRNA transcriptional fusion was constructed by ligating a 1335 bp fragment containing the 16S rRNA gene into the plasmid pGEM7Zf(+) (Promega Corp., Madison, Wisconsin, USA). Details of the construction are shown in Figure 3.12. This construct was confirmed by sequence analysis from both the T7 and SP6 priming sites (Figure 3.13). Sequence data obtained were compared with the published sequences of pGEM7ZF(+) and the *E. coli* rRNA operon gene (Genbank accession number JO1695). Sequence alignment showed that pCT123 contained an 1335 bp insert of *E. coli* 16S rRNA. Sequence data gave excellent agreement with that of the assembled sequence information.

Subclones pCT124 and pCT125 were constructed in order to sequence all of the inserted fragment. pCT124 was constructed by deleting 681 bp of the plasmid pCT123. Details of the construction are shown in Figure 3.14. pCT123 was digested with *Eco*RI and the fragment recovered from a GTG agarose gel (FMC Bioproducts, Rockland, ME, USA). DNA was purified using Geneclean (Bio 101, Inc., CA, USA) and the compatible ends religated. The vector was transformed into the *E. coli* strain DH5α and screened on Ap agar. pCT125 was constructed by deleting a 661 bp *Bgl*II, *Bam*HI fragment from plasmid pCT123 (Figure 3.14). pCT123 was digested with *Bam*HI, *Bgl*II and the fragment

recovered from a GTG agarose gel. DNA was purified using Geneclean and compatible ends re-ligated. The vector was transformed into the *E. coli* strain DH5α and screened on Ap agar. Alignment of sequence data with that of assembled sequence information confirmed pCT123 contained an 1335 bp insert of *E. coli* 16S rRNA. The two were in good agreement.

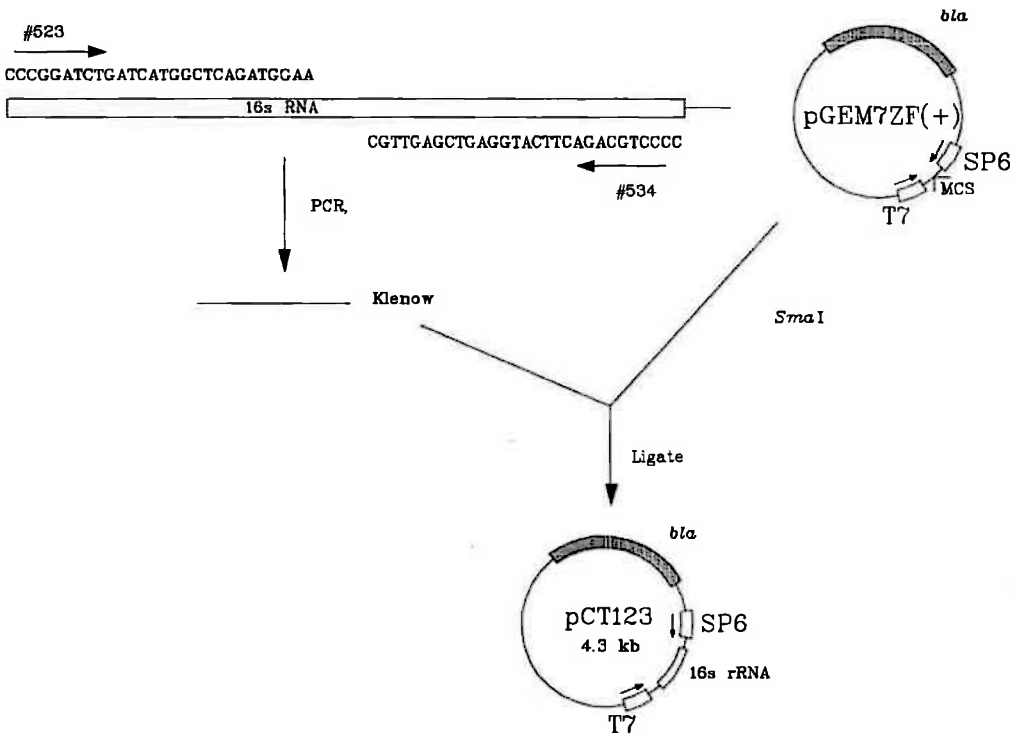

Figure 3.12. Construction of pCT123.
Plasmid pCT123 was constructed by ligating a 1335 bp PCR fragment encoding *E. coli* 16S rRNA, to plasmid pGEM7Zf(+). The PCR fragment encoding the 16S rRNA gene was amplified from chromosomal DNA isolated from the *E. coli* strain DH1, using two primers (#523 and #534). The generated PCR fragment was end-filled with Klenow and cloned into the *Sma*I site of the vector pGEM7Zf(+). The cloned vector was transformed into the strain DH5α and plated onto X-gal/Ap agar. Positive colonies were identified as white β-galactosidase -ve, colonies. The resulting plasmid is designated pCT123.

```
pCT123                                  GGGCG ATTGGGCCCGACGTCGCATGCTCCTCTAGACTCGAGGAATTCG
pGEM7ZF(+)/16S rRNA                     GGGCGAATTGGGCCCGACGTCGCATGCTCCTCTAGACTCGAGGAATTCG
                                        ***** ****************************************

                                  SmaI/PCR
pCT123                          GTACCCCCCCGGATCTGATCATGGCTCAGATGGAACGCTGGCGGCAGGCCT
pGEM7ZF(+)/16S rRNA             GTACCCCCCCG-ATCTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCCT
                                ********** ******************* ******************

pCT123                          AACACATGCAAGTCGAACGGTAACAGGAAGAAGCCTGCTTCTTTGCTGAC
pGEM7ZF(+)/16S rRNA             AACACATGCAAGTCGAACGGTAACAGGAAGAAGCTTGCTTCTTTGCTGAC
                                ********************************** ***************

pCT123                          GAGTGGCGGACGGGTGAGTAATGTCTGGGAAACTGCCTGATGGAGGGGGA
pGEM7ZF(+)/16S rRNA             GAGTGGCGGACGGGTGAGTAATGTCTGGGAAACTGCCTGATGGAGGGGGA
                                **************************************************

pCT123                          TAACTACTGGAAACGGTAGCTAATACCGCATXACGTCGCAAGACCAAAAA
pGEM7ZF(+)/16S rRNA             TAACTACTGGAAACGGTAGCTAATACCGCATAACGTCGCAAGACCAAAGA
                                ******************************** ***************** *

pCT123                          GGGGGACCTTCGGGCCTCTTXCCATCGGATGTGCCCAGATGGGATTAGCT
pGEM7ZF(+)/16S rRNA             GGGGGACCTTCGGGCCTCTTGCCATCGGATGTGCCCAGATGGGATTAGCT
                                ******************** ******************************

pCT123                          AGTAGGTXXGGTAAXCGGCTCACCTAGGAGACGATCCCTAGCTGGTCTGA
pGEM7ZF(+)/16S rRNA             AGTAGGTGGGGTAAC-GGCTCACCTAGGCGACGATCCCTAGCTGGTCTGA
                                ******* *****  ************ **********************

pCT123                          GAGGATGACCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGG
pGEM7ZF(+)/16S rRNA             GAGGATGACCAGCCACACTGGAAXTGAGACACGGTCCAGAXTCCTACGGG
                                ************************ ***************** *********

pCT123                          AGGCAGCAGTGGGGAATATTGCACAATGGGCGCAAGCCTGATGCAGC CA
pGEM7ZF(+)/16S rRNA             AGGCAGCAGTGGGGAATATTGCACAATGGGCGCAAGCCTGATGCAGCCCA
                                ********************************************** **

pCT123                          TGCCGCGTGTATGAAGAAGCCTTCGGGTTGTAAAGTACTTTCAGCGGGGA
pGEM7ZF(+)/16S rRNA             TGCCGCGTGTATGAAGAAG-CTTCGGGTTGTAAAGTACTTTCAGCGGGGA
                                ******************** ******************************

pCT123                          GGAAGGGAGTAAAGTTAATACCTTTGCTCATTGACGTTACCCGCAGAAGA
pGEM7ZF(+)/16S rRNA             GXAAGGGAGTAAAGTTAATACCTTTGCTCATTGACGTTACCCGCAGAAGA
                                * ************************************************

pCT123                          AGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAG
pGEM7ZF(+)/16S rRNA             AGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGGTGCAAG
                                **************************************************

pCT123                          CGTTAATCGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTTTGTTAAGT
pGEM7ZF(+)/16S rRNA             CGTTAATCGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTTTGTTAAGT
                                **************************************************

pCT123                          CAGATGTGAAATCCCCGGGCTCAACCTGGGAACTGCATCTGATACTGGCA
pGEM7ZF(+)/16S rRNA             CAGATGTGAAATCCCCGGGCT AACCTGGGAACTGCATCTGATACTGGCA
                                ********************** ****************************

pCT123                          AGCTTGAGTCTCGTAGAGGGGGGTAGAATTGGG CCCGACGTCGCATGCT
pGEM7ZF(+)/16S rRNA             AGCTTGAGTCTCGTAGAGGGGGGTAGAATTGGGGCCCGACGTCGCATGCT
                                ********************************** ***************

pCT123                          CCTCTAGACTCGAGGAATTCCAGGTGTAGCGGTGAAATGCGTAGAGATCT
pGEM7ZF(+)/16S rRNA             CCTCTAGACTCGAGGAATTCCAGGTGTAGCGGTGAAATGCGTAGAGATCT
                                **************************************************
```

```
pCT123                  GGAGGAATACCGGTGGCGAAGGCGGCCCCCTGG-ACGAAGACTGACGCTC
pGEM7ZF(+)/16S rRNA     GGAGGAATACCGGTGGCGAAGGCGGCCCCCTGGGACGAAGACTGACGCTC
                        ********************************* ***************

pCT123                  AGGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCAC
pGEM7ZF(+)/16S rRNA     AGGTGCGAAAGCGTGGGGAGCAAACXGGATTAGATACCCTGGTA-TCCAC
                        ************************ ***************** *****

pCT123                  GCCGTAAACGATGTCGACTTGGAGGTTGTGCCCTTGAGGCGTGGCTTCCG
pGEM7ZF(+)/16S rRNA     GCCGTAAACGATGTCGACTTGGAA-TTGTGCCCTTGAXGCGTGGCTTCCG
                        ********************** ********** ***********

pCT123                  GAGCTAACGCGTTAAGTCGACCGCCTGGGGAGTACGGCCGCAAGGTTAAA
pGEM7ZF(+)/16S rRNA     GAGCTAACGCGTTAAXTCGACCGCCTGGGGAXTACGGCCGCAAXGTTAAA
                        *************** *************** ********** ******

pCT123                  ACTCAAATGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTA
pGEM7ZF(+)/16S rRNA     ACTCAAATGAATTGACGGGGGGCCGCACAAGCGGTGGAGC-TGTGGTTTA
                        ******************** ***************** *********

pCT123                  ATTCGATGCAACGCGAAGAACCTTACCTGGTCTTGACATCCACGGAAGTT
pGEM7ZF(+)/16S rRNA     ATTCGATGCAACGCGAAGAAC-TTACCTGGTCTTGACATCCACGGAA-TT
                        ******************** ************************ **

pCT123                  TTCAGAGATGAGAATGTGCCTTCGGGAACCGTGAGACAGGTGCTGCATGG
pGEM7ZF(+)/16S rRNA     TTCAGAG-TGAGAATGTGCCTTCGGGAACC-TGAGACAGGTGCTGCATGG
                        ******* ********************** ******************

pCT123                  CTGTCGTCAGCTCGTGTTGTGAAATGTTGGGTTAAGTCCCGCAACGAGCG
pGEM7ZF(+)/16S rRNA     CTGTCGTCAGCTCGTGTTGTGAAATGTTGGGXTAAGTCCCGCAACGAGCG
                        ******************************** *****************

pCT123                  CAACCCTTATCCTTTGTTGCCAGCGGTCCGGCCGGGAACTCAAAGGAGAC
pGEM7ZF(+)/16S rRNA     CAACCCTXATCCTTTGTTGCCAGCGGTCCGGCCGGGAACTCAAAGGAGAC
                        ******* ****************************************

pCT123                  TGCCAGTGATAAACTGGAGGAAGGTGGGGATGACGTCAAGTCATCATGGC
pGEM7ZF(+)/16S rRNA     TGCCAATGATAAACTGGGAGAAGGTGGGGATGACGTCAAXTCATCATGGC
                        ***** ********************************** **********

pCT123                  CCTTACGACCAGGGCTACACACGTGCTACAATGGCGCATACAAAGAGAAG
pGEM7ZF(+)/16S rRNA     CCTTACGAXCAGGGCTACACACGTGCTACATGGGCGCTTACAAAGAGAAG
                        ******** ******************** ****** ************

pCT123                  CGACCTCGCGAGAGCAAGCGGACCTCATAAAGTGCGTCGTAGTCCGGATT
pGEM7ZF(+)/16S rRNA     CGACCTCGCGAGAGCAAACGGACCTCATAAAGTGCGTCGTAGTCCGGATT
                        ************************************************

                                                    PCR/SmaI
pCT123                  GGAGTCTGCAACT-CGACTCCATGAAGTCTGCGGGGGGGTTCGAAATCGA
pGEM7ZF(+)/16S rRNA     GGAGTCTGCAACTTCGACTCCATGAAGTCTGCGGGGGGGTTCGAAATCGA
                        ************* ************************************
```

Figure 3.13 Sequence analysis of pCT123.
Alignment of actual sequence data derived from pCT123 with sequence information describing the constructed T7/SP6::*cat* transcriptional fusion, based on the sequence data for pGEM7Zf(+) and the *E. coli* rRNA operon gene (Genbank accession number JO1695). Identities are indicated by an asterisk, gaps are shown as a hyphen. In order to sequence the whole insert two sub-clones were made (pCT124 and pCT125), in which a fragment of the gene encoding the 16S rRNA had been cut out, and the vector re-ligated (Figure 3.14). Rows labelled pCT123 represent sequence data whereas pGEM7ZF(+)/16S rRNA represents published sequence data. Vector-derived nucleotides are presented in smaller fonts.

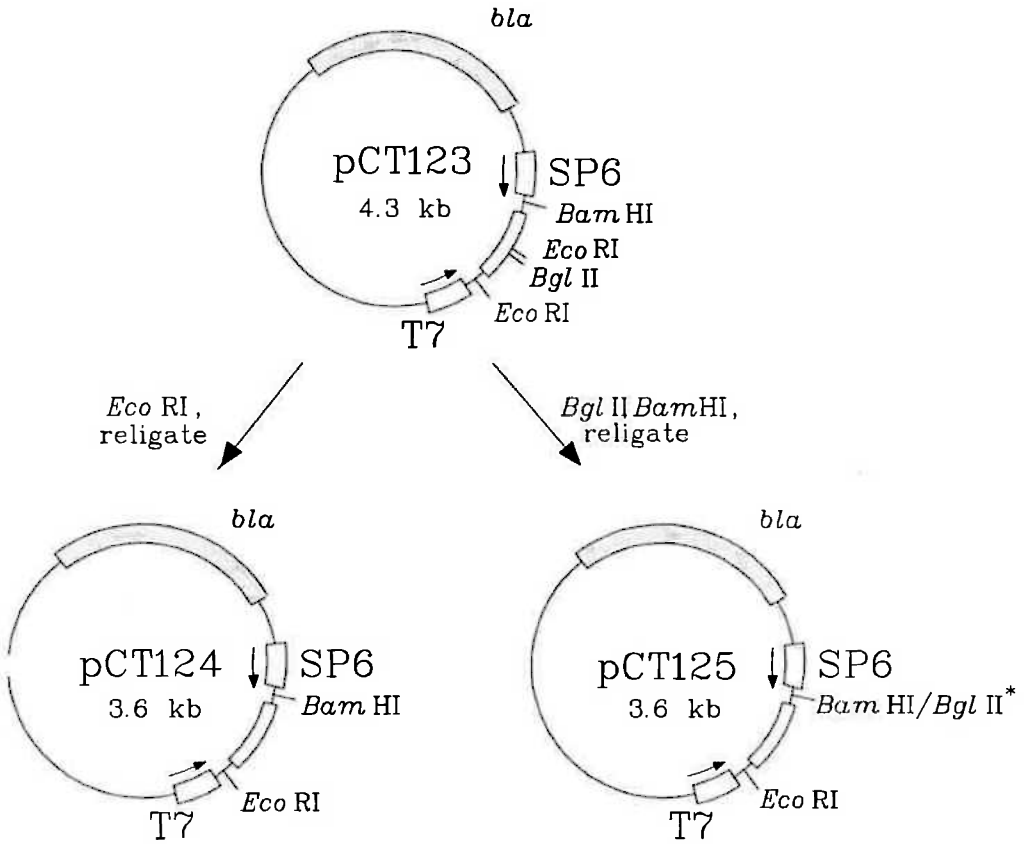

Figure 3.14 Construction of pCT124 and pCT125.
Plasmid pCT124 was constructed by cutting plasmid pCT123 with the restriction endonuclease *Eco*RV and ligating the vector. Plasmid pCT125 was constructed by cutting the plasmid pCT123 with the restriction endonucleases *Bgl*II and *Bam*HI and ligating the vector.

Chapter 4

Optimisation and validation of assays

4.1 Introduction

Considerable attention in this thesis is placed on optimisation, validation and development of analytical assays. Much published work has characterised bacterial expression systems qualitatively. In contrast, the use of accurate estimates allow quantitative characterisations to be carried out. Extensive research to prove that estimates are in fact accurate is therefore needed. Furthermore, a simple mathematical model describing CAT protein synthesis is used in this thesis (Chapters 5 and 6) to establish the relative efficiencies of transcription and translation. Mathematical modelling is only useful if based on reliable measures of variables.

In this chapter, the optimisation, validation and development of analytical methods to quantify relevant variables are described. Detailed studies of the recovery of CAT protein, total soluble protein, and total RNA are provided. Sample integrity during storage and accurate detection of CAT protein, *cat* mRNA, 16S rRNA, and plasmid DNA are also examined.

4.2 Results and discussion

4.2.1 Recovery of CAT protein from *E. coli*

Meaningful comparison of bacterial expression systems relies on a disruption method that completely releases the intracellular reporter protein without degradation. Techniques available for laboratory-scale disruption have previously been reviewed (Hughes *et al.*, 1971), and include thermal, enzymatic, chemical and mechanical methods. The most efficient method for cell disruption will totally release the desired product without degrading or denaturing it.

Chemical treatment with acetone and SDS can release intracellular proteins from *Escherichia coli* in equivalent amounts to sonication and bead-mill agitation (Bhaduri and Demchick, 1983). The presence of SDS makes this method unsuitable for the preparation

of extracts for enzyme assays, because enzyme activity may be inhibited. For example, detergents such as Triton X-100 and SDS inhibit CAT enzyme activity (Lu and Jiang, 1992). Combined chemical and enzymatic treatments are also effective, but results depend on growth phase (Dean and Ward, 1992). This complicates meaningful comparison of promoter strength for organisms in different growth phases.

Mechanical disruption methods such as ultrasonication and French pressing are often used at the laboratory scale to overcome the problems of incomplete and variable release associated with chemical and enzymatic disruption methods. However, sonication has also been reported to inactivate enzymes (Desai, 1968) and results in variable release (Desai, 1968; Feliu and Villaverde, 1994). Furthermore, the efficiency of the sonication treatment is affected by both sample volume (Feliu and Villaverde, 1994) and power output (Caldeira and Cabral, 1994). A comparison of the methods showed that the French press was more efficient at releasing proteins than sonication (Schmitt, 1976). However, it was not established that complete disruption was obtained.

In this section the effectiveness of sonication and the French press for cell disruption and release of CAT in both rich and minimal media were compared. Actual cell disruption was determined and compared by measurement of total CAT release. The effect of adding the detergent Triton X-100 following disruption by sonication and French pressing was examined. Furthermore, the accuracy of the CAT-ELISA assay was determined and the efficiencies of the various disruption methods were tested statistically.

The experimental work was carried out as follows:
Cell suspensions were prepared by cultivating *E. coli* JM101 (*TonA.*) transformed with the plasmid pCT102, in either nutrient broth supplemented with Ap (100 μg/mL) or minimal medium supplemented with thiamine (40 μg/mL) and Ap (100 μg/mL). CAT expression was induced at OD_{600} = 0.5 (rich media and minimal media) or 0.8 (minimal media) by adding IPTG to a final concentration of 0.4 mM.

Three hours post induction, the culture was washed and resuspended in buffer (10 mM Tris, 1 mM EDTA). Samples were stored, five times concentrated, at –70°C for later analysis. Cell samples were diluted to the original volume. Cell suspensions (2 mL) were

sonicated in 10 mL tubes using a Branson Sonifier, Model B-15. Other suspensions (10 mL) were disrupted using a French press operated at 100 MPa. Some of the disrupted samples were treated with detergent (Triton X-100, supplied with the CAT-ELISA kit). Cell disruption was quantified using an Applied Imaging disc centrifuge (Middelberg, 1992; Kleinig, 1997). CAT protein was measured as described in section 2.23.

Firstly, CAT protein recovery from cells cultivated in rich media was examined. Results are presented in Figure 4.1 CAT protein recovery from cells disrupted by sonication at four different power outputs was compared. Maximal CAT recovery was achieved at 35 Watts using two sonication cycles of 2 min. Increasing the power levels lead to decreased recovery of CAT as assessed by the CAT-ELISA method. This effect may be due to heat generation within the samples. Figure 4.2 compares CAT recovery using either sonication or French pressing. The power output for the sonicator was fixed to 35 Watts and an additional sonication cycle was included. All of the sonicated samples, (but only one set of the French-pressed samples) were treated with Triton X-100. Maximal CAT recovery was achieved after disruption by the French press followed by treatment with lysis buffer. Treatment with detergent was clearly necessary to achieve maximal CAT recovery possibly because CAT partitions to the cellular debris. Maximal CAT recovery after disruption by sonication was achieved after three cycles of 2 min. Pulsed sonication reduced recovery.

The previous experiment was repeated except minimal media was used. Figure 4.3 compares CAT recovery using either sonication or French pressing. Maximum CAT recovery from bacteria cultivated in minimal media was achieved after cell disruption by French pressing (2 passes) followed by treatment with Triton X-100. Maximal CAT recovery after disruption by sonication was achieved with 25 Watts after two cycles of 2 min. Increased power did not increase the recovery. As mentioned in the above paragraph, this was possibly due to heat generation within the sample. Previous experiments showed that maximal CAT recovery after sonication from bacteria cultivated in rich broth occurred for 35 Watts and three sonication cycles. Less vigorous disruption was required to release CAT protein from cells cultivated in minimal media. This is most likely due to the formation of a more rigid cell membrane after proliferation in rich broth. This is confirmed

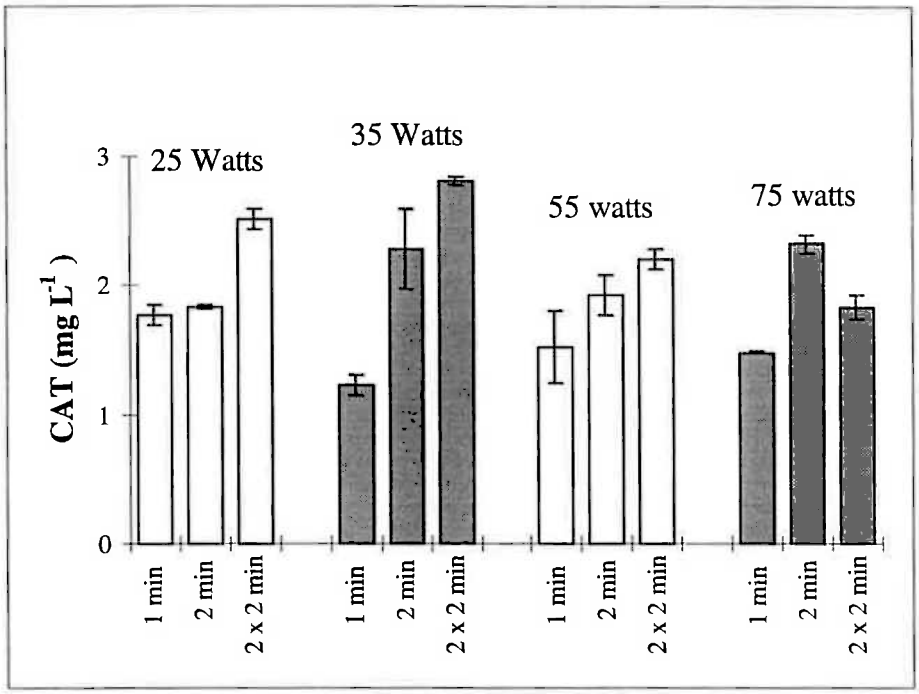

Figure 4.1. Recovery of CAT from *E. coli* cultivated in rich media after disruption by ultrasonication at increasing power levels. (Values shown as mean ± SD).

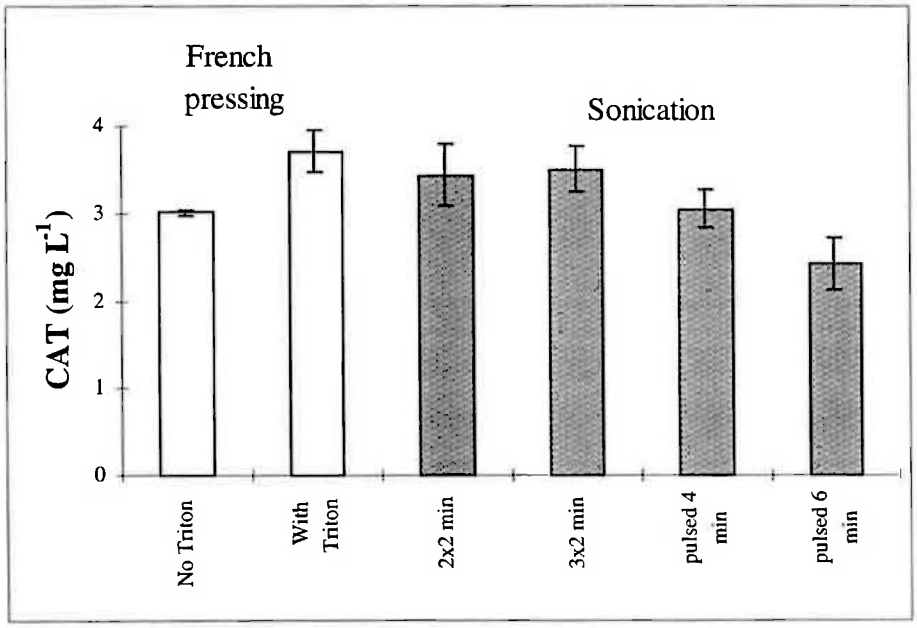

Figure 4.2. Recovery of CAT from *E. coli* cultivated in Rich media after disruption by French pressing or sonication. Sonicated samples were all treated with Triton X-100. (Values shown as mean ± SD)

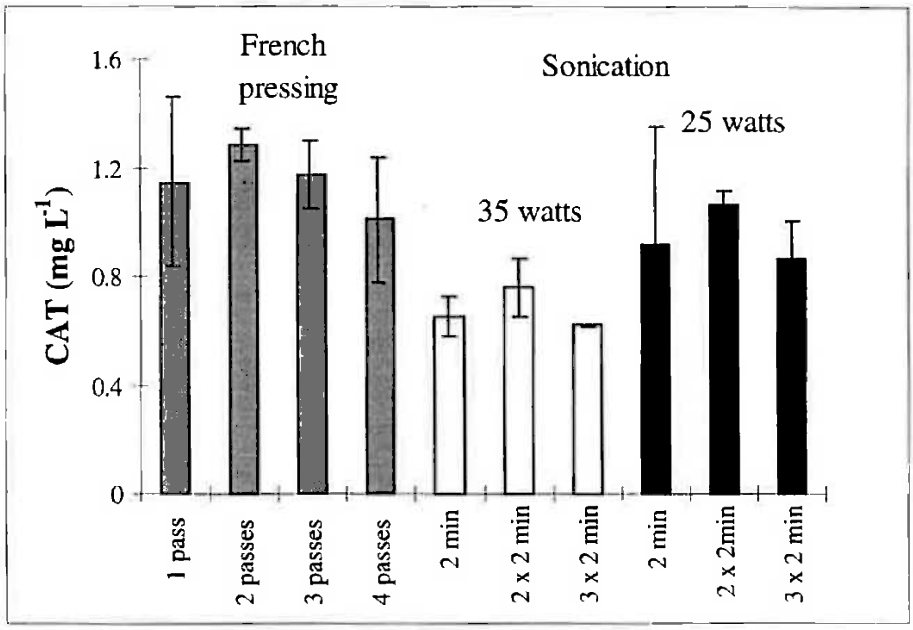

Figure 4.3. Recovery of CAT from *E. coli* cultivated in minimal media after disruption by French pressing or sonication at 25 W or 35 W. All samples were treated with Triton X-100 after disruption. Cell samples had been stored at −70°C before analysis (Values shown as mean ± SD)

by results by Leduc *et al.*, (1989) who showed that the peptidoglycan layer in stationary phase cells were 8.8 ±1.8 mm versus 6.6 ± 1.5 mm for cell in exponential growth phase. Table 4.1 illustrates the effect of Triton X-100 on CAT recovery from cells disrupted by French pressing. As shown earlier (Figure 4.2), detergent is necessary to achieve maximal CAT recovery. A comparison of Table 4.2 and Figure 4.3 suggests that the comparatively low recovery of CAT following sonication may be attributable to incomplete cell breakage. The French press totally disrupted the cell wall of *Escherichia coli*, whereas only a fraction of cells were disrupted following sonication. One pass through the French press achieved 99% disruption compared with 42% disruption for a single two minute sonicator treatment. A longer sonication cycle (3x2 min) only increased disruption to 80%. Data following sonication for 2 min show that less than half of the cells were disrupted but more than half the CAT protein was released. This observation might be due to partial disruption, where the cells appear intact but are sufficiently porous for CAT to be released. Similar observations have been reported by Kaback and Deuel (1969). They found that sonication resulted in product release, but that intact and sonicated samples scanned in an electron microscope appeared identical.

CAT recovery using French pressing followed by treatment with Triton X-100 was compared with recovery after chemical treatment alone (Table 4.3). Treatment with Triton X-100 alone is clearly a very inefficient method for cell disruption as only 19 % recovery was achieved compared to two passes in the French press followed by treatment with Triton X-100.

The finding that French pressing provided better disruption and release of CAT than sonication and that treatment with a solubilising agent such as Triton X-100 was needed in order to achieve maximal recovery, was tested statistically. Results are shown in Table 4.4. A Student's t-test showed a significant difference between CAT recovery following disruption by French pressing and sonication ($p = 0.001$). Similarly, treatment with Triton X-100 following disruption released significantly more CAT protein ($p = 0.005$).

