MOLECULAR EXPRESSION SYSTEM DESIGN: THEORETICAL AND EXPERIMENTAL CHARACTERIZATION OF A NOVEL CROSS-REGULATION SYSTEM AND ITS APPLICATION AS A METABOLIC SWITCH

Thesis by

Wilfred Chen

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

Department of Chemical Engineering

California Institute of Technology

Pasadena, California U.S.A.

1993

(Submitted September 8, 1992)

ACKNOWLEDGEMENTS

I would like to express my gratitude and appreciation to my advisor, Dr. Jay Bailey, for introducing me into a new exciting area of biotechnology, and for his constant encouragement and financial support during my time at Caltech.

It has been a pleasure to work in the Bailey group in all these years and to meet friends from different parts of the world. I would like to especially thank Juan Diaz-Ricci for introducing me around the lab and Pauli for showing me the art of molecular biology. He has been both a great teacher and friend. Much thanks to Pete, Girish, Phil, Pauli and Wolfgang, which made up the morning coffee break group, all of your friendships have been the most enjoyable.

Lastly, special thanks to my family for their patience, understanding and support which made it possible for me to achieve my goals and aspirations.

ABSTRACT

A novel cross-regulation system was described in this work for regulated recombinant proteins expression. Molecular-level mathematical models of the *lac* and λP_R promoter-repressor systems were used in example calculations and the effectiveness of the cross-regulation system was evaluated by comparing with two other repressor control configurations (constitutive repressor synthesis and autogenous regulation). Simulation results suggested that this system offered the best control of transcription over a broad range of copy number in the uninduced state and also provided the highest overall transcription rate in the induced state.

The validity of the cross-regulation system in an actual experimental setting was also examined. Because of their desired properties, the *tac-lacI* and λP_L-*cI* promoter-repressor systems were used to construct vectors for regulated expression of chloramphenicol acetyltransferase (CAT). The polymerase chain reaction (PCR) technique was employed to generate the *lacI* and *cI* structural genes. Functional assays were performed to ensure active products from these PCR fragments. Induction results matched well with model prediction indicated that CAT expression from the cross-regulation system is two fold higher than the control (constitutive repressor synthesis configuration) with a very similar low basal expression. By using different copy number plasmids, the cross-regulation system has been shown to be equally applicable over copy numbers from 50-150.

Various factors influencing the recombinant protein yield from the cross-regulation system in *E. coli* were studied. It was determined that the optimum yield can be obtained by induction at 2-3 hr into the exponential growth and by using an IPTG concentration exceeding 0.5 mM. A limitation at the transcription level was determined to

be the most crucial factor for CAT expression in a batch fermentation mode. A cease in the CAT production was coincided with a 10 fold decrease in the CAT mRNA level after transition into the stationary phase. This bottleneck, however, can be eliminated by extending cell growth either by using a fed-batch fermentation mode or by using an unmutagenized *E. coli* strain.

In view of the interesting transcriptional property of the cross-regulation system, it was applied as a metabolic switch to provide a novel mean for the redirection of metabolic flux. The validity of the metabolic switch was illustrated by an alternation in the *Vitreoscilla* hemoglobin (VHb) and CAT expression patterns before and after induction. A practical application to alternate glycogen synthesis and degradation was examined and results indicated a five fold increase in glycogen synthesis and a 40% increase in glycogen degradation.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Regulation of cloned-gene expression

The remarkable advent of recombinant DNA technology in the past decade has opened up new opportunities to genetically modify microorganisms in order to biologically produce products of interest. Some of the most important examples included the production of therapeutic proteins,[4] antibiotics,[13] and organic solvents and acids.[3] All of these processes involve the insertion of DNA fragments that encoded for some specific proteins. Some of these proteins are themselves the products of interest, whereas some are expressed to facilitate the synthesis of the desired products. In the former case, the overall objective is to maximize the production of the desired proteins. However, in the latter case, our goal is to control the expression of the different proteins in order to achieve the optimum product yield.

Expression of a cloned gene can be constitutive or regulated depending upon the promoter employed. Examples of promoters which can drive constitutive gene expression are the β-lactamase promoter of *E. coli*[12] and the glyceraldehyde phosphate dehydrogenase[10] and the enolase[18] promoters of *S. cerevisiae*. Promoters which can be utilized for the regulated expression of cloned gene in *E. coli* include the *lac*[17], *trp*[16], *tac*[2], λP_L[11], and T7[15] phage promoters.

Each of these expression systems has its strong points and its weak points. When viewed form a fermentation process standpoint, constitutive systems are simple. No chemical addition or temperature shift is needed to induce expression. However, the constitutive production of a recombinant protein can interfere with cell growth, resulting in a decreased specific growth rate or even cell death. The problems associated with constitutive expression can be reduced by utilizing a regulated expression system. The fermentation strategy employed is to grow cells to a high cell density without any cloned-

gene expression. When a high cell density is achieved, high expression of the cloned gene is obtained by the addition of an inducer, the depletion of an inhibitor or corepressor, or by a shift in temperature, depending upon the nature of the promoter.

1.2 Motivation for this work

Recombinant protein production is usually governed by several factors including transcription, mRNA stability, translation, and protein degradation. In order to maximize production, one must carefully consider the compromise between recombinant protein synthesis and maintaining host cell growth. In fact, the most critical factor in determining the yield of recombinant proteins is the ability to regulate cloned-gene expression. Most of the currently available systems provide regulation at the transcription level and the transcription activity usually depends upon interactions with a specific repressor protein. The only exception is in the case of the T7 phage promoters in which the transcription activity is influenced by the presence of the T7 RNA polymerase. However, each of these systems has their own limitations. Some of these limitations include: 1) A loss of transcriptional control before induction/derepression beyond a certain threshold gene copy number. In this situation, multiple copies of the promoter titrate out the host repressor protein as in the case of the *lac* and *trp* promoter systems.[5,14] 2) A poor expression even under fully induced/derepressed conditions. In this case by including extra copies of the gene encoding for the repressor protein, a better control of basal expression is obtained in return for a tradeoff in the maximal induced expression level.[14] 3) For the temperature induced system, an increase in the culture temperature is usually accompanied by an increase in protein degradation.[1]

Modern recombinant DNA and genetic techniques have enabled us to explore different genetic strategies in order to address these limitations. All of the current

approaches usually focused on addressing only one aspect of the problems but failed to provide an universal solution for all these limitations. Such an expression system should has the ability to maintain a low level of basal expression but at the same time provides a high level of maximal expression. Often the design of such a new expression system is facilitated by the use of mathematical models. Many models for different promoter systems have been developed in the past years.[6-9] Models can allow us to determine the overall performance of a new system easily which is often difficult to approach using experimental means alone. Use of models to predict results which might not otherwise have been established without tedious experimentation and the experimental confirmation of these predictions is a great challenge.

1.3 Scope of Thesis

The aim of this thesis is to characterize a novel cross-regulation system for regulated cloned-gene expression and its application as a metabolic switch for the redirection of metabolic flux.

Chapter 2 describes the discovery of this novel system. The properties of this system are compared with three other modes of regulation by using molecular-level mathematical models of the *lac* and λP_R promoter-repressor systems as example. The transcription efficiency before and after induction for all these different systems are examined.

Experimental verification and characterization of the cross-regulation system are described in Chapter 3. A particular experimental realization of this system is constructed using the *tac-lacI* and λP_L-*cI* promoter-repressor systems. Using chloramphenicol acetyltransferase (CAT) as a model protein, the experimental characteristics of the cross-regulation system are compared to model predictions. Various different properties of this

In Chapter 4, factors influencing the recombinant protein yield using the cross-regulation system in *E. coli* are investigated. Examples such as the effect of induction and mRNA level on the recombinant protein production are examined. mRNA analyses are applied to obtain a better understanding of this system on the molecular level and to examine the rate-limiting factor for cloned-protein production.

Results from Chapter 2 suggest that the cross-regulation system can achieve a complete shift in the transcription steady state from one promoter to another upon induction. This behavior is subsequently demonstrated in Chapter 3 and 4. In Chapter 5, a novel way of redirecting metabolic flux is described by applying the cross-regulation system as a metabolic switch. Demonstration of this new concept is illustrated by a model system using *Vitreoscilla* hemoglobin (VHb) and CAT as the markers. A practical application of the metabolic switch in manipulating glycogen synthesis and degradation is described.

In another project, we have functionally expressed *Vitreoscilla* hemoglobin (VHb) in *Saccharomyces cerevisiae*. The effect of intracellular expression of VHb on yeast aerobic metabolism is investigated in Appendix A.

1.4 References

1. **Baker, S.A., Grossman, A.D. & Gross, C.A.** (1984). A gene regulating the heat shock response in *Escherichia coli* affects proteolysis. *Proc. Nat. Acad. Sci. U.S.A.* **81**, 6779-6783.
2. **de Boer, H.A., Comstock, L.J. & Vasser, M.** (1983). The *tac* promoter: a functional hybrid derived form the *trp* and *lac* promoters. *Proc. Natl. Acad. Sci. USA*. **80**, 21-25.
3. **Ingram, L.O., Conway, T., Clark, D.P., Sewell, G.W. & Preston, J.F.** (1987). Genetic engineering of ethanol production in *Escherichia coli*. *Appl. Environ. Microbiol.* **53**, 2420-2431.
4. **Johnson, I.S.** (1983). Human insulin from recombinant DNA technology. *Science*, **219**, 632-637.
5. **Latta, M., Philit, M., Maury, I., Soubrier, F., Denefle, P. & Mayaux, J-F.** (1990). Tryptophan promoter derivatives on multicopy plasmids: A comparative analysis of expression potentials in *Escherichia coli*. *DNA and Cell Biology*, **9**, 129-137.
6. **Lee, S.B. & Bailey, J.E.** (1984). Genetically structured models for *lac* promoter-operator function in the *Escherichia coli* chromosome and in multicopy plasmids: *lac* operator function. *Biotechnol. Bioeng.* **26**, 1372-1382.
7. **Lee, S.B. & Bailey, J.E.** (1984). Genetically structured models for *lac* promoter-operator function in the *Escherichia coli* chromosome and in multicopy plasmids: *lac* promoter function. *Biotechnol. Bioeng.* **26**, 1383-1389.
8. **Lee, S.B. & Bailey, J.E.** (1984). A mathematical model for *λdv* plasmid replication: analysis of wild-type plasmid. *Plasmid*, **11**, 151-165.

9. **Lee, S.B. & Bailey, J.E.** (1984). A mathematical model for *λdv* plasmid replication: analysis of copy number mutants. *Plasmid*, **11**, 166-177.

10. **McAlleer, W.J., Buniack, E., Maighette, R.Z., Wample, E. Miller, W.J. & Hilleman, M.P.** (1984). *Nature*, **307**, 178-180.

11. **Remaut, E., Stanssens, P. & Fiers, W.** (1981). Plasmid vectors for high-efficiency expression controlled by the P_L promoter of coliphage lambda. *Gene*, **15**, 81-89.

12. **Russell, D.R. & Bennett, G.N.** (1981). Characterization of the β-lactamase promoter of pBR322. *Nucleic Acids Res.* **9**, 2517-2533.

13. **Stanzak. R. Matsushima, P. Baltz, D.H. & Roa, R.N.** (1986). Cloning and expression in *Streptomyces lividans* of clustered erythromycin biosynthesis genes from *Streptomyces erythreus*. *Bio/Technology*, **4**, 229-232.

14. **Stark, M.J.R.** (1987). Multicopy expression vectors carrying the *lac* repressor gene for regulated high level expression of genes in *Escherichia coli*. *Gene*, **51**, 255-267.

15. **Tabor, S. & 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.

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

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

18. **Zabriskiem D.W. & Arcuri, E.J.** (1986). Factors influencing productivity of fermentations employing recombinant microorganisms. *Enzyme Micro. Technol.* **8**, 706-717.

CHAPTER 2

MOLECULAR DESIGN OF EXPRESSION SYSTEMS: COMPARISON OF DIFFERENT REPRESSOR CONTROL CONFIGURATIONS USING MOLECULAR MECHANISM MODELS

Source: W. Chen, S.B Lee and J.E. Bailey, *Biotechnol. Bioeng.*, **38**, 679-687, 1991

2.1 Abstract

Molecular-level mathematical models have been used to evaluate the effectiveness of eight different configurations of repressor synthesis control on the regulation of cloned product gene expression initiated from a promoter-operator sequence. Both single and dual-repressor situations were considered, employing genetically structured models for the *lac* and λP_R promoter-operator in example calculations. Simulation results suggest that the most effective mode of cloned gene expression control is a cross-regulation configuration carried on a multicopy plasmid. This system was able to control cloned product gene transcription in the uninduced state over a broad range of plasmid copy number and also provided the highest overall transcription rate in the induced state. The general strategies suggested by these simulations should be applicable for other repressor-operator-promoter systems in diverse hosts.

2.2 Introduction

Maximizing the amount of protein produced from a cloned gene in a recombinant organism requires careful consideration of the tradeoffs involved between cloned-gene expression and host cell growth and biosynthetic activity. High levels of cloned gene expression have been shown to reduce host cell growth rate and, concomitantly, overall protein synthesis activity.[8,23] This occurs because of competition between plasmid-directed and host cell-directed activity for common pools of precursors, chemical energy and electrons, activator species, repressor molecules, transport apparatus, and enzymes and other catalytic assembles. Under some cases, high expression levels can even cause cell death. Moreover, it has been observed that induction or derepression of cloned gene expression accentuates this inhibitory effect.[24] Therefore, it is very important to regulate the expression of cloned genes in the manufacture of valuable protein using recombinant cells.

In order to maximize product synthesis, cells are typically grown to high densities with minimal cloned gene expression followed by a production phase in which high expression results from induction or derepression of cloned gene transcription.[24,17] Several different promoter-operator systems, including *lac*, *trp*, and λP_L, have been cloned and utilized to regulate transcription of cloned genes.[4,25] In each of these systems, the transcription activity of the promoter-operator depends upon interaction with a corresponding specific repressor protein. This interaction may be influenced by adjusting the temperature (ts regulatory mutant) or the composition (inducer or corepressor concentration) of the growth medium in order to switch from low to high promoter activity.

Limitations of many current regulated promoter-operator systems are loss of control of cloned gene expression if the number of cloned gene copies per cell exceeds

some threshold or poor expression under fully activated conditions. In order to explore genetic strategies for addressing these limitations, it is useful to consider the different types of repressor expression configurations which can be established using recombinant DNA and modern genetic techniques.

Three different modes of repressor synthesis can be identified: (1) Constitutive synthesis, which proceeds at a constant level independent of the repressor content and the repressor configuration; (2) autogenous regulation, in which the synthesis of repressor is influenced by the content of repressor in the cell; and (3) cross regulation, in which the synthesis of the repressor is controlled by the presence of a second repressor molecule for a different promoter. With current methods for genetic manipulation, two different types of locations for repressor genes are possible. Repressor can be provided by repressor gene(s) encoded in the chromosome, or the repressor may be formed by the expression of a repressor gene included in an extrachromosomal element. Combining these possibilities, eight different combinations of the repressor gene locus and regulation of the repressor gene transcription can be identified. These are illustrated schematically in Figure 1.

In this work, molecular level mathematical models are used to evaluate the effectiveness of each of these different configurations. The preferred system is the one which, for a range of vector copy numbers, retains control of the promoter under the uninduced condition and provides the highest transcription level under the induced state. In general, any promoter-repressor system can be used for these different configurations. However, detailed models for the *lac* promoter-operator system[12,13] and the λP_R promoter-autorepressor system[14,15] which have been formulated and tested extensively will be used here to evaluate different modes of providing and regulating repressor in the host cell. These models allow simulation of various

situations in which the host cell contains many copies of the promoter-operator system and multiple copies of the repressor gene. This study provides an illustration of the application of molecular mechanism models to compare alternative molecular designs of recombinant DNA molecules and host genotype. In this context, such models serve as a guide for the genetic design of the host-vector expression system.

2.3 Model Descriptions

As described previously,[12-15] repressor synthesis can be represented by the following cellular material balance equations on the repressor message and protein, respectively:

$$\frac{d[mRNA]_R}{dt} = k_p^o \eta_R [G]_R - k_d [mRNA]_R - \mu [mRNA]_R, \tag{1}$$

$$\frac{d[R]_o}{dt} = k_q^o \xi_R [mRNA]_R - k_e [R]_o - \mu [R]_o. \tag{2}$$

Similarly, material balances on the cloned gene message and protein can be written as follows:

$$\frac{d[mRNA]_P}{dt} = k_p^o \eta_P [G]_P - k_d [mRNA]_P - \mu [mRNA]_P, \tag{3}$$

$$\frac{d[P]}{dt} = k_q^o \xi_P [mRNA]_P - k_e [P] - \mu [P], \tag{4}$$

where the transcription rate constant (k_p^o) and the translation rate constant (k_q^o) are the average rates per template and, as such, incoprorate the RNA polymerase elongation rate and the polypeptide chain elongation rate, respectively.[16] Substantial deviations from the cell-average value of the frequency of transcription initiation as

well as the regulatory effect on transcription have been incorporated into the transcription efficiency term (η). An analogous definition also applies for the translation efficiency (ξ).

In equations (1) through (4), subscripts R and P denote repressor and product protein, respectively. To minimize the number of different parameters in these calculations, it has been assumed that the transcription rate constant (k_p^o), the translation rate constant (k_q^o), the mRNA decay rate constant (k_d) and the protein decay rate constant (k_e) are the same for the repressor and the product. The translation efficiencies for the repressor and the product, ξ_R and ξ_P, respectively, could normally be influenced by the ribosome binding site nucleotide sequence (Shine-Delgarno sequence), and the distance between the initiation codon (ATG or GTG) and the S-D sequence. These two parameters may be different but are here assumed to be constant unaffected by the plasmid copy number or the design of repressor synthesis regulation. Clearly, these two parameters and other could be adjusted by the corresponding changes in nucleotide sequence to the appropriate values for a particular system in which significant differences occur between the repressor and the product kinetic properties.

The major attention in this work is focused primary upon the product gene transcription efficiency η_P. The overall transcription efficiency of a promoter is influenced by the interaction of the promoter-operator region of the DNA with regulatory proteins such as repressor, activator, and RNA polymerase. In this paper, the expression of the product protein will be assumed to be controlled by the *lac* promoter-operator system. This particular promoter-operator is considered as an example here because detailed genetically structured models for the *lac* operon have been formulated previously.[12,13] The modeling approach presented here could

be readily applied to other promoter-operator systems. The objective is to elucidate central qualitative features of different repressor expression configurations.