Table 4.1. Recovery of CAT from *E. coli* cultivated in minimal media after disruption by French pressing. Values represents the average of two repeats.

With Triton X-100			No Triton X-100		
Passes	**CAT protein (mg L^{-1})**	**Standard deviation**	**Passes**	**CAT protein (mg L^{-1})**	**Standard deviation**
1	1.15	.31	1	.73	.33
2	1.28	.06	2	.86	.08
3	1.17	.12	3	.97	.07
4	1.01	.23	4	.79	.38

Table 4.2. Disruption of cells after French pressing or sonication as determined by analytical disc centrifuge.

French pressing		Sonication		
Passes	**Disruption (%)**	**Time (min)**	**Power (Watts)**	**Disruption (%)**
1	98.8	2	25	42
2	100	2×2	25	79.4
3	100	3×2	25	79.5
4	100	2	35	38.6
		2×2	35	62.2
		3×2	35	85.4
		4 (pulsed)	35	43.6
		6 (pulsed)	35	36.8

Table 4.3. Recovery of CAT from *E. coli* cells cultivated in minimal media after disruption by French pressing or detergent alone. Values represents the average of five repeats.

French pressing, 2 passes + Triton X-100		**Triton X-100**	
CAT protein (mg L^{-1})	**Standard deviation**	**CAT protein (mg L^{-1})**	**Standard deviation**
4.32	0.36	0.82	0.04

Table 4.4. Statistical analysis of recovery of CAT after disruption by French pressing or sonication.

	CAT protein (mg L^{-1})	**Standard deviation**	**Repeats**
French pressing + Triton X-100	1.095	0.178	10
French pressing	0.846	0.125	10
Sonication	0.666	0.156	10

4.2.2 Recovery of total soluble protein from *E. coli*

In this section the effectiveness of sonication and French pressing for release of total soluble protein from *E. coli* cultivated in either rich broth or minimal media are compared. Furthermore, the accuracy on the protein assay and the efficiency of various disruption methods are tested statistically.

The experimental work was carried out as in section 4.2.1. Total soluble protein are measured as described in section 2.24.

Results from cells cultivated in rich media are shown in Figure 4.4. Total soluble protein recovery from cells disrupted by different sonication treatments are compared. Maximal protein recovery was achieved at 25 Watts using two sonication cycles of 2 min. This is surprising as it previous was found that maximal CAT protein recovery was achieved at 35 Watts for cells cultivated in rich media (section 4.2.1). However, as was the case for CAT protein release (section 4.2.1), increasing the power levels lead to decreased recovery of protein. Figure 4.5 compares total soluble protein recovery using sonication or French pressing. Maximal soluble protein recovery was achieved after disruption by the French press followed by treatment with lysis buffer. Protein recovery after disruption by sonication was comparable for all treatments.

The previous experiment was repeated except minimal media was used. Figure 4.6 shows a comparison of total soluble protein recovery using sonication or French pressing. Maximal soluble protein recovery from bacteria cultivated in minimal media was achieved after cell disruption by French pressing (2 passes) followed by treatment with Triton X-100. Maximal protein recovery after disruption by sonication followed by treatment with Triton X-100 was achieved for 25 Watts with either two or three cycles of 2 min.

The finding that French pressing provides better disruption and release of total soluble proteins than sonication was tested statistically. Samples in this experiment were

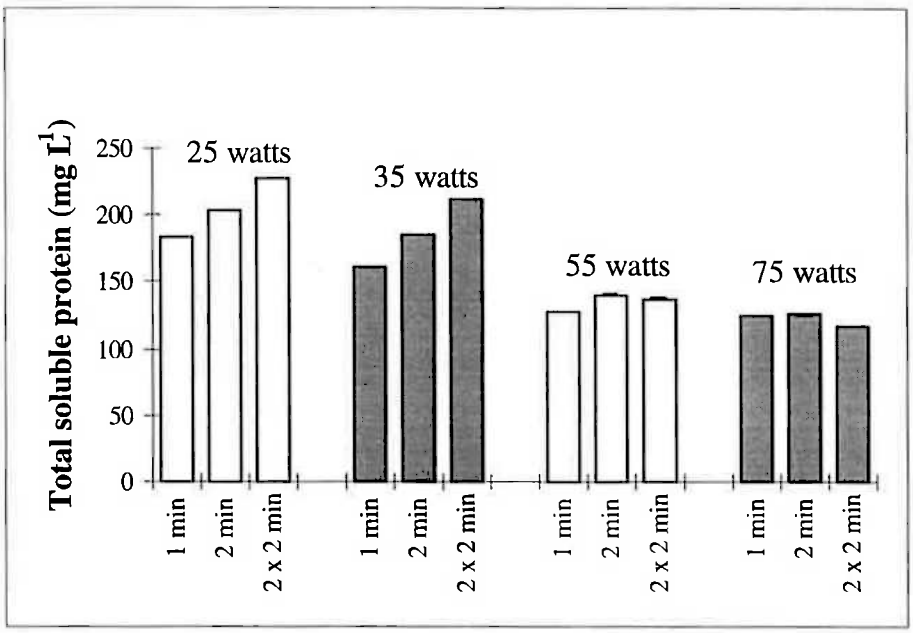

Figure 4.4. Recovery of total soluble protein from *E. coli* cultivated in rich media after disruption by ultrasonication at increasing power levels. (Values shown as mean ± SD).

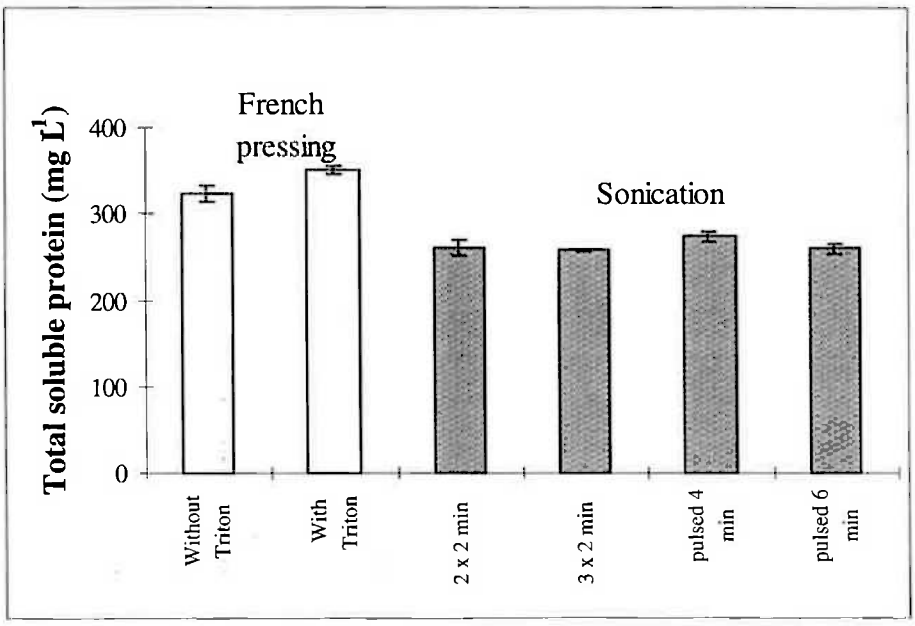

Figure 4.5. Recovery of total soluble protein from *E. coli* cultivated in rich media after disruption by French pressing or sonication. Sonicated samples were all treated with Triton X-100. (Values shown as mean ± SD).

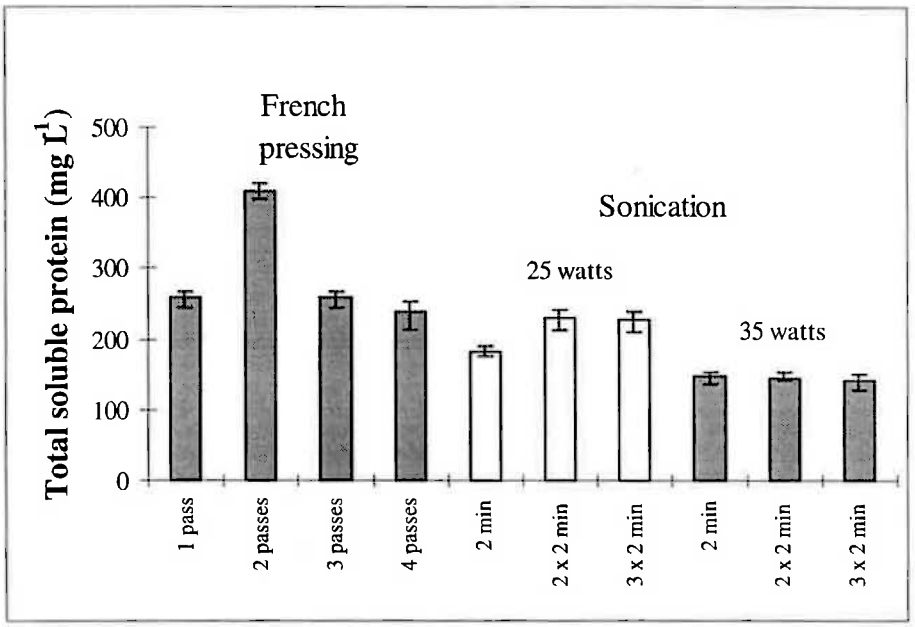

Figure 4.6. Recovery of total soluble proteins from *E. coli* cultivated in minimal media after disruption by French pressing or sonication at 25 or 35 W. All samples were treated with Triton X-100 after disruption. (Values shown as mean ± SD).

Table 4.5. Statistical analysis of recovery of soluble protein after disruption by French pressing or sonication.

	Total soluble proteins (mg L^{-1})	**Standard deviation**	**Repeats**
French pressing	595.07	37.83	10
Sonication	205.70	10.55	10

not treated with Triton X-100. Results are shown in Table 4.5. A Student's t-test confirmed a difference between protein recoveries following disruption by French pressing and sonication ($p << 0.01$).

4.2.3 CAT and total soluble protein integrity during sample storage

In a large study where many samples are collected during the course of an experiment, it is often necessary to store samples before analysis can be completed. From these constraints, reliable results can only be achieved if sample integrity is maintained during storage. The aim of the work described in this section was to examine whether cellular protein levels as measured by CAT-ELISA and standard protein assay changed during storage.

The experimental work was carried out as in section 4.2.2 except cells were cultivated in minimal media. CAT recovery after French pressing followed by treatment with Triton X-100 was estimated on twelve samples before and after storage for 22 h at –70°C.

Results are shown in Tables 4.6 and 4.7. Freezing affected CAT recovery (Table 4.6). Maximal CAT recovery was achieved after 1 pass before storage and after 2 passes after storage. This was contrary to expectations as freezing disrupts a significant fraction of cells according to the analytical data presented in Table 4.8 Differences in pressure release under disruption could explain this behaviour. However, this is unlikely as total soluble protein release does not follow the same trend (Table 4.7). It is possible that CAT associates with other cell components (eg. cell wall, proteins etc.) upon storage and therefore requires 2 passes after freezing. This could be due to the simple fact that freezing displaces water bringing proteins closer together.

There was no difference in total soluble protein recovery upon storage (Table 4.7). Maximal levels were achieved after 1 pass in the French press followed by treatment with

Table 4.6. Recovery of CAT from *E. coli* cultivated in minimal media before and after storage for 22 hours at –70°C. All samples were treated with Triton X-100 after disruption by French pressing. Results represents an average of three repeats from 3 samples (nine repeats).

Before storage			After storage (22 h)		
French Press passes	**CAT protein (mg L^{-1})**	**Standard deviation**	**French Press passes**	**CAT protein (mg L^{-1})**	**Standard deviation**
1	1.66	.01	1	1.47	.1
2	1.59	.04	2	1.6	.1
3	1.58	.10	3	1.35	.2
4	1.32	.01	4	1.12	.08

Table 4.7. Recovery of total soluble protein from *E. coli* cultivated in minimal media before and after storage for 22 hours at –70°C. All samples were treated with Triton X-100 after disruption by French pressing. Results represents an average of three repeats from 3 samples (nine repeats).

Before storage			After storage (22 h)		
French Press passes	**Total soluble proteins (mg L^{-1})**	**Standard deviation**	**French Press passes**	**Total soluble proteins (mg L^{-1})**	**Standard deviation**
1	306.1	29.7	1	302.4	39.5
2	283.9	7.4	2	274.7	5.4
3	275.6	7.4	3	292.2	11
4	264.5	14.4	4	265.5	30.8

Table 4.8. Disruption of *E. coli* cultivated in minimal media after disruption by French pressing before and after storage at –70°C determined by analytical disc centrifuge.

Before storage		**After storage (22 h)**	
French Press passes	**Disruption (%)**	**French Press passes**	**Disruption (%)**
0	0	0	50
1	96	1	99
2	100	2	100
3	100	3	100
4	100	4	100

Triton X-100. Disruption of cells before freezing was 96 % after one pass in the French press and 100% after two passes compared with 99 % after 1 pass for stored samples (Table 4.8).

4.2.4 Total RNA isolation

There are several methods available for isolating cellular RNA. Guanidinium isothiocyanate lyses cells, is very effective at inactivating ribonucleases, and is often used when extracting RNA from ribonuclease-rich tissue such as the pancreas (Chirgwin *et al.*, 1979). However, methods using this chemical are often not suitable for procaryotic cells such as *E. coli,* as it is not possible to obtain high-quality RNA. Methods specific for procaryotes are the "Hot Phenol Method" as described by Aiba *et al.* (1981), the DEPC method (Summers, 1970), and the method of Gilman (1989). The disadvantage of these latter methods is that they require speed and organisation to avoid RNA degradation. The aim of this experiment was to compare two methods (Gilman 1989 and Aiba *et al.* 1981) for the recovery of total RNA from *E. coli.* The effect of total cell loading on extraction was also tested for the best method.

The experimental work in this section was carried out as follows. M9 minimal medium, supplemented with thiamine (40 µg/mL) and Ap (100 µg/mL) was inoculated with a single colony of *E. coli* JM101 (*TonA*) transformed with the plasmid pCT103, and cultivated at 37°C. At OD_{600} = 0.8, CAT expression was induced by addition of IPTG to 0.025 mM. After induction for 3 h (OD_{600} = 1.69), the culture was chilled on ice and 0.9 mL samples were recovered. Cells were sedimented (12000×g, 1 min) and stored at –70°C for later analysis. Total RNA was extracted from 5 mL of cells as described in section 2.13.2.1. Total RNA was extracted from another 5 mL of cells using the method of Gilman (1989) as described in section 2.13.2.2. Two filters were prepared. RNA from each of the extractions was transferred to separate nylon membranes in increasing amounts, using a slot-blotting manifold as described in section 2.13.3. Filters were hybridised and detected as described in section 2.13.4 and 2.13.5.

a volume of cells equivalent to 3, 4, 5 and 6 mL of culture with an OD_{600} of 1 (3.33 mL; 4.44 mL; 5.56 mL and 6.67 mL) using the modified method of Aiba *et al.* (1981).

The intensity of signals produced from *cat* mRNA as a function of total RNA extracted from *E. coli* using the methods of Aiba *et al.* (1981) and Gilman (1989) are shown in Figure 4.7. A linear relationship was found for both samples but with different slopes (α (hot phenol) = 0.050, R^2 = 0.96; α (Gilman) = 0.033, R^2 = 0.88.), suggesting that more *cat* mRNA is extracted using the hot phenol method than the method of Gilman (1989).

Total detected RNA in *E. coli* cells as a function of mL of culture equivalent to an OD_{600} of 1 are shown in Figure 4.8. There is a linear relationship between total RNA extracted and the amount of cells from which the RNA was extracted, for a specific range of cell loadings (OD_{600} = 2 to 5).

4.2.5 RNA integrity during sampling and storage

Samples for RNA extractions were routinely stored at –70°C. As RNA samples were prepared and analysed at different time points and thereafter compared, it was critical to know that RNA integrity was maintained.

RNA samples are often snap-frozen in dry ice/ethanol or liquid nitrogen at the time of sampling. This is done to avoid fast degradation of RNA species. For example, most messenger RNA has a half-life of between 2 and 4 min (Belasco, 1993). However, previous results showed that freezing disrupts up to 50 % of cells (Table 4.8). RNA loss may therefore occur following snap-freezing of the cell suspensions. An alternative method to recover RNA was sought. Four experiments were conducted. The first experiment examined the loss of total RNA from a cell suspension after snap-freezing on dry ice/70% Ethanol. The second and third experiments compared recovery from cell precipitates before and after storage at –70°C for two days. These two experiments differed in that the t = 0 (zero) samples in experiment 2 were not snap-frozen before RNA extraction. The fourth experiment examined the effect of quick cooling of cell suspensions

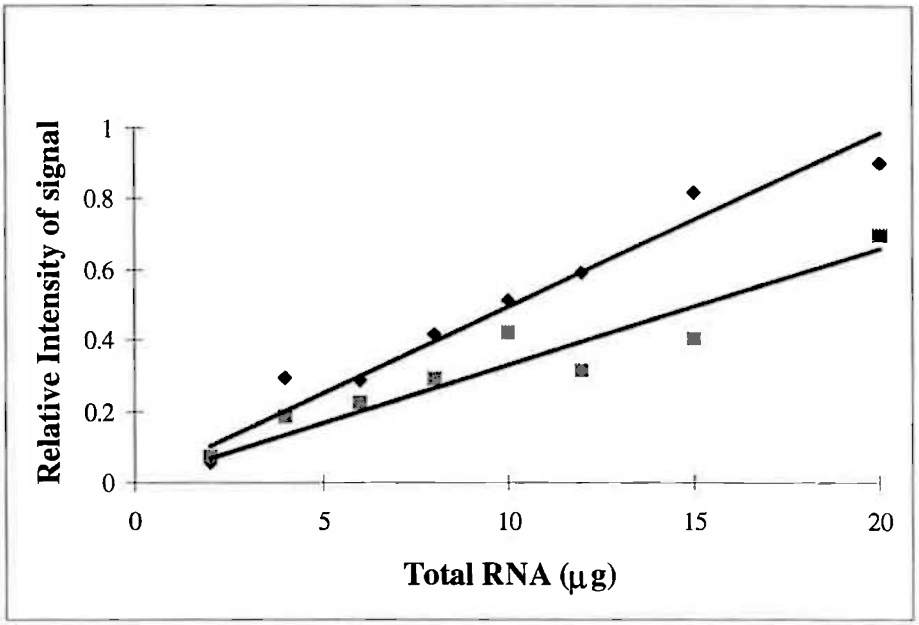

Figure 4.7. *cat* mRNA detected in total RNA extracted from *E. coli* JM101 (pCT103) using the method of Gilman (1989) (squares) or the hot phenol method (Aiba *et al.*, 1981) (diamonds).

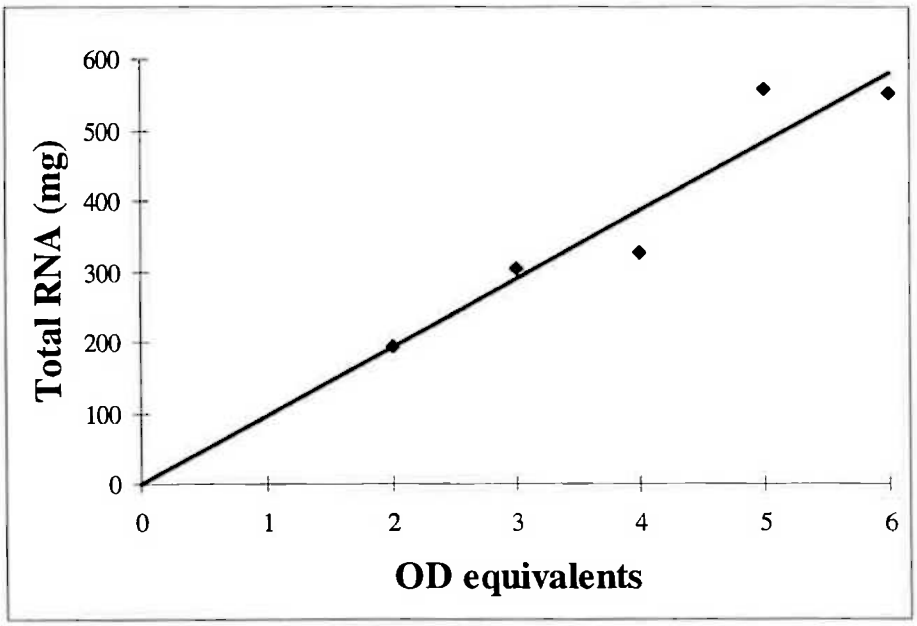

Figure 4.8. Effect of total cell loading on RNA extracted from logarithmic phase cultures of *E. coli* using the hot phenol method (Aiba *et al.*,1981).

in a sugar solution before centrifugation. Cell suspensions were chilled either a) on ice or b) by diluting (1:1) with ice-cold PBS/sucrose solution (20%).

Cell suspensions were prepared as follows. *E. coli* JM101 (*TonA*) transformed with the plasmid pPM1641 was cultivated in M9 minimal medium supplemented with thiamine (40 µg/mL) and Ap (100 µg/mL), at 37°C. Final OD_{600} was 0.5 (exp. 1) , 0.7 (exp. 2), 1.1 (exp 3) or 1.2 (exp 3). Cell samples (4 mL for experiments 1, 2 and 3 or 3 mL for experiment 4) were either sedimented (9800×g, 5 min, 4°C), snap-frozen in dry ice/70% (v/v) ethanol or quickly cooled down by mixing the samples with an equal volume of ice-cold PBS (1.21 g L^{-1} K_2HPO_4, 0.34 g L^{-1} KH_2PO_4, 8 g L^{-1} NaCl, pH 7.3)/20% sucrose. All samples were centrifuged before RNA extraction (9800×g, 5 min, 4°C). RNA was extracted as described in section 2.13.2.1.

Results from experiment 1 are presented in Table 4.9. Significantly more RNA was extracted from *E. coli* cells at time 0 than from cells that had been stored as a suspension. A Student t-test confirmed this ($p = 0.002$). Experiment 1 confirmed that some RNA was lost in the liquid phase during centrifugation of the cells, after storage. In the next experiment, RNA was extracted as previous from cells before and after storage. However, cells were this time centrifuged before freezing. Table 4.9 shows that more RNA was recovered from cell samples after storage than before storage (Table 4.9). A Student t-test showed that this difference was significant ($p = 0.008$). This suggested that the freezing process facilitated the release of cellular RNA. To examine this further, a third experiment was carried out. In this experiment, all cell samples were snap-frozen on dry ice/ethanol straight after centrifugation. Total RNA was extracted from half of the samples at time = 0 and from the other half after two days storage. There was no evidence of loss of RNA integrity upon storage when this procedure was followed ($p = 0.13$) (Table 4.9). Finally, a method to rapidly chill cell suspensions before centrifugation was explored (experiment 4). Total RNA extracted from cells mixed with an equal volume of ice-cold PBS/sucrose was compared to RNA extracted from cells chilled on ice. Equal amounts of RNA were recovered regardless how cells were chilled ($p = 0.42$). It was concluded that optimal recovery of RNA is achieved by collecting cell samples as quickly as possible, chilling them on ice, centrifuging (cold) and then snap-freezing the cells. This method is employed throughout the remainder of this thesis.

Table 4.9. Recovery of Total RNA from *E. coli* [pPM1641] cultivated in minimal media before and after storage at –70°C. Experiment 1 examined the loss of total RNA from a cell suspension after snap-freezing. Experiments 2 and 3 compared recovery from cell centrifugates. Experiment 4 examined the effect of quick cooling of cell suspensions in a sugar solution before centrifugation. Values represents an average of ten repeats.

	Before storage		**After storage**		
Experiment	**Total RNA (μg)**	**Standard deviation**	**Total RNA (μg)**	**Standard deviation**	**t-test p =**
1. Snap-frozen cell suspension after storage vs cell suspensions in PBS/sucrose (t=0).	269.6	50.7	177.9	25.8	0.002
2. Centrifuged cells before and after storage.	250.9	31.6	291.8	19.8	0.008
3. Centrifuged cells, snap-frozen before and after storage.	233.9	29.9	218.0	20.7	0.13
4. Cell suspension cooled a). on ice or b). in ice-cold PBS/sucrose.	a) 354.9 b) 393.1	a) 60.1 b) 68.0			0.42

4.2.6 Quantitation of *cat* mRNA

Hybridisation of a labelled probe to nucleic acids immobilised onto a solid support is widely used for the determination of specific RNAs. This method is at best semi-quantitative, but commonly applied. For example, a major problem is loss of sample during transfer for Northern analysis, resulting in underestimation of actual RNA levels. Further, autoradiographs prepared from agarose gels often show smears behind the bands (Stoeckle and Guan, 1993) also making quantitation problematic. This can be overcome by separating RNA species on polyacrylamide gel electrophoresis (PAGE) sequencing gels (Stoeckle and Guan, 1993), but accurate quantitation is still hampered by loss during transfer and by imprecise loading onto the gel. Using the slot-blotting technique, a transcript of interest is immobilised directly onto the membrane avoiding any loss of RNA. As a quantitative method, the slot-blotting procedure should therefore be more accurate than Northern transfer.

RNA immobilised onto a solid support using slot-blotting can be visualised after hybridisation to a labelled probe and exposure to X-ray film. The generated signal is quantified by scanning densitometry. mRNA levels are given as relative measures by comparing individual samples (Pease and Wolf, 1994; White and Bancroft, 1982) or as absolute levels by relating the signal to that of a pure standard. Immobilised RNA can be detected with either radioactive or non-radioactive labels. Recent advances in non-radioactive probe detection systems have significantly improved sensitivity to the extent that they now offer comparable performance to radiolabelled probes (Furuta *et al.*, 1990). The advantages of non-radioactive probes, apart from the safety and environmental aspects, include reduced exposure times, improved probe stability, and often lower background signal.

cat mRNA levels expressed by *E. coli* have been determined by Northern blotting (deFranco and Schottel, 1989; Chen *et al.*, 1995) and by comparison with an *in vitro* produced standard (Meyer and Schottel, 1991) using the dot-blotting technique. As described above, transfer of nucleic acids from a gel onto a solid support underestimates the true concentration. Furthermore, it was discovered during this work that quantitation of

cat mRNA by slot-blotting is highly sensitive to the total amount of RNA co-immobilised onto the solid support (see experiment 2). Quantitation of message by comparing signal intensity to that of an *in vitro* produced standard results in gross underestimation of RNA concentration. Steric hindrance occurring from total cellular RNA possibly limits the probe's ability to anneal target mRNA.

In this section, the feasibility of isolating *cat* mRNA prior to immobilisation using Dynabeads M-280 Streptavidin is investigated (experiment 1). Furthermore, an improved method to quantify *cat* mRNA in *Escherichia coli* using slot-blotting is presented (experiment 2).

Experiment 1: A 720 bp or 100 bp DNA fragment encoding all or part of the *cat* gene were synthesised by PCR (section 2.11.2). The primary strand DNA probe was attached to the Dynabeads (section 2.18). *cat* mRNA was isolated from total bacterial RNA by solution hybridisation (direct capture as described in the Dynabead manual) with the Dynabeads, and eluted.

Different hybridisation solutions were examined: 1) 6×SSC (0.9 M NaCl, 0.09 M Na-citrate), 2) 2×SSC (0.3 M NaCl, 0.03 M Na-citrate, 3) 6×SSPE (0.9 M NaCl, 0.06 M NA_2PO_4, pH = 7.4, 0.006 M EDTA), 4) 2×SSPE (0.3 M NaCl, 0.02 M NA_2PO_4, pH = 7.4, 0.002 M EDTA) or 5) 6×SSC, 0.1% SDS. Hybridisation temperatures from 37°C – 42°C were examined.

Different elution strategies were examined: 1.) 2 mM EDTA (95°C, 5 min) 2.) high-salt elution buffers (6×SSPE) 3.) high-salt elution buffers (6×SSPE) containing 50% formamide or 4.) TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0), or 5.) 0.15 M NaCl. Temperatures from 50°C to 75°C were examined.

cat mRNA was transferred to a positively charged nylon membrane using a slot-blotting manifold as described in section 2.13.3 and quantified by densitometry following hybridisation and detection by chemiluminescence as described in sections 2.13.4 and

2.13.5. Binding capacity, stability of *cat* mRNA in hybridisation solution, hybridisation time, and elution from beads were examined.

Quantitation of *cat* mRNA by isolation using Dynabeads and detection by slot-blotting was not feasible. The binding capacity of the double stranded DNA fragment to the Dynabeads could not be established as only half of the secondary strand was recovered by melting. Furthermore, *cat* mRNA levels in solution decreased with time under hybridisation, but subsequently increased in the hybridisation step to levels above 100%. The fact that mRNA levels in the hybridisation solution decreased initially meant that it most likely was captured by the secondary strand attached to the Dynabeads. However, the subsequent increase in the hybridisation solution suggested that mRNA was quickly eluted again. The fact that mRNA was detected in the hybridisation solution in levels above 100% most likely means that the secondary strand, that was not removed during the binding step, was eluted during the hybridisation. Very little message was recovered during the elution step, confirming that message had possibly been eluted already in the hybridisation step. Based on these results it was concluded that it is not possible to quantitatively isolate *cat* mRNA from crude RNA extracts using Dynabeads.

Experiment 2: *cat* mRNA standard and the complementary DIG-labelled RNA probe were synthesised as described in sections 2.14 and 2.15. RNA was transferred to a positively-charged nylon membrane using a slot-blotting manifold as described in section 2.13.3. *cat* mRNA was quantified by densitometry following hybridisation and detection by chemiluminescence as described in sections 2.13.4 and 2.13.5. Total RNA was extracted from a shake-flask culture of *E. coli* JM101 (New England Biolab), as described in section 2.13.2.1.

Known amounts of standard *cat* mRNA were immobilised and hybridised with the DIG-labelled RNA probe to determine assay linearity (Figure 4.9). The same amount of standard was immobilised in combination with total RNA extracted from *E. coli* JM101 (Figure 4.9). A linear relationship for both samples was obtained, but with very different slopes. To exclude the possibility that this effect was due to degradation of *cat* mRNA caused by nucleases in the *E. coli* extract, the standard was immobilised in combination

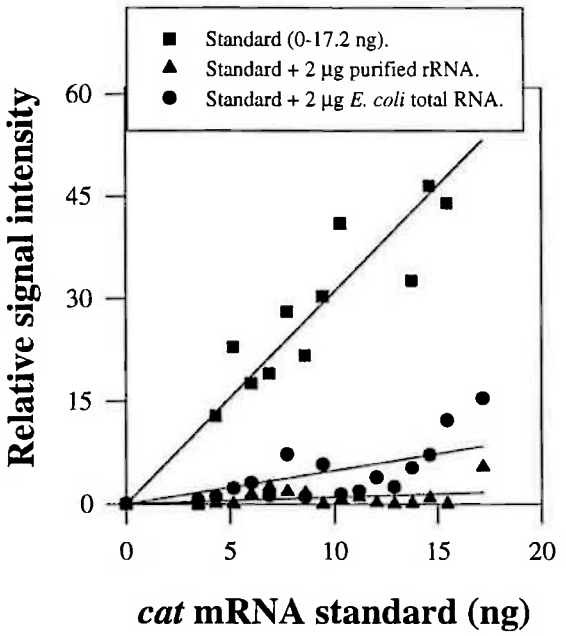

Figure 4.9. *cat* mRNA standard curve for pure *in vitro* produced standard, and in combination with increasing levels of either total RNA extracted from *E. coli* or commercially obtained 16/23S rRNA.

with commercially-purified 16S/23S ribosomal RNA (Figure 4.9). A similar reduction in response was observed. It was concluded that the quantitation of *cat* mRNA by filter hybridisation is affected by the presence of unrelated RNA, possibly due to steric effects. The quantitation of *cat* mRNA levels in crude RNA extracts determined by relating signal intensity to that of a pure standard is therefore highly inaccurate. The above result demonstrates the need to account for ribosomal RNA levels by constructing a three-dimensional calibration curve.

An example of the quantitation of *cat* mRNA levels in crude extract is demonstrated by Figures 4.10 and 4.11. A standard calibration curve combining five levels of *cat* mRNA (5-35 ng) with six levels of commercially-purified 16/23S rRNA (0-1 μg) is constructed (Figure 4.10). Ribosomal RNA level is plotted on the x-axis and intensity of the produced signal is plotted on the y-axis. The five contours correspond to five different levels of purified *cat* mRNA. The distance between contours reduces with increasing amounts of message as assay saturation is approached. In the same experiment, a sample with an unknown level of *cat* mRNA is loaded in increasing amounts to establish the linear range of the assay (Figure 4.11). Assay linearity is confirmed below 0.5 μg of total RNA loaded. The amount of *cat* mRNA corresponding to each point on the graph is determined by transferring the (x,y)-point to the standard calibration curve (Figure 4.10). The total amount of *cat* mRNA immobilised from the sample can then be determined by simple linear interpolation between contours. This can then be converted to a concentration by rationing with the total RNA level.

It was also found that the level of produced signal varied for different experiments (results not shown). All attempts to reduce variability were unsuccessful, and it was concluded that the variability was due to subtle differences in the probes. This effect was also found when using probes prepared as a single batch. Clearly, variability in both the levels of contaminating RNA and probe efficiency means that the assay must be calibrated each time it is used.

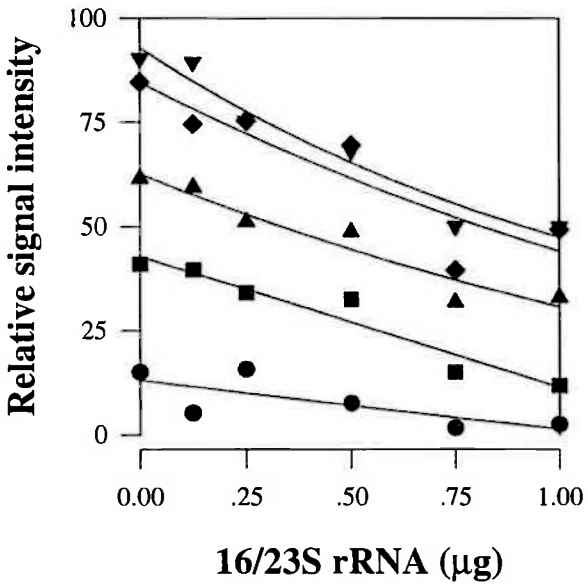

Figure 4.10. Standard calibration curve combining five levels of *cat* mRNA (5 (●), 10 (■), 15 (▲), 25 (◆), and 35 (▼) ng) and five levels of 16/23S ribosomal RNA. X-ray film exposure time was 30 seconds.

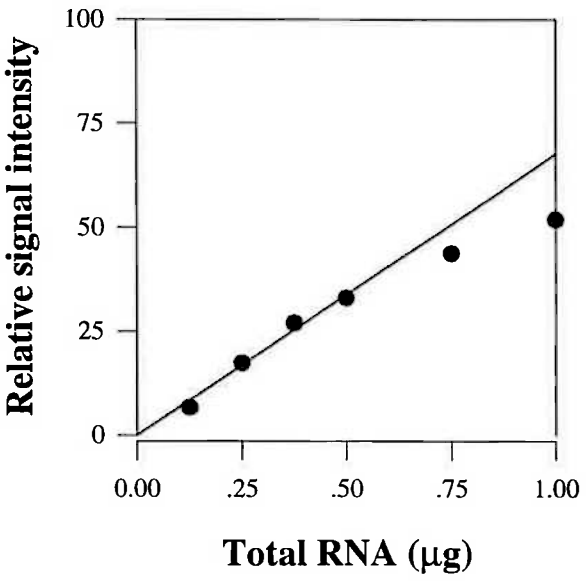

Figure 4.11. Relative signal intensity versus total RNA immobilised for a sample extracted from an induced *E. coli* shake flask culture. X-ray film exposure time was 30 seconds.

4.2.7 *cat* mRNA integrity during extraction.

The aim of this experiment was to examine *cat* mRNA integrity during the extraction of cellular RNA. An induced culture of *E. coli* producing CAT protein was spiked with a known amount of *cat* mRNA standard (195 ng). The theoretical amount of standard that was recovered after extraction was calculated as 20.8 ng. This was based on 195 ng of standard initially with 400 µL recovered from each 700 µL sample at each extraction (assuming nil degradation):

$$195\ \text{ng} \times \left(\frac{400\ \mu\text{L}}{700\ \mu\text{L}}\right)^4 = 20.8\ \text{ng}$$

This calculated amount was mixed with total cellular RNA extracted from a non-recombinant *E. coli* strain. Finally, *cat* mRNA was extracted from an induced, CAT-producing culture of *E. coli*. The rationale behind this strategy was that previous results showed that the quantitation of *cat* mRNA by slot-blotting in crude RNA extracts is highly sensitive to the amount of total RNA co-immobilsed onto the solid support (section 4.2.6). Individual samples therefore need to contain the same amount of total RNA. *cat* mRNA itself contributes this steric effect. Hence the need to spike induced cell samples with an *in vitro* produced *cat* mRNA standard.

The experimental work was conducted as follows. Two shake-flasks containing 11 mL of LB medium were prepared. The first flask was inoculated with a single colony of *E. coli* JM101 (New England Biolab), transformed with plasmid pPM1641, and cultivated at 37°C. At OD_{600} = 0.85, CAT expression was induced by the addition of IPTG to 0.4 mM for 2.5 h. Final OD_{600} after induction was 0.8. Ten samples (1 mL) were chilled and sedimented (12000×g, 1 min) and stored overnight as a pellet at –70°C for later analysis. The cell samples were defrosted and 195 ng of *in vitro* produced *cat* mRNA standard (section 2.14) was added to five of them. Total RNA was extracted from all of the ten samples as described in 2.13.2.1, with the modification that only 400 µL of the 700 µL of the top layer was recovered after each phenol extraction (four extractions total).

A second shake flask was inoculated with a non-recombinant *E. coli* JM101 and cultivated 37°C until OD_{600} = 0.8. Five samples were stored and total RNA was extracted as above. *cat* mRNA standard (20.8 ng) was combined with RNA extracted from 1 mL of the non-CAT producing *E. coli* culture. RNA was transferred to a positively charged nylon membrane using a slot-blotting manifold as described in section 2.13.3. *cat* mRNA was quantified by densitometry following hybridisation and detection by chemiluminescence as described in sections 2.13.4 and 2.13.5. Assay linearity was confirmed by including standards. Results are presented in Table 4.10. The results show that signal intensity from standard (20.8 ng) in combination with total RNA plus signal intensity from *cat* mRNA extracted from an induced culture in combination with co-extracted RNA equals signal intensity from *cat* mRNA extracted from an induced culture as well as from 195 ng of standard:

$$\text{Signal intensity}_{\text{Standard 20.8 ng + non-rec. culture}} + \text{Signal intensity}_{\text{Induced culture}} = \text{Signal intensity}_{\text{Standard 195 ng + induced culture}}$$

From the above results it was concluded that no significant amount of *cat* mRNA was lost during total RNA extraction, using the hot phenol method.

4.2.8 Determination of half-life ($t_{1/2}$) for *cat* mRNA

The aim of this next experiment was to examine *cat* mRNA stability in *E. coli.*

The experimental work in this section was carried out as follows. M9 minimal media supplemented with thiamine (40 µg/mL) and Ap (100 µg/mL), 140 mL, was inoculated with a single colony of *E. coli* JM101 (*TonA*) transformed with the plasmid pPM1641 and cultivated at 37°C. At OD_{600} = 0.9 CAT was induced by adding IPTG to a final concentration of 0.1 mM for 2 h. At this point rifampicin was added to a final concentration of 200 µg/mL to block mRNA synthesis. Samples (4 mL) were collected after 0, 1, 2, 4, 6 and 8 min into 4 mL of ice-cold PBS containing 20% sucrose. Cells were sedimented and RNA was extracted as described in section 2.13.2.1.

Table 4.10. *cat* mRNA integrity during the extraction from total cellular RNA. *cat* mRNA was detected by slot-blotting from a) total RNA extracted from recombinant *E. coli* spiked with a CAT standard b) CAT standard mixed with total RNA extracted from non-recombinant *E. coli* c) total RNA extracted from recombinant *E. coli*. Values represent an average of five repeats.

	Signal intensity	**Standard deviation**	**Number of samples (n)**
Standard (20.8 ng) + non-recombinant culture	973	33	5
Induced culture	4373	386	5
Standard (195 ng) + induced culture	5311	341	5

Chemical half-life is measured by blocking mRNA synthesis. Rifampicin binds to RNA polymerase and prevents transcriptional initiation. The decay of mRNA is assumed to follow first-order kinetics:

$$\ln\left(\frac{[mRNA]_t}{[mRNA]_0}\right) = -k \cdot t \qquad (4.1)$$

The rate constant k is obtained from the slope of a semilog plot of mRNA as a function of time (Figure 4.12). The slope was – 0.5657 min^{-1} and hence k = 0.6 min^{-1}. This gives a half-life $t_{1/2}$ = 1.16 min. Meyer and Schottel (1991) measured the *cat* mRNA half-life in MOPS/succinate medium and obtained $t_{1/2}$ = 0.9 min, which is comparable to the measured value in this work. A simple mathematical model to establish the relative efficiencies of transcription and translation are used in Chapters 5 and 6. The estimated half-life $t_{1/2}$ = 1.16 min was used as a model parameter in those Chapters

4.2.9 Plasmid copy number

The aim of this experiment was to establish the linear detection range for plasmid DNA.

The experimental work in this section was carried out as follows. M9 minimal media supplemented with Ap (100 µg/mL) was inoculated with a single colony of *E. coli* JM101 (*TonA*) transformed with the plasmid pPM1641 and cultivated at 37°C. At OD_{600} = 0.56, 1.8 mL of culture was collected (equivalent to 1 mL of cells with an OD_{600} of 1.0) and DNA was extracted as described in section 2.25. DNA was resuspended in water and linearised with *Eco*RV. The final volume was adjusted to 660 µL. Increasing volumes (10-40 µL) of plasmid DNA corresponding to 0.015-0.060 equivalents were loaded into a TBE agarose gel (Figure 4.13). One equivalent describes the unknown amount of DNA

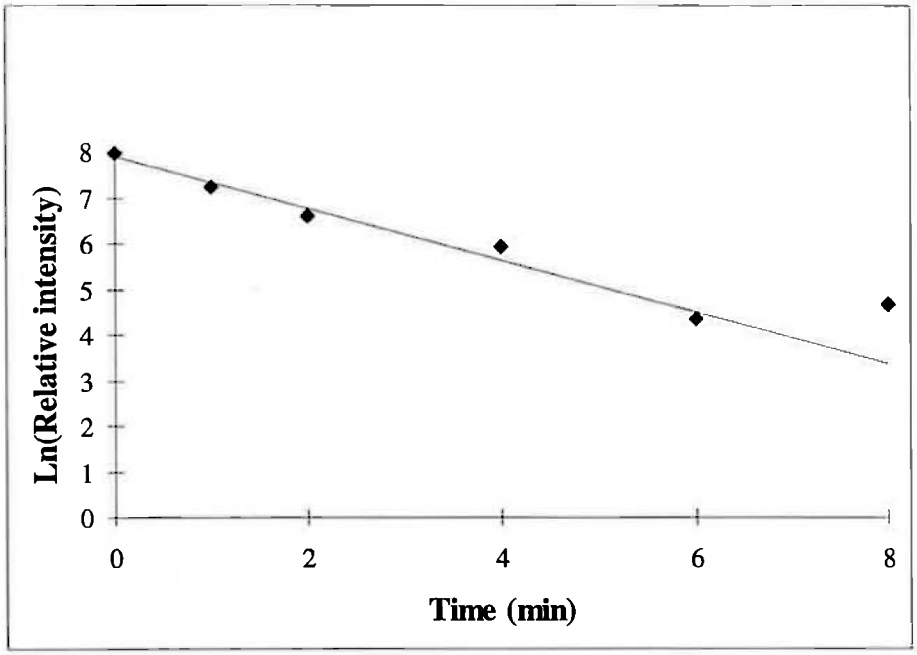

Figure 4.12. Determination of chemical half-life for *cat* mRNA. *cat* mRNA synthesis was blocked by adding rifampicin to a growing culture of *E. coli* (time = 0). *cat* mRNA was quantitated in total RNA extracts at various time-points.

extracted from 1 mL of cells with an OD_{600} of 1. In a second experiment, cells were cultivated to an OD_{600} of 0.84. 1.19 mL of cells were collected. Extracted DNA was linerised with *Eco*RV and the final volume adjusted to 445 µL. DNA sample was diluted 50/75, 50/100, 50/125 and 50/150 with water. A constant volume of 20 µL was loaded for all of the dilutions, corresponding to 0.045, 0.030, 0.225, 0.018, and 0.015 equivalents (Figure 4.14). Known amounts of DNA standard (pBR322 cut with *Hin*dIII) were also loaded in both experiments. The gels were electrophoresed, stained, photographed and scanned as described in section 2.25.

Unknown levels of DNA were estimated by comparing signal intensities to those of the standard. Figures 4.13 and 4.14 show that there is a linear relationship between signal intensity and plasmid DNA equivalent to DNA recovered from 0.030/0.045 mL of cells with an OD_{600} of 1.

4.2.10 Cell count

The aim of this experiment was to establish the linear cell counting range using the Coulter Counter.

The experimental work was carried out as follows. Cell number was measured as the particle number using a Coulter counter (section 2.19.1) Isoton II was used for all dilutions. Threshold was determined from counting latex spheres (1.1 µm). Assay linearity was examined by counting dilutions (40,000×; 20,000×; 13,333×; 10,000×; 8,000×; 4,000× and 2,000×) of an *E. coli* culture cultivated to OD_{600} = 3.40.

The minimum threshold that excluded background counting was 4 (Figure 4.15). Particle count for thresholds between 4 and 10 showed a plateau where particle count is mostly cells. A threshold of 8 was used in all experiments. The linear counting limit using the Coulter counter was found to be above 79000 (R^2 = .9989) (Figure 4.16). Higher counts were not tested as the orifice blocked easily.

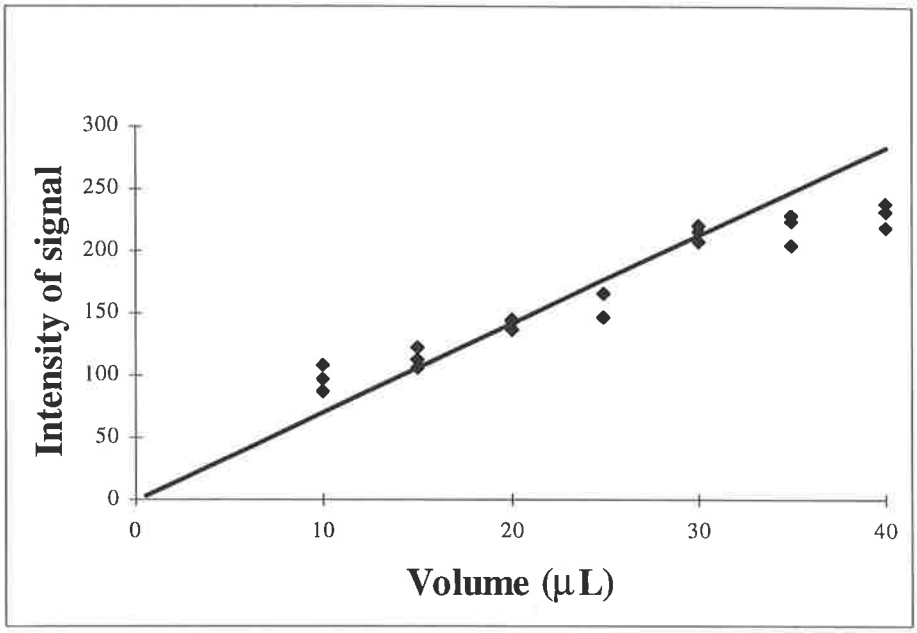

Figure 4.13. Quantitation of plasmid DNA by TBE gel electrophoresis and scanning densitometry. Increasing amounts of plasmid DNA were loaded onto the gel by increasing the sample volumes.

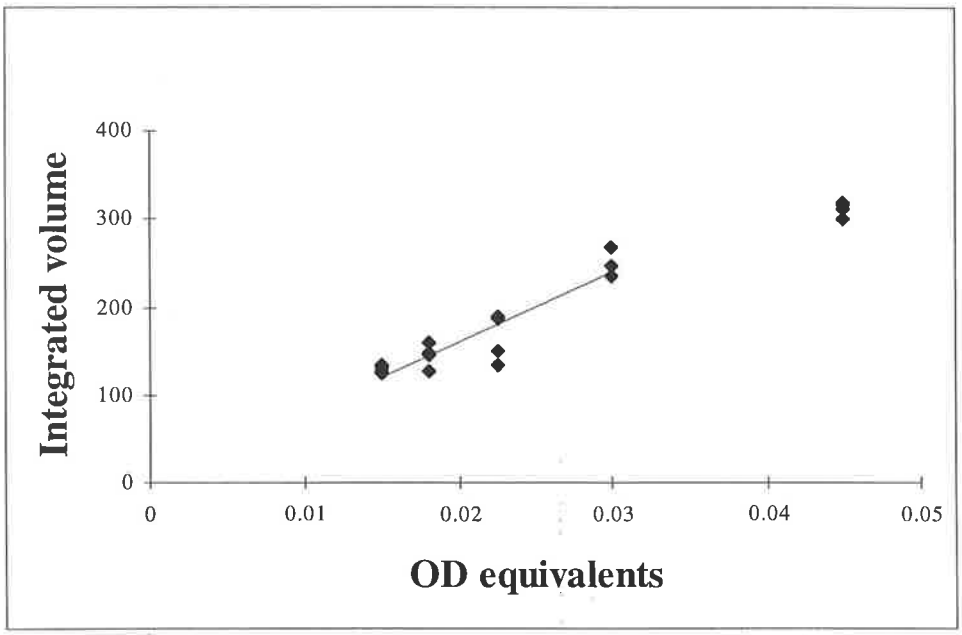

Figure 4.14. Quantitation of plasmid DNA by TBE gel electrophoresis and scanning densitometry. Increasing amounts of plasmid DNA were loaded onto the gel in a constant volume of 20 µL.

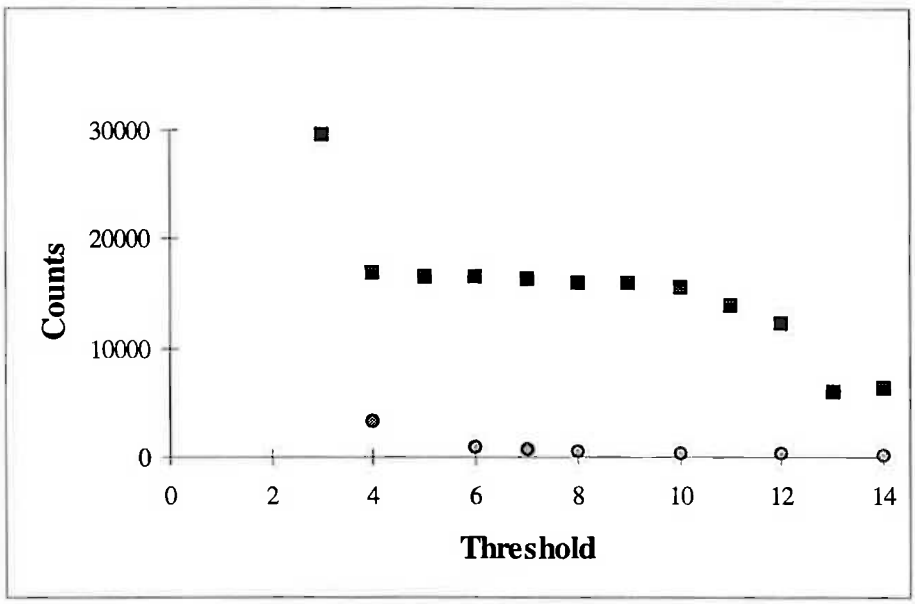

Figure 4.15. Determination of threshold for particle counting using the Coulter counter. Latex sphere count (squares) and Isoton II background (circles).

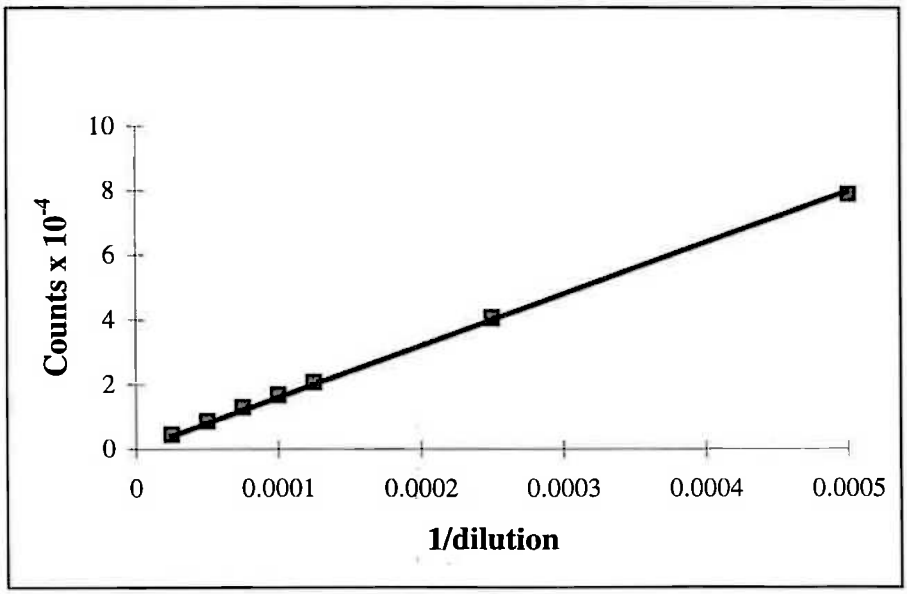

Figure 4.16. Determination of linear detection range for *E. coli* JM101 in the Coulter counter when diluted with Isoton II.

4.3 Conclusion

In this chapter, analytical methods to quantify CAT and total soluble protein recovery, and methods to isolate total RNA and plasmid DNA were developed, optimised, and validated. A method to quantitate *cat* mRNA was also developed. Assay linearity was established for cell counts, total RNA isolation, *cat* mRNA and plasmid DNA detection. Furthermore, the stability of *cat* mRNA in cells and also during extraction was examined. Finally, integrity studies of CAT, total soluble protein, and RNA during storage were described.

The use of the French press resulted in more efficient cell disruption and CAT release than sonication. French pressing totally disrupted bacteria whereas only a fraction of cells was disrupted following sonication. However, sonication apparently made undisrupted cells sufficiently porous for some release of CAT protein. Even when cells were fully disrupted only a fraction of total protein was released. CAT presumably associated with cellular debris. It was therefore necessary to treat homogenates with a solubilising agent such as the detergent Triton X-100. The need for additional treatment coupled with the inhibitory effect of detergents confirmed that quantitative measurement of total CAT provided a more accurate measure of promoter strength compared to quantitation by enzyme assays.

Detection of *cat* mRNA by slot-blot hybridisation was hampered by the presence of unrelated RNA. The intensity of the signal produced from an *E. coli* sample depended strongly on the amount of total RNA co-immobilised onto the filter. Quantitation of message by comparison of actual signal intensity with that of a pure standard resulted in significant underestimation. It is necessary to include a standard calibration curve for each assay, by combining pure standard at different concentrations with commercially-purified rRNA. As assay response varied between experiments, the calibration must be repeated every time it is employed.

The integrity of RNA, CAT, and total soluble protein were not compromised by storage. However, freezing affected CAT recovery, so that more vigorous disruption was required.

In contrast to much published work characterising bacterial expression systems qualitatively, the use of accurate estimates of data allow quantitative characterisations to be carried out. In this thesis, CAT protein production under control of either IPTG or stationary-phase inducible promoters is examined, and the two types of expression systems are compared for their ability to produce recombinant proteins. Furthermore, a simple mathematical model describing CAT protein synthesis is used to establish the relative efficiencies of transcription and translation. Accurate estimates of key parameters makes it possible to conduct these tasks reliably.

Chapter 5

Recombinant protein production under control of IPTG inducible promoters

5.1 Introduction

A variety of bacterial expression systems have been designed for high-level protein production. The most efficient systems offer tight control of transcription and translation, as obtained after promoter sequence optimisation (Brosius *et al.*, 1985) or optimal vector design (Chen *et al.*, 1991). The most widely used promoters, including *tac*, pL and T7, are induced either chemically (for example by isopropyl-β-thiogalactopyranoside (IPTG)) or by a temperature increase.

It is well known that a high expression of recombinant protein after induction can lead to a marked decrease in overall biomass yield (Andersson *et al.*, 1996) and can have detrimental effects on both host and recombinant protein. This is supported by previous studies showing that CAT protein production is dependent on the time and mode of induction (Bentley *et al.*, 1991). Furthermore, elevated expression of foreign proteins in *E. coli* can result in increased activation of proteases, and therefore product degradation, due to cellular stress responses (Ramirez and Bentley, 1995). Over-expression of proteins such as CAT, which is considered endogenous to some strains of *E. coli*, can lead to enhanced production of proteases identical to those induced by the stringent response (Harcum and Bentley, 1993). Rapid increases in CAT concentration on de-repression by 5 mM IPTG were followed by an even faster decline. Cells induced with lower IPTG concentrations did not exhibit this phenomenon. It is clear that induction of protein expression with IPTG requires careful optimisation.

A comprehensive model to explain expression of recombinant proteins in *E. coli* has been proposed (Schuler *et al.*, 1979; Schuler and Domach 1982). This model has been improved by Peretti and Bailey (1986) to account for translation initiation and the distribution of RNA polymerase between promoters for stable RNA, mRNA, and plasmid-encoded mRNA. A further extension by Peretti and Bailey (1987) characterised the effect of plasmid insertion on metabolic activity on the cell and on gene expression. It included the effect of copy number, promoter strength, and RBS. The assumptions underlying this model were later modified by Laffend and Schuler (1994) to include details of mRNA synthesis and comparmentalisation of RNA species. At increased translational activities, rRNA synthesis is induced thus stimulating ribosomal protein synthesis. As the demand

increases, ribosomal protein becomes limiting due to competition between ribosomal protein mRNA and plasmid mRNA. mRNA must compete with plasmid mRNA and translational efficiency is reduced (Laffend and Schuler, 1994). Translational limitation of protein production as described by the Laffend and Schuler model was confirmed by Vind *et al.* (1993). They showed that efficiency of translation following induction by IPTG is not directly related to functional β-galactosidase mRNA levels for the *lacZ* gene on a high copy number plasmid. During periods of high transcriptional activity, protein production is limited by the concentration of free ribosomes. Under these circumstances, mRNA is degraded. This inevitably leads to sub-optimal expression of recombinant proteins.

The translational yield of mRNA is determined by many factors. These factors include the ability of the Shine-Dalgarno (SD) sequence to base pair to the 3' end of the 16S rRNA (Hui and de Boer, 1987, Jacob *et al.*, 1988), the distance between the ATG initiation codon and the SD sequence, and the potential for secondary structure formation within the translational control region. As an example of the first, a natural RBS of the *lacZ* gene was substituted for more efficient ones, originating from genes for very abundant proteins (Vind *et al.*, 1993). This provided a 2.5 fold improvement in yield and increased functional half lives four fold. Similarly, a point mutation in the RBS of the *lac* gene under control of the *tac* promoter resulted in 12.2-fold increase in protein yield and a 4.3-fold increase in specific mRNA, indicating that transcription of recombinant proteins is coupled to translation (Mattanovich *et al.*, 1996). However, even when the SD sequences are identical, the ability to compete for free ribosomes differs by more than a factor of two (Vind *et al.*, 1993). The *in-vivo* efficiency of an RBS might therefore be modulated by additional translational sequences.

Altering the distance from 7 to 13 bases between the ATG initiation codon and the SD sequence for the CAT protein did not significantly affect the yield of CAT (Schottel *et al.*, 1984). By contrast, reducing the potential for secondary structures within the translation control region altered CAT levels ten-fold.

There is indirect evidence that the overall frequency of translation initiation in *E. coli* is adjusted depending on the availability of precursors for protein synthesis (Gold, 1988). Translation may therefore be improved by amino acid addition. For example, the

addition of phenylalanine to an *E. coli* culture producing CAT, which is high in aromatic acids, resulted in an approximately two-fold increase in specific activity. Phenylalanine may improve the translational process by assuring an ample supply of amino-acetylated transfer RNAs for protein assembly. Alternatively, as CAT is significantly higher in aromatic amino acids than the average *E. coli* protein, IPTG induction might deplete the amino acid pool and elicit the stringent control. Stringent control would replenish amino acid pools by inducing degradation of non-essential proteins, possibly including CAT (Ramirez and Bentley, 1995). Phenylalanine addition may therefore "relax" the cells and reduce the cellular stress response.