The transcription efficiency for the complete wild-type *lac* operon is determined by the interaction of the *lac* promoter-operator regions with three regulatory molecules[3,18]: *lac* repressor protein, cyclic AMP receptor protein (CRP) complexed with cyclic AMP (cAMP), and RNA polymerase (RNP) activated by sigma factor. However, for the L8UV5 *lac* promoter-operator and the *lac* portable promoter which have been widely used for the regulation of cloned gene expression, the CRP-cAMP complex does not affect the overall transcription efficiency.[2] Such a situation is assumed here. Since the major objective of this study is to determine the influence of different types of repressor synthesis on cloned gene expression, it is assumed that the interaction between RNA polymerase and the *lac* promoter are unaffected by any of the different repressor synthesis design discussed here. Furthermore, according to previous results,[13] it has been estimated that RNP binding activity varies only slightly with copy number up to 100. Thus, effects of copy number on RNP binding are neglected in these calculations.

First, product transcription efficiency (η_P) will be considered. Assuming that only the *lac* repressor protein is involved in the regulation of cloned gene transcription, η_P can be expressed as

$$\eta_P \approx \Phi_P = (1 - \Psi_P), \tag{5}$$

where Ψ_p is the binding probability of the repressor molecules to the operator sequence. As shown previously,[12] the binding probability of the repressor to the operator is given by the equation

$$\Psi_P = \frac{1}{2}\left[\left(1 + \frac{[R]_o}{[O]_o} + \frac{1}{\alpha[O]_o}\right) - \left\{\left(1 + \frac{[R]_o}{[O]_o} + \frac{1}{\alpha[O]_o}\right)^2 - \frac{4[R]_o}{[O]_o}\right\}^{\frac{1}{2}}\right], \tag{6}$$

where

$$\alpha = \frac{K_{A1} + K_{C1}K_{D1}[I]}{1 + K_{B1}[D]_o + K_{C1}[I] + K_{C1}K_{E1}[D]_o[I]}. \tag{7}$$

$[R]_o$, $[O]_o$, $[D]_o$, and $[I]$ are the total repressor concentration, the operator concentration, the concentration of nonspecific binding sites for the repressor, and the inducer concentration, respectively.

The transcription efficiency for repressor synthesis (η_R) is dependent upon the cases considered (see Figure 1). When the repressor transcription is constitutive as in Cases I and II, η_R is not influenced by the repressor level and thus is assumed to be constant. η_R for a constitutive promoter deviates from unity depending on the extent to which its promoter strength (frequency of initiation) deviates from the cell-average value. For autogenous repressor synthesis, the transcription of both the repressor and the product are controlled by the *lac* promoter-operator; therefore, η_R is equal to η_P in Cases III and IV.

For Cases V to VIII, the transcription of the *lac* repressor is controlled by a second controllable promoter. In this study, the λP_R promoter will be considered in particular, again because of a good prior modeling framework for this particular promoter.[14,15] Again, it should be noted that this approach to genetic design is not restricted to this particular choice of a second regulated promoter. Transcription of the λP_R promoter is influenced by interactions between RNA polymerase and the cro repressor. By assuming that the concentration of the operator is much lower than that of the repressor ($[CRO] \approx [CRO]_o$), the transcription efficiency for the λP_R promoter is derived as follows:[14]

$$\eta_\lambda = \frac{1 + K_3[CRO]_o}{1 + a[CRO]_o + b[CRO]_o^2 + c[CRO]_o^3}, \tag{8}$$

where

$$a = K_1 + K_2 + K_3, \tag{9}$$

$$b = K_1K_2 + K_2K_3 + K_3K_1, \tag{10}$$

$$c = K_1K_2K_3. \tag{11}$$

K_1, K_2, and K_3 represent the binding affinity of the cro repressor at the three different operator sites, respectively, and $[CRO]_o$ represents the total cro repressor concentration. However, in this study the concentration of the operator will frequently exceed the concentration of the repressor especially for those cases where the repressor gene is included in the plasmid. Therefore, the actual free repressor concentration must be considered in determining the transcription efficiency. As shown in detail in the Appendix, the free repressor concentration, $[CRO]$, can be determined as:

$$[CRO]_o = [CRO] + \frac{K_1[P_RO_R]_o[CRO]}{1 + K_1[CRO]} + \frac{K_2[P_RO_R]_0[CRO]}{1 + K_2[CRO]} + \frac{K_3[P_RO_R]_o[CRO]}{1 + K_3[CRO]}. \tag{12}$$

where $[P_RO_R]_o$ is the total concentration of the λP_R operator. Since this is a fourth order polynomial equation for $[CRO]$, four solutions will exist. However, in all cases, there is only one physically feasible solution (other solutions either complex or negative). As expected, the concentration of free repressor is very close to that of total repressor when $[CRO]_o$ is much greater than $[P_RO_R]_o$. On the other hand, $[CRO]$ deviates greatly from $[CRO]_o$ at the other end. With $[CRO]$ determined, the transcription efficiency can be obtained by replacing $[CRO]_o$ by $[CRO]$ in equation 8.

The gene concentration $[G]$ is given by

$$[G] = \frac{G}{V_cN_A}, \tag{13}$$

where, for the chromosome,[12]

$$G^b = exp\Big[\int_{t-C-D}^{t-fC} \mu(\tau)\, d\tau\Big], \tag{14}$$

and for a multicopy vector

$$G^p = G^b\overline{N}. \tag{15}$$

$\overline{N}$ in equation (15) is the average vector copy number.

Depending upon which of the cases shown in Figure 1 is considered, the definitions and equations used for calculations must be adjusted as summarized in Table 1.

2.4 Model Parameters

Most of the parameters appearing in the model equations can be determined from information in the literature. A summary of the parameter values used for these simulation are listed in Table II. The transcription rate constant and the translation rate constant have been estimated from the relationship between the growth rate and these rate constants established previously.[16] The doubling time of *E. coli* is assumed to be one hour. The parameters characterizing binding affinities for the *lac* repressor (K_{A1} to K_{E1}) and the cro repressor (K_1 to K_3) to their respective operator sites are assigned values corresponding to the wild type system. Concentrations of the bacterial chromosome ($[G]^b$), the operator ($[O]_o^b$), and the non specific binding sites ($[D]_o^b$) have been calculated by using equations 13 and 14 with C=40 min, D=20 min, and f=0.5.[12]

In the presence of cloning vectors, the total operator concentration ($[O]_o$), and the total non-specific binding site concentration ($[D]_o$) are given as

$$[O]_o = [O]_o^b + [O]_o^p = [O]_o^b(1 + \overline{N}), \tag{16}$$

$$[D]_o = [D]_o^b + [D]_o^p, \tag{17}$$

For the case in which the promoter-operator site is not present in the chromosome,

$$[O]_o = [O]_o^b \overline{N}. \tag{18}$$

For non-specific binding, it follows by definition that the probability of protein binding is indepentent of nucleotide sequence. Accordingly, the number of non-specific binding sites increases with the amount of DNA. Therefore, $[D]_o$ can be represented as

$$[D]_o = [D]_o^b(1 + c_1\overline{N}), \tag{19}$$

where c_1 is the ratio of non-specific binding site contents of the plasmid to those of the chromosome. The size of the cloning vector, which influences the parameter c_1, has been chosen to be 0.1% of the *E. coli* chromosome. This corresponds, for example, to a vector such as pBR322.[13]

2.5 Simulation Results

This work is concerned only with the steady-state solutions of the model described above. However, because these equations are highly nonlinear and not amenable to analytical solution, it is convenient to determine the steady -state conditions numerically by integrating the system of differential equations using

step-size Runge-Kutta algorithm with predictor-corrector until time-invariant values are approached asymptotically by all intracellular concentrations. This method was used to simulate a number of different scenarios as summarized next. Simulation results are numerically stable independence of initial conditions.

Transcription Efficiency under Uninduced Condition

With the current advance in genetic manipulation techniques, it is possible to modify the translation efficiency by altering the ribosome binding site nucleotide sequence. For example, a ribosome binding site nucleotide sequence (TTAAAAT $\overline{\text{TAAGGAGG}}$, barred portion represents the Shine-Delgarno sequence) with very high translation efficiency was synthesized and applied successfully to the production of cloned gene products up to 10-20% of the total intracellular protein.[7] In Figures 2 and 3, the cloned operator activity Φ_P, which is directly related to the cloned gene transcription efficiency (eq.5), is plotted as a function of the vector copy number $\overline{N}$ for two different translation efficiencies ($\xi_P = \xi_R$). Trends for all of the different repressor synthesis and regulation designs illustrated in Figure 1 are similar for both translation efficiencies.

For Case I, in which repressor is provided by constitutive expression of a chromosomal gene, regulation of the cloned promoter-operator is lost as the vector copy number increases. For Case II, in which repressor is obtained from constitutive expression of a repressor gene included in the expression plasmid, the transcription efficiency decreases as a function of the vector copy number.

Different relationships between plasmid copy number and the transcription efficiency is obtained for the autogenous regulation cases (III and IV). When the repressor modulates its own promoter-operator in a chromosomal gene as in Case III,

there is modest loss of control with increase in vector content of the cell. However, η_P increases only gradually over the whole range of copy number. For Case IV, in which the repressor gene with its autoregulated promoter-operator is included in the plasmid, the transcription efficiency is essentially independent of the vector copy number.

For Cases V to VIII, expression of the *lac* repressor is controlled by cross regulation with the λP_R promoter. However, for Case V and Case VI, the autorepressor (CRO) gene for the λP_R promoter is encoded with the *lac* repressor gene. In the final two cases, the autorepressor gene is encoded with the product gene. For Case V, the transcription efficiency remains under control of the repressor up to a critical copy number. Beyond that critical copy number, the repressor gradually loses control of transcription. Similar to the behavior of the previous case, Case VII also retains control of transcription up to a critical copy number. However, in this case the transcription efficiency changes rapidly from one extreme to another (i.e., changing from complete control to no control of transcription). Local stability analysis reveals that the low transcription efficiency state remains stable until the critical copy number where one of its eigenvalues changes sign (data not shown). Conversely, the opposite is true for the high transcription efficiency state. Finally, for Cases VI and VIII, the system is able to control product transcription over the whole range of copy number with η_P decreasing with increasing copy number.

Although Φ_P is independent of ξ_P for the first six cases, it should be noted that cases VII and VIII could indeed be dependent on ξ_P. Simulation results reveal that trends are similar for all ξ_P/ξ_R ratio less than 1. Only when this ratio increases to around 30 will trends be reversed. That is, instead of controlling expression for the product gene, the expression of the repressor gene is being controlled. This

situation can always be avoided, however, by appropriate genetic manipulation of, for example, the ribosome binding sites on product and/or repressor message.

Transcription Efficiency under Induced Condition

As mentioned in the introduction, the production of cloned gene products usually involves two stages. A growth phase in which cloned gene transcription is uninduced or repressed follows by a production phase in which cloned gene expression is induced or derepressed. Therefore, not only are we interested in understanding the transcription efficiency of cloned gene under uninduced conditions, but it is also important to know how the transcription efficiency behaves under induced conditions for all cases considered. Inducer (IPTG) concentration for the *lac* promoter employed in all calculations was 10^{-3}M.

In Figure 4, the overall transcription rate ($\overline{N}\eta_P$) is plotted as a function of copy number for all cases. Increasing copy number usually increases transcription because of increased cloned gene dosage. This trend holds for all cases considered except Case II in which the overall transcription rate tends to increase at low copy number, reaches a maximum, and then decreases. This implies that at high copy number, the repressor level is too high for the presence of inducer to completely derepress cloned gene expression. A similar phenomenon is observed for Case IV, where the overall transcription rate is always lower than for the other cases. In this case, the addition of inducer not only induces expression of the product, but also induces expression of the repressor. Therefore, high repressor concentration counteracts the derepression effect of the inducer and keeps the overall transcription rate low.

For Case VII and VIII, in which repressor is controlled by cross regulation, the overall transcription rate is the highest. The addition of inducer increases

expression not only of the product but also of the cro repressor. This represses expression of the *lac* repressor and further increases product expression. Case I and Case V have slightly lower transcription rates and are followed by Case VI. For all these configurations, the repression level is essentially independent of the inducer concentration. Therefore, the lower the repressor level, the higher the transcription rate (compare Case V and Case VI). An intermediate transcription rate is observed for Case III. As in Case IV, addition of inducer also induces repressor expression. Since the repressor gene is only present in a single copy in the chromosome, repressor level is not high enough to completely counteract the derepression effect of the inducer.

2.6 Discussion

Several promoter-operator sequences, such as those cloned from the *lac*, *trp*, and λP_L operons are widely used for regulation of cloned gene expression. The corresponding repressors are the lacI, trpR, and cI, respectively. Of these, only lacI gene expression is constitutive. Transcription of the trpR, and cI genes are regulated by an autorepression system.[9,21] Previous reports based upon experimental studies of the efficiency of cloned *lac* promoter-operator sequences and *trp* promoter-operators in maintaining control of cloned gene expression vary significantly. Using common cloning vector with copy numbers in the range of 20-30, a cloned *lac* promoter-operator provides transcription at nearly the induced level without any inducer presence in the medium in a wild-type *E. coli* host.[1,6,22] Use of host strains containing various lacI mutations giving overproduction of the *lac* repressor have been employed to provide retention of cloned product gene transcription repression in these system at modest copy number.[20] In the framework developed here,

this corresponds to the host-vector system designated Case I. This system has the disadvantage that control of cloned gene transcription will be eventually lost at some copy number. Furthermore, the level of repressor is constant independent of whether inducer is present or not. One solution for providing adequate repressor concentrations in order to maintain control of cloned gene transcription is inclusion of the repressor gene in the cloning vector as in Case II. Although this system retains control of transcription, the overexpression of repressor leads to a low transcription rate after the addition of inducer. Therefore, both systems have deficiencies as a regulatory system for cloned gene expression.

By contrast, previous experimental studies show that cloned *trp* promoter-operators provide effective control of cloned gene transcription for multicopy vectors in the autorepression configuration shown in Case III.[10,11] Here, when the trpR gene product is titrated as the number of vectors in the cell increases, the autorepression system provides increased transcription of the chromosomal repressor gene, adapting the host-vector system to a variety of vector-host relationships.

The simulation results obtained here agree qualitatively with these previous experimental reports, although the data are limited to date. Also, such comparison must be qualified by the comment that *trp* promoter-operator transcription regulation includes some features that are different from those found in the *lac* promoter-operator system on which the present model and calculations are based. However, it is reasonable to expect that the trends shown here apply equally well to the *trp* system.

Although for Case IV the system is able to retain control of transcription over the entire range of copy number (η_P independent of copy number), its overall

transcription rate is relatively low in the induced state. This is not a desired feature since a central goal is to maximize product protein production after induction of cloned gene expression.

Other alternatives considered here for controlling cloned gene transcription involve cross regulation of repressor expression. As of today, no experimental data are available to verify our simulation results. However, bearing in mind the qualitative success of the results from the first four cases as compared to the experiment, it is reasonable to expect that the actual system will follow the trends calculated for the cross regulation system. Although many different types of promoter-repressor can be used for cross regulation, the λP_R promoter-autorepressor system was chosen for these example calculations since its kinetic model is well tested. The simulation results of cases V and VI are not significantly different from those of cases I and II; however, they are included in order to complete consideration of all different possbilities. Moreover, such results may not be easy to anticipate without model simulation.

From our simulation results, the best expression control system are Case VII and Case VIII. The Case VIII configuration is not only able to retain control of cloned product gene transcription in the uninduced state, it also provides the highest overall transcription rate in the induced state. Case VII also has attractive control characteristics, with the drawback of abrupt loss of regulation at a critical copy number. Additional simulations (results not shown) have shown that increasing the number of repressor gene copies in the chromosome will shift the critical copy number to higher values. In particular, by including two tandem copies of the repressor gene on the chromosome in Case VII, this system is able to control cloned product gene transcription up to a copy number of 120 (data not shown),

while retaining the high transcription rate in the induced state. An advantage for this configuration is a substantial reduction in the amount of repressor molecules synthesized compared to Case VIII.

Although it is true that the cross-regulation configuration shows the highest efficiency in this study, we cannot exclude the possibility that a single-repressor system might be more effective than the dual-repressor system with different sets of parameter values (for example, changes in translation efficiency or transcription rate constant by appropriate genetic manipulations). Since the transcription efficiency depends on the level of repressor content, such changes will undoubtedly alter the transcription efficiency by changing the amount of repressor. For example, if the overall transcription rate (OTR) values for different configurations are normalized to those at $\overline{N} = 1$, then the OTR for Case III becomes as high as that for Case VII or Case VIII. However, considering the reasonable case in which significant interactions between transcription and translation do not occur, the *trend* for the induction ratio (i.e., η_P(Induced)/η_P(Uninduced)) as a function of copy number is independent of the values of parameters such as k_p^o, k_q^o, ξ_P, and ξ_R. Figure 5 shows a plot of the induction ratio verus copy number for all different cases. Case VIII shows the highest induction ratio amongst all cases with induction ratio increasing verus copy number. On the other hand, for Case III, the induction ratio approaches one as copy number increases. Therefore, allowing high level induction of transcription for such a system would result in a loss of transcription control under uninduced conditions.

This study illustrates how a kinetic model of recombinant systems formulated at the molecular level can be used to explore the qualitative and eventually some quantitative features provided by different genetic designs. Given the extremely

capabilities available now for precise design of the promoter-operator region at the nucleotide sequence level, it is important to utilize such quantitative and systematic modeling relationships to assess complex interactions in recombinant systems and to compare the effects of many alternative molecular designs which are in principle available. Current capabilities for molecular genetic manipulation provides so many options that empirical trial and error approaches are inadequate for optimization. Instead, quantitative, systematic models are necessary for rational and efficient expression system design.

The main objective of this work is to illustrate the major qualitative features of the alternative arrangements for repressor synthesis as summarized in Figure 1. However, in order to optimize the design of the product gene promoter-operator and the production and regulation of the associated repressor, one must take into account different demands on the host cell metabolic and biosynthetic capacity posed by the different repressor expression design and the effects of these demands on product gene expression level. These considerations must be merged with the economic criteria for the process and the available process environmental manipulations in order to evaluate and to optimize the host genotype and the molecular design of the plasmid. In this context, combining our current models with the kinetic models for product formation of recombinant culture such as those developed by Lee, Seressiotis, and Bailey[18] and Betenbaugh and Dhurjati[5] should be useful.

The success of the present study in simulating trends for alternative regulation configurations for repressor design is quite encouraging and suggests that this framework might be applicable for other promoters. However, these trends are not necessarily universal, and full analysis of any particular system (either experimental or theoretical with appropriate model parameters) is recommended.