Another factor responsible for a change in translational activity is growth conditions. Indirect evidence suggests mRNA's are generally translated more frequently as the growth rate increases (Bremer and Dennis, 1987). Furthermore, a change in growth conditions differentially affects RBS efficiency (Jacques *et al.*, 1992). For example, translation of *lacZ* mRNA was reduced ten-fold when comparing gene expression in acetate minimal media and rich media (Jacques *et al.*, 1992). Furthermore, the rate of translation elongation decreased upon a rich to minimal medium change (Pedersen, 1984).

In this chapter, CAT protein levels obtained at different IPTG concentrations are compared with *cat* mRNA levels, and the best inducer concentration for a batch system is established. Furthermore, the rate limiting steps in protein synthesis, under IPTG induced conditions, are explored by relating CAT protein to both *cat* mRNA and 16S rRNA. Other factors such as the effect of oxygen availability on expression of CAT, growth phase of the inoculum, and the starting OD_{600} for the fermentation are also examined. Experiments are conducted in both shake-flasks and fermenters to determine whether the results from each system are comparable. The systems are further characterised in mathematical models. The model described by Laffend and Schuler (1994) is very complex and requires monitoring of an extensive number of variables. It was im practical to collect this amount of data in the current study. A simple model describing CAT protein synthesis is used here to establish the relative efficiencies of transcription and translation in the system studied.

Experimental work

Shake-flasks

Shake-flasks (5 L) containing 500 mL minimal media (section 2.2.2.1) were inoculated 1:100 with an overnight culture of *E. coli* JM101, transformed with plasmid pCT103, also grown in minimal media, and cultivated at 37°C. CAT protein production was induced at OD_{600} = 0.8 or 2.4 by addition of IPTG at various concentrations. Samples of culture were withdrawn for analysis at several time points after induction.

Fermentations

An overnight culture of *E. coli* JM101, grown in minimal media (section 2.2.2.1), transformed with plasmid pCT103, was inoculated into 2 L fermenters (Applikon, Schiedam, The Netherlands) containing 1.2 L of minimal media to give a starting OD_{600} of either 5×10^{-4} or 5×10^{-5} in the fermenter. Fermentation conditions are described in chapter 2. CAT protein production was induced at OD_{600} = 0.8 or 2.4 by addition of IPTG at various concentrations. Samples of culture were withdrawn for analysis at several time points after induction.

CAT assay, RNA extractions, slot-blotting, hybridisation, RNA detection by chemiluminescence, and *cat* mRNA quantitation were performed as described in sections 2.23, 2.13.2.1, 2.13.3, 2.13.4, 2.13.5 and 4.2.6. *cat* mRNA standards and the DIG-labelled complimentary RNA probe were constructed as described in sections 2.14 and 2.15. 16S rRNA standards and the DIG-labelled complimentary RNA probe were constructed as described in sections 2.16 and 2.17. Microscopic observation of cells, and analysis of homogenates by analytical disc centrifugation, confirmed that CAT did not form as inclusion bodies. Growth rate was calculated using equation (5.1), where X is the dry cell mass weight.

$$\mu = \frac{1}{X}\cdot\frac{dX}{dt} \qquad (5.1)$$

5.2 Results and discussion

In this section, CAT protein levels obtained at different IPTG concentrations are compared with *cat* mRNA and 16S rRNA levels and the optimal inducer concentration is also established. Experiments are conducted in both shake-flasks and fermenters to determine whether the results from each system are comparable. Furthermore, the effects of oxygen availability on expression of CAT, the growth phase of inoculum, and the starting OD for the fermentation are also examined.

5.2.1 Shake-flask experiments

CAT protein levels and the corresponding mRNA levels from the first set of shake-flask experiments are shown in Figures 5.1A and 5.2A (Result are presented as per mL of culture to allow direct comparison with results shown in Figure 5.1B, for which cell number data were unavailable. Figure 5.1B is discussed later in the text). Protein production was initiated at OD_{600} = 0.8 by adding IPTG to a final concentration of 0, 0.1, 0.4 or 1 mM IPTG. CAT protein production was monitored for the next three hours. *cat* mRNA and protein were detected almost immediately following induction with IPTG. The highest levels of *cat* mRNA were obtained using IPTG inducer at a concentration of 1 mM. However, *cat* mRNA levels decreased significantly 1 or 2 h after induction when inducer concentrations exceeded 0.1 mM. By contrast, CAT protein accumulated steadily over the 3 h of the experiment for all inducer concentrations. However, the rate of accumulation decreased at higher inducer concentrations. The maximum level of CAT protein accumulation was detected at 0.1 mM IPTG.

A similar set of shake-flask experiments was performed next, however protein production was initiated at higher OD (OD_{600} = 2.4). IPTG concentration was varied from 0 to 1 mM. The resulting CAT protein levels and corresponding mRNA levels are expressed as per cell basis in Figures 5.1E and 5.2E. Results are also shown as total levels per mL (Figures 5.1C and 5.2C) to allow direct comparison with results in Figures 5.1A and 5.2A. The highest levels of both CAT protein and *cat* mRNA were obtained using an IPTG concentration of 0.4 mM, on both a total and a per-cell basis. However, both CAT

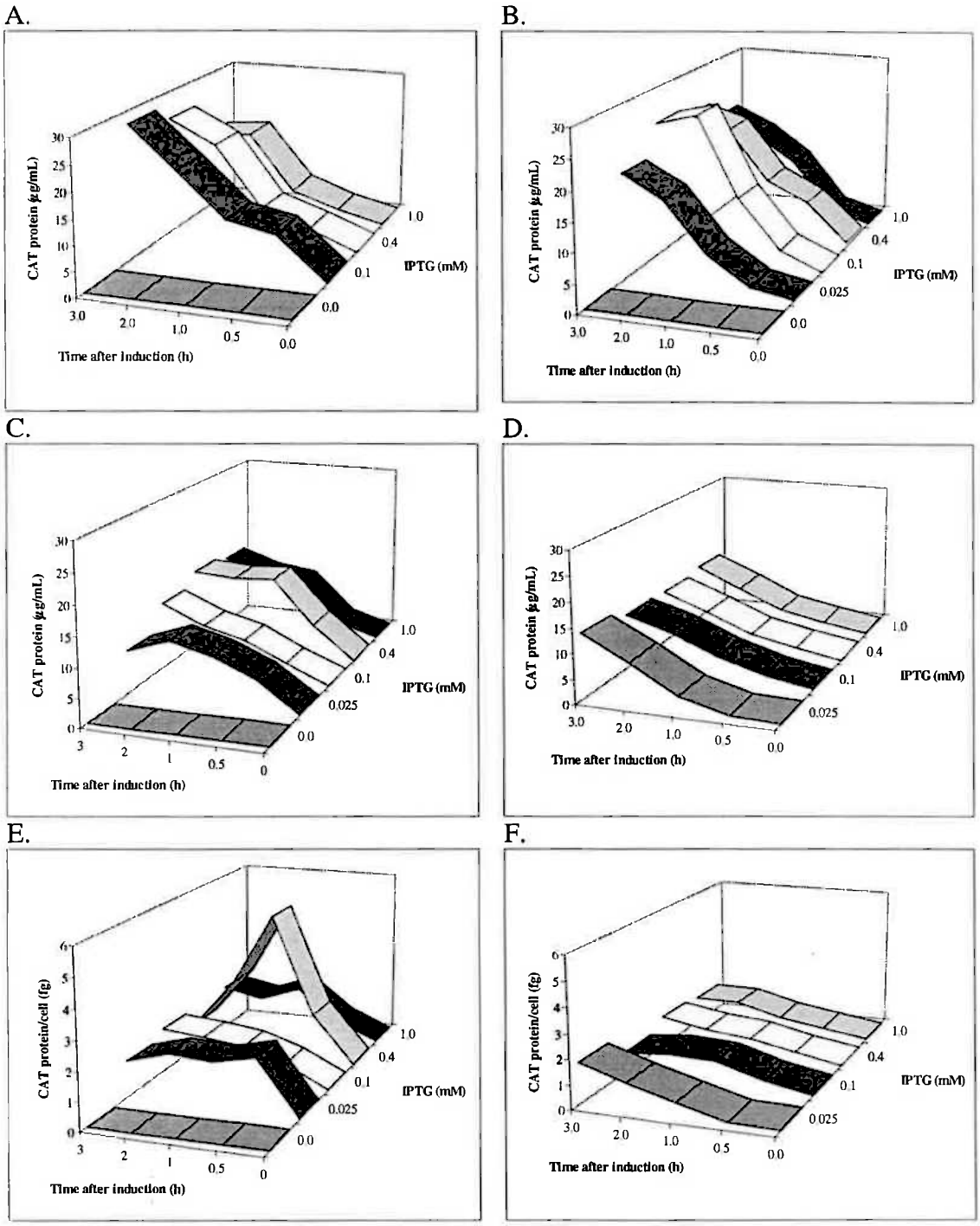

Figure 5.1. CAT protein under control of the *tac* promoter.

A. Shake-flask experiments. Protein production was induced at OD = 0.8 by addition of IPTG (0,0.1, 0,4, 1 mM).

B. Fermentation experiments. Protein production was induced at OD = 0.8 by addition of IPTG (0, 0.025, 0.1, 0,4, 1 mM).

C+E. Shake-flask experiments. Protein production was induced at OD = 2.4 by addition of IPTG (0, 0.025, 0.1, 0,4, 1 mM).

D+F. Fermentation experiment. Protein production was induced at OD = 2.4 by addition of IPTG (0.025, 0.1, 0,4, 1 mM).

Experiments were conducted in minimal media.

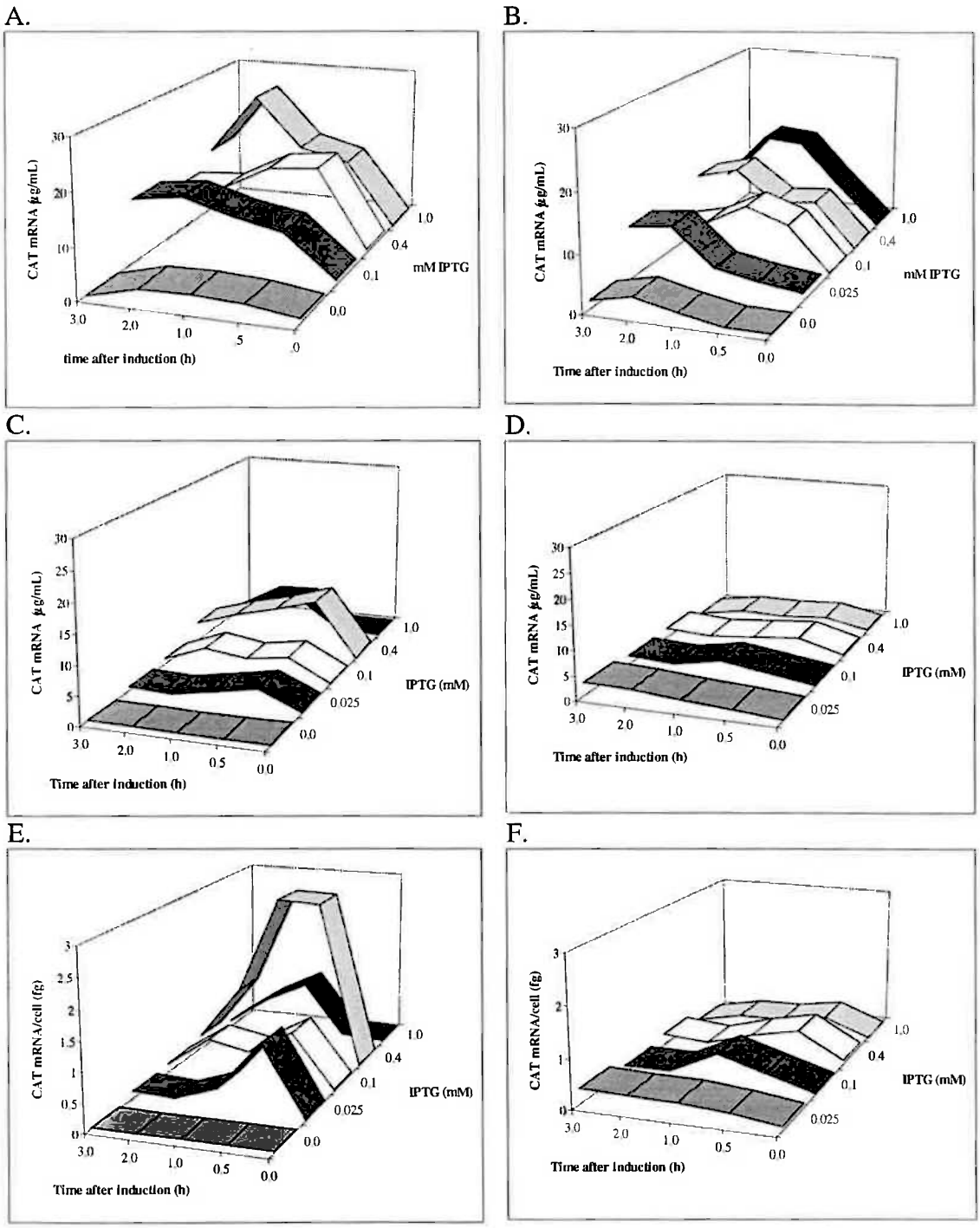

Figure 5.2 *cat* mRNA under control of the *tac* promoter.

A. Shake-flask experiments. Protein production was induced at OD = 0.8 by addition of IPTG (0,0.1, 0,4, 1 mM)

B. Fermentation experiments. Protein production was induced at OD = 0.8 by addition of IPTG (0, 0.025, 0.1, 0,4, 1 mM)

C+E. Shake-flask experiments. Protein production was induced at OD = 2.4 by addition of IPTG (0, 0.025, 0.1, 0,4, 1 mM)

D+F. Fermentation experiment. Protein production was induced at OD = 2.4 by addition of IPTG (0.025, 0.1, 0,4, 1 mM)

Experiments were conducted in minimal media.

protein and *cat* mRNA in the cells decreased dramatically after 1 h of induction. Induction at higher or lower concentration resulted in significantly lower accumulation of both CAT protein and *cat* mRNA. When comparing both CAT protein and *cat* mRNA after induction at two different cell densities (OD_{600} = 0.8 or 2.4), (expressed as total levels per mL in Figures 5.1A and C and 5.2 A and C), it is clear that levels are higher for induction at the lowest OD_{600}.

For comparison, protein production in shake-flask experiments induced at either low (OD_{600} 0.8) or high (OD_{600} 2.4) cell density are graphed in Figure 5.3. Cell count, CAT protein, *cat* mRNA and 16S rRNA levels per cell are shown. Optimal IPTG concentration (0.1 mM or 0.4 mM) was used for induction. A non-induced culture was included as a control. Assuming that productivity of the individual cell is identical for all cell densities, one would expect total protein yield to increase linearly with cell density. However, Figure 5.3 demonstrates that this is not necessarily correct. CAT protein and *cat* mRNA increased steadily after induction at OD_{600} = 0.8. By contrast, both CAT protein and *cat* mRNA per cell decreased significantly 1 h after induction at OD_{600} = 2.4.

The fact that the highest levels of CAT protein accumulation were achieved at low levels of IPTG inducer (0.1 mM), following induction at OD_{600} = 0.8, and that an increase in IPTG concentration resulted in a concomitant increase in mRNA level but not CAT protein, suggests that CAT protein production is severely limited at the translational level. This is expected, since protein production will be limited at the translational level because of competition for free ribosomes during periods of high transcriptional activity, brought about by induction of expression of proteins from a strong promoter, such as *tac*. However, protein expression in shake-flasks following induction at a high cell density (OD_{600} = 2.4) exhibited a different regulation (Figure 5.1E and 5.2E). In these experiments, maximum protein levels occurred at t = 1 h and decreased again thereafter. *cat* mRNA followed a similar pattern as CAT protein with peaks preceding periods of CAT accumulation. Much lower levels of both *cat* mRNA and CAT protein were observed for induction with 1 mM IPTG. Induction at OD_{600} = 2.4 is near the onset of stationary phase (OD_{600} = 3.1). Induction so close to stationary phase was possibly accompanied by either stimulated nuclease activity or transcriptional down-regulation and thereby lower levels of

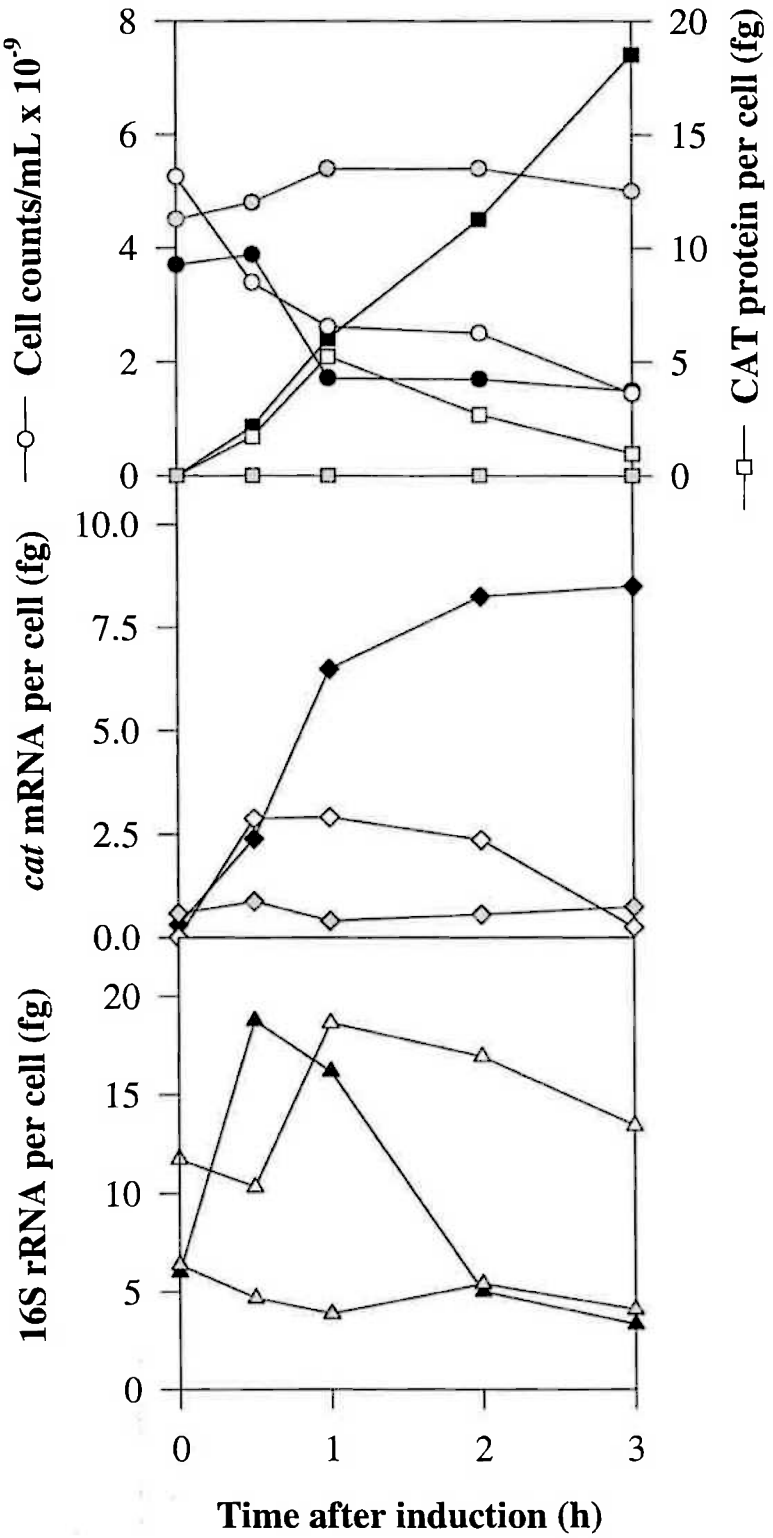

Figure 5.3. Total cell counts (●), and expression levels for CAT protein (■), *cat* mRNA (◆) and 16S rRNA (▲) per cell under control of the *tac* promoter in *E. coli* shake-flask. CAT protein production was initiated at OD_{600} = 0.8 by addition of 0.1 mM IPTG (black) or OD_{600} = 2.4 by addition of 0 (grey) or 0.4 mM IPTG (white).

CAT. A similar reduction in CAT protein after induction just prior to stationary phase has been reported by Bentley *et al.* (1991).

The concentration of ribosomes was estimated indirectly by measuring 16S rRNA. Cells responded to induction by increasing the 16S rRNA concentration (Figure 5.3) and therefore probably the number of ribosomes, thus increasing the translational capacity. By contrast, the 16S rRNA levels for non-induced cells remained constant. However, the high level of 16S rRNA was not sustainable when cells produced relatively high levels of CAT protein (induction at OD_{600} = 0.8). A plausible explanation is ribosomal competition between ribosomal protein mRNA and plasmid mRNA resulting in exhaustion of the ribosomes. This hypothesis is supported by the fact that 16S rRNA levels remain roughly constant when lower levels of CAT were obtained (induction at OD_{600} = 2.4) (Figure 5.3).

5.2.2 Fermentation experiments

The previous shake-flask experiments were repeated using fermenters to determine whether the two systems were comparable. Induction of CAT protein at both low and high cell density was examined.

In the first set of fermentations, CAT protein production was initiated at OD_{600} = 0.8 at IPTG concentrations of 0, 0.025. 0.1, 0.4 or 1 mM. The results for *cat* mRNA and corresponding protein levels are shown for ease of comparison in Figures 5.1B and 5.2B. Results are presented as per mL of culture as cell numbers were not determined. Expression of *cat* mRNA and protein was initiated immediately after induction of *cat* with IPTG. The highest levels of *cat* mRNA were obtained using IPTG inducer at a concentration of 1 mM. However, the levels decreased 2 h after induction. Lower inducer concentration resulted in either decreased or constant levels of *cat* mRNA after 1-2 h. Maximal CAT protein levels were achieved for 0.1 mM IPTG. All inducer concentrations resulted in a steady accumulation of CAT protein over the 3 h of the experiment. These results are similar to shake-flask experiments, induced under the same conditions. No significant increase in cell mass was observed following induction for either the shake-flask or fermenter systems.

In a second set of fermentations, CAT protein production was initiated at OD_{600} = 2.4. The cell density was expected to be three times that at OD_{600} = 0.8. Hence, the glucose concentration was increased threefold (from 3.3 g/L to 10 g/L) to avoid induction prior to stationary phase. In this experiment, four fermenters were run in parallel. CAT protein production was induced at OD_{600} = 2.4 with final IPTG concentrations of 0.025, 0.1, 0.4 or 1 mM. CAT protein and *cat* mRNA per cell are shown in Figures 5.1F and 5.2F. Results are also shown as total levels per mL (Figures 5.1D and 5.2D) to allow direct comparison with results in Figures 5.1B and 5.2B. The highest levels of *cat* mRNA and CAT protein per cell occurred at an inducer concentration of 0.025 mM. However, differences in expression levels were minimal for all inducer concentrations, both on a total and a per-cell basis. Maximal CAT levels in shake flask experiments (induction at OD_{600} = 2.4) exceeded those in fermentation experiments (four-fold) (Figures 5.1E and F). Furthermore, when comparing fermentation results for relatively high cell density (OD_{600} = 2.4) with low cell density (OD_{600} = 0.8), cells were much better at producing CAT when induced at low cell density. These results suggest that CAT protein expression is dependent on the oxygen tension and that oxygen may have an inhibitory effect on production.

The possible effect of oxygen on CAT protein production was explored in the next set of experiments. Again, four fermenters were operated in parallel. These were inoculated to a starting OD_{600} of 5×10^{-5}, using the same inoculum. CAT protein production was initiated in two of the reactors at OD_{600} = 2.4 by adding IPTG to a final concentration of 0.4 mM. This concentration was chosen as it previously provided the highest protein levels (in shake flask experiments). The two remaining reactors were run as controls. After 1 hour of induction, the oxygen supply in two of the fermenters (induced and control) was terminated. Cell counts, CAT protein per cell, and corresponding mRNA levels are presented in Figure 5.4. Cell numbers increased for the non-induced culture when cultivated in an oxygen-rich environment. By contrast, cell numbers remained constant for induced cultures or when oxygen was restricted. Both *cat* mRNA and CAT per cell were monitored after induction. The maximum level of *cat* mRNA was detected 1 hour after induction. Maximum CAT protein per cell was obtained for t = 2 h and was constant during the next two hours of the experiment. There was no significant difference between CAT protein and *cat* mRNA following induction, irrespective of the oxygen

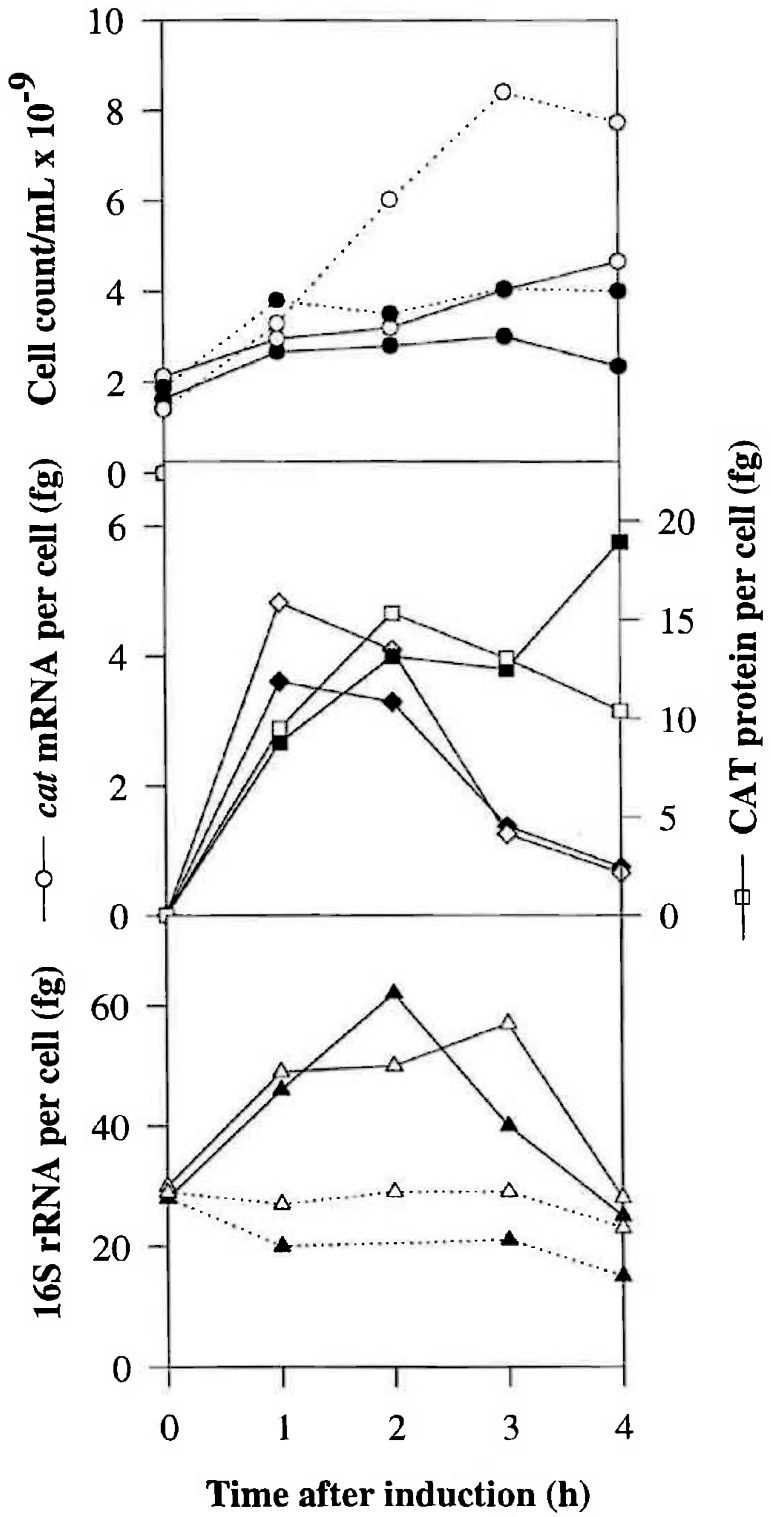

Figure 5.4. Total cell counts (●), and expression levels for CAT protein (■), *cat* mRNA (◆) and 16S rRNA (▲) per cell under control of the *tac* promoter in *E. coli* batch fermentation. CAT protein production was initiated at OD_{600} = 2.4 by addition of 0 (····) or 0.4 mM IPTG (—). Oxygen supply was kept above 30% (empty) or terminated after 1 h of induction (filled).

concentration. CAT and *cat* mRNA for non-induced cultures are not plotted but were close to zero throughout the experiment. These results failed to confirm any toxic or inhibitory effect from oxygen on CAT accumulation. The experiment was repeated (results not shown) to confirm this conclusion.

Previous shake-flask experiments indicated CAT protein expression was limited at the translational level (Figures 5.1 and 5.2). In these fermentation results (Figure 5.4), 16S rRNA levels increased after induction with IPTG to cope with the increased metabolic burden imposed on the cells, as found previously. The fact that CAT protein levels remained constant even after 16S rRNA levels decreased (t = 2 or 3 h), supports the argument of translational limitation of protein synthesis presented above. This is further supported by a decrease in *cat* mRNA which was not protected by ribosomes and therefore degraded.

CAT protein levels per cell varied between runs even when experiments were carried out under "identical" conditions. An attempt was made to elucidate the cause of this change. Possible causes included:

1. differences in inocula (growth phase);
2. differences in starting OD_{600} in the fermenter (plasmid instability).