2.7 Acknowledgement

This work was supported by the Catalysis and Biocatalysis Program of the Advanced Industrial Concepts Division of the U.S. Department of Energy. S.B.L was supported in part by the Korea Science and Engineering Foundation (KOSEF).

2.8 Nomenclature

C	Time required for chromosome replication
c_1	Number of nonspecific binding sites per plasmid/number of nonspecific binding sites per chromosome
$[CRO]$	cro repressor concentration
D	Time between completion of a round of replication and subsequent cell division
$[D]$	Concentration of nonspecific binding sites for repressor
G	Number of DNA molecules
f	A fractional distance of gene from the origin
$[G]$	DNA concentration
$[I]$	Inducer concentration
k_d	mRNA decay rate constant
k_e	Protein decay rate constant
k_p^o	Overall transcription rate constant
k_q^o	Overall translation rate constant
K_1, K_2, K_3	Binding affinities for cro repressor
$K_{A1}, K_{B1}, K_{C1}, K_{D1}, K_{E1}$	Binding affinities (see Ref.12)
$[mRNA]$	mRNA concentration
$\overline{N}$	Average plasmid copy number (number of plasmid molecules/number of chromosomes)
N_A	Avogadro's number

$[O]$	Operator concentration
$[P]$	Cloned gene product concentration
$[P_R O_R]$	λP_R operator concentration
$[R]$	Repressor concentration
$[R_2]$	Second repressor concentration
t	Time
V_c	Cell volume
Greek	
η	Overall transcription efficiency
ξ	Overall translation efficiency
μ	Specific growth rate
Ψ_P	Probability of repressor binding to its specific binding site
Φ_P	Fraction of operators without bound repressor
Subscript	
P	Cloned gene product
R	Repressor
o	Total concentration
Superscript	
b	Bacterial chromosome
p	plasmid

2.9 Appendix

Assuming that non-specific binding of the cro repressor is not important, the cro repressor can bind to each of the three operator sites on the λP_R promoter operator region ($P_R O_R$).

$$CRO + O_1 \rightleftharpoons O_1CRO; \qquad K_1 = [O_1CRO]/[CRO][O_1], \tag{A1}$$

$$CRO + O_2 \rightleftharpoons O_2CRO; \qquad K_2 = [O_2CRO]/[CRO][O_2], \tag{A2}$$

$$CRO + O_3 \rightleftharpoons O_3CRO; \qquad K_3 = [O_3CRO]/[CRO][O_3]. \tag{A3}$$

Total balance on cro repressor and each of the three operators:

$$[CRO]_o = [CRO] + [O_1CRO] + [O_2CRO] + [O_3CRO], \tag{A4}$$

$$[O_1]_o = [O_1] + [O_1CRO], \tag{A5}$$

$$[O_2]_o = [O_2] + [O_2CRO], \tag{A6}$$

$$[O_3]_o = [O_3] + [O_3CRO]. \tag{A7}$$

Combining equations A5-A7 and A1-A3 yield:

$$[O_iCRO] = \frac{K_i[O_i]_o[CRO]}{1 + K_i[CRO]}. \tag{A8}$$

The index i refers to operator sites 1, 2, or 3. Furthermore, the total concentration of each of the three operator sites is the same; therefore,

$$[O_1]_o = [O_2]_o = [O_3]_o = [P_R O_R]_o. \tag{A9}$$

Substituting eq.(A8) and (A9) into (A4) yields eq.(12).

2.10 References

1. Backman, K., Ptashne, M., Gilbert, W. 1976. Construction of plasmids carrying the cI gene of bacteriophage λ. Proc. Natl. Acad. Sci. USA **75**: 4174-4178.
2. Beckwith, J., Grodzicker, T., Arditti, P.R. 1972. Evidence for two sites in the *lac* promoter region. J. Mol. Biol. **69**: 155-160.
3. Beckwith, J.R., Ziper, D. (ed.) 1970. The Lactose Operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
4. Bernard, H.V., Helinski, D.R. 1980. Bacterial plasmid cloning vehicles. p. 133-168 In: J. K. Setlow and A. Hollander (ed.), Genetic Engineering: Principles and Methods, vol 2. Plenum, New York.
5. Betenbaugh, M.J., Dhurjati, P. 1990. A comparison of mathematical model predictions to experimental measurements for growth and recombinant protein production in induced cultures of *Escherichia coli.* Biotechnol. Bioeng. **36**: 124-134.
6. Clewell, D.B., Helinski, D.R. 1972. Effect of growth conditions on the formation of the relaxation complex of supercoiled ColE1 deoxyribonucleic acid and protein in *Escherichia coli.* J. Bacteriol. **110**: 1135-1146.
7. Crowel, R., Seamans, C., Lomedico, R., McAndrew, S. 1985. Versatile expression vectors for high-level synthesis of cloned gene products in *Escherichia coli.* Gene **38**: 31-38.
8. Guarante, L., Yocum, R., Gilford, F. 1982. GAL10-CYC1 hydrid yeast promoter identifies the GAL4 regulatory region as an upstream site. Proc. Natl. Acad. Sci. USA **79**: 7410-7414.
9. Gunsalus, R.P., Yanofsky, C. 1980. Nucleotide sequence and expression of *Escherichia coli trpR*, the structural gene for the trp aporepressor. Proc. Natl. Acad. Sci. USA **77**: 7117-7121.

10. Hallewell, R.A., Emtage, S. 1980. Plasmid vectors containing the tryptophan operon promoter suitable for efficienct regulation expression of foreign genes. Gene **9**: 27-47.

11. Hershfield, V., Boyer, H.W., Yanofsky, C., Lovett, M.A., Helinski, D.R. 1974. Plasmid ColE1 as a molecular vehicle for cloning and amplification of DNA. Proc. Natl. Acad. Sci. USA **71**: 3455-3459.

12. Lee, S.B., Bailey, J.E. 1984. Genetically structured models for *lac* promoter-operator function in the *Escherichia coli* chromosome and in multicopy plasmids: *lac* operator function. Biotechnol. Bioeng. **26**: 1372-1382.

13. Lee, S.B., Bailey, J.E. 1984. Genetically structured models for *lac* promoter-operator function in the chromosome and in multicopy plasmids: *lac* promoter function. Biotechnol. Bioeng. **26**: 1383-1389.

14. Lee, S.B., Bailey, J.E. 1984. A mathematical model for λdv plasmid replication: analysis of wild-type plasmid. Plasmid **11**: 151-165.

15. Lee, S.B., Bailey, J.E. 1984. A mathematical model for λdv plasmid replication: analysis of copy number mutants. Plasmid **11**: 166-177.

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

17. Lee, S.B., Ryu, D.D.Y., Siegel, R., Park, S.H. 1988. Performance of recombinant fermentation and evaluation of gene expression efficiency for gene product in a two-stage continuous culture system. Biotechnol. Bioeng. **31**: 805-820.

18. Lee, S.B., Seressiotis, A., Bailey, J.E. 1985. A kinetic model for product formation in nstable recombinant populations. Biotechnol. Bioeng. **27**: 1699-1709.

19. Miller, J.H., Reznikoff, W.S. (ed.) 1978. The Operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

20. O'Farrell, P.H., Polisky, B., Gelfand, D.H. 1978. Regulated expression by readthrough translation from a plasmid-encoded β-galactosidase. J. Bacteriol. **134**: 645-654.

21. Ptashne, M. 1971. Repressor and its action. p. 221-238 In: A.D. Hershey (ed.), The Bacteriophage Lambda. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

22. Sadler, J.R., Tecklenburg, M., Bete, J.L. 1980. Plasmids containing many tandem copies of a synthetic lactose operator. Gene **8**: 279-300.

23. Seo, J.H., Bailey, J.E. 1985. Effect of recombinant plasmid content on growth properties and cloned gene product formation in *Escherichia coli.* Biotechnol. Bioeng. **27**: 1668-1674.

24. Siegel, R., Ryu, D.D.Y. 1985. Kinetic study of instability of recombinant plasmid pPLc23trpA1 in *E. coli* using two-stage continuous culture system. Biotechnol. Bioeng. **27**: 28-33.

25. Thompson, R. 1982. Plasmid and phage M13 cloning vectors. p. 1-52 In: R. Williamson (ed.), Genetic Engineering 3. Academic Press, London.

2.11 Table Captions

Table I. A summary of model parameters and equations used for calculations for the different repressor expression designs as indicated in Fig. 1.

Table II. Summary of model parameter values.

Table I

Case	η_R	η_P	$[G]_R$	$[G]_P$	$[O]_o$	$[D]_o$
I	Constant	Φ_P	$[G]^b$	$[G]^b\overline{N}$	$[O]_o^b\overline{N}$	$[D]_o^b(1+c_1\overline{N})$
II	Constant	Φ_P	$[G]^b\overline{N}$	$[G]^b\overline{N}$	$[O]_o^b\overline{N}$	$[D]_o^b(1+c_1\overline{N})$
III	Φ_P	Φ_P	$[G]^b$	$[G]^b\overline{N}$	$[O]_o^b(1+\overline{N})$	$[D]_o^b(1+c_1\overline{N})$
IV	Φ_P	Φ_P	$[G]^b\overline{N}$	$[G]^b\overline{N}$	$[O]_o^b\overline{N}$	$[D]_o^b(1+c_1\overline{N})$
V	η_λ	Φ_P	$[G]^b$	$[G]^b\overline{N}$	$[O]_o^b\overline{N}$	$[D]_o^b(1+c_1\overline{N})$
VI	η_λ	Φ_P	$[G]^b\overline{N}$	$[G]^b\overline{N}$	$[O]_o^b\overline{N}$	$[D]_o^b(1+c_1\overline{N})$
VII	η_λ	Φ_P	$[G]^b$	$[G]^b\overline{N}$	$[O]_o^b\overline{N}$	$[D]_o^b(1+c_1\overline{N})$
VIII	η_λ	Φ_P	$[G]^b\overline{N}$	$[G]^b\overline{N}$	$[O]_o^b\overline{N}$	$[D]_o^b(1+c_1\overline{N})$

Table II

Parameter	Value	Ref/Source
k_p^o	5.0 M mRNA/M DNA-min	16
k_q^o	10.0 M protein/M mRNA-min	16
k_d	0.5 min^{-1}	16
k_e	0.01 min^{-1}	16
K_1	1.25×10^8 M^{-1}	14
K_2	1.25×10^8 M^{-1}	14
K_3	1.0×10^9 M^{-1}	14
K_{A1}	2×10^{12} M^{-1}	12
K_{B1}	1×10^3 M^{-1}	12
K_{C1}	1×10^7 M^{-1}	12
K_{D1}	2×10^9 M^{-1}	12
K_{E1}	1.5×10^4 M^{-1}	12
$[G]^b$	4×10^{-9} M	12
$[O]_o^b$	4×10^{-9} M	12
$[D]_o^b$	4×10^{-2} M	12
c_1	0.001	12

2.12 Figure Captions

Figure 1. Schematic illustration of the eight different configurations for the regulation of cloned gene transcription considered in this paper. These cases differ with respect to their modes of repressor synthesis and their source of repressor. For cases I, III, V, and VII, repressor gene is encoded in the chromosome, whereas, for cases II, IV, VI, and VIII, the repressor gene is included in the plasmid. For Case I and Case II, repressor synthesis is constitutive. For Case III and Case IV, the content of repressor influences its synthesis. Finally, for cases V to VIII, the content of a second repressor influences the repressor synthesis. However, for cases V and VI, the second repressor gene is fused to form an operon with the first repressor gene ($[mRNA]_{R2} = [mRNA]_R$). For the last two cases, the second repressor gene is fused to form an operon with the product gene ($[mRNA]_{R2} = [mRNA]_P$).

Figure 2. Model calculation of normalized uninduced transcription efficiency as a function of vector copy number for all the different cases with $\xi = 0.02$. For cases I and II, $\eta_R = 0.1$, in order to match the repressor concentration synthesized from the *lacZ* promoter as reported.[12]

Figure 3. Model calculation of normalized uninduced transcription efficiency as a function of vector copy number for all the different cases with $\xi = 0.1$. This is to illustrate that the same trend is observed for the transcription efficiency with different ξ values.

Figure 4. Model calculation of the normalized overall transcription rate as a function of vector copy number for all the different cases with $\xi = 0.02$. Inducer concentration $[I]$ =1×10^{-3} M for all cases.

Figure 5. Model calculation of induction ratio for all the different cases with $\xi = 0.02$ and parameter values listed in Table II.

Figure 1

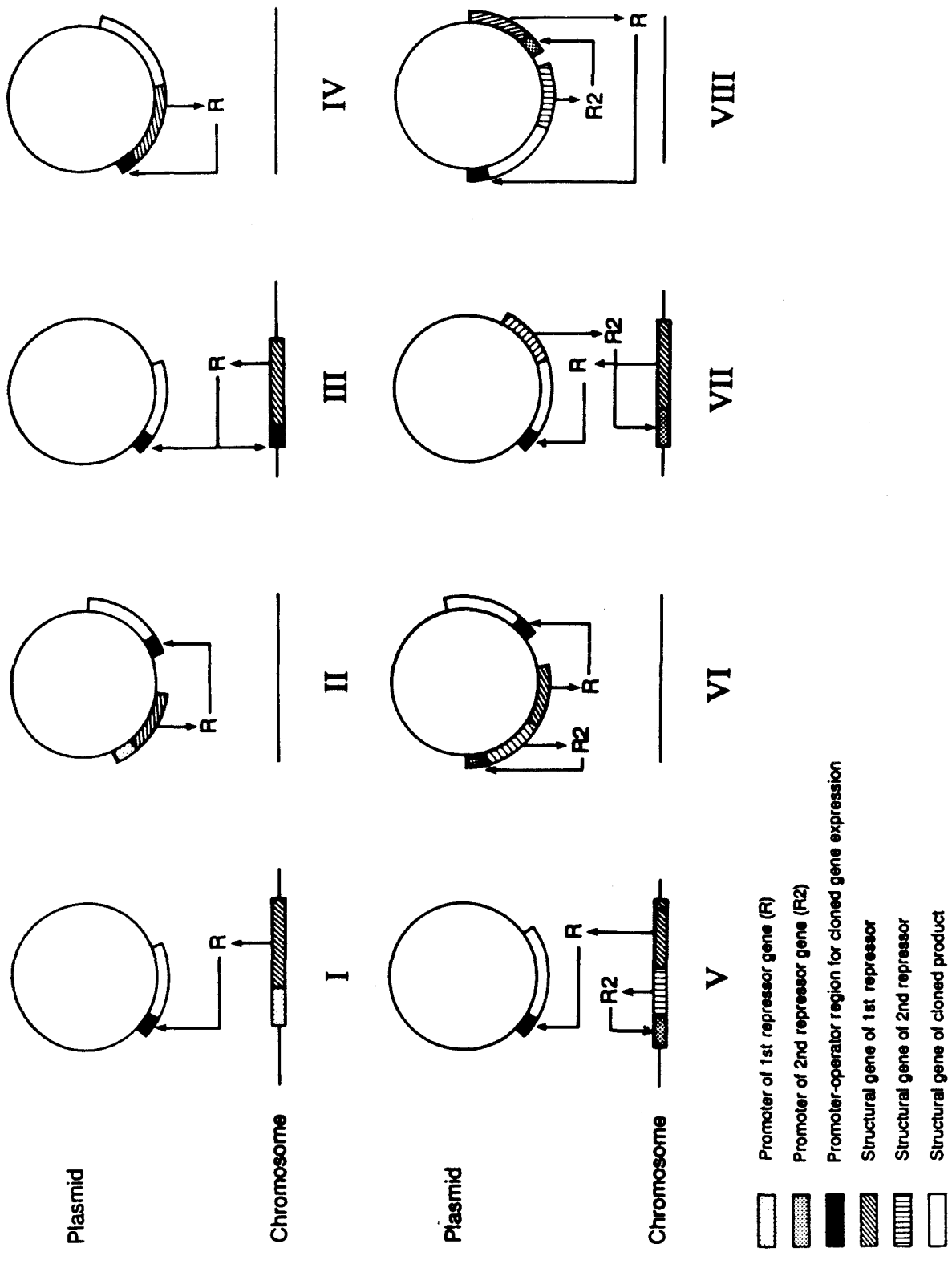

Plasmid
Chromosome
I
II
III
IV
V
VI
VII
VIII
R
R2
Promoter of 1st repressor gene (R)
Promoter of 2nd repressor gene (R2)
Promoter-operator region for cloned gene expression
Structural gene of 1st repressor
Structural gene of 2nd repressor
Structural gene of cloned product

Figure 2

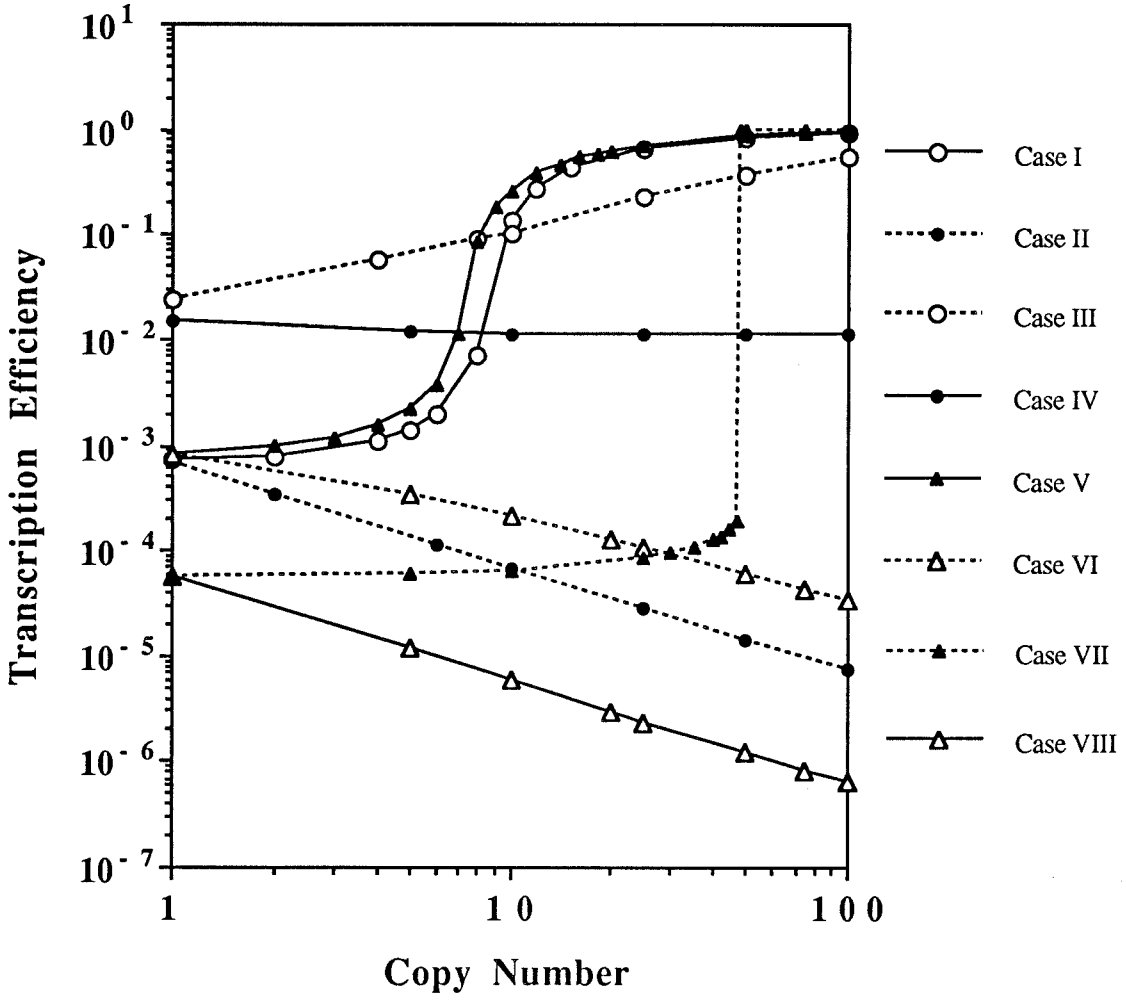

Transcription Efficiency
10 1
10 0
10 - 1
10 - 2
10 - 3
10 - 4
10 - 5
10 - 6
10 - 7
1
10
100
Copy Number
Case I
Case II
Case III
Case IV
Case V
Case VI
Case VII
Case VIII

Figure 3

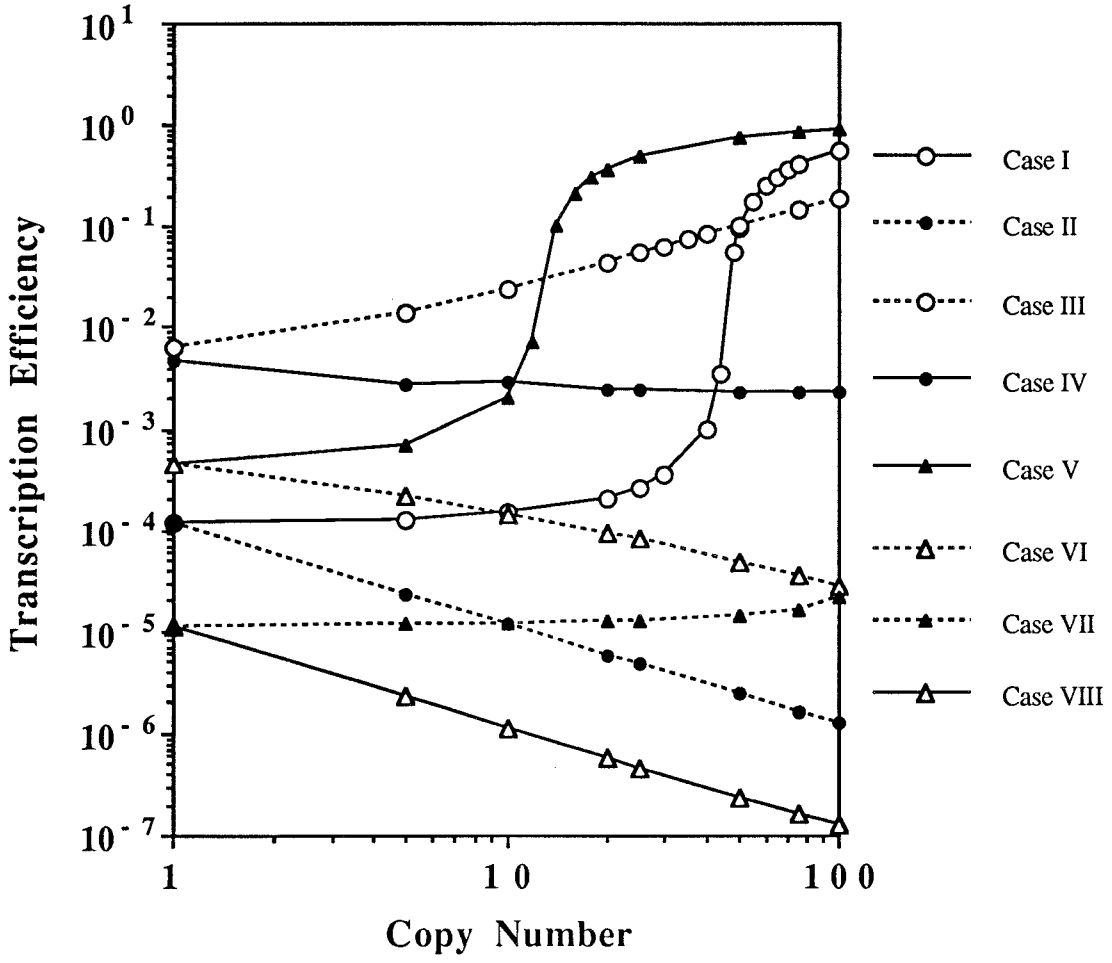

Transcription Efficiency
10 1
10 0
10 - 1
10 - 2
10 - 3
10 - 4
10 - 5
10 - 6
10 - 7
1
10
100
Copy Number
Case I
Case II
Case III
Case IV
Case V
Case VI
Case VII
Case VIII

Figure 4

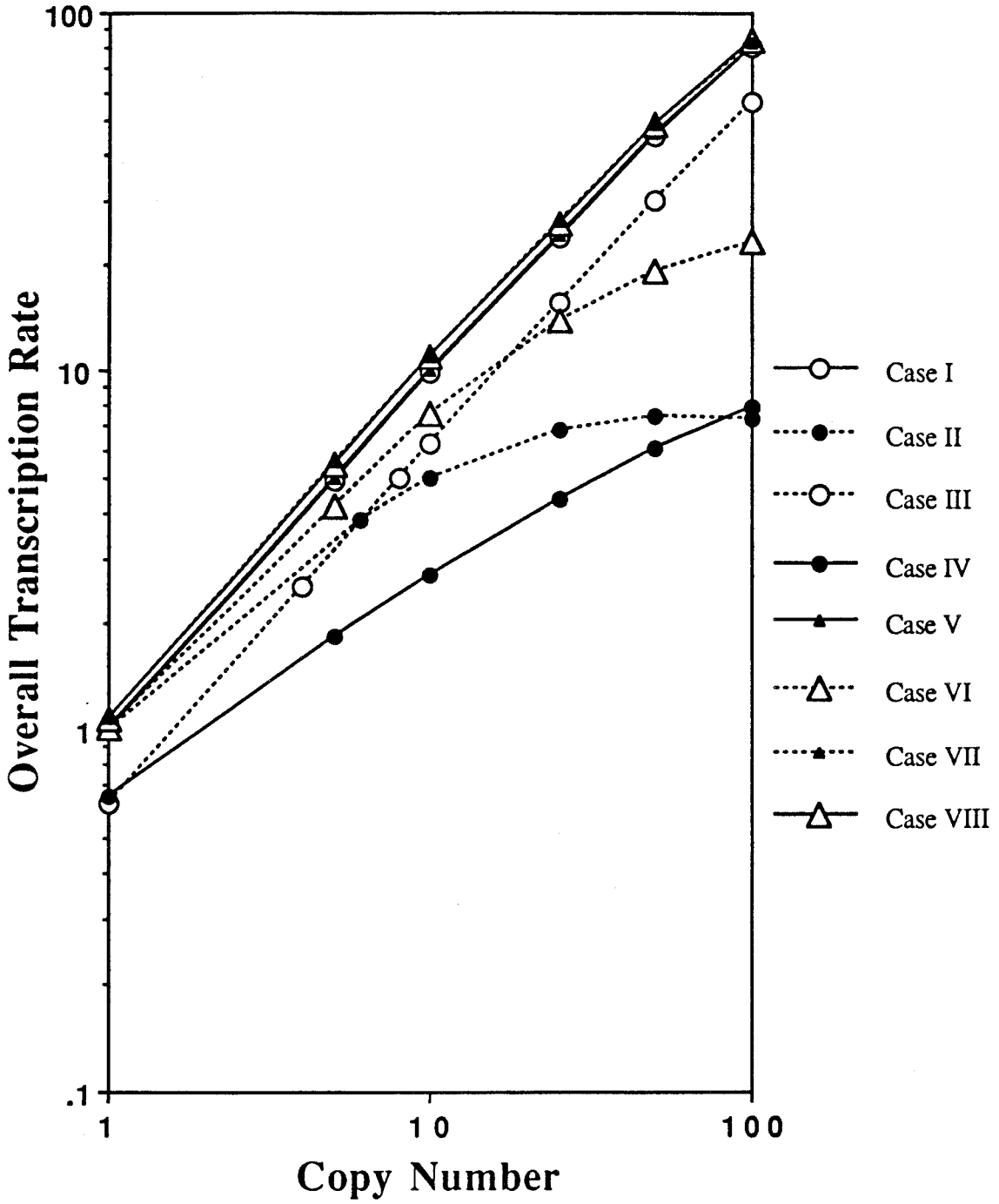
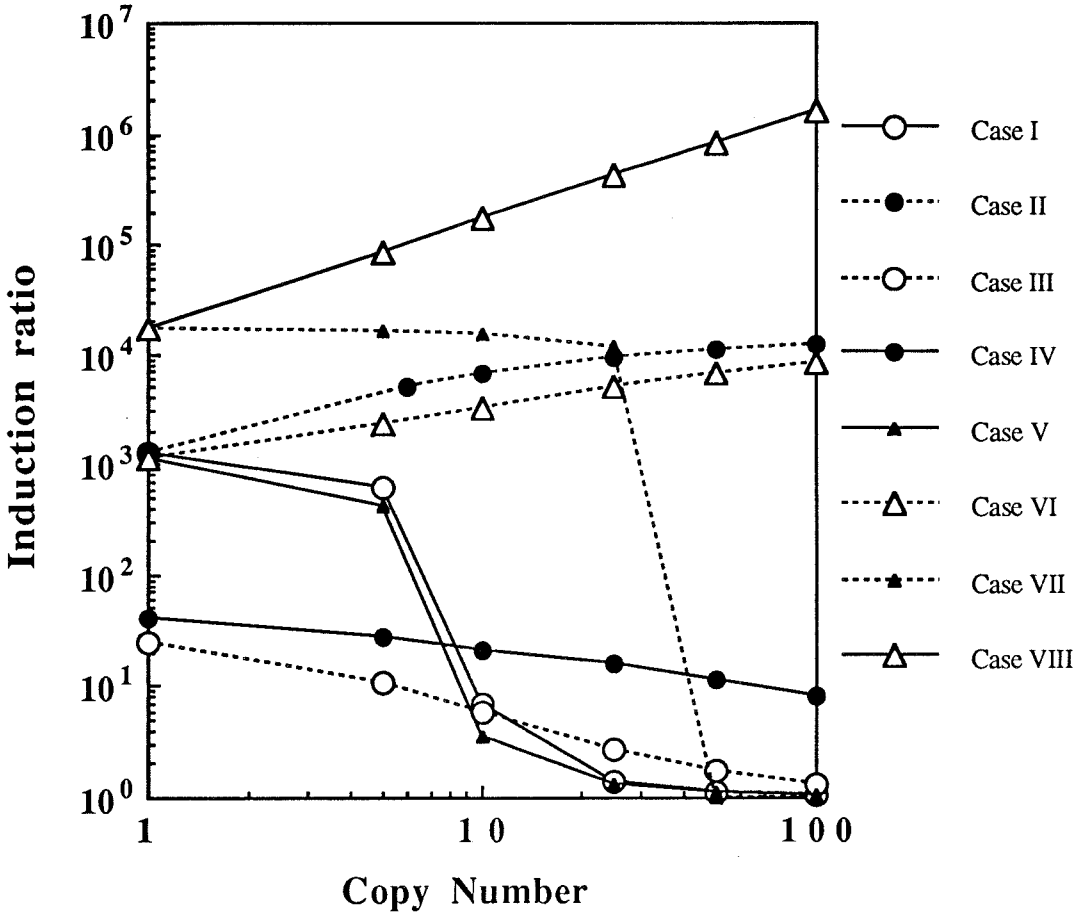

100
10
1
.1
Overall Transcription Rate
1
10
100
Copy Number
Case I
Case II
Case III
Case IV
Case V
Case VI
Case VII
Case VIII

Figure 5


Induction ratio
10^7
10^6
10^5
10^4
10^3
10^2
10^1
10^0
1
10
100
Copy Number
Case I
Case II
Case III
Case IV
Case V
Case VI
Case VII
Case VIII

CHAPTER 3

CONSTRUCTION AND CHARACTERIZATION OF A NOVEL CROSS-REGULATION SYSTEM FOR REGULATING CLONED GENE EXPRESSION IN *ESCHERICHIA COLI*

Source: W. Chen, P.T. Kallio and J.E. Bailey, *Gene.*, Submitted, 1992.

3.1 Abstract

A novel cross-regulation system employing a dual repressor control configuration was constructed using the *tac-lacI* and λP_L-*cI* promoter-repressor systems. Expression of a reporter protein, chloramphenicol acetyltransferase (CAT), using the cross-regulation system was compared to a commonly used constitutive repressor synthesis system. Induction results using different copy number vectors indicate that the CAT expression levels are at least two-fold higher using the cross-regulation system which has a very low basal expression. These results match well with previous mathematical modeling predictions indicating excellent control of basal expression and also higher cloned-gene expression post-induction over a broad range of copy numbers for a cross-regulation control configuration. Induction of the cross-regulation system both up-regulated the activation pathway and down-regulated the inhibition pathway shifting the system steady state away from *lac* repressor expression into CAT and *cI* repressor expression . The control strategy presented here should be equally applicable to regulate other promoter-repressor systems in diverse hosts.

3.2 Introduction

With recent advances in genetic techniques, it is now possible to produce various recombinant proteins in different organisms. Production of large quantities of these cloned proteins can be achieved by combining gene amplification and strong promoters. However, high expression of recombinant proteins has been shown to reduce host cell growth rate and, concomitantly, overall protein synthesis capability (Bentley et al., 1990; Park et al., 1990). It is likely that competition between chromosomal-directed, host metabolic activity and plasmid-directed metabolic activity is a general contributor to this detrimental effect (Peretti and Bailey, 1987). Good regulation of cloned gene expression is therefore essential for industrial production of recombinant proteins. Preferentially, production is delayed until the end of the growth phase when high expression of the cloned gene is achieved by induction or derepression (Park et al., 1991).

Cloned gene expression is governed by the rate of transcription, the stability of the mRNA, the rate of translation, and the stability of the protein . In particular, regulation of cloned gene transcription in *E. coli* has been thoroughly studied for different promoter-operator systems including *lac* (Makoff and Oxer, 1991), *trp* (Latta et al., 1990), *tac* (Yaffe et al., 1988), and λP_L (Mott et al., 1985). In each of these systems, transcription activity depends upon interactions with a specific repressor protein. For example, gene expression controlled by the λP_L promoter is usually regulated by the temperature sensitive *cI* repressor (ts-*cI*) (Remaut et al., 1983). Induction of cloned gene expression requires shifting the temperature from 28-30^{0}C to 42^{0}C (inactivates the ts-*cI* repressor). The main disadvantage of the temperature-induced system is the requirement for growth preinduction at a suboptimum growth temperature (28-30^{0}C), resulting in reduced growth rate and increased contamination risks. Furthermore, temperature shift to

42^{0}C activates the heat-shock response which results in significant changes in the protein composition of the cell, including elevated level of protease La which is implicated in the initial proteolytic attack on abnormally folded proteins.(Baker et al., 1984; Goff & Goldberg, 1986; Kosinski & Bailey, 1991).

For expression systems utilizing the *lac, tac,* and *trp* promoters on multicopy vectors, a high basal level of expression is often observed in a wild-type strain in which only a single copy of the corresponding repressor gene is included in the chromosome (Stark, 1987; Latta et al., 1990). In this situation, multiple copies of these operators titrate out the wild-type level of repressor. This is particularly undesirable if the cloned product is toxic to the host. In order to alleviate this problem, a single copy of the repressor gene is often included in the expression vector, providing sufficient repressor to prevent expression before induction or derepression. Unfortunately, due to a high level of repressor molecules produced, the maximal induced levels of expression are typically reduced in such constructs (Stark, 1987).

Recently, a new expression system design was proposed (Chen et al., 1991) based upon molecular-level mathematical modeling of the involved intracellular interactions. This novel cross-regulation system employed a dual repressor control configuration to regulate cloned-gene expression. In this arrangement, the repressor gene for the second promoter and the product gene are fused together with the first promoter to form an operon, while the expression of the first repressor gene is controlled by the second promoter. Product gene transcription is induced by addition of inducer which inactivates the first repressor. A schematic of the system is shown in Figure 1. Simulation results using *lac* (first promoter) and λP_R (second promoter) promoter-repressor systems as examples indicated that this system retains good control of cloned-gene transcription before induction and also provides the highest transcription level post-induction (Chen et

al., 1991). In this situation, it was postulated that the addition of inducer not only increases transcription of the product but also that of the second repressor. This represses expression of the first repressor and further increases product expression. Thus, the addition of inducer increases transcription of the product gene in two ways: (i) provides a positive effect with the formation of a repressor-inducer complex which is not able to bind to the operator site of the first promoter and, (ii) subsequently decreases a negative effect by reducing the synthesis of the first repressor.

This paper describes the construction and characterization of a particular experimental realization of the cross-regulation system in *E. coli*. Because of their useful properties and the availability of the required genetic elements, the *tac-lacI* and λP_L-*cI* promoter-repressor systems were chosen. Chloramphenicol acetyltransferase (CAT) was used as a model cloned protein because of its stability in *E. coli* and existence of a sensitive assay for its activity. In particular, the transcription of the *cat* gene and the fused *cI* gene is regulated by the *tac* promoter, while the transcription of the *lacI* gene is controlled by the λP_L promoter (Figure 4). In order to compare this new system with a currently employed transcription control configuration, a control expression system utilizing the constitutive repressor synthesis configuration was also constructed. In this construct, the *cat* gene is under the control of the *tac* promoter, and the *lac* repressor is provided from a single copy of the $lacI^q$ allele included in the plasmid. Applicability of the cross-regulation system over a broad range of plasmid copy numbers is illustrated by transferring the expression cassette into a series of closely related copy number mutant plasmids.