The latter is plausible as the cells undergo more generations of growth before induction. As a consequence, the chance of plasmid segregational instability is increased. Again, a simple set of four fermenters were run in parallel. Two were inoculated to a starting OD_{600} of 5×10^{-5} using an inoculum in stationary phase (OD_{600} = 3) and the remaining two were inoculated to a staring OD_{600} of 5×10^{-4}, using an inoculum in logarithmic growth phase (OD_{600} = 1.1). CAT protein production was initiated at OD_{600} = 2.4 by adding IPTG to a final concentration of 0.4 mM. The results (specific growth rate and CAT protein per cell) are presented in Figure 5.5. Clearly, CAT protein levels increased during the four hours of the experiment, in the first pair. By contrast, in the second pair, the highest levels of CAT protein were achieved at 1 or 2 h after induction, followed by a slight decrease. High CAT protein levels were accompanied by a decrease in growth rate (μ) and *vice versa* (Figure

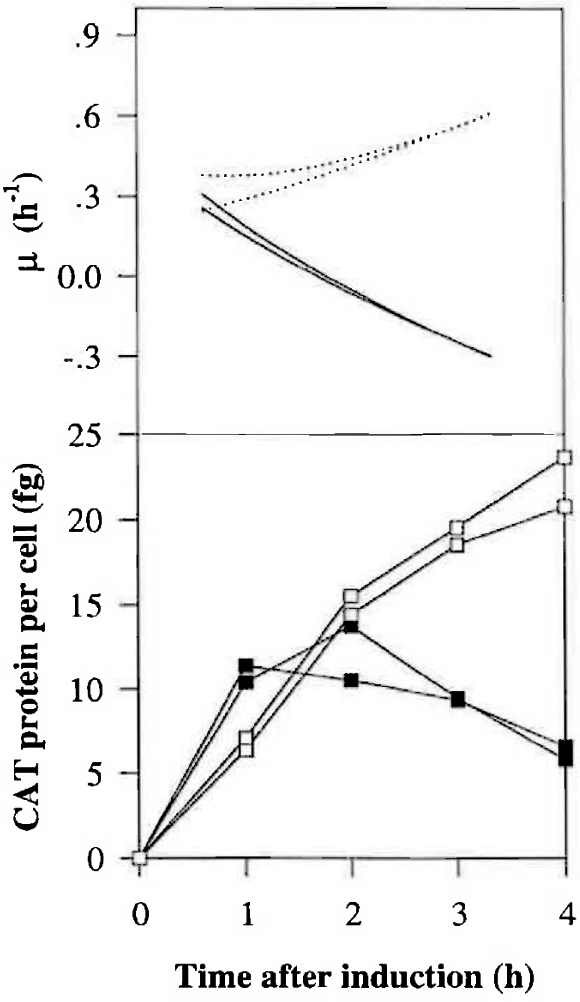

Figure 5.5. Specific growth rate and expression levels for CAT protein (■) per cell under control of the *tac* promoter in *E. coli* batch fermentation. CAT protein production was initiated at OD_{600} = 2.4 by addition of 0.4 mM IPTG. Starting OD_{600} after inoculation was 5×10^{-5} using a stationary phase culture (empty, —) or 5×10^{-4} using a logarithmic phase culture (filled, ·····).

5.5). This suggests that energy was channelled to recombinant protein production rather than biomass production. This provides evidence of tight regulation. However, it was unclear why there was a difference between expression levels in the experiments. The most plausible explanation remains the differences in inoculums. The highest levels of CAT was found for the fermenters inoculated with an inoculum in stationary phase. This may be explained by the ability for *E. coli* to adapt to changes in their environment. When cells enter stationary phase, new proteins are synthesised to make the cell more resistant to environmental stresses. It is quite possible that cells are able to 'remember' previous growth conditions so that when subcultured into fresh media and cultured to a new stationary phase, cells are much more resistant than bacteria reaching stationary phase the first time. Imposing protein production on bacteria is a major stress factor because of the increased metabolic burden.

The above experiments clearly showed that variations in results between fermentation experiments inoculated with the same culture were minimal (Figure 5.5). It is therefore valid to make comparisons between such experiments. Alternatively, the variation between results obtained from fermenters inoculated with different cultures were significant. Clearly, one needs to be cautious when comparing results from individual experiments.

It was also noted that the highest levels of CAT corresponded to the lowest starting OD_{600} (5×10^{-5}). Plasmid segregational instability was therefore an unlikely factor affecting the CAT protein levels. Furthermore, previous fermentation experiments, conducted as above did not show any evidence of plasmid segragational instability. This was evaluated by comparing viable cell counts on nutrient plates versus Ap selection plates.

The most plausible explanation for CAT variability between runs remained the differences in growth phase for inoculums. However, the two inoculums used in the above experiment were produced from separate single colonies. To exclude variability due to differences in the individual colonies, another experiment was conducted. Also, variability due to differences in starting OD_{600} was further investigated. An overnight culture of JM101 [pCT103] was subcultured into two shake-flasks and cultivated at 37°C to achieve

either stationary phase (OD_{600} = 3) or logarithmic phase inoculums (OD_{600} = 0.5) after 12 generations. Again four fermenter were run in parallel. Two were inoculated to a starting OD_{600} of either 5×10^{-3} or 5×10^{-5} using the stationary phase inoculum. The other two fermenters were inoculated to these same starting OD_{600}s using the logarithmic phase inoculum. CAT protein production was initiated at OD_{600} 2.4 by adding IPTG to a final concentration of 0.4 mM. Measured CAT protein levels are shown in Figure 5.6. Again, the highest levels of CAT was found for the fermenters inoculated with an inoculum in stationary phase. The results also indicated that an increase in starting OD_{600} of 10^2 (almost 7 generations) resulted in higher CAT values. This may suggest some plasmid segragational instability, in contrast to results above. However, CAT protein levels decreased some time after induction and differences became minimal after 4 h of induction. This decrease is in contrast to results in Figure 5.5. IPTG may have some toxic effect on the protein production and is possibly responsible for this effect. These results are therefore only indicative.

5.3 Model description

In order to optimise high-level expression of recombinant proteins, it is necessary to understand the rate-limiting steps. The use of an appropriate mathematical model to describe the process should aid in optimising vector design.

Simplistically, biosynthesis depends upon transcription of an encoding gene and translation of the resulting mRNA. The initiation frequencies of these events is encoded by the plasmid. Thus, the overall rate of plasmid gene expression depends upon the interaction of plasmid regulatory functions and host-cell biosynthetic activity. The material balance on the cloned gene message and protein can be described by the following differential equations (Lee and Bailey, 1984):

$$\frac{d(mRNA)_p}{dt} = k_{p_o}\eta(G)_p - k_d(mRNA)_p - \mu(mRNA)_p \qquad (5.2)$$

$$\frac{d(P)}{dt} = k_{qo}\xi_p(mRNA)_p - k_e(P) - \mu(P) \qquad (5.3)$$

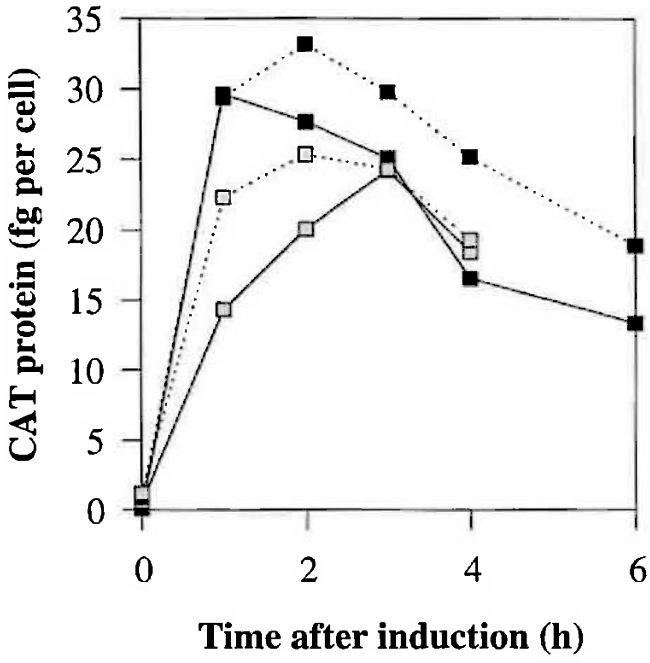

Figure 5.6. Expression levels for CAT protein per cell under control of the *tac* promoter. Fermenters were inoculated to a starting OD_{600} = 5×10^{-3} (black) or 5×10^{-5} (grey) using a stationary phase inoculum (····) or a logarithmic growth phase inoculum (—). CAT protein production was initiated at OD_{600} = 2.4 by addition of 0.4 mM IPTG

G describes the amount of plasmid (grams) in the cell and was estimated experimentally. Plasmid levels did not vary significantly after induction. mRNA concentration is influenced by the transcription rate, the efficiency of transcription (measured by the transcription efficiency factor (η)), and by the rate of decay (k_d). Furthermore, intracellular concentrations are used in this model. Consequently, the effect of dilution by growth is included. Protein synthesis is determined by the translation rate, the efficiency of translation (characterised by the translation efficiency constant (ξ)), the rate of decay (k_e), and the effect of dilution by growth.

5.3.1 Estimation of model parameters

The transcription rate constant (k_{po}) and the translation rate constant (k_{qo}) have been calculated as described by Lee and Bailey (1984) (see appendix A.1). The protein decay rate constant for CAT protein (k_e) was assumed to be the same as for a repressor protein (Lee and Bailey, 1984). The same assumption was made previously when modelling β-galalactosidase protein production (Chen *et al.*, 1991). The *cat* mRNA decay rate constant (k_d) was measured to be 0.6 min^{-1} (this thesis, section 4.2.8) and was assumed to be independent of growth rate as reported by Meyer and Schottel (1991).

The coupled differential equations were integrated using a fourth-order Runge-Kutta algorithm. The transcription efficiency (η) and translation efficiency (ξ) were estimated by minimisation of sum squared error (SSE) between experimental values and model prediction using Newton's search algorithm. This was performed using a spreadsheet package (Microsoft Excel, version 7), using the solver function.

5.3.2 Model regressions

Fermentation data for CAT expression after induction at OD_{600} = 2.4 by the addition of 0.4 mM IPTG were regressed to equations 5.2 and 5.3. Typical results are shown in Figure 5.7. Regressed and experimental data points are plotted against time. The plot confirms that there was good reproducibility between fermentation experiments initiated with the same inoculum. The plot also confirms that equations 5.2 and 5.3

Table 5.1. Model parameter values.

Parameter	Value	Reference
k_p	4.26 M mRNA/M DNA min	Appendix 5.1 (Lee and Bailey 1984)
k_d	0.6 min^{-1}	This thesis, section 4.2.8.
η	estimated by minimisation of sum squared error (SSE)	Table 5.2 and Figure 5.7
k_q	13.6 M protein/M mRNA min	Appendix 5.1 (Lee and Bailey 1984)
k_e	0.01 min^{-1}	Lee and Bailey 1984
ξ	estimated by SSE	Table 5.2 and Figure 5.6

describe the experimental data well. A parity plot comparing the regressed CAT protein values with experimental values is shown in Figure 5.8. Clearly, equation 5.3 provides a reasonable estimate of CAT protein levels. However, the model has a tendency to over predict lower values while under predicting higher values. The bias in the system suggests that the model may be too simplistic and additional parameters may have to be measured to provide improved prediction. More comprehensive models to describe recombinant protein expression in *E. coli* have been developed (Schuler and Domach, 1982; Perretti and Bailey, 1987; Laffend and Schuler, 1994). However, the aim of this work was to compare transcriptional and translational efficiencies to determine the limiting level of protein synthesis. This, together with the reasonable correlation between experimental and regressed data in this work using two differential equations, compared to twelve in Laffend and Schuler's model (1993), suggests that the latter is unnecessarily complicated for this system.

As outlined earlier, transcriptional (η) and translational (ξ) efficiencies were estimated by minimisation of an objective function OF given by:

$$OF = \sum_{i=0}^{N} (mRNA_{1,\exp} - mRNA_{i,pred})^2 + (P_{1,\exp} - P_{i,pred})^2 \qquad (5.4)$$

Values of these parameters for various runs are listed in Table 5.2. Fermentations initiated with the same culture are grouped and boxed by solid lines. It was clear that both η and ξ are consistent within each of the boxes. In contrast, there is a significant variation between the boxes. This follows the statement made earlier about variations in results between fermentation experiments, while variation between experiments inoculated with the same culture was minimal. From Table 5.2 it is shown that the highest values of maximum CAT protein correspond to the highest values of ξ (translational efficiency). Clearly, efficient translation means high protein levels. Interestingly, the highest values of ξ correspond to the lowest values of η (transcriptional efficiency). This means that lower transcriptional efficiency corresponds to better translation and hence higher CAT protein levels. These results suggest, that improving transcriptional activity in the cell does not

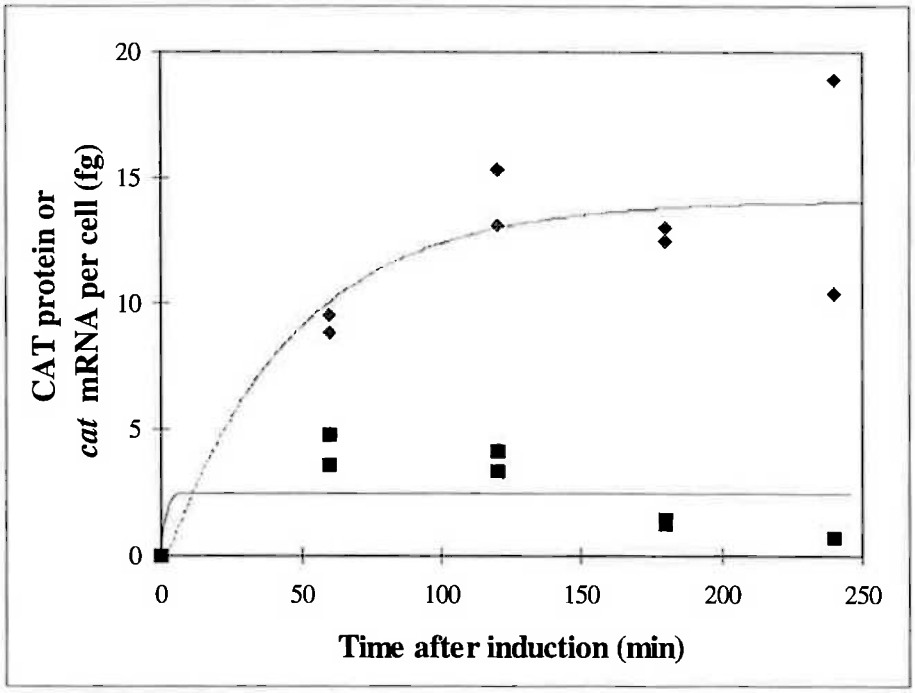

Figure 5.7. CAT protein per cell (squares) and *cat* mRNA per cell (diamonds) against time. CAT protein production was induced at OD600 = 2.4 by adding IPTG to a final concentration of 0.4 mM. η = 0.18, ξ = 0.009.

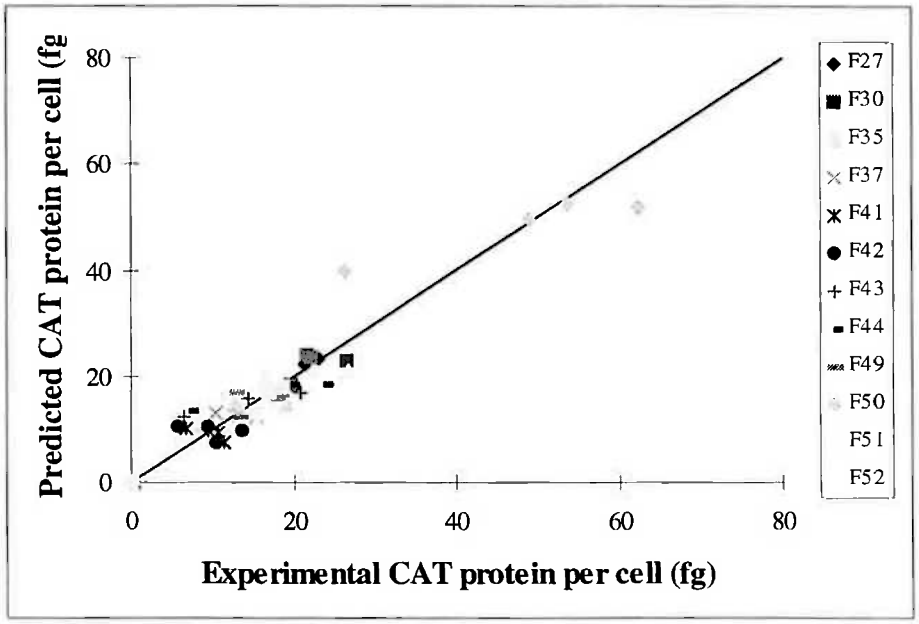

Figure 5.8. Parity plot comparing experimental data from fermentation experiments induced at OD_{600} = 2.4 by adding IPTG to a final concentration of 0.4 mM.

Table 5.2. Predicted η and ξ for *E. coli* batch fermentation. CAT protein production was induced at OD600 = 2.4 by adding IPTG to a final concentration of 0.4 mM.

$\frac{SSE_{CAT}}{SSE_{total}}$	$\frac{SSE_{mRNA}}{SSE_{total}}$	SSE_{total}	η	ξ
0.45	0.55	12.6	0.17	0.010
0.61	0.39	40.1	0.19	0.010
0.68	0.32	39.2	0.20	0.008
0.73	0.27	22.3	0.16	0.009
0.41	0.59	72.8	0.45	0.003
0.38	0.62	121.6	0.45	0.002
0.53	0.47	105.8	0.61	0.003
0.41	0.59	176.3	0.68	0.003
0.11	0.89	219.8	0.63	0.003
0.93	0.07	268.7	0.81	0.003
0.03	0.97	48.6	0.37	0.008
0.54	0.46	123.6	0.49	0.005

make cells better recombinant protein producers. It is possible that some cellular signals down-regulate translation during high transcriptional activities. Transcriptional efficiencies varied from 16-81%, while the translational efficiencies varied from 0.2-1%, (i.e., considerably more DNA is transcribed than mRNA is translated). The high transcriptional activity was expected as CAT protein expression was under control of the strong *tac* promoter. The relatively low translational efficiency suggests that CAT protein expression is limited at the translational level, as previously proposed and discussed.

These results imply that further improvement of transcription is unlikely and may even have a negative effect on expression. However, higher protein levels may be achieved with improved translational activity.

5.4 Conclusion

The aim in this chapter was to optimise recombinant CAT protein production under the control of the IPTG-inducible *tac* promoter. Furthermore an attempt was made to identify the rate limiting step in recombinant protein synthesis during deregulation. Other factors influencing CAT protein production that were considered included the effect of oxygen availability, the growth stage of the inoculum, and the starting OD_{600} of the fermentations.

The optimal inducer concentration was 0.1 mM when inducing cells at relatively low cell density (OD_{600} = 0.8), and 0.4 mM for higher cell density (OD_{600} = 2.4). Maximum CAT protein expression was independent of oxygen tension. However, CAT protein production was highly dependent on the time of induction. Induction close to stationary phase produced lower levels of CAT compared to induction in logarithmic phase. This is due to either increased nuclease activity or a transcriptional regulation. Furthermore, CAT protein expression levels were dependent on the inoculum. Some evidence suggested that cultures inoculated with a stationary-phase culture were superior in terms of CAT protein expression.

Although it was clear that different inoculums produced different expression levels, it was not possible to determine the real cause of this variation. However, the

reproducibility between fermentation experiments inoculated with the same culture was very good.

The mRNA "log jam" effect produced when genes are expressed from multi-copy recombinant plasmids was illustrated by this work. The ribosomal machinery cannot effectively meet the metabolic and energy requirements necessary to effectively process mRNAs for maintenance and recombinant protein needs. Transcriptional activities exceeded translational activities in *E. coli* producing CAT, after induction with IPTG. This was confirmed by mathematical models which described CAT protein biosynthesis reasonably well. The model showed that the highest values of CAT corresponded to the highest values of translational efficiency (ξ). Interestingly, the highest levels of ξ corresponded to the lowest levels of transcriptional efficiency (η). This is an important observation. It suggests that an optimisation of transcription can have a negative effect on translation when over-expressing proteins.

Despite an extensive body of literature focusing on bacterial translation, most workers concentrate on transcription when optimising recombinant-protein expression. This work shows that increased attention should be paid to optimising translation during over-expression of recombinant proteins in *E. coli*.

Chapter 6

Recombinant protein production under control of stationary phase inducible promoters

6.1 Introduction

E. coli is widely used for the over-expression of proteins for commercial or research purposes. High-level expression of recombinant protein can lead to a marked decrease in overall biomass yield (Andersson *et al.*, 1996) and can have a detrimental effect on both the host and recombinant protein. Various regulatory systems have been developed to separate cell growth from gene expression. The most widely used promoters, namely *tac* (de Boer *et al.* 1983), pL (Remaut *et al.*, 1981) and T7 (Studier *et al.*, 1990; Tabor and Richardson 1985) are induced either chemically (for example, by use of isopropyl-β-thiogalactopyranoside (IPTG)) or by a temperature increase. Chemical induction incurs both cost and environmental penalties for the process whilst induction by a temperature increase can lead to the production of heat-shock proteins, including proteases. Recently, other systems have been investigated including an oxygen-responsive promoter (Dikshit *et al.*, 1990 and Khosla *et al.*, 1990), a pH-responsive promoter system (Chou *et al.*, 1995) and a carbon-starvation inducible promoter (Tunner *et al.*, 1992). In these systems, gene expression is controlled by the restriction of oxygen, the alteration of pH, and glucose starvation, respectively.

The ultimate goal in recombinant-gene expression is to ensure maximum transcriptional and translational activities at high cell density. Continued growth diverts nutrient into biomass production rather than product formation. Protein production in a non-growing dense population is therefore ideal. The use of starvation promoters renders this proposition feasible. A problem encountered in a non-growing population is a decrease in overall protein synthesis as a result of stringent control of rRNA (Cashel and Rudd, 1987). A feeding strategy must therefore be established to ensure a sufficient supply of energy and hence adequate levels of rRNA, and to eliminate stress responses such as increased protease production.

E. coli bacteria possess the ability to sense and react to changes in their environment. Some thirty to fifty new proteins are induced in stationary phase following carbon starvation (Groat *et al.*, 1986). These proteins are involved in alteration of cell morphology and increased resistance to environmental stresses such as heat and increased osmolarity. Genes induced by carbon starvation are divided into two groups, based on

whether cAMP is involved in regulation of expression. The *cst* genes require cAMP for their induction whereas *pex* genes are either independent of, or negatively regulated by, cAMP (Schultz *et al.*, 1988). Consequently, expression from the *cst* gene promoters in the presence of glucose will be repressed. By contrast, at high cell density with correspondingly low nutrient levels, cAMP and hence *cst* gene promoter activity increase. The use of *cst* gene promoters for protein production has been examined previously (Matin 1991, Tunner *et al.*, 1992). However, cloning the *cstA* promoter in front of human growth hormone only produced a four-fold protein increase following transition from growth phase into the stationary phase (Matin, 1991). β-galactosidase synthesis under control of the *cst-1* promoter has been examined in both batch and fed-batch fermentations (Tunner *et al.*, 1992). Induction of the promoter followed glucose exhaustion in the batch fermentation mode. Addition of acetate in the fed-batch mode resulted in a further three-fold increase in β-galactosidase after induction. This suggested that acetate provided an energy source for protein expression. The above result is not surprising as acetate metabolism is known to promote high intracellular cAMP levels (Buettner *et al.*, 1973) which in turn stimulate promoter activity.

Although the *pex* genes are induced by carbon starvation, they are either independent of, or are negatively regulated by, cAMP (Schultz *et al.*, 1988). There is strong evidence to suggest that some *pex* genes are induced by starvation for other nutrients as well as a variety of stresses (McCann *et al.*, 1991) and this is clearly shown by studies of *katE* and *katF*, which are representative of the *pex* genes. *katF* encodes a sigma factor which is a global regulator of other stationary-phase induced genes (McCann *et al.*, 1991; Lange and Hengge-Aronis, 1991). The *katE* gene encodes catalase HPII (Loewen *et al.*, 1985). *katE* expression is regulated primarily at the level of transcription (Schellhorn and Stones, 1992) by KatF (Lange and Hengge-Aronis, 1991). In the presence of benzoate, *katE* is expressed at only a third of the normal level in stationary-phase cultures when *katF* is fully induced (Mulvey *et al.*, 1990). Hence, other factors may also be involved in the regulation. There is conflicting evidence as to the nature of these factors. For example, acetate has been reported to stimulate expression from the *katE* and *katF* promoters (Schellhorn and Stones, 1992). However, this finding contradicts that of Mulvey *et al.* (1990). *katF* expression is controlled both transcriptionally, translationally, and by other factors such as KatF protein stability (Lange and Hengge-Aronis, 1995). Furthermore,

expression from both the *katE* and *katF* promoters is dependent on media composition (Mulvey *et al.*, 1990). When *E. coli* was cultivated in LB media containing glucose, levels of the reporter protein β-galactosidase were low in the logarithmic phase but increased during the stationary phase (Mulvey *et al.*, 1990). Cultivation in LB or in glucose-based minimal media resulted in high levels of the reporter protein in logarithmic phase (Mulvey *et al.*, 1990). This suggested that both glucose and amino acids are required to keep expression low during logarithmic growth phase. In contrast, Schellhorn and Stones (1992) reported low levels of β-galactosidase during logarithmic phase with increasing levels during the stationary phase, for cells cultivated in LB. Weak acids such as o-hydroxybenzoate and p-aminobenzoate have a stimulatory effect on both the *katE* and *katF* expression (Mulvey *et al.*, 1990). Similarly, both benzoate and propionate are reported to be effective inducers of both genes (Schellhorn and Stones, 1992).

In this chapter, the use of stationary-phase inducible promoters (*katE* and *katF)* for expression of the recombinant protein CAT is examined. It is clear that contradictory findings on the regulation of these promoters have been reported in the literature. Consequently, a detailed analysis of fermentation studies using these promoters to drive the expression of CAT is presented. Several batch-mode fermentations were conducted. Intracellular protein levels on a per-cell basis are related to mRNA and 16S ribosomal RNA concentrations. The effect of glucose, acetate, and o-hydroxybenzoate as putative inducers of transcription from the *katE* and *katF* gene promoters is examined. Expression of CAT in fed-batch mode is also characterised. CAT protein expression under control of the *katE* promoter is further characterised by the use of a simple mathematical model.

Experimental work

An overnight culture of *E. coli* JM101 (ATCC no: 33876) transformed with plasmid pCT100 or pCT101 grown in glucose minimal media (2.2.2.1) was subcultured into fresh media and further cultivated at 37°C. Fermenters (Applikon, Schiedam, The Netherlands) containing 1.2 L or 1.5 L of glucose minimal media were inoculated to provide a starting OD_{600} of 5×10^{-4}. Fermentation conditions are described in section 2.26. CAT assay, D-Glucose and acetate determination, RNA extractions, slot-blotting,

hybridisation, RNA detection by chemiluminescence and *cat* mRNA quantitation were undertaken as described in sections 2.23, 2.27, 2.28, 2.13.2.1, 2.13.3, 2.13.4, 2.13.5 and 4.2.6. *cat* mRNA standards and the DIG-labelled complimentary RNA probe were constructed as described in sections 2.14 and 2.15. 16S rRNA standards and the DIG-labelled complimentary RNA probe were constructed as described in sections 2.16 and 2.17. Cell count and viable cell counts were performed as described in section 2.19.

The glucose and acetate results represent an average of two sample repeats, while protein and most of the RNA results represent an average of four repeats. Standard errors were not included on the presented graphs to improve clarity.

6.2 Results and Discussion

The expression levels of CAT resulting from the two promoters *katE* and *katF* in *E. coli* cultivated in glucose minimal media are compared in Figure 6.1. Comparable levels of CAT protein per cell were produced as cells entered stationary phase, irrespective of the type of promoter used to drive the expression. The levels remained constant in late stationary phase for both systems (Figure 6.1). *cat* mRNA levels were low for the *katE* system in logarithmic phase and increased in stationary phase, indicating higher transcriptional activity (Figure 6.2). *cat* mRNA levels for the *katF* system tended to be higher in logarithmic phase than in stationary phase. Translational down regulation in growth phase together with constant protein levels during late stationary phase suggested that the *katE* system was worthy of further investigation. The concentration of ribosomes in the cells was inferred indirectly by measuring cellular 16S rRNA (Figure 6.2). The results suggest that ribosomal levels decreased following the transition from logarithmic growth phase to stationary phase, probably due to the stringent control as described by Cashel and Rudd (1987). This raises the issue of translational limitation due to a limited number of ribosomes.

Table 6.1. Cultivation media and addition of "inducers" to batch-mode fermentations.

Set	Plasmid(s)	Genotype	Media	Additions
1	pCT100 pCT101	*katE*::*cat* *katF*::*cat*	Glucose minimal media	None
2	pCT100	*katE*::*cat*	Glucose minimal media	1. Glucose 4.2 g/L 2. None
3	pCT100	*katE*::*cat*	Glucose minimal media (1,2,3) or Luria Broth (4) (Miller 1972)	1. Acetate (20 mM) 2. o-hydroxybenzoate (10 mM) 3. None 4. None

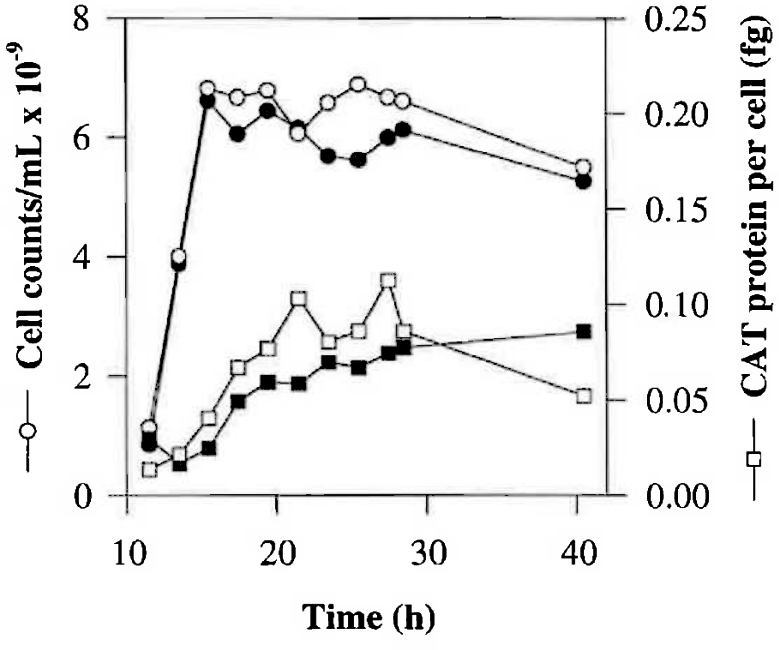

Figure 6.1. Cell count (●) and expression levels for CAT protein per cell (■) from *E. coli* in batch fermentation in glucose minimal media under control of the *katE* promoter (filled symbols) and the *katF* promoter (empty symbols).