3.3 Materials and Methods

Bacterial strains and plasmids

Escherichia coli strains DH5α (F^-, *endA1*, *hsdR17*($r_k^-m_k^+$), *supE44*, *thi-1*, λ^-, *recA1*, *gryA96*, *relA1*, *ϕ80dlacAm15*) (Bethesda Research Laboratories) and HB101 (*supE44*, *hsdS20*($r_B^-m_B^-$), *recA13*, *ara-14*, *proA2*, *lacY1*, *galK2*, *rpsL20*, *xyl-5*, *mtl-1*) (Bolivar & Backman, 1979) were used in all cloning experiments. Strains CY15050 (W3110*tnaAZ,DlacΔ169*(lTLF)) (Kelly & Yanofsky, 1982) and CGSC808 (*lacI22*, λ^-, *relA1*, *spoT1*, *thi-1*) (*E. coli* Genetic Stock Center) were used for functional test of the *cI* and *lacI* PCR gene products. Plasmid pTCAT was a gift of C. Khosla (Hughes et al., 1989) containing a *tac-cat* fusion suitable for this study. Plasmid pMJ1560 (Amersham) was used as the source for the *lacI* gene. The *cI* gene was obtained from plasmid pKB252 (Backman & Ptashne, 1978). Plasmids pKK223-3 and pPL-Lambda (Pharmacia) were used as the source of the *tac* promoter and the λP_L promoter, respectively. Plasmid pKQV4 (Strauch et al., 1989) is essentially the same as pKK223-3 except it also contained the $lacI^q$ gene isolated form pMJR1560. A series of different copy number plasmids pDM246, pDM247, and pFH118 were described before (Moser & Campbell, 1983). Plasmid pSL1180 (Pharmacia) which carried a super polylinker was used for most of the subcloning steps.

Media and growth conditions

LB medium containing 10 g/l Difco tryptone, 5 g/l Difco yeast extract, 10 g/l NaCl, 3 g/l K_2HPO_4 and 1 g/l KH_2PO_4 (pH 7.0) was used for all growth experiments. Ampicillin was added to a concentration of 50 mg/l for selection. Shake flask experiments

were carried out at 250 rpm in a New Brunswick INNOVA 4000 shaker at 37^{0}C. The temperature-shift experiment was conducted in a New Brunswick GYROTORY water bath shaker model G76 at 250 rpm. For induction of the *tac* promoter, IPTG was added to 1mM concentration unless otherwise described.

Amplification of *lacI* and *cI* genes using polymerase chain reaction (PCR)

The polymerase chain reaction (PCR) technique (Saiki, et al., 1988) was used to synthesize both the *lacI* and *cI* structural gene using plasmid pMJR1560 and plasmid pKB252 as templates, respectively. Primers I and II (Figure 2a) were used to amplify the *lacI* gene. The native *lacI* transcriptional termination signal which overlapped into the *lac* promoter region was replaced by a strong *trpA* transcriptional termination signal (Christie et al., 1981) as indicated in Figure 2a. In addition, an extra translational stop codon was inserted in frame after the original one. The Shine-Dalgarno sequence (S-D) for the *lacI* gene was also included in primer I. *Bsp*EI sites were created at both ends of the primers to facilitate subcloning of the amplified PCR fragment. To synthesize the structural gene of the *cI* repressor, primers III and IV were used (Figure 2b). The S-D sequence and the transcriptional termination signal were not included in these primers because they will be supplied from plasmid pKVQ4 used for subcloning the amplified *cI* fragment. Restriction sites *Eco*RI and *Pst*I were created at the 5' end of primer III and IV, respectively. All primers were synthesized at the Caltech Applied Microchemical Facility.

The PCR reaction was carried out in a 50 µl final reaction volume containing 2.5 ng of each respective template DNA, 5 µl of 10X reaction buffer (100mM Tris-HCl pH. 8.3, 500 mM KCl, 15 mM $MgCl_2$, 0.01%(W/V) gelatin), 8 µl of dNTPS mix (1.25 mM each), 2 mM each of the primers, and 0.5 µl (1U/µl) of Taq DNA polymerase (Cetus). The amplification reaction was carried out for 36 cycles in a DNA thermal cycler (Perkin

Elmer-Cetus). The DNA was denatured at 92^{0}C for 1 min, annealed at 42^{0}C for 2 min, and extended at 72^{0}C for 3 min. The amplified products were analyzed on a 1.2% agarose gel in order to verify that they have the correct sizes.

Construction of the constitutive repressor synthesis vectors (pCS series)

Plasmid pKC2 was constructed by inserting a 1.5kb *Bsp*HI fragment from pTCAT containing the *tac-cat* fusion into the corresponding complementary *Nco*I site in plasmid pSL1180. The resulting plasmid retained the entire *tac-cat* fusion including the strong *rrnB* termination sequence (this sequence is included in plasmid pKK223-3) but did not extend into the *bla* gene region.

Subcloning of the *lacI^q* fragment into pKC2 was accomplished by cleaving plasmid pMJR1560 with *Kpn*I and *Pst*I. A 1.2 kb fragment containing the entire *lacI^q* allele including its own promoter and transcription termination sequence was then subcloned into the corresponding sites in pKC2 to create pKC6 (Figure 4). This construct contains the *lacI^q* and the *tac-cat* fusion facing in the opposite direction in order to minimize the possibility that any transcription initiated elsewhere on the plasmid could be extended into this region.

To transfer this expression cassette into different copy number plasmids, pKC6 was cleaved with *Bst*BI and *Sph*I and a 3.1 kb fragment was ligated into the *Sac*I and *Sph*I sites of plasmids pDM246, pDM247, and pFH118 to give pCS246, pCS247, and pCS118, respectively. The *Bst*BI-*Sac*I sites were made blunt with Klenow fragment and T4 DNA Polymerase prior to ligation.

Construction of the cross-regulation vectors (pCRR series)

The 736 bp *cI* fragment obtained from PCR amplification was purified by using the Geneclean kit (Bio101), digested with *Eco*RI and *Pst*I nucleases, and ligated into pKQV4 previous opened with the same enzymes. The resulting plasmid pTCI carried the *cI* gene under the control of the *tac* promoter. Primer III used for the synthesis of this *cI* fragment was designed such that the distance between the start codon of the *cI* gene and the ribosome binding site (SD sequence) on the plasmid is 10 base pairs long (recommended 10-15 base pairs for effective translation initiation according to Pharmacia; see Figure 2c). Plasmid pSIAT was derived by replacing a 622 bp *Mlu*I-*Pst*I fragment carrying the entire *tac* promoter from pKC2 by a 922 bp *Bam*HI-*Pst*I fragment containing the *tac-cI* fusion with the *Mlu*I-*Bam*HI end rendered blunt by "filling in" (using Klenow fragment of DNA polymerase), thus creating a *tac-cI-cat* operon (Figure 4). Both *Bam*HI and *Mlu*I sites were regenerated after ligation.

The *lacI* PCR product was isolated on a 1.2% agarose gel and purified with a Geneclean kit. Proteinase K treatment was applied to the purified DNA fragment in order to improve the subcloning efficiency (Crowe et al., 1991). It was then subjected to phenol-chloroform extraction and ethanol precipitation. The purified 1169 bp *lacI* fragment was then ligated into the unique *Sma*I site of pUC18 to generate pUC18-lacI. To construct plasmid pλ-lacI, a 1.2 kb *lacI* fragment was obtained after cleaving pUC18-lacI with *Eco*RI and *Sph*I and inserted into the *Bsp*EI and *Sph*I sites of pPL-Lambda with blunt end ligation at the *Eco*RI-*Bsp*EI sites (both ends were made blunt by filling in with the Klenow fragment of DNA polymerase). The resulting construct expressed the *lac* repressor under the control of the λP_L promoter.

Plasmids pλ-lacI and pSIAT were cleaved with *Nru*I and *Bam*HI and a 1.9 kb

fragment from pλ-lacI was inserted into pSIAT, replacing the smaller fragment from the latter to yield pKC7. The resulting construct has the *tac-cI-cat* operon and the λP_L-*lacI* fusion facing in the opposite orientation (Figure 4). Since transcription termination signals were included at the 3' end of the gene, it is unlikely that transcription initiated within the plasmid can be extended into this region. Subsequently, it is likely that the expression of these genes would be under control of their own promoters. To transfer the expression cassette into the different copy number plasmids, pKC7 was digested with *Sac*II and *Nco*I and a 3.4 kb fragment carrying the entire cross-regulation cassette was ligated into the same sites of the pCS series to obtain plasmids pCRR118, pCRR246, and pCRR247.

Protein and enzyme assays

Cells were disrupted by sonication. The soluble fraction was used for protein and CAT analysis. Total protein concentration was determined using a Sigma kit (No. P5656). CAT activity was determined with ^{14}C-labeled butyryl coenzyme A (New England Nuclear) according to recommended protocols (Newman et al., 1987). For this assay, 1 ml of cell extract was added to a 7 ml glass miniscintillation vial (Kimble) containing sufficient 100 mM Tris-HCl (pH 7.8) to give a total volume of 50 µl. 200 µl of freshly prepared 1.25 mM chloramphenicol in 100 mM Tris-HCl (pH 7.8) and 0.1 µCi of ^{14}C-Butyrl CoA were added. 5 mL of a scintillation fluor (Econofluor) was gently overlaid on top of the reaction mixture. The vial was counted in a liquid scintillation counter (Beckman model LS5801) and the CAT activity was calculated from the slope of the cpm versus incubation time data. CAT activity is expressed in units of CAT per milligram of total soluble protein. Total CAT content was measured by a CAT ELISA kit obtained from 5 Prime, 3 Prime, Inc.

β-galactosidase was assayed at 30°C using *O*-nitrophenol-β-D-galactopyranoside as the substrate. Change in A_{420} was monitored by a rate assay in a thermostatted spectrophotometer (Shimadzu UV260). Specific activity is expressed as Miller units; 1 Miller unit=100 $\Delta A_{420}/min/OD_{600}$ (Miller, 1972). SDS polyacrylamide gel electrophoresis (SDS-PAGE) was performed according to the method of Laemmli (Laemmli, 1970). For SDS-PAGE analysis, cell lysate was boiled for 5 min in gel loading buffer (10% glycerol, 5% 2-mercaptoethanol, 3.3% SDS and 0.5 M Tris, pH 6.8) and then electrophoresed on a 12.5% polyacrylamide gel.

Chemicals, reagents, and DNA Manipulations

All restriction endonucleases, modifying enzymes (T4 DNA polymerase, Klenow fragment, T4 DNA Ligase) and Proteinase K were purchased form New England BioLabs and Boehringer Mannheim Biochemicals. Taq DNA Polymerase was obtained from Cetus Corp. ^{14}C-labeled butyrl coenzymeA was obtained form New England Nuclear. All DNA manipulations were done according to standard methods (Sambrook et al. 1989). DNA fragments were eluted from agarose gel using a Geneclean Kit (Bio 101).

Plasmid copy number determinations

Plasmid copy number was determined essentially by the method of Projan et al. (1983). The cell pellet was suspended and incubated for 30 min at 37°C in 50 μl lysis buffer containing 20 mM tris-HCl, pH 8.0, 10 mM EDTA, 100 mM NaCl, 20% sucrose, 75 mg lysozyme, and 2U/ml RNase. Fifty microliters of 2% SDS was added, and the samples were vortexed at maximum setting for 1 min. After the samples were freeze-thawed two times (-70 to 20°C), 2 mg proteinase K was then added, and the samples were incubated at 37°C for 30 min. Twenty five microliters of loading buffer (50%

glycerol, 1 mM EDTA, pH 8.0, and 1% bromophenol blue) were added, and 15 ml samples were loaded on a 0.9% agarose gel. The gel was electrophoresed for 3 hr at 80V and was subsequently stained for 40 min with 1 mg/L ethidium bromide and destained 2 X 20 min with water. The gel was illuminated with a Chromato-vue transilluminator (Ultraviolet Products, INC.), photographed with Polaroid type 665 film, and the negatives were developed according to the manufacturer's instructions. The negatives were scanned with a LKB 2200 scanning densitometer. The plasmid copy number was calculated by multiplying the peak height ratio by the base number ratio of chromosome to plasmid.

3.4 Results

Criteria for choosing the first and second promoter-repressor systems

Although many different types of promoter-repressor systems can be used for the cross-regulation system, there are at least two criteria that must be satisfied. These two criteria were determined from previous modeling efforts (Chen & Bailey, unpublished results), and they are presented here to serve as guidelines for future cross-regulation system design. There is no strict requirement for choosing the first promoter-repressor system as long as it is a fairly strong promoter. On the other hand, in choosing the second promoter-repressor system, satisfaction of the following characteristics are important: (1) the sensitivity of repression of this promoter by its repressor must be less than that of the first promoter. This is necessary such that a higher repressor concentration is needed to achieve the same repression level for the second promoter as that achieved for the first promoter; (2) the second promoter must be a stronger promoter than the first promoter. These two criteria are required to guarantee that the first promoter will be turned off in the pre-induction state. In general many systems can satisfy these two requirements. For this

work, the *tac* promoter has been picked for the first promoter because it is a conveniently available strong promoter and has been used to express many recombinant proteins in *E. coli* (De Boer et al., 1983; Amann et al., 1983). The λP_L promoter is used as the second promoter because it is reported to be stronger than the *tac* promoter and has also been shown to satisfy the first criterion (Johnson et al., 1981).

The *lacI* and *cI* PCR fragments produce fully functional gene products

It is well known that DNA amplification using the PCR technique is susceptible to error due to the absence of the 3' to 5' proofreading exonuclease function in the Taq DNA polymerase (Kohler et al., 1991). Normally, DNA sequencing is the most reliable method to ensure that no error is incorporated into the amplified DNA fragment. However, in this work we are only interested in obtaining gene fragments that produce functional active gene products. A sensitive test for the presence of functional *lac* repressor is the ability to repress β-galactosidase expression in a *lacI*- mutant that cannot provide its own *lac* repressor. Mutant CGSC808 lacking the *lac* repressor forms blue colonies on an X-gal plate due to the presence of β-galactosidase. On the other hand, this mutant forms white colonies on an X-gal plate carrying pUC18-lacI, demonstrating that the *lac* repressor produced from the PCR *lacI* fragment is indeed functional (data not shown).

A similar experiment was carried out for the *cI* PCR fragment. In this experiment, plasmid pUC18-cI, which carries the *cI* PCR fragment inserted into the *EcoRI* and *PstI* sites of pUC18, was transformed into E. coli CY15050. This strain carries a λ prophage and a temperature-sensitive *cI* repressor. Shifting temperature from 30°C to 42°C destroys the temperature-sensitive *cI* repressor activity. The prophage DNA is detached from the host chromosome changing the prophage from lysogenic to lytic state, and

eventually the host lyses. Figure 3 shows that strain CY15050 carrying pUC18-cI can remain intact and continue growing after temperature shift, indicating that the PCR *cI* fragment produces active *cI* repressor to complement the host's temperature-sensitive *cI* repressor.

The cross-regulation system functions as predicted from simulation

For plasmid pKC7 (Figure 4), the expression of the *cat* gene is under cross-regulation control. The X-gal plating behavior of CGSC808/pKC7 (only white colonies were observed) made it likely that the *lacI* gene under control of the λP_L promoter would supply enough *lac* repressor for complete repression even in the presence of multiple copies of the *lac* operator (both from the *tac* promoters on the plasmid and the *lac* promoter on the chromosome). This is in good agreement with simulation predictions indicating the pre-induction repression of the first promoter, in this case the *tac* promoter (Chen et al., 1991). This is very important because, if this trend is reversed, the system will fail to control basal expression from the *tac* promoter, controlling instead the λP_L promoter. Data from another experiment supporting the claim of adequate *lac* repressor synthesis in the cross-regulation construct is shown in Table I. In this table, the β-galactosidase activity was measured for strains CGSC808 carrying plasmids pUC18, pMJR1560 and pλ-lacI. As indicated, the strain carrying pUC18 shows a high level of β-galactosidase activity as opposed to the other two constructs. Examining these data further, the β-galactosidase activity is lower for the pλ-lacI construct compared to that for pMJR1560. This is an excellent indication that the amount of *lac* repressor provided from the λP_L promoter is higher than that from its native promoter.

In order to quantify the levels of expression, we measured CAT production from the constitutive repressor synthesis configuration (DH5α/pKC6) and also that from the

cross-regulation configuration (DH5α/pKC7). Before induction, both strains show a low level of CAT activity, with that of DH5α/pKC7 slightly lower (Figure 5). Upon induction by the addition of IPTG, the CAT activity increases significantly. However, the CAT activity for DH5α/pKC6 eventually levels off while that for DH5α/pKC7 continues to increase to approximately two times the level in DH5α/pKC6 after three hours of induction. Since both systems utilize the same S-D sequence for the *cat* gene, the difference in expression is expected to be contributed solely from the difference in transcription. Results from this experiment match well with our model predictions which indicated excellent control of expression prior to induction and higher expression post-induction for the cross-regulation construct.

The dynamics of the cross-regulation system is illustrated in Figure 6. Before induction, no band corresponding to the CAT protein is visible for either construct (Figure 6, lanes 1 and 6). After IPTG addition, the CAT band immediately appears. For the cross-regulation system, bands for both *cI* repressor and CAT appear at the same time, showing that the synthetic operon constructed for the cross-regulation system is functional (Figure 6, lane 7). Furthermore, there is a major difference in the intensity of the *lac* repressor band. For the cross-regulation system, the intensity of this band post-induction is much weaker than that for the constitutive repressor synthesis configuration (see Figure 6, lanes 4 and 9). This shows that the expression of the *lac* repressor from the λP_L promoter is almost completely turned off after induction.

Although more CAT protein accumulated in three hours for the cross-regulation system, the initial increase after induction is delayed about 1 hour. This delay presumably reflects the time needed to overcome the higher concentration of *lac* repressor produced in the cross-regulation system at the time of induction. However, higher final expression was eventually achieved due to a decrease in the *lac* repressor concentration because

production of the *lac* repressor is turned off after induction.

Applicability of the cross-regulation system over a broad range of plasmid copy number

In order to demonstrate that this new configuration works equally well over a broad range of plasmid copy number, the expression cassettes for both the cross-regulation system and the constitutive repressor synthesis system were transferred to a series of very similar but different copy number vectors. These vectors only differ in their source of RNAI which is an inhibitor of replication for the ColE1 type replicon (Moser & Campbell, 1983).

Plasmid content measurements (Table II) obtained from densitometry scan of photographic negatives are in qualitative agreement with those observed previously (Seo and Bailey, 1985). As indicated in previous experiments with these different copy number vectors, the specific growth rate of *E. coli* carrying these vectors decreases with increasing copy number (Seo & Bailey, 1985). The same trends were also observed here for both configurations as can be seen in Table II. Approximately the same decrease in growth rate is observed for both configurations before induction, indicating basal expression of the CAT protein does not play an important role in the reduction of growth. On the other hand, growth is much more severely reduced for the cross-regulation configuration after the addition of IPTG reflecting the redirection of the cell resources towards CAT production. This reduction increases with increasing copy number as shown in Figure 7.

Induction experiments reveal that the cross-regulation system consistently produces at least 2 times the CAT level compared to the constitutive regulation system independent of copy number (Table III). Most importantly, the basal levels of CAT expression are comparable to those obtained from the control system (in the range of

0.01% of total protein). Results from all these induction experiments exhibit very similar trends with time post induction as those observed with the pKC6 and pKC7 constructs. For the cross-regulation system, CAT level continues to increase up to 7 hr post-induction (data not shown).

IPTG dosage response

In order to understand the response of the cross-regulation system to induction, a series of shake-flask experiments were performed in which the IPTG concentration was varied from 0 to 2 mM. These results are presented in Figure 8. For IPTG concentrations in the range of 0 to 0.5 mM, the CAT levels are essentially identical for both the cross-regulation and constitutive inhibition systems. This suggests that the cross-regulation system is only partially induced at these concentrations and that expression of the *lac* repressor from the λP_L promoter is not completely turned off. On the other hand, the CAT level increases to a higher level for the cross-regulation system with IPTG concentrations beyond 0.5 mM. This indicates a complete shift of steady state is achieved at such higher IPTG concentrations. Only a small increase in CAT level is observed beyond an IPTG concentration of 2 mM (data not shown).

3.5 Discussion

The ability to regulate expression of a cloned-gene is of great importance. This is particularly true if the cloned-gene product is toxic to the host cell. If such a protein is expressed prematurely, it may lead to selection against plasmid-containing cells and low final cell density. On the other hand, a high level of cloned-gene expression is eventually desirable. This could be accomplished by expressing the cloned gene with a strong promoter carried on a multicopy vector. Unfortunately, cells are usually not equipped

with the machinery to regulate the expression of multiple gene copies except for a few exceptions such as rRNAs and tRNAs. Controlling basal expression levels from these expression vectors becomes a potentially important problem. Normally, the only solution considered to obtain excellent control of basal expression is by elevating constitutive expression of the corresponding repressor protein. This strategy has been applied for a wide variety of different promoter systems (Mieschendahl et al., 1986). Such an approach, however, leads to overproduction of repressor protein and full expression of the cloned-gene is difficult to achieve. A previous report has shown that the induced level of protein synthesized from such a configuration was around 30% less than that obtained from an expression vector which did not include the repressor gene (Stark, 1987). This is an excellent indication that the transcription potential of this system is far from saturated. Therefore, increasing the transcription activity of the promoter at the induced state is important.

The novel cross-regulation system constructed in this work has the ability to address both problems at the same time. This system not only retains excellent control of basal expression, but it also achieves a higher induced level of cloned protein production. Induction enables this system to both up-regulate its activation pathway and down-regulate its inhibition pathway. Although no such system has been reported in bacterial cells, a very similar type of dual cross-regulation control was recently reported for the hormone-sensitive adenylyl cyclase system in rat cells (Hadcock et al., 1990 and 1991).

Our results can be summarized in a model presented in Figure 9. The addition of IPTG is believed to have two major effects. First, the formation of a IPTG-*lac* repressor complex results in an increase in the transcription activity from the *tac* promoter. This is demonstrated by an increase in CAT and *cI* repressor level. Increasing the intracellular *cI* repressor concentration then cross-regulates the λP_L promoter and turns off transcription

initiating there. The net result is the reduction in *lac* repressor concentration as demonstrated in Figure 6. This type of behavior is extremely desirable since the inhibitory pathway (in this case the *lac* repressor) is no longer necessary once induction is desired. Moreover, we have shown that the level of repression when the *lac* repressor is provided from the λP_L promoter is greater than when using its native promoter affording improved control of basal expression prior to induction.

It should be noted that the response of the cross-regulation system to induction is slightly more sluggish. This possibly derives from a higher *lac* repressor concentration at the time of induction which is necessary for better control of basal expression. Once an initial lag is completed, the expression level for the cross-regualtion system continues to increase up to 7 hr post-induction. Similar behavior is also observed in the IPTG dosage experiment. The response of the cross-regulation system at low IPTG concentration is almost identical to that of the control system. For this situation, only a small reduction in the *lac* repressor concentration is expected to occur. For this reason, no major difference is observed from the constitutive repressor synthesis situation. In contrast, substantial increases in cloned gene expression can be achieved with higher IPTG concentration. Using higher IPTG concentration shifts the system into a new steady state providing cloned gene instead of first repressor expression.

Various types of systems have been applied for regulated high level expression of cloned genes. One of the most efficient systems involve the use of bacteriophage T7 RNA polymerase which is specific only for a T7 promoter. However, even a small amount of T7 RNA polymerase can direct most of the cell resources towards cloned-gene expression. A series of attempts have been made including using a combined T7 promoter with the *lac* operator and the creation of an autogene in order to control the basal expression of T7 RNA polymerase such that even toxic gene products can be expressed

(Dubendorff & Studier, 1991a, 1991b; Studier, 1991). In this context, the cross-regulation system offers another excellent alternative for controlling T7 RNA polymerase expression.

Results presented in this study introduce a new kind of expression control concept. It is reasonable to expect that this type of configuration can be applied using a wide variety of different promoter systems and hosts and is not restricted to the example illustrated in this study. The success of these experiments also validate the utilization of molecular level mathematical models for discovering novel genetic designs with practical utility. Work is currently in progress in our laboratory to determine the factors affecting recombinant protein yield using the cross-regulation system.

3.6 Acknowledgements

This work was supported by the Advancement Industrial Concepts Division of the U.S. Department of Energy. W. Chen was supported in part by a Predoctoral Training Grant in Biotechnology from the National Institute of General Medical Sciences (National Research Service Award 1 T32 GM 08346-01, Pharmacology Science Program). We wish to thank Dr. Keqin Chen for technical assistance in performing the PCR reaction. We also thank Dr. C. Yanofsky , Dr. M.A. Strauch and Dr. M. Ptashne for supplying strain CY15050 and plasmids pKQV4 and pKB252, respectively.

3.7 References

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

Backman, K. & Ptashne, M. (1978). Maximizing gene expression on a plasmid using recombination *in vitro*. *Cell*, **13**, 65-71.

Baker, S.A., Grossman, A.D. & Gross, C.A. (1984). A gene regulating the heat shock response in *Escherichia coli* affects proteolysis. *Proc. Nat. Acad. Sci. U.S.A.* **81**, 6779-6783.

Bentley, W.E., Morkajalili, N., Anderson, D.C., Davis, R.H. & Kompala, D.S. (1990). Plasmid-encoded protein: The princpal factor in the "Metabolic burden" associated with recombinant bacteria. *Biotechnol. Bioeng.* **35**, 668-681.

Bolivar, F. & Backman, K. (1979). Plasmids of *Escherichia coli* as cloning vectors. *Methods Enzymol.* **68**, 245-251.

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

Christie, G.E., Farnham, P.J. & Platt, T. (1981). Synthetic sites for transcription termination and a functional comparison with tryptophan operon termination sites *in vitro*. *Proc. Nat. Acad. Sci. U.S.A.* **78**, 4180-4184.

Crowe, J.S., Copper, H.J., Smith, M.A., Sims, M.J., Parker, D. & Gewert, D. (1991). Improved cloning effeciency of polymerase chain reaction (PCR) products after proteinase K digestion. *Nucl. Acids. Res.* **19**, 184-184.

De Boer, H. Comstock, H.L. & Vasser, V. (1983). The *tac* promoter: a

functional hybrid derived form the *trp* and *lac* promoter. *Proc. Nat. Acad. Sci. U.S.A.* **80**, 21-25.

Dubendorff, J.W. & Studier, F.W. (1991a). Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with the *lac* repressor. *J. Mol. Biol.* **219**, 45-59.

Dubendorff, J.W. & Studier, F.W. (1991b). Creation of a T7 autogene. Cloning and expression of the gene for bacteriophage T7 RNA polymerase under control of its cognate promoter. *J. Mol. Biol.* **219**, 61-68.

Goff, S.A. & Goldberg, A.L. (1986). Production of abnormal protein in *E. coli* stimulates transcription of *lon* and other heat shock genes. *Cell,* **41**, 587-595.

Hadcock, J.R., Ross, M., Watkins, D.C. & Malbon, C.C. (1990). Cross-regulation between g-protein-mediated pathways. *J. Biol. Chem.* **265**, 14784-14790.

Hadcock, J.R., Port, J.D. & Malbon, C.C. (1991). Cross-regulation between g-protein-mediated pathways. *J. Biol. Chem.* **266**, 11915-11922.

Hughes, D.E., Curtis, J.E., Khosla, C. & Bailey, J.E. (1989). A new oxygen-regulated promoter for the expression of proteins in *Escherichia coli. Biotechniques*, **7**, 1026-1028.

Johnson, A.D., Poteete, A.R., Lauer, G., Sauer, R.T., Ackers, G.K. & Ptashne, M. (1981). λ repressor and cro-components of an efficient molecular switch . *Nature*, **294**, 217-223.

Kelly, R.L. & Yanofsky, C. (1982). *Trp* aporepressor production is controlled by autogenous regulation and inefficient translation. *Proc. Nat. Acad. Sci. U.S.A.* **79**, 3120-3124.

Kohler, S.W., Provost, G.S., Fieck, A., Kretz, P.L., Bullock, W.O., Sorge, J.A., Putman, D.L. & Short, J.M. (1991). Spectra of spontaneous and

mutagen-induced mutations in the *lacI* gene in transgenic mice. *Proc. Nat. Acad. Sci. U.S.A.* **88**, 7958-7962.

Kosinski, M. & Bailey, J.E. (1991). Temperature and induction effects on the degradation rate of an abnormal β-galactosidase in *Escherichia coli*. *J. Biotech.* **18**, 55-68.

Laemmili, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-685.

Latta, M., Philit, M., Maury, I., Soubrier, F., Denefle, P. & Mayaux, J-F. (1990). Tryptophan promoter derivatives on multicopy plasmids: A comparative analysis of expression potentials in *Escherichia coli*. *DNA and Cell Biology*, **9**, 129-137.

Makoff, A.J. & Oxer, M.D. (1991). High level heterologous protein expression in *E. coli* using mutant forms of the *lac* promoter. *Nucl. Acids Res.* **19**, 2417-2421.

Mieschendahl, M. Petri, T. & Hanggi, U. (1986). A novel prophage independent *trp* regulated lambda P_L expression system. *Bio/Technology*, **4**, 802-808.

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

Moser, D.R. & Campbell, J.L. (1983). Characterization and complementation of pMB1 copy number mutants: effect of RNA1 gene dosage on plasmid copy number and incompatibility. *J. Bacteriol.* **154**, 809-818.

Mott, J.E., Grant, R.A., Ho, Y.S & Platt, T. (1985). Maximizing gene expression from plasmid vectors containing the λP_L promoter: Strategies for overproducing transcription termination factor ρ. *Proc. Nat. Acad. Sci. U.S.A.* **82**, 88-92.

Newman, B.M., Morency, C.A. & Russian, K.O. (1987). A novel rapid assay

for chloramphenicol acetyltransferase gene expression. *Biotechniques*, **7**, 444-447.

Park, S., Ryu, D.D.Y. & Kim, J.Y. (1990). Effect of cell-growth rate on the performance of a 2-stage continuous culture system in a recombinant *Escherichia coli* fermentation. *Biotechnol. Bioeng.* **36**, 493-505.

Park, T.H., Seo, J-H. & Lim, H.C. (1991). Two-stage fermentation with bacteriophage λ as an expression vector in *Escherichia coli*. *Biotechnol. Bioeng.* **37**, 297-302.

Peretti, S.W & Bailey, J.E. (1987). Simulations of host-plasmid interactions in *Escherichia coli*: Copy number, promoter strength, and the ribsome binding site strength effects on metabolic activity and plasmid gene expression. *Biotechnol. Bioeng.* **24**, 316-328.

Projan, S.J., Carleton, S. & Novic, R.P. (1983). Determination of plasmid copy number by fluorescence densitometry. *Plasmid*, **9**, 182-190.

Remaut, E., Tsao, H. & Fiers, W. (1983). Improved plasmid vectors with a thermoinducible expression and temperature-regulated runaway replication. *Gene*, **22**, 103-133.

Saili, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. & Erlich, H.A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, **239**, 487-491.

Sambrook, J., Frilsch, E.F. & Maniatis, F.E. (1989). Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Seo, J-H. & Bailey, J.E. (1985). Effect of recombinant plasmid content on growth properties and cloned gene product formation in *Escherichia coli*. *Biotechnol. Bioeng.* **27**, 1668-1674.

Siegel, R. & Ryu, D.D.Y. (1985). Kinetic study of instability of recombinant

plasmid pPLc23trpA1 in *E. coli* using two-stage continuous culture system. *Biotechnol. Bioeng.* **27**, 28-33.

Stark, M.J.R. (1987). Multicopy expression vectors carrying the *lac* repressor gene for regulated high level expression of genes in *Escherichia coli. Gene*, **51**, 255-267.

Strauch, M.A., Spiegelman, G.B., Perego, M., Johnson, W.C., Burbulys, D. & Hoch, J.A. (1989). The transition state transcription regulator *abrB* of *Bacillus subtilis* is a DNA binding protein. *EMBO J.* **8**, 1615-1621.

Studier, F.W. (1991). Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *J. Mol. Biol.* **219**, 37-44.

Yaffe, B.M., Levison, B.S., Szasz, J. & Sternlicht, H. (1988). Expression of a human alpha-tubulin properties of the isolated subunit. *Biochemistry*, **27**, 1869-1880.

3.8 Tables

Table I. β-galactosidase activity of strain CGSC808 carrying plasmids pUC18, pλ-lacI and pMJR1560. Cells were grown in shake flasks at 37°C and 275 rpm.

Strain	Source of *lac* repressor	β-gal activity, Miller unit
GSC808/pUC18	no	3000
GSC808/pMJR1560	$lacI^q$	60
GSC808/p λ-lacI	λ-lacI	<0.1

Table II. Comparison of specific growth rates for strain HB101 carrying the pCRR and pCS vectors grown in LB medium under non induced conditions at 37°C

Vector	copy number	Specific growth rate; pCRR series, hr^{-1}	Specific growth rate; pCS series, hr^{-1}
247	50	0.772	0.774
246	90	0.770	0.781
118	150	0.704	0.711

Table III. CAT expression level from strain HB101 carrying the pCRR and pCS vectors.

Vector	CAT level from the pCRR vectors, g CAT/g protein	CAT level from the pCS vectors, g CAT/g protein
247	0.0192	0.0085
246	0.0339	0.0096
118	0.0435	0.0190

Cells were grown at 37 ^{0}C until O.D.$_{600}$=1 and the culture was then split into two halves. To one, IPTG was added to 1 mM concentration for induction.

3.9 Figure Captions

Figure 1. Schematic representation of the cross-regulation system. The crossing of control occurs because the gene product from the first promoter, the second repressor R2, controls expression from the second promoter, while the gene product from the second promoter, the first repressor R, controls the first promoter.

Figure 2. (a) PCR primers I and II used for the synthesis of the *lacI* gene. The bold letters indicate the region of homology with the DNA template. The shadow letters indicate the S-D sequence of the *lacI* gene and the trpA transcriptional termination signal . (b) PCR primers II and IV used for the synthesis of the *cI* gene. The bold letters indicate the region of homology with the DNA template. (c) Spacing between the S-D sequence of the *cI* gene and the start codon on plasmid pTCI is indicated.

Figure 3. Functional test for the PCR *cI* fragment. Strain CY15050 carrying plasmids pUC18 and pUC18-cI were grown at 30^{0}C. IPTG was added to 1mM concentration at 1.5 hr to induce production of the *cI* repressor from plasmid pUC18-lacI. Half of each culture was then transferred to 42^{0}C. Optical density was measured at 660nm with a Spectronic 21 spectrophotometer (Milton Roy). Symbols: (O), pUC18 at 37^{0}C; (●), pUC18-lacI at 37^{0}C; (❑), pUC18 at 42^{0}C; (■), pUC18-lacI at 42^{0}C.

Figure 4. Maps of the plasmids pKC6 and pKC7 carrying the constitutive repressor

synthesis system and the cross-regulation system, respectively. Key: ptac:*tac* promoter, pL:λP_L promoter, MCS:multi-cloning sites, Tm:Termination sequence.

Figure 5. Comparison of growth and CAT production for DH5α/pKC6 and DH5α/pKC7. Cells were grown to approximately the same cell density ($O.D._{660}$=0.6) before the addition of IPTG. Symbols: (O), O.D. for pKC7; (Δ), O.D. for pKC6; (●), CAT activity for pKC7; (▲), CAT activity for pKC6.

Figure 6. Expression of CAT from plasmids pKC6 and pKC7. SDS-PAGE analyses of total protein from cells carrying plasmid pKC6 (lanes 1-4) and plasmid pKC7 (lanes 6-9) are shown for the time immediately before and 1, 2 and 3 hr after the addition of IPTG, respectively. Protein MW markers (87kd, 66kd, 42kd, 31kd and 21kd) are shown in lane 5. Positions of the *lacI*, *cI* and *cat* gene products are indicated.

Figure 7. Growth curves for the pCRR vectors carrying the cross-regulation system. A: pCRR247; B: pCRR246; C: pCRR118. Cells were grown at 37^{0}C to $O.D._{600}$=1.0 and the culture was then split equally into two flasks. To one, IPTG was added to a concentration of 1mM to induce CAT production. (O) indicates non-induced culture and (●) indicates induced culture.

Figure 8. IPTG dosage response of the cross-regulation system. CAT production as a function of IPTG concentration from HB101 carrying plasmids pCRR118 (❑) and pCS118 (O). Cells were grown at 37°C to $O.D._{600}$=1.0 before the addition of various amount of IPTG.

Figure 9. A model representing the effect of IPTG addition to the cross-regulation system.

Figure 1.