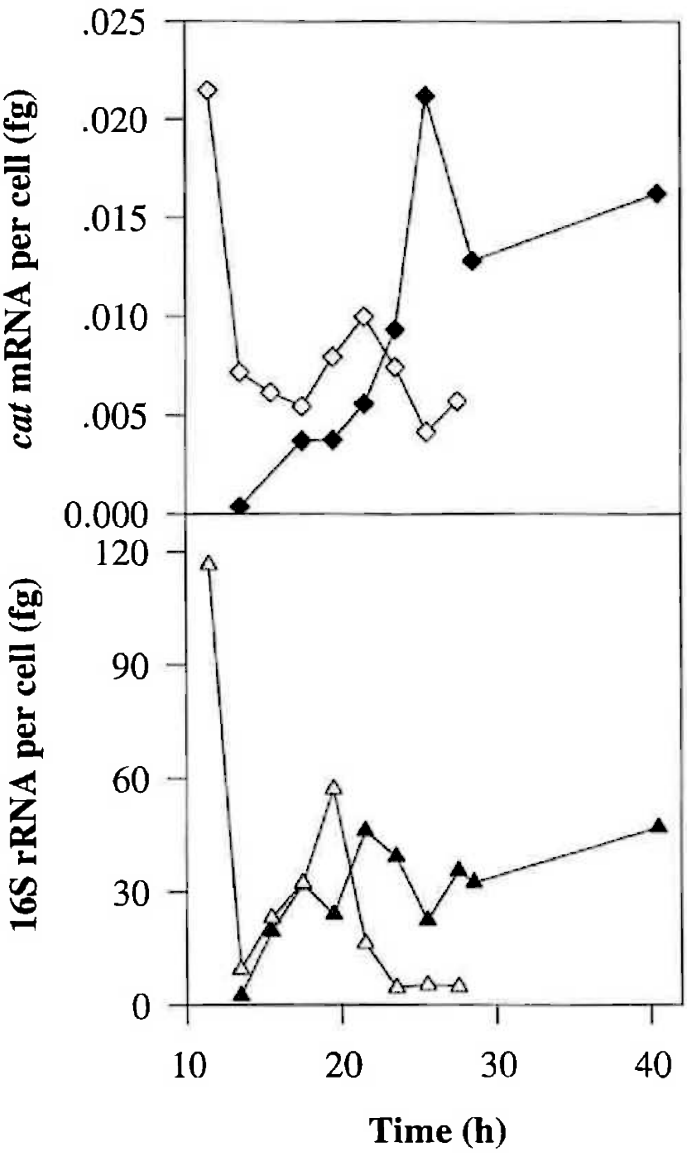

Figure 6.2. *cat* mRNA (◆) and 16S rRNA (▲) per cell from *E. coli* in batch fermentation in glucose minimal media under control of the *katE* promoter (filled symbols) and the *katF* promoter (empty symbols).

Glucose and amino acids act in concert to repress expression from the *katE* gene promoter during the logarithmic growth phase (Mulvey *et al.*, 1990). This assertion was tested by conducting simultaneous batch fermentations in glucose minimal media and Luria broth (LB). The results are summarised in Figure 6.3. Higher CAT protein levels were observed during the growth phase for cultivation in LB, as reported by Mulvey *et al.* (1990). *cat* mRNA displayed a similar trend, suggesting high transcriptional activity in logarithmic phase in LB. In contrast, observed *cat* mRNA levels were low during logarithmic phase in minimal media. Hence, the addition of amino acids was unnecessary to reduce promoter activity in this media. Levels of 16S rRNA, and hence ribosomes, were lower in cell cultivated in LB compared to minimal media. This is a surprising result as the number of cellular ribosomes per unit cell mass has been reported to increase with growth rate (Jinks-Robertson and Nomura, 1987).

Interestingly, at t = 24 h both *cat* mRNA and 16S rRNA levels increased while the CAT protein levels remained constant in glucose minimal media (Figure 6.3). Although both message and ribosomes are available (assuming that 16S rRNA is a reliable measure of number of ribosomes) the message is not translated. These result suggest that lack of sufficient energy is a limiting factor in translation. It is therefore necessary to supply energy.

Having established the possibility of separating the growth and production phases using the *katE* promoter system in glucose minimal media, the next objective was to determine whether glucose, acetate or o-hydroxybenzoate acted as inducers. An inducer is employed in this instance to improve product yield and also to extend the production period. These inducers may therefore also act as energy sources. Obviously, starved cells cannot sustain the production of recombinant proteins. A feeding strategy must be devised to produce a metabolic state where non-growing cells manufacture the product. Initially glucose was examined as a possible indirect inducer. Cellular levels of cAMP depend on cellular levels of glucose. Furthermore, since *katE* belongs to the *pex* class of genes, it is most probably either negatively regulated by or expressed independently of cAMP. To determine whether the presence of glucose increased levels of CAT expressed by an *E. coli* strain carrying a *katE::cat* transcriptional fusion, JM101[pCT100] was grown in batch fermentation using glucose minimal media. Glucose was added to a final concentration of

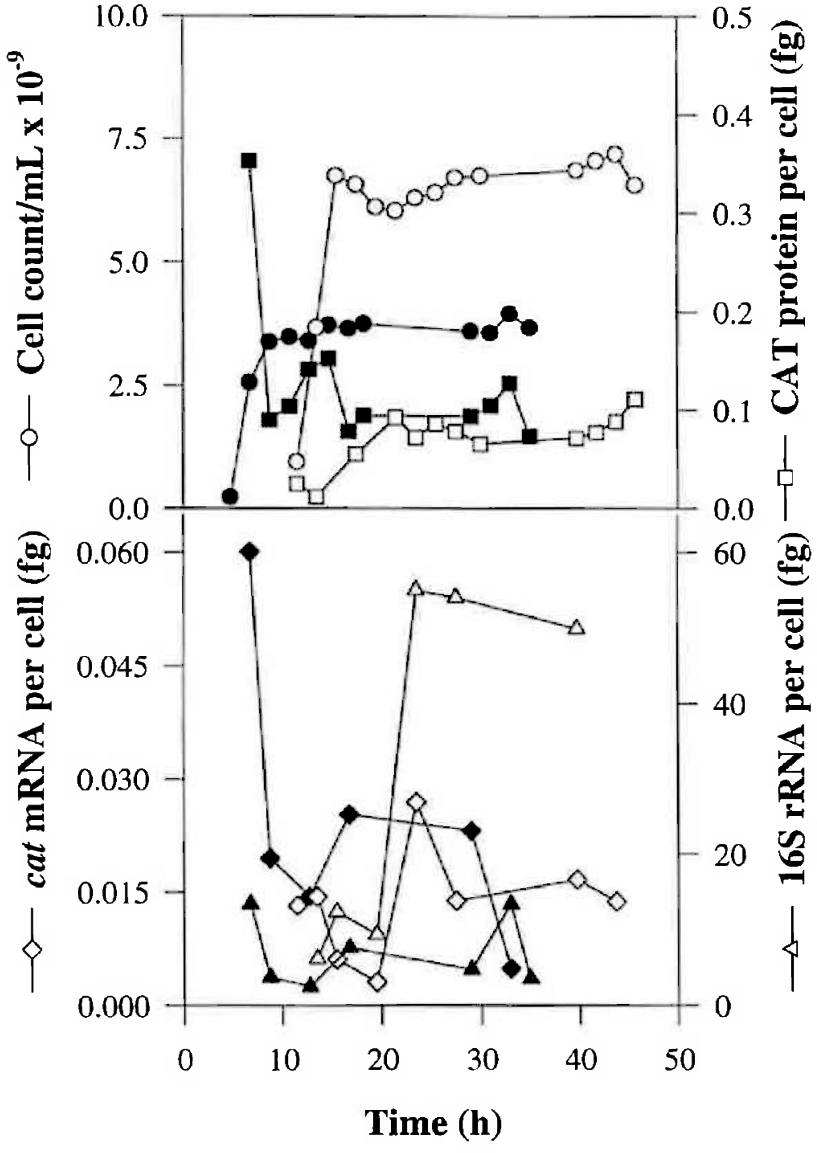

Figure 6.3. Cell counts (●) and expression levels for CAT protein (■), *cat* mRNA (◆) and 16S rRNA (▲) per cell under control of the *katE* promoter in *E. coli* batch fermentation in Luria Broth (filled symbols) or glucose minimal media. (empty symbols).

4.2 g/L, four hours after exhaustion of the initial glucose charge (at t = 17.5 h) (Figure 6.4). A negative control (with no glucose addition) was conducted in parallel. Although OD_{600} values in the culture increased upon glucose addition (results not shown), cell counts were approximately the same with and without additional glucose. Cells therefore most likely utilised the added glucose to increase their cell and probably ribosomal mass. CAT protein levels per cell increased in both fermenters as the culture approached stationary phase. Glucose addition resulted in higher CAT protein levels per cell compared to the control (no additional glucose). As well, additional glucose stimulated *cat* mRNA accumulation (Figure 6.4) suggesting a negative regulatory role for cAMP on the *katE* promoter. However, increasing levels of 16S rRNA per cell and the possible increase in ribosomal mass after glucose addition suggested that the improvements in CAT protein levels occuring in late stationary phase were due to increased translational activity. This was confirmed by a decrease in 16S rRNA levels in late stationary phase as CAT protein concentration stabilised.

The possible regulatory effect of cAMP on the *katE* promoter was further investigated. cAMP is known to increase during acetate metabolism (Buettner, 1973). Furthermore, conflicting reports exist as to the role of acetate as an inducer of the *katE* promoter. To study this effect, total CAT protein levels expressed in an *E. coli* batch fermentation in glucose minimal media with or without added acetate were compared. The results are presented as Figure 6.5. Acetate was added in the growth phase at time = 13.5 h to a final concentration of 20 mM. Figure 6.5 demonstrates complete consumption of acetate in stationary phase. Acetate levels in the control accumulated to approximately 10 mM during the growth phase. This excess was consumed during stationary phase. Optical density (OD_{600}) of the two cultures were similar (results not shown) whereas cell counts increased significantly when acetate was added. This observation, together with the fact that total CAT protein per cell remained constant even after acetate addition, implies that the acetate was probably utilised for cell division and maintenance. These results suggested that elevated cAMP levels brought about by acetate metabolism were unable to regulate expression from the *katE* promoter. Furthermore, this observation confirmed that increased CAT protein levels after addition of glucose in the previous experiment was due to increased translational activity. Increases in 16S rRNA levels during early stationary

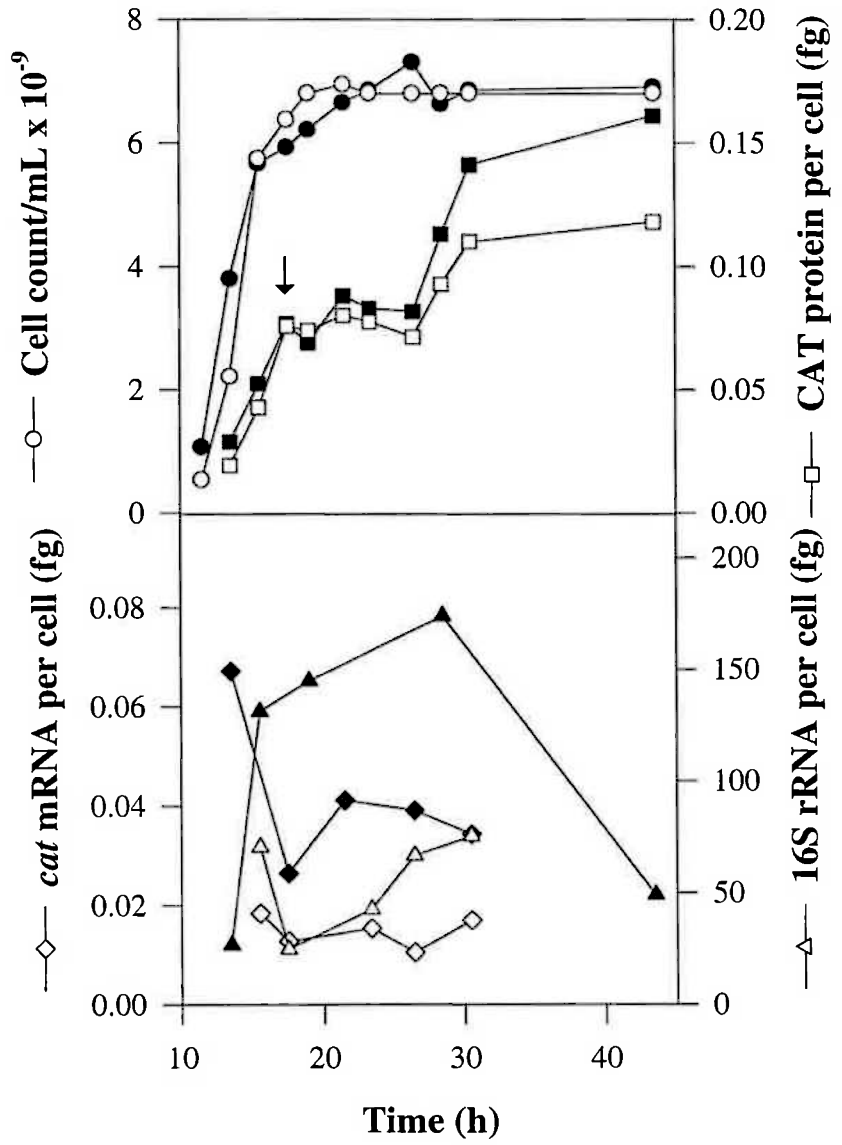

Figure 6.4. Cell counts (●) and expression levels for CAT protein (■), *cat* mRNA (◆) and 16S rRNA (▲) per cell under control of the *katE* promoter in *E. coli* batch fermentation in glucose minimal media. Additional glucose was added two hours after exhaustion to a final concentration of 4.2 g/L (filled symbols) or 0 g/L (negative control, empty symbols) at time = 17.5 h (indicated by an arrow).

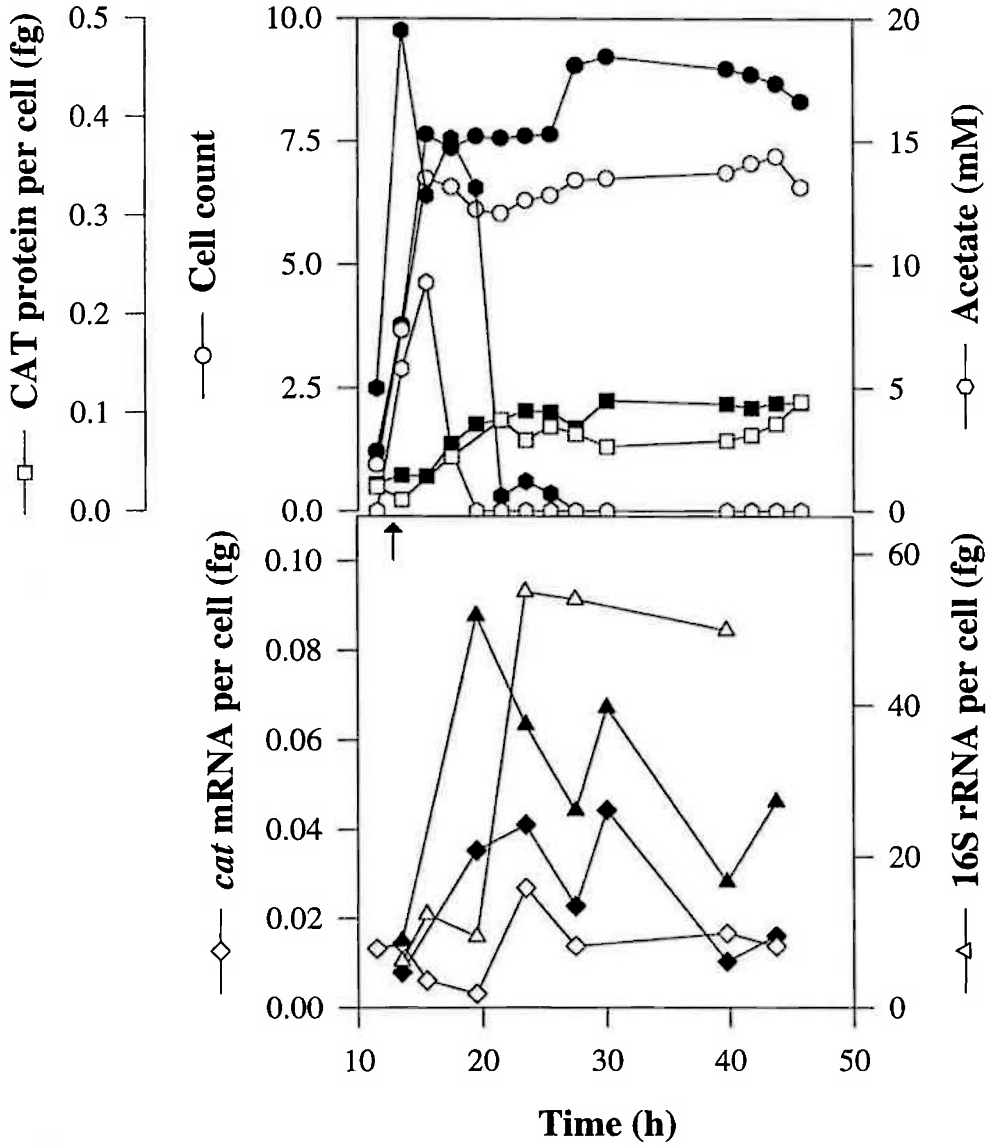

Figure 6.5. Cell counts (●), acetate (●) and expression levels for CAT protein (■), *cat* mRNA (◆) and 16S rRNA (▲) per cell under control of the *katE* promoter in *E. coli* batch fermentation in glucose minimal media. Media was supplemented with 20 mM sodium acetate (filled symbols) or 0 mM sodium acetate (empty symbols) at time 13.5 h (indicated with an arrow). Acetate levels reached zero for both fermentations experiments after t = 28 h.

phase, followed by decreased levels in late stationary phase, may result from increased nuclease activity.

Another reported observation is the effect of weak acids such as benzoate, o-hydroxybenzoate, p-aminobenzoate and propionate. It is suggested that these acids stimulate *katE* promoter activity in rich media (Mulvey *et al.*, 1990; Schellhorn and Stones, 1992). O-hydroxybenzoate was chosen as a possible inducer of the *katE* promoter. This acid was added to an *E. coli* [pCT100] culture grown in glucose minimal media at OD_{600} = 8 (t = 16 h) to achieve a final concentration of 10 mM. CAT expression levels were then compared with those from a batch culture with no acid addition (Figure 6.6). Comparable increases in cell number and CAT protein per cell with time were observed in both systems. Acetate was exhausted in early stationary phase for control cultures, whereas it accumulated in the presence of o-hydroxybenzoate. No difference in *cat* mRNA was detected. However, 16S rRNA levels were much higher for control fermentations. This suggested inhibition by o-hydroxybenzoate of the cell's translational apparatus. Clearly, this acid fails to stimulate promoter activity in glucose minimal media and may actively interfere with translation.

Finally a fed-batch fermentation was conducted to characterise the behaviour of the *katE* promoter under glucose-limiting conditions (Figure 6.7). In this experiment, *E. coli* cells [pCT100] were cultivated to an OD_{600} of 10. The initial charge of glucose was exhausted at this point. Glucose feeding then commenced (described in section 2.26) and was maintained for 5.9 h. Glucose levels fell below 0.06 g/L during the period of feeding (results not shown). Cell numbers, determined by viable cell counts on ampicillin plates and compared to plates without selection, confirmed insignificant plasmid loss during all stages of the fed-batch fermentation. Acetate was rapidly consumed following glucose exhaustion at OD_{600} = 10. Acetate levels were observed to increase during the early stages of feeding. Accumulated acetate was consumed as the cells stopped multiplication. The CAT profile is interesting. Total CAT protein levels rose slowly until t = 4 h. At this point, the levels increased sharply, coinciding with the drop in acetate concentration. Between 5 and 8 hours, total CAT protein increased slowly as the acetate accumulated. At t = 8 h acetate was exhausted and CAT protein levels again rose. The maximum CAT

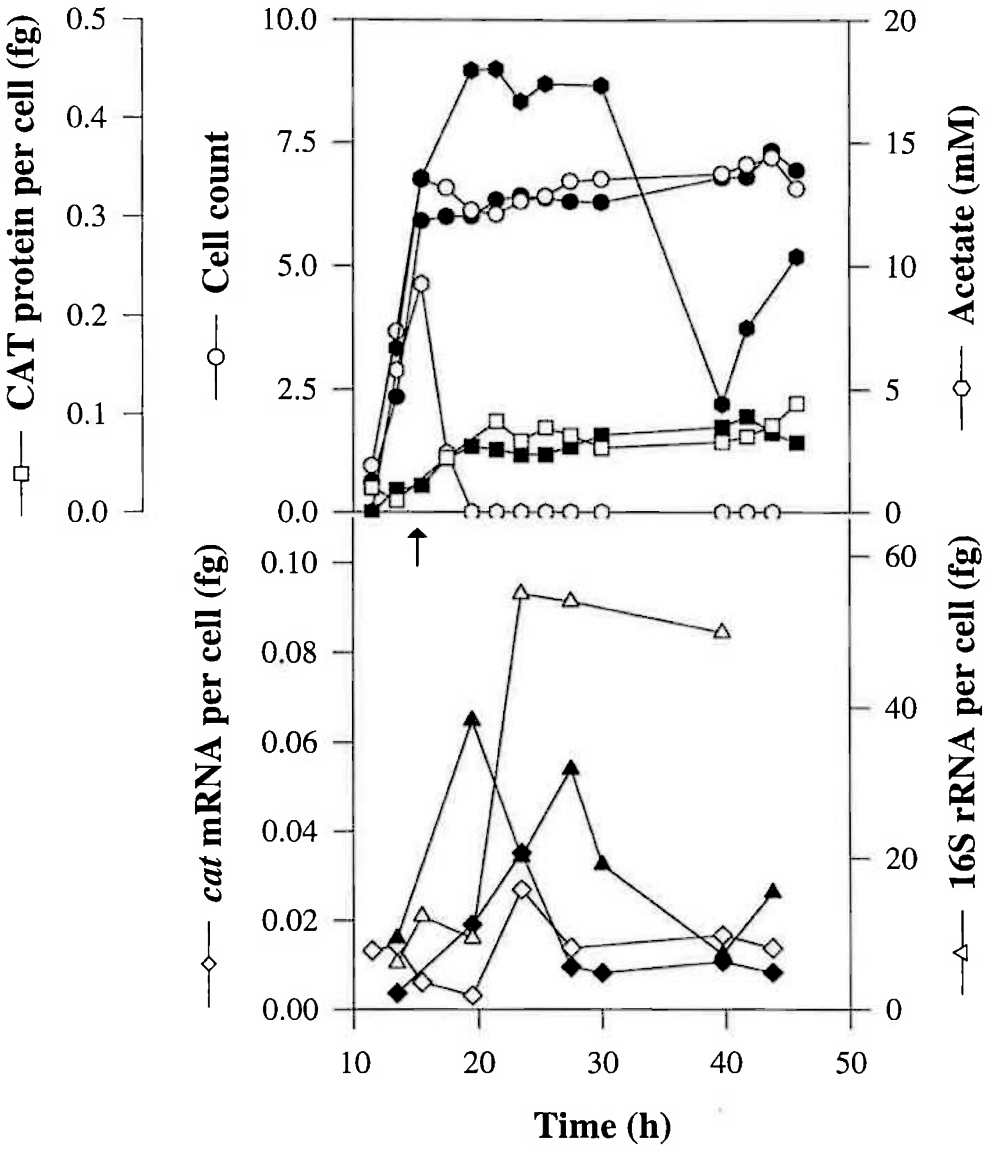

Figure 6.6. Cell counts (●), acetate (●) and expression levels for CAT protein (■), *cat* mRNA (◆) and 16S rRNA (▲) per cell under control of the *katE* promoter in *E. coli* batch fermentation in glucose minimal media. Media was supplemented with 10 mM o-hydroxybenzoate (filled symbols) or 0 mM o-hydroxybenzoate (empty symbols) at OD_{600} = 8 (t=16 h) (indicated with an arrow).

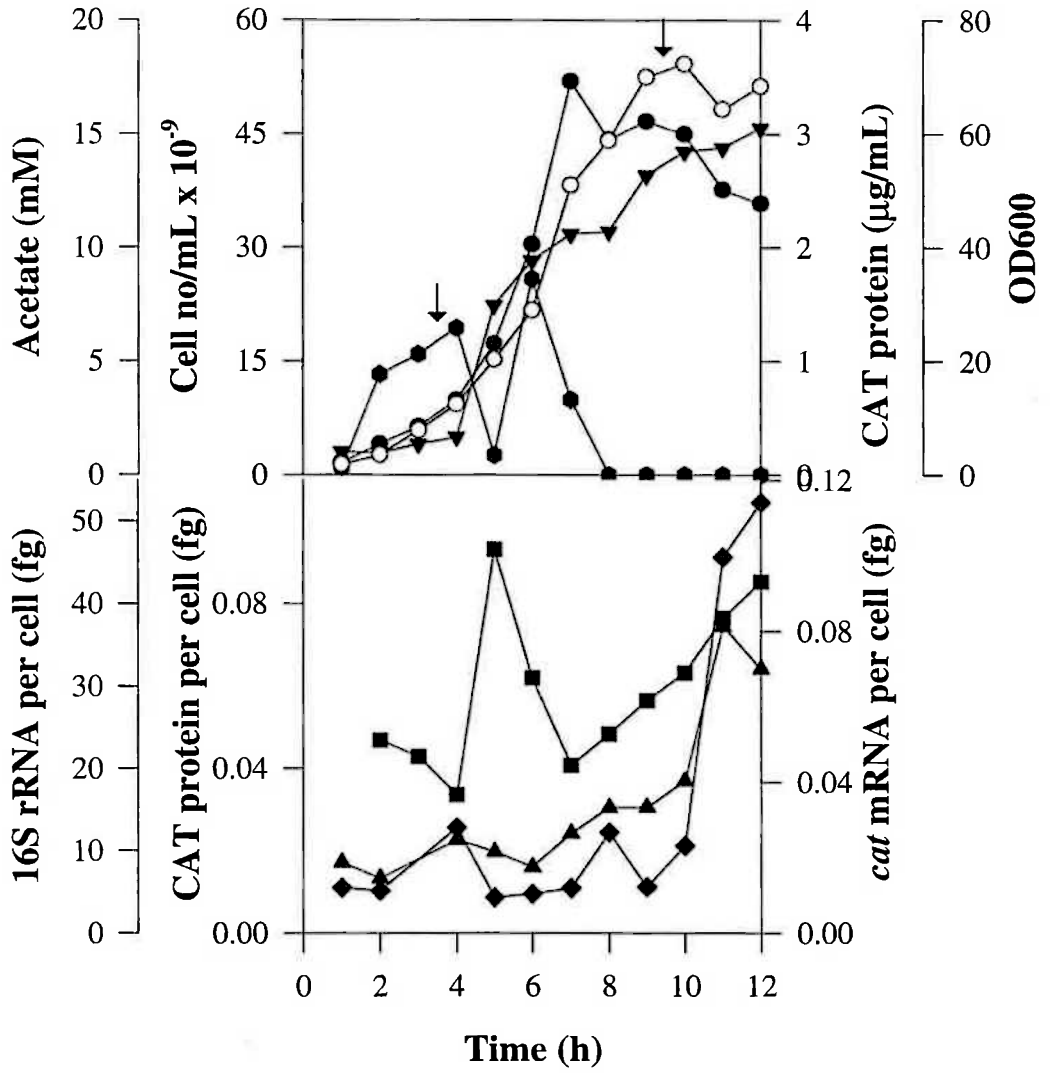

Figure 6.7. Cell density (○), cell counts (●), total CAT protein (▼), acetate (⬢) and expression levels for CAT protein (■), *cat* mRNA (◆) and 16S rRNA (▲) per cell under control of the *katE* promoter in *E. coli* fed-batch fermentation in glucose minimal media. Feeding was commenced upon exhaustion of initial charge of glucose at time = 3.6 h (OD_{600} = 10) and ceased at time = 9.5 h (indicated with arrows).

protein levels per cell were achieved at t = 5 h when the concentration of acetate was almost zero. The levels then fell (due to a dilution effect as cells were dividing) but increased 1.5 h before feeding was ceased (again corresponding to zero acetate). These results appear to suggest that CAT protein expression is sensitive to the presence of acetate. However, *cat* mRNA levels per cell closely followed acetate levels up to 5 h, achieving a maximum at 4 h and a minimum at 5 h. The exact opposite occurred for t = 8 h and acetate fell to a minimum whereas *cat* mRNA reached a second maximum. Because of this divergence in behaviour, neither acetate nor cAMP (synthesised by acetate metabolism) appear likely regulators of the *katE* promoter. At t = 7 h, cell division ceased as the energy required exceeded that available from feeding. Acetate was consumed and the cells entered stationary phase. 16S rRNA levels per cell increased during this stationary phase, confirming that the feeding of glucose was sufficient to alleviate the stringent control. Interestingly, *cat* mRNA levels increased dramatically at t = 9.5 h, when feeding was stopped. This indicated that the rate of feeding was to fast for maximal transcriptional activities. Nevertheless, CAT protein production continued in the non-dividing cells. These results are very encouraging and demonstrate the feasibility of establishing a growth-limited culture with selective protein production. Careful definition of a critical growth rate, below which the stringent control is repressed, may permit the *katE* promoter to be used for high-density fermentation. Determination of the critical growth rate permits optimisation of a feeding strategy to ensure a sufficient supply of energy. rRNA synthesis, tRNA synthesis, and the synthesis of several other proteins which function in the transcriptional or translational machinery are subject to stringent control (Jensen and Pedersen, 1990). Starved ribosomes will stall at codons where a substrate shortage exists and commence synthesis of large quantities of ppGpp. In turn, this will reduce the speed of both stable RNA and mRNA chain elongation. Consequently, the rate of RNA chain termination and therefore the concentration of free RNA polymerase will fall rapidly (Jensen and Pedersen, 1990).

The fact that culture cell density continued to increase without cell division in the final stage of feeding suggested that cell size varied. Cell volume increases in *E. coli* after shifting to richer media (Kubitschek, 1990). Cell volume change within the same media was here further investigated. A plot of cell numbers versus cell dry weight (CDW) is

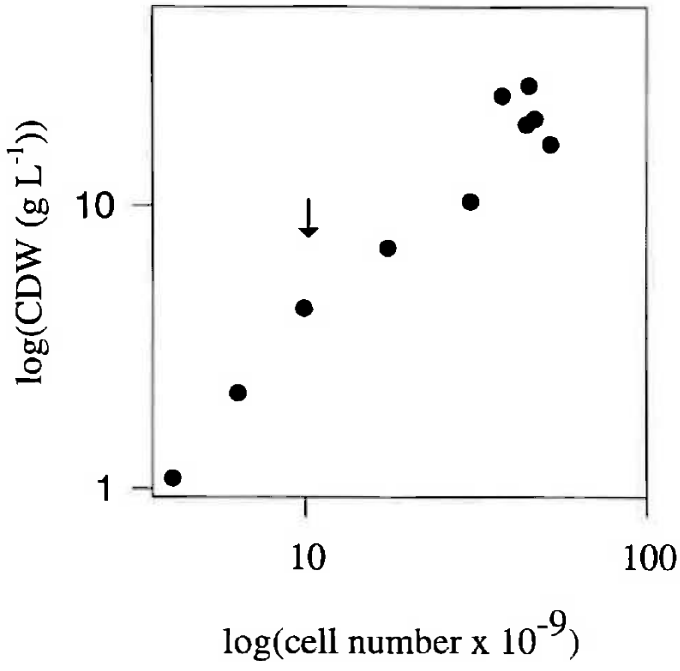

Figure 6.8. Correlation of CDW and cell number for *E. coli* Fed-batch fermentation in glucose minimal media. Feeding was commenced upon exhaustion of the initial charge of glucose. Cell number at this point is indicated with an arrow.

shown in Figure 6.8. A correlation was identified. However, this correlation altered as feeding was initiated (indicated by the arrow); consequently CDW per cell was less compared with the batch phase (before feeding). Effectively, cells "fatten" as fast as they can (specific growth rate of cells = 0.69 h^{-1}), when the bulk glucose concentration exceeds maintenance requirements. As the carbohydrate source becomes limiting, the cells are "slimming down" (specific growth rate of cells = 0.44 h^{-1}) while dividing at the same rate.

The fact that CDW increased whilst cell numbers remained constant, suggests an increased number of ribosomes per cell. This argument is supported by ribosomes constituting a large fraction of the cell mass (eg. 45% for growth rate = 2.7 h^{-1} (Jinks-Robertson and Nomura, 1987)). Translational activities were therefore expected to be better, as confirmed by increased levels of CAT protein (Figure 6.7). However, *cat* mRNA levels remained relatively low in the absence of cell division. When feeding was stopped the message levels increased dramatically (ten-fold). To summarise this observation; the feeding rate was sufficient to overcome the stringent response but was too fast for optimal stimulation of transcription. This demonstrates the importance of an optimised feeding strategy.

The reported protein levels under control of the *katE* promoter are lower than previously documented. Schellhorn and Stones (1992) found a fifty-seven-fold increase in the reporter gene β-galactosidase under control of the *katE* promoter as cells made the transition from growth phase to stationary phase, and a twenty-seven-fold increase under control of the *katF* promoter in rich media. In this study, an approximate five-fold increase was achieved. However, it is difficult to compare the two sets of data for four reasons. Firstly, the levels of total CAT protein are presented here on a per cell basis, whereas Schellhorn and Stones (1992) express results as β-galactosidase enzyme activity (Miller units). Secondly, experiments done by Schellhorn and Stones (1992) were conducted in LB media while these tests were done using glucose minimal media. LB may contain nutrients that directly stimulate/inhibit promoter activity. Thirdly, Schellhorn and Stones (1992) reported higher promoter activity during stationary phase than logarithmic phase in LB media whereas the opposite was found in this work using this media. Hence, promoter regulation is clearly not the same. Finally, a comparison of x-fold increases is a

comparison of the initial and final activities, and therefore dependent on how well the system is repressed before induction.

cat mRNA per cell in the batch experiments increased sixty-fold when cells made the transition from logarithmic growth phase to stationary phase. This, in relation to five-fold increase in CAT protein mentioned above, suggests a translational limitation. Batch results from *E. coli* fermentation in glucose minimal media showed that CAT protein levels per cell increased together with 16S rRNA levels in stationary phase. A plot of CAT protein against 16S rRNA (Figure 6.9) (R^2 = 0.9) confirms that there is a reasonable correlation between the two (R^2 = 0.7). Translational limitation due to the availability of free ribosomes is therefore likely in glucose minimal media. Other factors such as secondary mRNA structure, and the availability of initiation factors, charged tRNAs, and translation elongation factors may further contribute to this limitation. A good negative correlation between CAT protein and 16S rRNA was also found after addition of 10 mM o-hydroxybenzoate (Figure 6.9), confirming that this acid inhibits the translational machinery.

A possible factor explaining the observed *cat* mRNA increase is a change in the regulation of plasmid replication causing an increase in gene dosage. Results showed that plasmid copy number varied only slightly when comparing the logarithmic growth and stationary phases (results not shown). Thus, increased plasmid copy number could not account for the observed increase in *cat* mRNA.

Another possible explanation for increased mRNA levels without a proportional increase in CAT levels is rapid degradation of the protein either intracellularly or during the assay. Elevated expression of foreign proteins in *E. coli* can result in increased protease activation, and therefore product degradation due to cellular stress responses (Ramirez and Bentley, 1995). Over-expression of proteins such as CAT, which is considered endogenous to some strains of *E. coli*, can lead to enhanced protease production identical to that induced by the stringent response (Harcum and Bentley, 1993). In the present work, this will only be a concern if CAT protein epitopes recognised by the polyclonal antibody are degraded. This is because it is the protein expression that is of concern, and not the actual protein level.

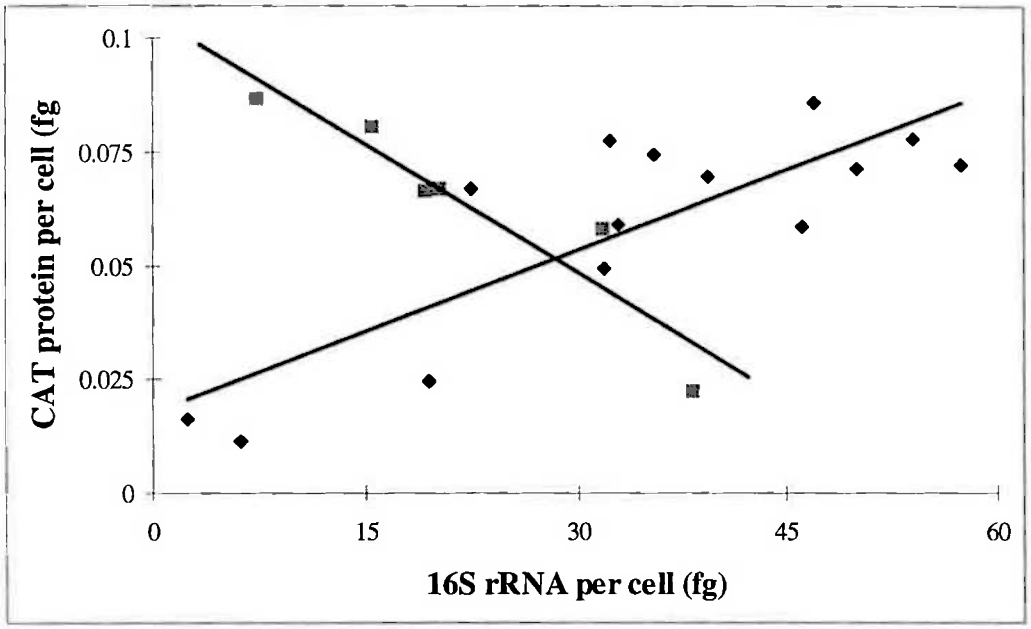

Figure 6.9. CAT protein per cell against 16S rRNA levels for *E. coli* JM101 (pCT100). Cells were cultivated in glucose minimal media. Media was supplemented with 0 mM (diamonds, $R^2 = 0.7$) or 10 mM o-hydroxybenzoate (squares, $R^2 = 0.9$).

A third possible explanation for increased *cat* mRNA without increased CAT levels is failure to detect the protein. GroEL expression is increased in carbon-starved cells (Groat *et al.*, 1986). It is considered to be a general starvation protein. Furthermore, GroEL interacts with CAT proteins (Bochkareva *et al.*, 1988) and may be involved in CAT assembly. This binding may mask epitopic sites and thus interfere with the CAT-ELISA assay.

6.3 Model description

Mechanistic models to predict stationary phase gene expression in *E. coli* during periods of nutrient limitation do not currently exist. The development of such models is complicated by the difficulty in maintaining a population in a steady state while expressing recombinant protein. Chemostat studies are hampered by growth on vessel walls at low dilution rates (Heijnen *et al.*, 1992). Furthermore, nutrients may not be uniformly distributed at low dilution rates eg. Stouthamer *et al.* (1990). These problems may be overcome by changing the media composition. For example, the amino acids valine, alanine and leucine greatly decrease growth rate, possibly as a result of feed-back inhibition of other biosynthetic pathways (Yee and Blanch, 1993).

It was earlier demonstrated that equations 5.1 and 5.2 provide a reasonable estimate of CAT protein levels after induction using IPTG (section 5.3.1). This model was investigated for its ability to describe protein expression in stationary phase. The model was modified to omit the effect of dilution by growth. Stationary phase is characterised by cell growth being equal to cell death, resulting in an overall dilution of zero. The material balance on the cloned gene message was thereby described by the following equations.:

$$\frac{d(mRNA)_p}{dt} = k_{p_o}\eta(G)_p - k_d(mRNA)_p \qquad (6.1)$$

$$\frac{d(P)}{dt} = k_{qo}\xi_p(mRNA)_p - k_e(P) \qquad (6.2)$$

6.3.1 Estimation of model parameters

The characterisation and estimation of model parameters is explained in section 5.3.

6.3.2 Model regressions

Fermentation data of CAT expression in stationary phase were modelled using equations 6.1 and 6.2. Values of the two parameters η and ξ (transcriptional and translational efficiencies) were obtained by minimising an objective function as described in Chapter 5 (equation 5.3). A spreadsheet (Microsoft Excel, version 7) was used to perform regressions of the model to experimental data.

A comparison of experimental and regressed results from *E. coli* batch fermentations conducted in glucose minimal media is presented in Figure 6.10. The fermentations were inoculated using the same culture. Initial conditions (t = 0) correspond to the commencement of stationary phase. Figure 6.10 confirms that the experimental results are highly reproducible.

Transcriptional (η) and translational (ξ) efficiencies were estimated as outlined by minimisation of SSE. Estimates are summarised Table 6.2. Translational efficiencies varied from 0.7 - 0.8 %. Transcriptional efficiencies were much lower and varied from 0.04 - 0.06 %, suggesting a severe transcriptional limitation of CAT expression. However, this does not correspond to the observation that *cat* mRNA increased 60-fold by transition from logarithmic growth phase to stationary phase, while CAT protein only increased 5-fold, which suggests a translational limitation.

Table 6.2. Predicted η and ξ for *E. coli* batch fermentation induced upon entry into stationary phase.

SSE CAT	SSE mRNA	SSE total	η	ξ
0.001066	0.000305	0.001371	0.0006	0.007
0.00503	0.00068	0.00571	0.0004	0.008

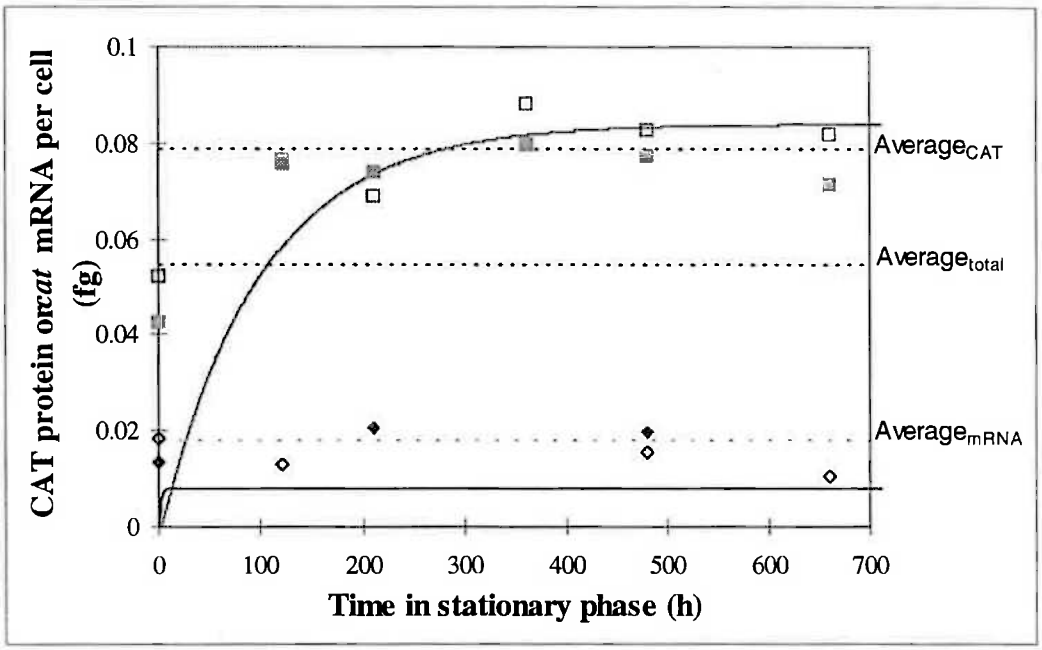

Figure 6.10. CAT protein per cell (squares) and *cat* mRNA per cell (diamonds) against time for *E. coli* batch fermentations in glucose minimal media. CAT protein production was induced upon entry into stationary phase. $\eta = 0.0006$, $\xi = 0.008$.

In light of this contradiction, the accuracy of the differential equation model describing the stationary phase was examined. The predictions from the differential model were compared to two simplistic models. In the first, the value of either CAT or mRNA was assumed equal to the mean of all readings ($\bar{y}$). The second model was perhaps more realistic and assumed that both CAT and mRNA remained constant at a value equal to the individual means ($\bar{y}_{cat}, \bar{y}_{mRNA}$). Sum of squares errors for a combined average (SSE_{av}) was calculated using equations 6.3-6.4. Sum of squares errors for individual averages ($SSE_{separate\ means}$) was calculated using equation 6.5.

$$\bar{y} = \frac{\Sigma y_{cat} + \Sigma y_{mRNA}}{n_{cat} + n_{mRNA} - 1} \tag{6.3}$$

$$SSE_{av.} = \Sigma\ (y_{cat} - \bar{y})^2 + \Sigma\ (y_{mRNA} - \bar{y})^2 \tag{6.4}$$

$$SSE_{separate\ means} = \Sigma (y_{cat} - \bar{y}_{cat})^2 + \Sigma\ (y_{mRNA} - \bar{y}_{mRNA})^2 \tag{6.5}$$

$$\frac{SSE_{total}}{SSE_{av.}} = 0.4077 \tag{6.6}$$

$$\frac{SSE_{total}}{SSE_{dual}} = 2.1511 \tag{6.7}$$

Comparison with the mean of all readings shows that the predicted error from the differential equations was 0.41 times that of the combined average. This indicates that the differential model is superior to a single-average model which is not surprising. However, comparison with the more realistic model of constant individual values shows that the predicted error from the differential equations was 2.15 times that for the simple model. Based on this result, it was concluded that the differential equation model poorly represents the stationary phase system. A contributing factor to this is the limited data at the time of transcriptional up-regulation, meaning that the system dynamics are not very well defined.

Consequently, the values of the transcriptional and translational efficiencies calculated from the model will have considerable error. The experimentally-suggested results of translational limitation is therefore reasonable.

As discussed earlier in this chapter, both transcriptional and translational activities are repressed by stringent control. A careful definition of the critical growth rate, below which the stringent control is induced, permits optimisation of a feeding strategy sufficient to maximise transcription and to supply adequate amounts of energy and ribosomes. It would therefore be better to model a continuous process rather than a batch process.

A Parity plot comparing the regressed CAT protein values with experimental values is shown in Figure 6.11. Clearly, equation 6.2 does not provide an accurate estimate of CAT protein levels in stationary phase. One reason for poor estimation is the fact that protein levels are relatively low. The effect of measurement error in the system is therefore large. However, it is clear that more comprehensive models to describe recombinant protein expression during stationary phase in *E. coli* need to be developed, and that high-integrity data near up-regulation are required.

6.4 Comparison of the *katE* gene or *tac* promoters

CAT protein expression levels in *E. coli* batch fermentation experiments under control of the IPTG inducible *tac* promoter or the stationary phase inducible *katE* gene promoter were compared.

Maximum CAT protein levels after IPTG induction in batch fermentation studies were 20-33 fg per cell (Figures 5.4, 5.5 and 5.6). By contrast, the maximal CAT protein level obtained by entry into stationary phase in batch fermentation experiments was around 0.08 fg per cell (Figures 6.3 and 6.5). The stationary-phase expression system is 250 times less efficient than the IPTG system. However, this is not really surprising, because of the stringent control in stationary phase. CAT protein levels could be improved by a factor of at least 2.6 by feeding glucose in fed-batch fermentation mode, as a result of improved

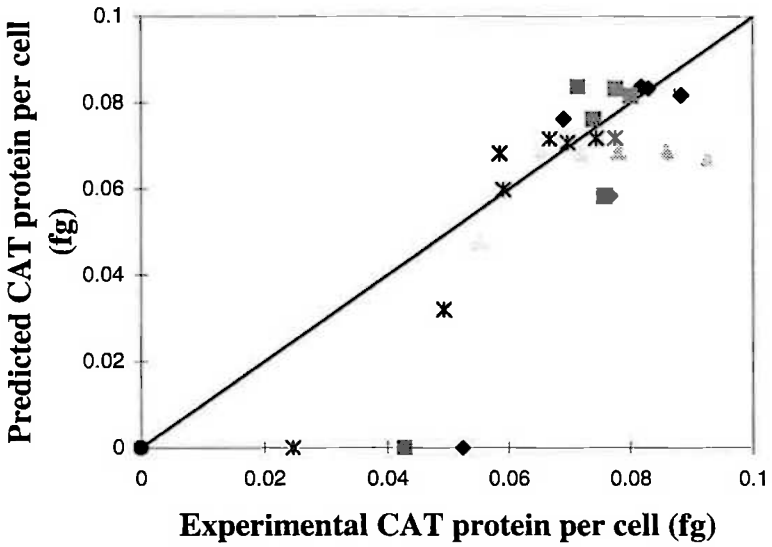

Figure 6.11. Parity plot comparing experimental data from fermentations induced by entry into stationary phase.

transcriptional and translational activities (Figure 6.7). In this experiment, transcriptional activities were improved 10-fold. Future work to establish a growth rate, by which stringent control is induced is expected to improve the performance of the stationary phase system.

6.5 Conclusions

The use of the stationary-phase inducible promoters *katE* and *katF* for recombinant protein expression in *E. coli* has been examined. These promoters were fused to the CAT reporter protein and promoter efficiency was measured.

Transcriptional activities for the *katF* system tended to be higher in logarithmic phase than in stationary phase for *E. coli* [pCT101] cultivated in glucose minimal media. This gene promoter was therefore not investigated further. By contrast, separation of cell growth and CAT production was possible using the stationary-phase-inducible promoter *katE* [pCT100] in glucose minimal media. Despite published reports indicating that the weak acids acetate and o-hydroxybenzoate stimulate *katE* promoter activity, the present study was unable to demonstrate this effect. Indeed it was found that o-hydroxybenzoic acid inhibited translational activity and acetate had no effect on *katE* promoter activity.

CAT protein level in batch experiments was limited at the translational levels. However, both transcriptional and translational activities were improved by continuously feeding glucose, thus establishing a growth-limited culture for CAT protein production. This suggests that the *katE* promoter can be used for high-density fermentations, thus avoiding the need for chemical inducers such as IPTG. Cost and environmental penalties are thereby avoided. However, future work is needed to define a critical growth rate below which the stringent control is induced.

The relative efficiency of the stationary-phase induced expression of recombinant CAT was evaluated by comparing expression levels with those for the IPTG-inducible *tac* promoter. Final levels of CAT protein are about 250 times higher in the IPTG-induced system for simple batch experiments. However, this is not really surprising because of the

stringent control in stationary phase. CAT protein levels under control of the stationary phase inducible *katE* gene promoter could be improved by a factor of at least 2.6, as proved by fed-batch experiments. Furthermore, transcriptional activity of the promoter was improved by a factor of 10. Future work to establish a growth rate, by which stringent control is induced, is expected to improve the performance of the stationary phase system.

Finally, the use of the *katE* gene promoter permits protein production without the addition of chemical inducers such as IPTG or the need to increase temperature, avoiding cost and environmental penalties as well as induction of heat-shock proteins. However, more information regarding regulation of the promoter is needed. This is discussed in Chapter 7. Evaluation of the usefulness of *katE* for commercial use would ultimately be based on an economic assessment.

Chapter 7

Summary and final discussion

The aim of this thesis was to characterise a bacterial expression system for recombinant production of proteins with relevance to industry. Commercial production of recombinant proteins is of great importance. Many proteins have been unavailable or only available in limited supply because of limited natural sources. With the development of DNA technology, it is now possible to produce large quantities of a wide range of proteins and also circumvent contamination problems such as those associated with source-located viruses and prions. In this thesis, recombinant protein expression under control of stationary-phase inducible promoters was characterised, and the expression levels were quantitatively compared with those under control of IPTG-inducible promoters

E. coli was chosen as the host to produce recombinant Chloramphenicol Acetyl Transferase (CAT). This organism is often preferred for commercial production of proteins. Numerous reasons reinforce this choice including extensive knowledge regarding the organism, rapid generation of biomass, the availability of low-cost culture media, and the large range of expression vectors available. The efficiency of a bacterial expression system may be monitored by measuring the total amount of a reporter protein produced. For example, CAT is a soluble protein that is simple to recover and detect. Quantitative estimates of the expression system's performance can therefore be achieved. CAT concentration may not accurately reflect expression for other recombinant proteins. In general, protein expression depends on the optimisation of many factors such as the binding of the RBS to the anti-Shine-Dalgarno sequence on the 16S rRNA, the distance between the RBS and ATG, the stability of mRNA, the cloned product and plasmid, the cell metabolic activity and the host-vector interaction (discussed in Chapter 1).

In this work, a number of bacterial expression systems were constructed using *promoter::cat* transcriptional fusions. These fusions were prepared by sub-cloning DNA fragments containing the IPTG inducible or stationary phase promoters, into a bidirectional promoter probe vector (Chapter 3). Constructs were all confirmed by sequence analysis (Chapter 3).

Recombinant protein expression in the engineered systems was quantified. Furthermore, simple mathematical models to establish the relative efficiencies of transcription and translation were developed. Consequently, considerable attention was

placed on the optimisation, validation and development of quantitative assays. Detailed studies to examine recovery, sample integrity during storage, and to confirm accurate detection of all species were conducted (Chapter 4).

Several findings resulted from this work. First, it was found that the French press was more efficient at disrupting cells and releasing CAT protein than sonication. French pressing totally disrupted bacteria whereas only a fraction of cells was disrupted following sonication. Between 96 and 99 % of cells were disrupted after 1 pass in the French press. By contrast, a maximum of 80 % of the cells were disrupted by sonication. A strict proportionality between the degree of disruption by sonication and the release of CAT was not found, as some cells became sufficiently porous for CAT to be released without being fully disrupted. The number of passes necessary to achieve maximal release, using French pressing, was a function of the culture media. Cells cultivated in rich media required more passes than cells cultivated in a defined media, probably because of the formation of a more rigid cell membrane after proliferation in rich broth. Furthermore, despite the fact that freezing disrupted a significant fraction of the cells more passes were required (storage at low temperature (-70°C)). This was explained by simple displacement of water during freezing, bringing cell proteins closer together. Also, treatment with detergent (e.g. Triton X-100) was necessary to achieve maximal CAT recovery, possible due to partitioning of CAT to the cellular debris. However, Triton X-100 has an inhibitory effect on enzyme activity confirming that quantitative measurement of total CAT by ELISA is more appropriate than enzyme activity measurement in the current study. Other findings were that sample integrity was not compromised by handling and storage. Storing at -70°C did not reduce CAT protein, total soluble protein, nor RNA levels in the cell samples.

Quantitation of *cat* mRNA by slot-blot hybridisation was compromised by the presence of unrelated RNA. Quantitation of message by comparing signal intensity to that of an *in vitro* produced standard resulted in gross underestimation of *cat* mRNA concentration. This was explained by steric hindrance occurring from total cellular RNA limiting the probe's ability to anneal to target mRNA. To overcome this problem, the feasibility of isolating *cat* mRNA from crude mixtures of RNA prior to immobilisation using Dynabeads M-280 Streptavidin was investigated. However, this experiment was not successful. The binding capacity of the double stranded DNA fragment to the Dynabeads

and the efficiency of *cat* mRNA hybridisation could not be established, as mass balances did not close. Fortunately, the problem of steric hindrance leading to underestimation of *cat* mRNA concentration by slot blotting was solved by constructing a three-dimensional calibration curve. This curve combined five levels of *cat* mRNA with six levels of commercially-purified 16/23S rRNA. In the same experiment, samples with unknown levels of *cat* mRNA were loaded. The actual amount of *cat* mRNA was determined by comparing signal intensities with those on the calibration curve. The standard calibration curve was included for each assay, as assay response varied between experiments because of probe variability. This finding confirms that many assays are at best semi-quantitative unless carefully developed and validated.

High level of recombinant protein expression after induction can lead to a marked decrease in overall biomass yield (Andersson *et al.*, 1996). This results in detrimental effects for both the host and recombinant protein. A variety of bacterial expression systems have been designed for high-level protein production to offer tight control of transcription, determined by the ability to regulate the promoter sequence. Widely-used promoters are induced chemically by isopropyl-β-thiogalactopyranoside (IPTG). The disadvantage of chemical induction is that both cost and environmental penalties are incurred. Such systems are therefore less attractive for the large-scale production of proteins. However, an IPTG-induced expression system was included in this thesis as a control system to compare the relative efficiency of a potential commercially relevant system (stationary phase induction). To make such comparisons useful, the control system with maximal performance had to be explored. In this work, the CAT reporter gene was under control of the *tac* promoter (Chapter 5). CAT protein levels obtained at different IPTG concentrations were compared with *cat* mRNA levels, and the best inducer concentration for a batch system was established. Furthermore, the rate limiting steps in protein synthesis under IPTG induced conditions were explored by relating CAT protein to both *cat* mRNA and 16S rRNA concentrations. Other factors such as oxygen availability, growth phase of the inoculum, and the starting OD for the fermentation were also examined. Experiments were conducted in both shake-flasks and fermenters to determine whether the results from each system were comparable. The systems were further characterised in a simple mathematical model to establish the relative efficiencies of transcription and translation in the IPTG induced system.

It was demonstrated that maximal protein expression was achieved for 0.1 mM IPTG after induction at OD_{600} = 0.8 in both shake-flask and fermentation experiments, whereas a concentration of 0.4 mM IPTG yielded maximal expression for induction at OD_{600} = 2.4. Maximum CAT protein expression levels were independent of oxygen tension. However, CAT protein production was highly dependent on the growth phase of the culture to be induced. Induction close to stationary phase produced lower levels of CAT compared to induction in logarithmic phase.

The reproducibility of CAT protein levels obtained from fermentations inoculated with the same culture was good. By contrast, levels varied between runs even when experiments were carried out under similar conditions. It was discovered that inoculation with a stationary-phase culture gave better CAT protein yield than cultures inoculated with logarithmic phase cultures. Similarly, fermenters inoculated to a higher starting OD_{600} resulted in higher CAT protein levels. However, differences were only measurable for large differences in starting OD_{600} (corresponding to almost 7 generations). Although this indicated some plasmid instability, the differences were not significant.

It was clear that CAT protein production under control of the *tac* promoter was limited at the translational level. For example, an increase in IPTG concentration beyond optimal resulted in a concomitant increase in mRNA level but not CAT protein. Furthermore, CAT protein levels remained constant for some time after induction while ribosomal (16S rRNA) levels decreased. This translational limitation was confirmed by a simple mathematical models to establish the relative efficiency of transcription (16-81 %) and translation (0.2-1 %). The translational limitation was explained by the ribosomal machinery not effectively meeting the metabolic and energy requirements necessary to effectively process mRNAs for maintenance and recombinant protein needs. Future work should focus on optimisation of translation for recombinant-protein expression in this system.

The IPTG-induced expression systems were compared with a system with commercial potential, using stationary phase promoters. Bacterial expression systems specifically designed for fermentation conditions have recently been investigated (Dikshit *et al.*, 1990; Khosla *et al.*, 1990; Chou *et al.*, 1995; Tunner *et al.*, 1992). Among these

systems are carbon-starvation inducible promoters (Tunner *et al.*, 1992). The use of starvation inducible promoters is attractive as they involve protein production in a slow or non-growing cell, thus avoiding the diversion of nutrients into biomass. Any waste of nutrients is avoided. The induction of stationary phase promoters is also an inexpensive process. A group of carbon starvation gene promoters (*cst* genes) have been investigated for their ability to uncouple the expression of recombinant β-galactosidase from cell growth (Tunner *et al.*, 1992). As *cst* gene promoters are positively regulated by cAMP (Schultz *et al.,* 1988), they are susceptible to glucose repression in logarithmic growth phase. This facilitates the separation of the biomass generation and expression phases during fermentation. However, a serious limitation to recombinant protein production in stationary phase is the loss of rRNA and ribosomes. The turnover of these macromolecules normally provides building blocks for the synthesis of starvation proteins, and maintenance energy which represents a significant drain on metabolic sources at low growth rates. When additional burden is placed on the cells by the introduction of heterologous genes, more energy is required, and a feeding strategy must be established to ensure adequate energy levels. In the work of Tunner *et al.* (1992), this energy need was partly provided by an acetic acid feeding strategy. ATP generation from acetate fuelling is much lower than from glucose conversion (Varma and Palsson, 1993). One would therefore expect that a glucose feeding strategy would result in higher protein levels. However, any glucose feeding strategy for recombinant protein production under control of the *cst* gene promoters will be difficult to implement in a controlled way.

Another group of starvation-inducible gene promoters, namely the *pex* gene promoters, are either independent of cAMP or are negatively regulated by it (Schultz *et al.,* 1988). One advantage of using these gene promoters is that glucose feeding during the production phase need not be strictly controlled. This is in contrast to the behaviour of the *cst* promoters, suggesting the *pex* promoters offer considerable advantage for large-scale work in poorly-mixed bioreactors. Glucose may also be regarded as an inducer of the *pex* gene promoters, and may therefore be employed to improve product yield and to extend the production period.

Given their apparent advantage, the *pex* gene promoters *katE* and *katF* were characterised in this work for their ability to uncouple the production of recombinant CAT

protein from *E. coli* cell growth. The use of these promoters is obviously compromised by promoter activity during the growth phase. It is therefore necessary to identify substances that inhibit promoter activity during the growth phase. Available evidence in the literature suggests that this is indeed possible (Mulvey *et al.*, 1990).