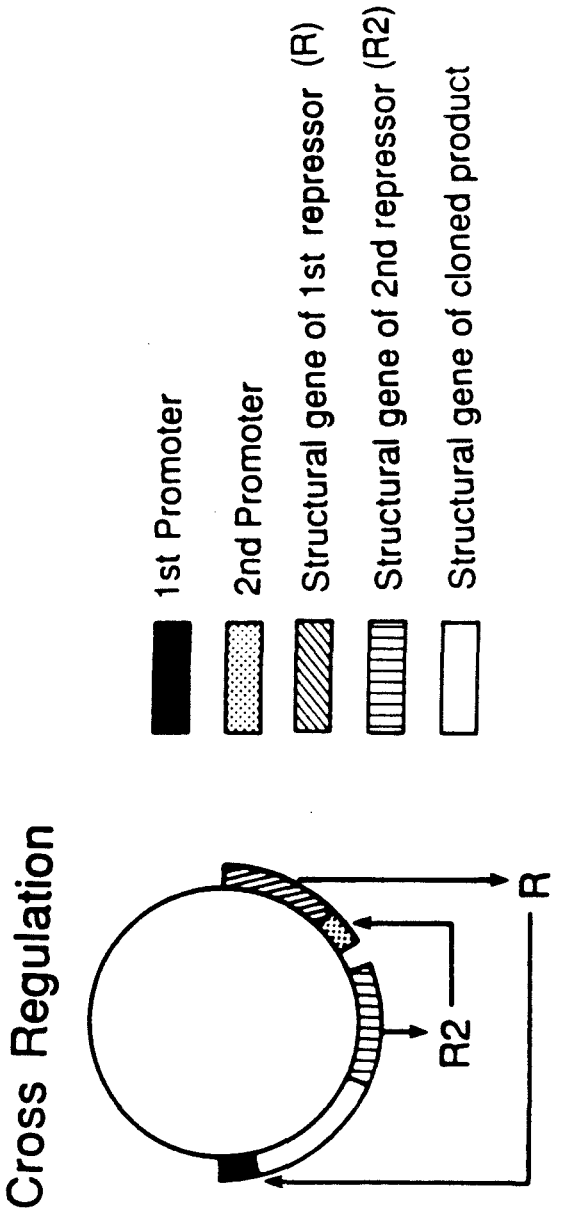

1st Promoter
2nd Promoter
Structural gene of 1st repressor (R)
Structural gene of 2nd repressor (R2)
Structural gene of cloned product
Cross Regulation
R
R2

Figure 2.

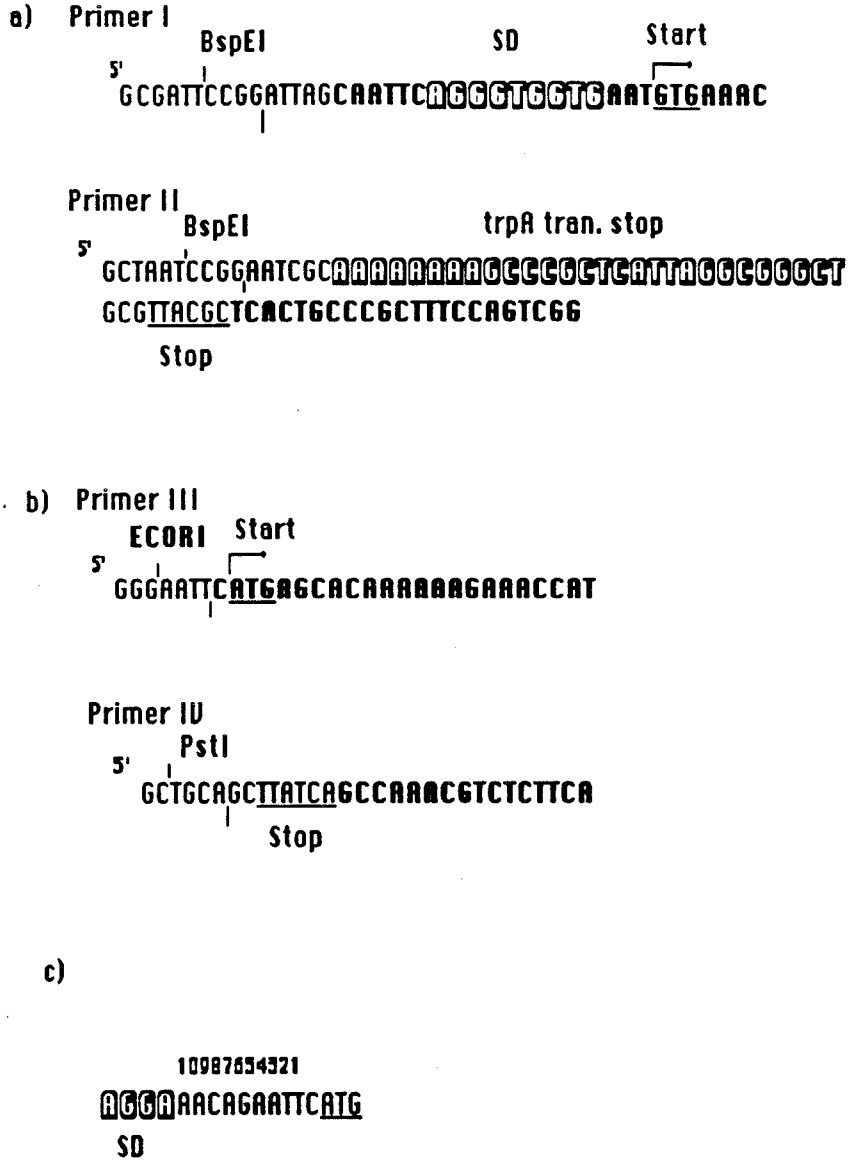

a) Primer I
BspEI
SD
Start
5'
GCGATTCCGGATTAGCAATTCAGGGTGGTGAATGTGAAAC
Primer II
BspEI
trpA tran. stop
5'
GCTAATCCGGAATCGCAAAAAAAAAGCCCGCTCATTAGGCGGGCT
GCGTTACGCTCACTGCCCGCTTTCCAGTCGG
Stop
b) Primer III
ECORI
Start
5'
GGGAATTCATGAGCACAAAAAAGAAACCAT
Primer IV
PstI
5'
GCTGCAGCTTATCAGCCAAACGTCTCTTCA
Stop
c)
10987654321
AGGAAACAGAATTCATG
SD
Spacing between the SD sequence and the start codon on pTCI

Figure 3.

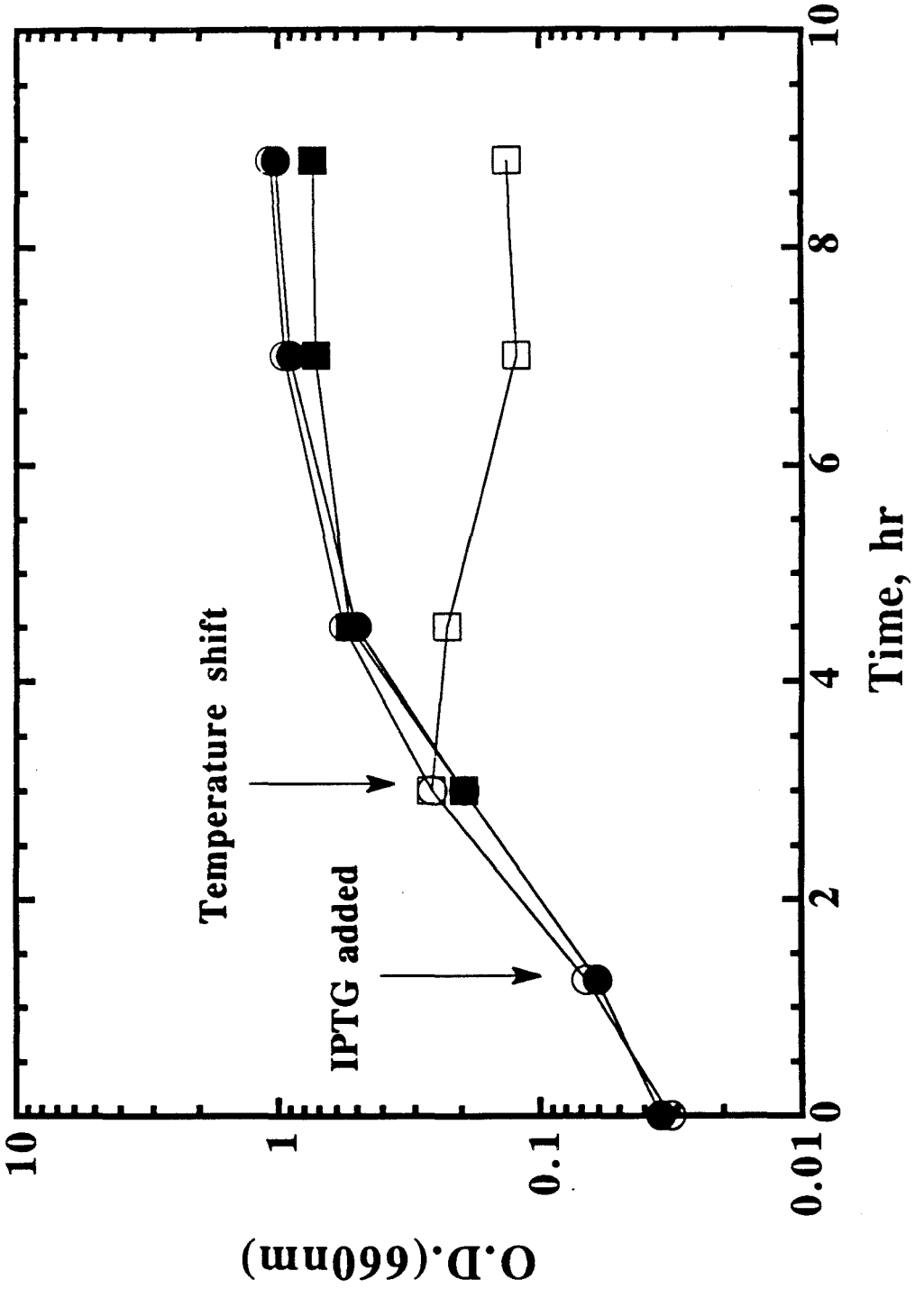

O.D.(660nm)
10
1
0.1
0.01
0
2
4
6
8
10
Time, hr
Temperature shift
IPTG added

Figure 4.

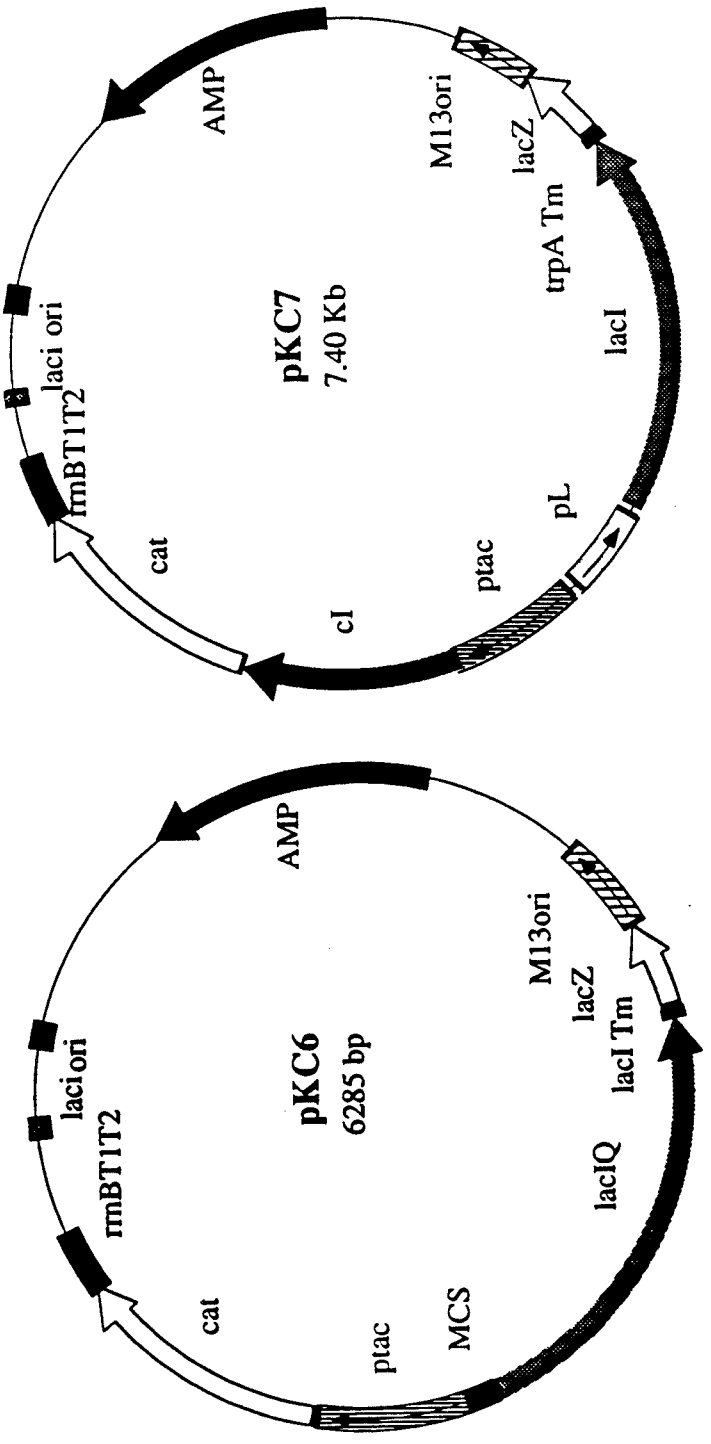

pKC7
7.40 Kb
AMP
M13ori
lacZ
trpA Tm
lacI
pL
ptac
cI
cat
rmBT1T2
laci ori
pKC6
6285 bp
AMP
M13ori
lacZ
lacI Tm
lacIQ
MCS
ptac
cat
rmBT1T2
laci ori

Figure 5.

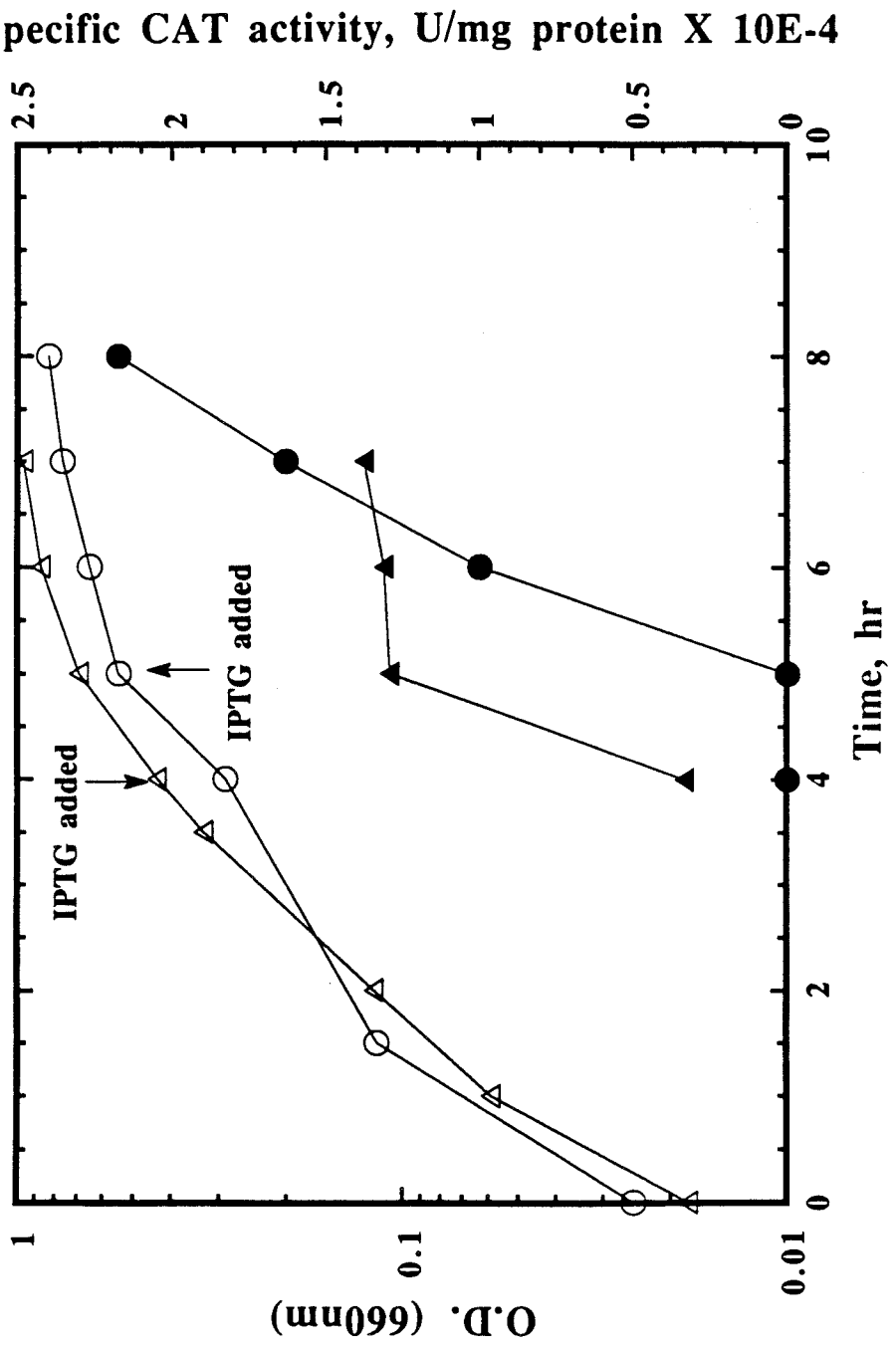

Specific CAT activity, U/mg protein X 10E-4
2.5
2
1.5
1
0.5
0
10
8
6
4
2
0
Time, hr
IPTG added
IPTG added
1
0.1
0.01
O.D. (660nm)

Figure 6.