Work undertaken in this thesis suggested that transcriptional activities for the *katF* system tended to be higher in logarithmic phase than in stationary phase, for *E. coli* [pCT101] cultivated in minimal media. This gene promoter was therefore not investigated further. By contrast, separation of cell growth and CAT production was possible using the stationary phase-inducible promoter *katE* in glucose minimal media. Despite published reports, no inducers of the *katE* gene promoter were identified. Acetate and o-hydroxybenzoate did not stimulate promoter activity as previously reported (Schellhorn and Stones, 1992; Mulvey *et al.*, 1990). This may be due differences in environmental factors. This work was conducted in a controlled environment (fermentation), whereas other work is generally conducted in uncontrolled shake-flasks, where key parameters (e.g., pH, pO_2) will vary. Interestingly, o-hydroxybenzoic acid actually inhibited translational activities. Glucose addition did improve CAT protein levels in late stationary phase, but this was believed to be due to increased translational activity.

CAT protein levels in batch experiments was limited at the translational level. However, both transcriptional and translational activities were improved by continuously feeding glucose, thus establishing a growth-limited culture for CAT protein production. This suggests that the *katE* promoter can be used for high-density fermentations, thus avoiding the need for chemical inducers such as IPTG. Cost and environmental penalties are thereby avoided. However, future work is needed to define a critical growth rate below which the stringent control is induced. Two reasons exist. The first is the need to define an optimal feeding rate. An optimal feeding rate will be sufficiently high to overcome the stringent response but will not exceed the threshold for optimal stimulation of transcription. The second reason for the need to define a critical growth rate is that such a development will aid the process of developing mechanistic mathematical models to predict stationary-phase-gene expression in *E. coli* under nutrient limitation.

A guide to establishing an optimal feeding rate can be obtained by comparison of optical density (OD_{600}) and cell numbers. This can be simply obtained using a Coulter counter and an on-line turbidity meter, during fed-batch or continuous fermentations. Constant cell number means that cells are in a true stationary phase, even though cell dry weight may increase. Cells "fatten up" when bulk glucose concentration exceeds maintenance requirements. Similarly, they "slim down" when the carbohydrate source becomes limiting. An increase in cell dry weight will result in increased cell density as measured by OD_{600}. Consequently, an increase in OD_{600} for constant cell number most likely means that ribosomal mass is increasing, and that the cells are not subject to stringent control. However, it also means that nutrient supply exceeds requirements, repressing transcriptional activities from stationary phase induced promoters. In summary, a growth limited culture with selective protein production will show constant cell count at constant cell density.

CAT protein expression levels in *E. coli* batch fermentation experiments under control of either IPTG or stationary phase inducible promoters were compared. Expression levels in the two systems were significantly different. Final levels of CAT protein were about 250 times higher in the IPTG-induced system. However, this is not really surprising because of the stringent control in stationary phase. Transcriptional activity of CAT under control of the stationary-phase inducible *katE* gene promoter can be improved by a factor of at least 10, as proved by fed-batch experiments. Future work to establish an optimal growth rate that prevents stringent control is expected to improve performance.

Future work

Before the *katE* gene promoter can be employed for commercial production of proteins, more information is needed. Not much is known about the regulation of *katE* gene expression. The *katE*::*cat* transcriptional fusion used in this work encoded 132 bp between the *katE* gene promoter and the AUG start codon for the *katE* gene. A regulatory DNA sequence between the two may exist. This regulation may be either positive or negative. The construction of precise transcriptional fusions will resolve this issue. The

transcriptional efficiency may be improved by optimisation of promoter strength and regulation of transcription. Similarly, translational efficiency may be improved by optimising binding of the RBS to the anti-Shine-Dalgarno sequence and by optimising the distance between the RBS and ATG. These are standard techniques in molecular biology, discussed in section 1.4. They were not examined in this thesis as the focus was on the effect of environmental parameters.

Work undertaken in this thesis showed that the transition from logarithmic growth phase to stationary phase stimulates *katE* expression. However, it also indicated that protein expression was limited at the translational level. The level of regulation could be shown experimentally. The construction of transcriptional and translational reporter gene fusions will allow both activities to be compared.

There is some evidence in the literature to show that *katE* expression is regulated by KatF. Other proteins may also be involved in this regulation. However, the control mechanism may be lost by the introduction of a multi-copy plasmid. This may simply be due to a dilution effect. Such a possibility should be explored and could be overcome by introducing the plasmid into a strain overproducing KatF constitutively.

Most importantly, the definition of a critical growth rate below which the stringent control is induced must be identified. This will allow the definition of an optimal feeding rate to supply sufficient energy for maximal translation. Secondly, it will aid the process of developing mechanistic mathematical models to predict stationary-phase-gene expression in *E. coli* under nutrient limitation.

Finally, the evaluation of the usefulness of *katE* for commercial use would ultimately be based on an economic assessment.

Appendix A.1

k_{po} and k_{qo} is calculated as described by Lee and Bailey (1984):

$$k_{po} = \frac{\left(\frac{1\ \text{mRNA molecule}}{z\ \text{ribonucleotides}}\right) \times \left(\frac{2400\ \text{ribonucleotides}}{\text{active RNA polymerase min}}\right) \times}{\left(\frac{z\ \text{deoxy ribonucleotides}}{1\ \text{plasmid DNA gene}}\right) \times \left(\frac{1\ \text{active RNA polymerase}}{d_p\ \text{deoxy ribonucleotides}}\right)} \quad (5.4)$$

$$k_{qo} = \frac{\left(\frac{1\ \text{protein molecule}}{y\ \text{amino acids}}\right) \times \left(\frac{1200\ \text{amino acids}}{\text{active ribosome min}}\right) \times \left(\frac{3y\ \text{nucleotides}}{1\ \text{plasmid gene transcript}}\right)}{\left(\frac{1\ \text{plasmid gene transcript}}{\text{mRNA molelcule}}\right) \times \left(\frac{1\ \text{active ribosome}}{d_r\ \text{nucleotides}}\right)} \quad (5.5)$$

d_p and d_r describe the average intermolecular distances between transcribing RNA polymerase molecules and translating ribosomes, respectively. From experimental results of Dennis and Bremer (1973 and 1974):

$d_p = 233\ \mu^{-2} + 78$ (nucleotides)

$d_r = 82.5\ \mu^{-1} + 145$ (nucleotides)

From equation 5.4 and 5.5:

$k_{po} = 4.26$ M mRNA/M plasmid DNA min

$k_{qo} = 13.6$ M protein/M mRNA min

Nomenclature

ξ	overall translation efficiency
η	overall transcription efficiency
μ	specific growth rate (min^{-1})
G	DNA concentration (fg)
k_p	transcription rate constant (M mRNA/M DNA min)
k_d	mRNA decay rate constant (min^{-1})
k_e	protein decay rate constant (min^{-1})
k_q	translation rate constant (M protein/M mRNA min)

Subscript

$_p$	plasmid

Bibliography

Aiba, H., Adhya, S., and de Crombrugghe, B. (1981). Evidence for two functional gel promoters in intact *E. coli*. *J. Biol. Chem.* **256:** 11905-11910.

Andersson, L., Yang, S., Neubauer, P. and Enfors, S. O. (1996). Impact of plasmid presence and induction on cellular response in fed batch cultures of *Escherichia coli*. *J. Bacteriol.* **46:** 255-263.

Amann, E., Brosius, J. and Ptashne, M. (1983). Vectors bearing a hybrid *trp-tac* promoter useful for regulated expression of cloned genes in *Escherichia coli*. *Gene* **25:** 167-178.

Amann, E. and Brosius, J. (1985). "ATG vectors" for regulated high-level expression of cloned genes in *Escherichia coli*. *Gene* **40**: 183-190.

Balbas, P., Soberon, X., Bolivar, F. and Rodriquez, R. L. (1988). Bacterial cloning vectors. In *Vectors*. (Editors, Rodriquez, R. L. and Denhardt, D. T.). Butterworth Publishers, Stoneham, MA)

Belasco, J. G. (1993). mRNA degradation in prokaryotic cells: An overview. In *Control of messenger RNA stability*, Eds. Belasco, J. G. and Brawerman, G. Academic Press, Inc. San Diego.

Bentley, W. E, Davis, R. H. and Kompala, S. (1991). Dynamics of induced CAT expression in *E. coli*. *Biotech. Bioeng.* **38:** 749-760

Berg, T., Dirla, U., Naumann, U., Heuft, H. G., Kuther, S., Lobeck, H., Schreier, E. and Hopf, U. (1996). Responsiveness to interferon alpha treatment in patients withchronic hepatitis C coinfected with hepatitis G virus. *J. Hepatol.* **25:** 763-768.

de Boer, H.A., Comstock, L. J. and Vasser, M. (1983). The *tac* promoter: a functional hybrid derived from the *trp* and *lac* promoter. *Proc. Natl. Acad. Sci. USA*. **80:** 21-25.

Bhaduri, S. and Demchick, P. H. (1983). Simple and rapid method for disruption of bacteria for protein studies. *Appl. Environ. Microbiol.* **46:** 941-943.

Bochkareva, E. S., Lissin, N. M. and Girshovich, A. S. (1988). Transient association of newly synthesized unfolded proteins with the heat-shock GroEL protein. *Nature* **336:** 254-257.

Boidol, W., Simmons, M., Topert, M. and Siewert, G. (1982). Recombinant plasmids with genes for biosynthesis of alkaline phosphatase of *E. coli*. *Mol. Gen. Genet.* **185:** 510-512.

Bolivar, F., Rodriguez, R. L., Greene, P.J., Betlach, M. C., Heyneker, H.L., Boyer, H. W. (1977). Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene* **2:** 95-113.

Bremer, H. and Dennis, P. P. (1987). Modulation of chemical composition and other parameters of the cell by growth rate in: *Escherichia coli and Salmonella typhimurium,*

(Neidhardt, F. C., Ingram J. L., Low K. B., Magsanik, B., Schaechter, M. and Unbarger, H. E. eds.). American Society for Microbiology, Washington, D.C., pp. 1527-1542.

Brosius, J. and Holy, A. (1984). Regulation of ribosomal RNA promoters with a synthetic *lac* operator. *Proc. Natl Acad. Sci. USA.* **81:** 6929-6933.

Brosius, J. L. (1984). Toxicity of an overproduced foreign gene product in *Escherichia coli* and its use in plasmid vectors for the selection of transcription terminators. *Gene* **27:** 161-172.

Brosius, J. L., Erfle, M. and Storella, J. (1985). Spacing of the -10 and -35 region in the *tac* promoter. *J. Biol. Chem.* **260:** 3539-3541.

Buettner, M. J., Spitz, E. and Rickenberg, H. V. (1973). Cyclic adenosine 3'.5'-monophosphate in *Escherichia coli. J. Bacteriol.* **114:** 1068-1073.

Caldeira, J. C. G. and Cabral J. M. S. (1994). Extraction of a steroid 1,2 dehydrogenase by sonication of *Arthrobacter simplex* cells. *Bioseparation* **4:** 271-278.

Cashel, M. and Rudd, K. E. (1987). The stringent response. In *Escherichia coli* and *Salmonella typhimurium* Cellular and Molecular Biology; Neidhardt F C, Ed.; Washington, D.C; pp 1410-1438.

Chen, W., Bailey, J. E. and Lee, S. B. (1991). Molecular design of expression systems: comparison of different repressor control configurations using molecular mechanism models. *Biotech. Bioeng.* **38:** 679-687.

Chen, W., Kallio, P. T. and Bailey, J. E. (1995). Process characterisation of a novel cross-regulated system for cloned gene production in *Escherichia coli. Biotech. Progr.* **11:** 397-402.

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. and Rutter, W. J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18:** 5294-5299.

Chou, C., Aristidou, A. A., Meng, S., Bennett, G. N. and San, K. (1995). Characterisation of a pH-inducible promoter system for high-level expression of recombinant proteins in *Escherichia coli. Biotech. Bioeng.* **47:** 186-192.

Damarau, K. and St. John, A. C. (1993). Role of protease subunits in degradation of carbon starvation proteins in *Escherichia coli. J. Bacteriol.* **175:** 53-63

Davis, B. D., Luger, S. M. and Tai, P. C. (1986). Role of ribosome degradation in the death of starved *Escherichia coli* cells. *J. Bacteriol.* **166:** 439-445.

Dawes, E. (1976). Endogenous metabolism and the survival of starved prokaryotes. In *The survival of Vegetative Microbes.* The Society for General Microbiology Symposium 26 ed. Gray, T & Postgate, J, pp. 19-53. Cambridge: Cambridge University Press.

Dean, C. R. and Ward, O. P. (1992). The use of EDTA or polymyxin with lysozyme for the recovery of intracellular products from *Escherichia coli*. *Biotech. Tech.* **6:** 133-138.

Desai, I. (1968). Cell preparation for the assay of threonine dehydratase in *Escherichia coli*. *Biochim. Biophys. Acta.* **167:** 467-469.

Dikshit, K. Dikshit, R. P. and Webster, D. (1990). Study of *Vitreoscilla* globin (*vgb*) gene expression and promoter activity in *E. coli* through transcriptional fusion. *Nucleic Acids Res.* **18:** 4149-4155.

Dykes, C. W. (1993). Molecular Biology and Biotechnology (3rd edition), pp 155-176. (Editors Walker, J. M. and Gingold, E.). The Royal Society of Chemistry, Cambridge.

Dynal (1995). Biomagnetic Techniques in Molecular Biology, 2nd ed. Dynal A/S. Oslo, Norway.

Feliu, J. X. and Villaverde A. (1994). An ultrasonication protocol for bacterial cell disruption and recovery of b-galactosidase fusion proteins. *Biotech. Tech.* **8:** 509-514.

De Franco, C. and Schottel, J. L. (1989). Terminal sequences do not contain the rate-limiting decay determinants of *E. coli cat* mRNA. *Nucleic Acids Res.* **17:** 1139-57.

Furuta, Y., Shinohara, T., Sano, K., Meguro, M. and Nagashima K. (1990). *In situ* hybridisation with digoxigenin-labelled DNA probes for detection of viral genomes. *J. Clin. Pathol.* **43:** 806-809.

Georgiou, G. and Bancyx, F. (1996). Bacterial host strains for producing proteolytically sensitive polypeptides. USA Patent no 5508192.

Georgiou, G., Telford, J. N., Schuler, M. L. and Wilson, D. B. (1986). Localisation of inclusion bodies in *Escherichia coli* cells overproducing b-lactamase or alkaline phosphatase. *Appl. Environ. Microb.* **52:**1157-1161.

Gerdes, K. (1988). The parB (hok/sok) locus of plasmid R1: A general purpose plasmid stabilisation system. *Bio/technology*. **6:** 1402-1405.

Gilman, M. In *Short protocols In Molecular Biology*; Ausubel, F. M.., Brent, R., Kingston, R. E., Moore, D. D. Seidman, J. G., Smith J. A., Struhl K. Eds; Greene Publishing Associates and Wiley-Interscience: New York, 1989; pp 146-148.

Goff, S. A. and Goldberg, A.L. (1985). Production of abnormal proteins in *E. coli* stimulates transcription of lon and other heat shock genes. *Cell* **41:** 587-595.

Gold, L. (1988). Post transcriptional regulatory mechanisms in *Escherichia coli*. *Annu. Rev. Biochem.* **57:** 199-233.

Gorman, C. M., Moffat, L. and Howard, B. H. (1982). Recombinant genomes which express chloramphenicol acetyl transferase in mammalian cells. *Mol. Cell. Biol* **2:** 1044-1051.

Groat, R.G., Schultz, J. E., Zychlinski, E., Bockman, A. T. and Matin, A. (1986). Starvation proteins in *Escherichia coli:* kinetics of synthesis and role in starvation survival. *J. Bacteriol.* **168:** 486-493.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166:** 557-580.

Harcum, S. W. and Bentley, W. E. (1993). Response dynamics of 26- 34- 39- 54- and 80-Kda proteases in induced cultures of recombinant *Escherichia coli. Biotech. Bioeng.* **42:** 675-685.

Heijnen, J. J., Terwisscha van Scheltinga, A. H. and Straathof, A. J. (1992). Fundamental bottlenecks in the application of continuos bioprocesses. *J. Biotechnol.* **22:** 3-22.

Higgins, C., Causton, H. C., Dance, G. S. C. and Mudd, E. A. (1993). The role of the 3'end in mRNA stability and decay. In *Control of messenger RNA stability.* Eds. Belaco, J. G. and Brawerman, G. Academic Press, Inc. San Diego. p. 13-30.

Hughes, D. E., Wimpenny, J. W. T. and Lloyd, D. (1971). *Methods in Microbiology.*, 5B, 1-54. Academic, New York.

Hui, A. and de Boer, H. A. (1987). Specialised ribosome system: Preferential translation of a single mRNA species by a sub-population of mutated ribosomes in *Escherichia coli. Proc. Nat. Acad. Sci. U. S. A.* **84:** 4762-4766.

Jacob, W. F., Santer, M. and Dahlberg, A. E. (1987). A single base change in the shine-Dalgarno region of 16S rRNA of *Escherichia coli* affects translation of many proteins. *Proc. Nat. Acad. Sci. U.S.A.* **84:** 4757-4761.

Jacques, N., Guillerez, J. and Dreyfus., M. (1992). Culture conditions differentially affect the translation of individual *Escherichia coli* mRNAs. *J. Mol. Biol.* **226:** 597-608.

Jensen, K. F. and Pedersen, S. (1990). Metabolic growth rate control in Escherichia coli may be a consequence of subsuration of the macromolecular biosynthetic apparatus with substrates and catalytic components. *Microbiol. Rev.* **54:** 89-100.

Jinks-Robertson, S. and Nomura, M. (1987). Ribosomes and tRNA. In *Escherichia coli* and *Salmonella typhimurium.* Cellular and Molecular Biology, Neidhardt F C, Ed, Washington D.C. pp: 1358-1385.

Jorgensen, L., Middelberg, A. P. J., O'Neill, B. K. and Thomas, C. J. (1996). Quantitation of chloramphenicol acetyl transferase (CAT) messenger RNA by filter hybridisation using a digoxigenin label. *Biotech. Tech.* **10:** 83-88.

Jorgensen, L., O'Neill, B. K., Thomas, C. J., Morona, R. and Middelberg, A. P. J. (1995). Release of chloramphenicol acetyl transferase from recombinant *Escherichia coli* by sonication and the French press. *Biotech. Tech.* **9:** 477-480.

Kaback, H. R. and Deuel, T. F. (1969). Proline uptake by disrupted membrane preparations from *Escherichia coli*. *Arch. Biochem. Biophys.* **132:** 118-129.

Khosla, C., Curtis, J. E. Bydalek, P, Swartz, J. R. and Bailey, J. (1990). Expression of recombinant proteins in *Escherichia coli* using an oxygen-responsive promoter. *Bio/Technology* **8:** 554-558.

King, R., Wells, J. R. E., Krieg, P., Snoswell M., Brazier, J., Bagley, C. J., Wallace, J. C., Ballard, F. J., Ross, M. and Francis, G. L. (1992). Production and characterisation of recombinant insulin-like growth factor-I (IGF-I) and potent analogues of IGF-I, with Gly or Arg substituted for Glu^3, following their expression in *Escherichia coli* as fusion proteins. *J. Mol Endocrin.* **8:** 29-41.

Kleinig, A. R. (1997). Cell disruption mechanics. PhD thesis. The University of Adelaide.

Kubitschek, H. E. (1990). Cell volume increase in *Escherichia coli* after shifts to richer media. *J. Bacteriol.* **172:** 94-101.

Laffend, L. and Schuler, M. L. (1994). Ribosomal protein limitations in *Escherichia coli* under conditions of high translational activity. *Biotech. Bioeng.* **43:** 388-398.

Lange, R. and Hengge-Aronis, R. (1991). Identification of a central regulator of stationary-phase gene expression in *Escherichia coli*. *Mol Microbiol.* **5:** 49-59.

Lange, R. and Hengge-Aronis, R. (1995). The cellular concentration of the σ^s sub-unit of RNA polymerase in *Escherichia coli* is controlled at the levels of transcription, translation and protein stability. *Genes Dev.* **8:** 1600-1612.

Laurence, D. R. and Bennett, P. N. (1992). Clinical Pharmacology 7^{th} ed. Edinburgh, Churchill Livingstone.

LaVallie, E. R., Di Blasio, E. A., Kovacic, S., Grant, K. L., Schendel, P. F. and McCoy, J. M. (1993). A Thioredoxin gene fusion expression system that circumvemts inclusion body formation in the *E. coli* cytoplasm. *Bio/Technology.* **11:** 187-193.

Leduc, M., Frehel, C., Siegel, E. and van Heijenoort, J. (1989). Multi-layered distribution of peptidoglycan in the periplasmic space of *Escherichia coli*. *J. Gen. Microbiol.* **135:** 1243-1254.

Lee, S. B. and Bailey, J. E. (1984). Analysis of growth rate effects on productivity of recombinant Escherichia coli populations using molecular mechanism models. *Biotech. Bioeng.* **26:** 66-73.

Liljenstrom, H. and Heijne, G. (1987). Translation rate modification by preferential codon usage: Intragenic position effects. *J. Theoret. Biol.* **124:** 43-55.

Loewen, P. C., Switala, J. and Triggs-Raine, B. L. (1985). Catalase HPI and HPII in *Escherichia coli* are induced independently. *Arch. Biochem. Biophys.* **243:** 144-149.

Lopez, P. J., Iost, I. and Dreyfus, M. (1994). The use of a tRNA as a transcriptional reporter: the T7 late promoter is extremely efficient in *Escherichia coli* but its transcript are poorly expressed. *Nucleic Acids Res.* **22:**1186-1193.

Lu, J. and Jiang, C. (1992). Detergents inhibit chloramphenicol acetyl transferase. *BioTechniques* **12:** 643-644.

Lubke, C., Boidol, W and Petri, T. (1995). Analysis and optimisation of recombinant protein production in Eschericia coli using the inducible *phoA* promoter of the *E. coli* alkaline phosphatase. *Enzyme Microb. Technol.* **17:** 923-928.

Maniatis, T., Fritsch, E. F and Sambrook, J. (1982). *Molecular cloning a laboratory manual.* Cold spring Harbour, N.Y. Cold spring Harbour Laboratory.

Matin, A. (1991). The molecular basis of carbon-starvation-induced general resistance in *Escherichia coli. Mol. Microbiol.* **5:** 3-10.

Mattanovich, D., Weik, R., Thim, S., Kramer, W., Bayer, K. and Katinger, H. (1996). Optimisation of recombinant gene expression in *Escherichia coli. Ann. N. Y. Acad. Sci.* **782:** 182-190.

McCann, M. P., Kidwell, J. P. and Matin, A. (1991). The putative σ factor katF has a central role in development of starvation-mediated general resistance in *Escherichia coli. J. Bacteriol.* **173:** 4188-4194.

McKenny, K., Shimatake, H., Court, D., Schmeissner, U., Brady, C., and Rosenberg, M. (1981). A system to study promoter and terminator signals recognised by *Escherichia coli* RNA polymerase. In Chirikjian, J. G. and Papas T. S., (eds.), Gene amplification and analysis, Vol II: Structural analysis of nucleic acids. Elsevier, New York, 1981, pp. 383-415.

Meyer, B. J. and Schottel, J. L. (1991). A novel transcriptional response by the *cat* gene during slow growth of *Escherichia coli. J. Bacteriol.* **173:** 3523-3530.

Middelberg, A .P. J. (1992). A model for the disruption of *Escherichia coli* by high-pressure homogenisation, PhD thesis. The University of Adelaide.

Miller J. H. (1972). Experiments in molecular genetics. Cold Spring Harbour Laboratory, Cold Spring Harbour, N.Y.

Mulligan, M. E., Brosius, B. and McClure, W. R. (1985). Characterisation *in vitro* of the effect of spacer length on the activity of *Escherichia coli* RNA polymerase at the *tac* promoter. *J. Biol. Chem.* **260:** 3529-3538.

Mulvey, M. R., and Loewen P. C. (1989). Nucleotide sequence of *katF* of *Escherichia coli* suggests katF protein is a novel σ transcription factor. *Nucleic Acids Res.* **17:** 9979-9991.

Mulvey, M. R., Switala, J., Borys, A. and Loewen, P. C. (1990). Regulation of transcription of *katE* and *katF* in *Escherichia coli. J. Bacteriol.* **172:** 6713-6720.

Novo Nordisk. URL: www.novo.dk/backgrou/history/history6.html.

Ossowski, I., Mulvey, M. R., Leco, P. A., Borys, A. and Loewen, P. C. (1991). Nucleotide sequence of *Escherichia coli katE*, which encodes catalase HPII. *J. Bacteriol.* **173:** 514-520.

Pease, A. and Wolf, J. R. (1994). Determination of the growth rate-regulated steps in expression of the *Escherichia coli* K-12 *gnd* gene. *J. Bacteriol.* **176:** 115-122.

Pedersen, S. (1984). *E. coli* ribosomes translate *in vivo* with variable rate. *EMBO* **3:** 2895-2898.

Perretti, S. W. and Bailey, J. E. (1986). Mechanistically detailed model of cellular metabolism for glucose-limited growth of *Escherichia coli* B/r-A. *Biotech. Bioeng.* **28:** 1672-1689.

Peretti, S. W. and Bailey, J. E. (1987). Simulations of host-plasmid interactions in *Escherichia coli*: Copy number, promoter strength and ribosome binding site strength effects on metabolic activity and plasmid gene expression. *Biotech. Bioeng.* **29:** 316-328.

Ramirez, D. M. and Bentley, W. E. (1995). Fed-batch feeding and induction policies that improve foreign protein synthesis and stability by avoiding stress response. *Biotech. Bioeng.* **47:** 596-608.

Reeve, C. A., Bockman, and Matin, A. (1984). Role of protein degradation in the survival of carbon-starved *Escherichia coli* and *Salmonella typhimutium*. *J. Bacteriol.* **146:** 476-483.

Remaut, E., Stanssens, P. and Fiers, W. (1981). Plasmid vectors for high-efficiency expression controlled by the P_L promoter of coliphage lambda. *Gene* **15:** 81-93.

Schellhorn, H. E. and Stones, V. L. (1992). Regulation of *katF* and *katE* in *Escherichia coli* K-12 by weak acids. *J. Bacteriol.* **174:** 4769-4776.

Schneider, K. and Beck, C. F. (1986). Promoter-probe vectors for the analysis of divergently arranged promoters. *Gene* **42:** 37-48.

Schottel, J. L., Sninsky, J. J. and Cohen S. N. (1984). Effects of alterations in the translational control region on bacterial gene expression: use of *cat* gene construct transcribed from the *lac* promoter as a model system. *Gene* **28:** 177-193.

Schuler, M. L. and Domach, M. M. (1982). *Mathematical models of the growth of individual cells*, in Foundations of Biochemical Engineering: Kinetics and thermodynamics in Biological systems., 93-133 American Chemical Society Publications.

Schuler, M., L., Leung, S. K. and Dick, C. C. (1979). A mathematical model for the growth of a single bacterial cell. *Ann N. Y. Acad. Sci.* **326:** 35-55.

Schultz, J. E., Latter, G. I. and Matin, A. (1988). Different regulation by cyclic AMP of starvation protein synthesis in *Escherichia coli*. *J. Bacteriol.* **170:** 3903-3909.

Schmitt, B. (1976). Pyruvate dehydrogenase in *Escherichia coli*: an active 17S species in crude extracts. *Biochemie*. **58:** 1405-1407.

Shirakawa, M., Tsurimoto, T., and Matsubara, K. (1984). Plasmid vectors designed for high-efficiency expression controlled by the portable *rec*A promoter-operator of *Escherichia coli*. *Gene* **28:** 127-132.

Sorensen, K. I. and Neuhard, J. (1991). Dual transcription initiation sites from the pyrC promoter control expression of the gene in *Salmonella typhimurium*. *Mol. Gen. Genet.* **225:** 249-256.

Stoeckle, M. Y. and Guan, L. (1993). Improved resolution and sensitivity of Northern blots using Polyacrylamide-Urea gels. *BioTechniques* **15:** 227-231.

Stouthamer, A. H., Bulthuis, B. A., van Verseveld, H. W. Energetics of growth at low growth rates and its relevance for the maintenance concept. In Microbial Growth Dynamics; Poole, R. K.; Bazin, M. J., Keevil, C. W., Eds; Society of General Microbiology, IRL Press: Oxford, England, 1990; pp 85-104.

Studier, F. W. and Moffatt, B. A. (1986). Use of Bacteriophage T7 RNA Polymerase to direct Selective High Level Expression of Cloned Genes. *J. Mol. Biol.* **189:** 113-130.

Studier, F.W., Rosenberg, A. H., Dunn, J.J. and Dubendorff, J.W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185:** 60-89.

Summers, W. C. (1970) A simple method for extraction of RNA from *E. coli* utilising diethyl pyrocarbonate. *Anal. Biochem.* **33:** 459-463.

Tabor, S. and Richardson, C. C. (1985). A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. USA* **82:** 1074-1078.

Tacon, W., Carey, N. and Emtage, S. (1980). The construction and characterisation of plasmid vectors suitable for the expression of all DNA phases under control of the E. coli tryptophan promoter. *Mol. Gen. Genet.* **177:** 427-438.

Tunner, J. R., Robertson, C. R., Schippa, S. and Matin, A. (1992). Use of glucose starvation to limit growth and induce protein production in *Escherichia coli*. *Biotech. Bioeng.* **40:** 271-279.

Varma, A. and Palsson, B. O. (1993). Metabolic capabilities of *Escherichia coli*. II. Optimal growth patterns. *J. Theor. Biol.* **165:** 503-522.

Vind, J., Sorensen, M. A., Rasmussen, M. D. and Pedersen, S. (1993). Synthesis of proteins in *Escherichia coli* is limited by the concentration of free ribosomes. *J. Mol. Biol.* **231:** 678-685.

Webb, M. (1993). Monoclonal antibodies. In *Molecular Biology and Biotechnology*, pp. 357-386. (Editors Walker, J. M. and Gingold, E. B.). Royal Society of Chemistry, Cambridge.

White, B. A. and Bancroft, F. C. (1982). Cytoplasmic dot hybridization. *J. Biol. Chem.* **257:** 8569-8572.

Williams, S. G. and Manning, P. (1991). Transcription of the *Vibrio Cholera* haemolysin gene, *hlyA*, and cloning of a positive regulatory locus, *hlyU*. *Mol. Micro.* **5:** 2031-2038.

Wood, T. K. and Perretti, S. (1990). Expression of protein synthetic capacity de to cloned-gene expression in *E. coli*. *Biotech. Bioeng.* **36:** 865-878.

Yanisch-Perron, C., Vieira, J. and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33:** 103-119.

Yee, L. and Blanch, H. W. (1993). Defined media optimisation for growth of recombinant *Escherichia coli* X90. *Biotech. Bioeng.* **41:** 221-230.