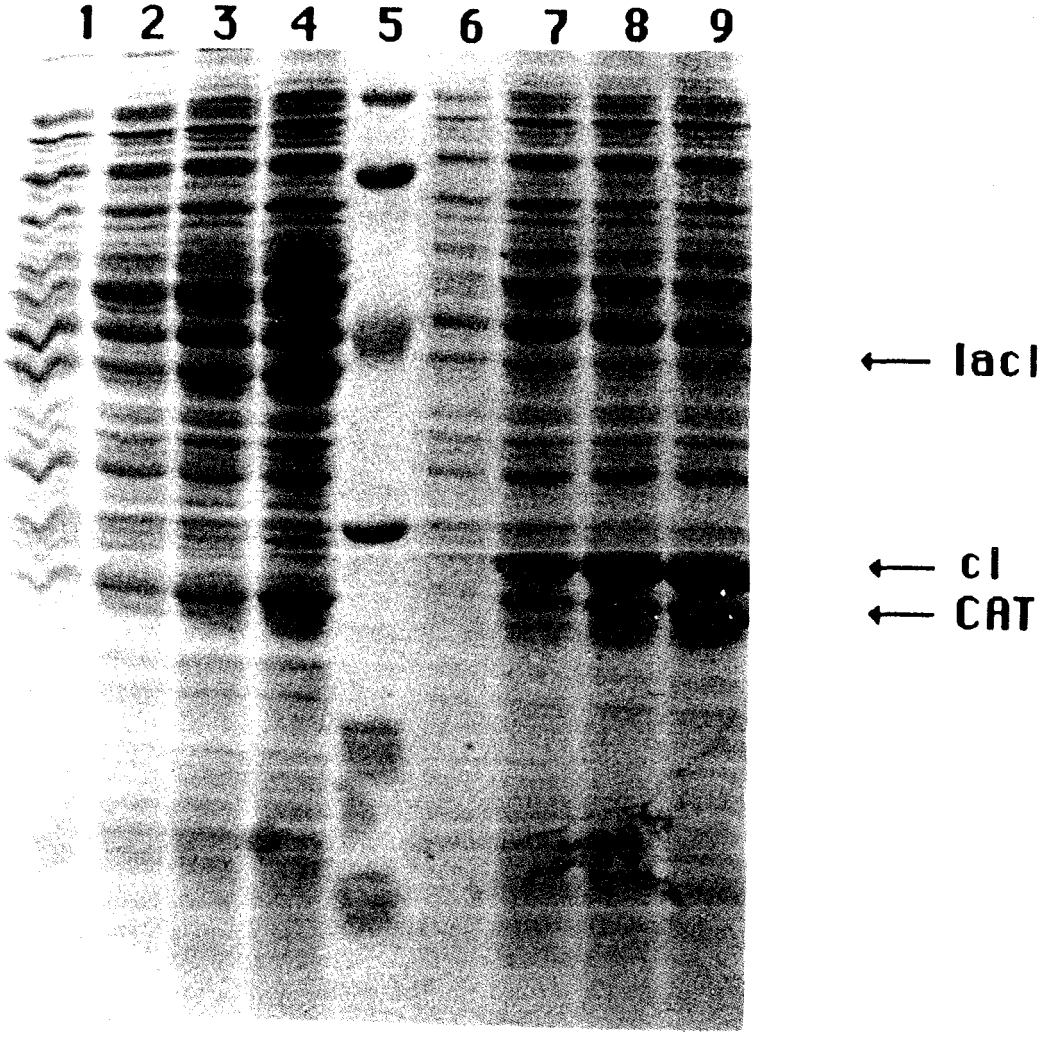


Figure 7.

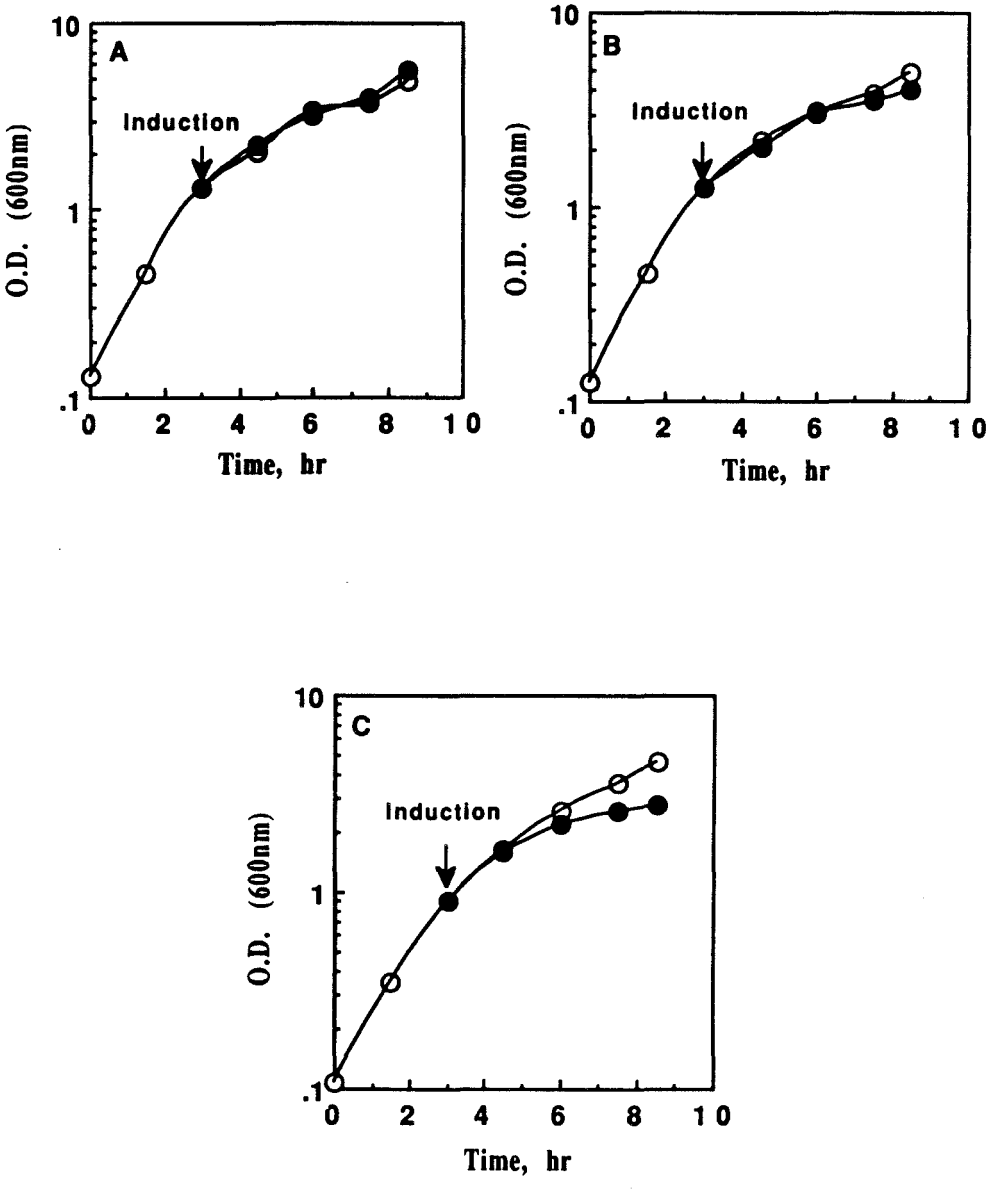


Figure 8.

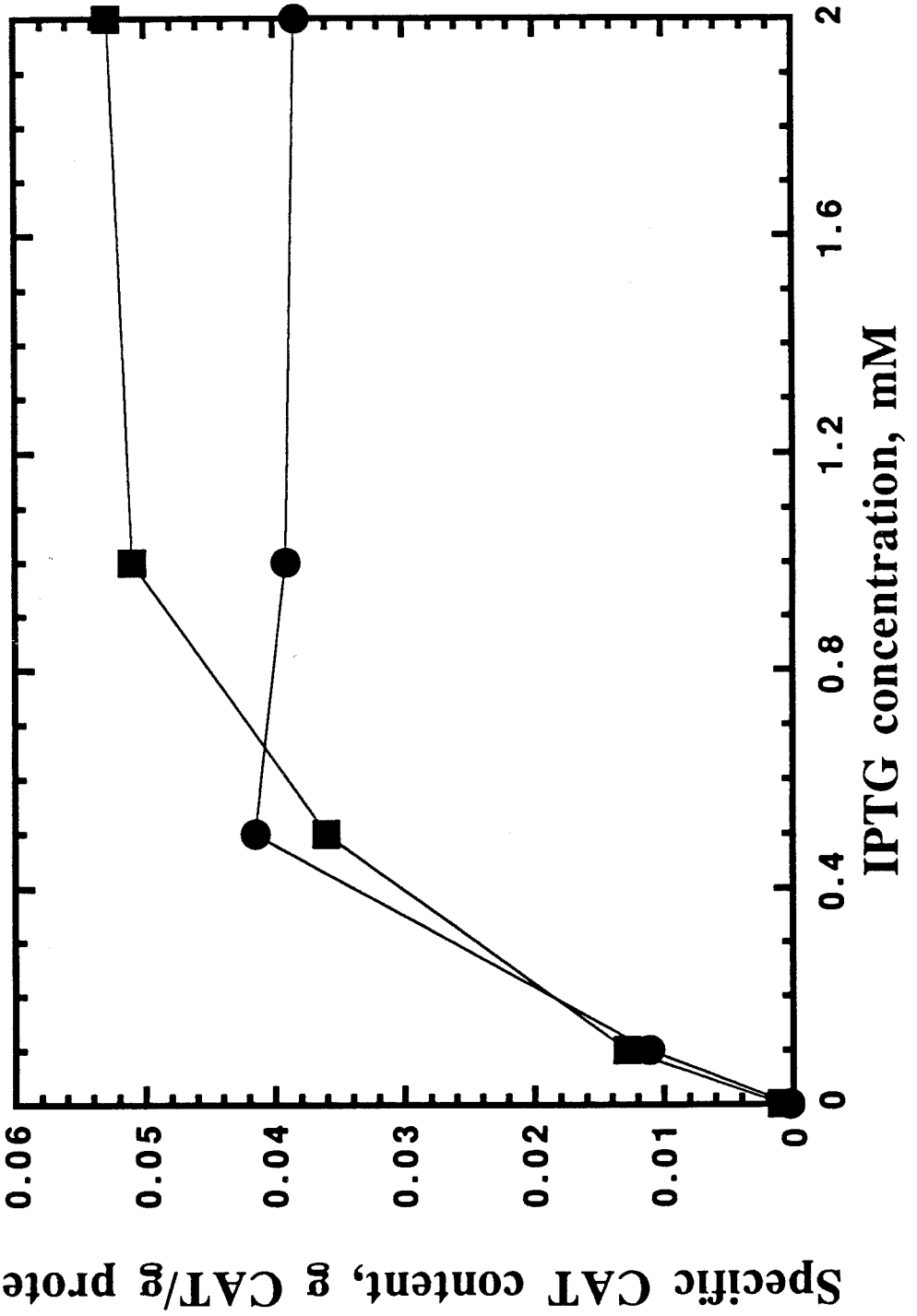
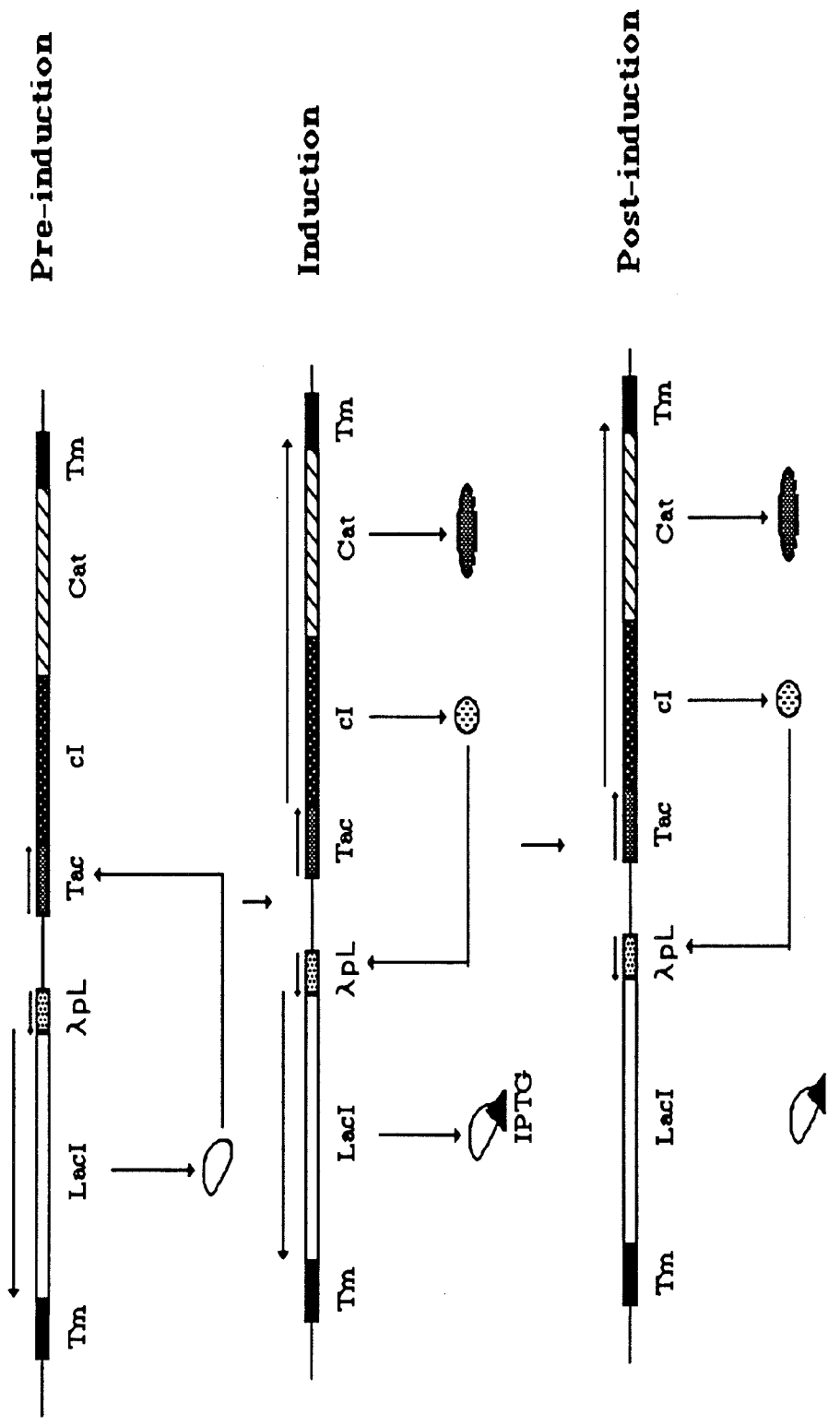

Specific CAT content, g CAT/g protein
0.06
0.05
0.04
0.03
0.02
0.01
0
0
0.4
0.8
1.2
1.6
2
IPTG concentration, mM

Figure 9.


Pre-induction
Induction
Post-induction
Tm
LacI
λpL
Tac
cI
Cat
Tm
IPTG

CHAPTER 4

FACTORS INFLUENCING RECOMBINANT PROTEIN YIELD IN *E.COLI* USING A NOVEL CROSS-REGULATION SYSTEM

4.1 Abstract

A novel cross-regulation expression system was employed for the production of a model recombinant protein, chloramphenicol acetyltransferase (CAT). In this study, the effect of induction on cell growth and the rate of CAT production is determined by inducing CAT expression with IPTG at various culture times. Our results suggest that induction at the mid-exponential growth phase provides the best compromise between cell growth and CAT production. Batch fermentation results indicate that the CAT production process is limited at the transcriptional level. A cease in the CAT production as culture growth shifts into the stationary phase coincides with a corresponding decrease in the CAT mRNA level. This limitation can be resolved by extending cell growth either by employing a fed-batch fermentation mode or by using an unmutagenized *E. coli* strain. Additional Northern blot analysis supports the previous hypothesis that a change in the transcriptional steady state is achieved in the cross-regulation expression construct. The *lacI* mRNA level decreases at least three fold after induction

4.2 Introduction

Since the early development of recombinant DNA technology, *E. coli* has been thoroughly used as a host for high-level expression of recombinant proteins. Some of the most important examples include human insulin[10] and human growth hormone[9]. Large-scale production of useful proteins is typically achieved with a two-stage process. In the first stage of such a process, cells are grown to a high cell density under reasonable balanced growth conditions. This requires that synthesis of the recombinant protein must be minimized. This growth stage is followed by a second stage in which high-level expression of the recombinant protein is achieved. The ability to maintain good regulation of recombinant protein expression is a crucial determinant of process productivity.

Several different promoter-repressor systems have been developed during the past decade to control the expression of recombinant proteins in *E. coli*. These include the *lac*[21], *trp*[19], *tac*[6], λP_L[15], and T7[18] phage promoters. In choosing a suitable promoter, one must consider the following characteristics: 1) The promoter must be regulated very tightly so that transcription is minimized until the promoter is intentionally switched on during the production phase. This is essential for expression of toxic proteins; 2) The promoter when activated should provide a high maximum level of transcription since the rate of protein synthesis is roughly proportional to the steady-state concentration of the corresponding mRNA.[7] Moreover, the inducibility of the promoter defined here as the ratio of the maximum to the minimum activity of the promoter must be maximized; 3) The method used to initiate transcription should be economical and easy for large-scale production.

Recently, a novel cross-regulation system employing a dual-repressor control

configuration has been proposed based on molecular-level mathematical modeling.[4] Simulation results indicated that this system has the ability to provide all these characteristics. This system not only retains excellent control of basal expression, but it also achieves a high level of induced expression. In addition, experimental characterization of the cross-regulation system using the *tac-lacI* and λP_L-*cI* promoter-repressor systems as examples has been conducted with results qualitatively consistent with simulation predictions.[5] In order to apply this new expression system for recombinant protein production in *E. coli*, we studied the factors influencing recombinant protein yield using the cross-regulation system. The correlation between message level and expression is investigated by determining the mRNA levels using Northern Blot analysis. Results reported here using chloramphenicol acetyltransferase (CAT) production as a model system show that the cross-regulation system provides a convenient method for tightly regulated, high level expression of recombinant proteins.

4.3 Materials and Methods

Bacterial strains and plasmid

Escherichia coli strains HB101 (*supE44, hsdS20*(r_B-m_B-), *recA13, ara-14, proA2, lacY1, galK2, rpsL20, xyl-5, mtl-1*) and MG1655 (λ-,F-) which is an unmutagenized *E. coli* K12 strain obtained from Cold Spring Harbor Laboratory were used in all experiments. Plasmid pCRR248 (Figure 1) which utilizes the cross-regulation system for regulated expression of chloramphenicol acetyltransferase (CAT) was used in all experiments.

Construction of plasmid pCRR248

Plasmid pCRR248 was constructed by transferring the cross-regulation expression cassette from plasmid pKC7.[5] This was accomplished by inserting a 3.4 kb *Sac*II/*Nco*I fragment from pKC7 into the same sites of plasmid pCS248[5] to create pCRR248.

Media and growth conditions

Both LB medium (10 g/L Difco tryptone, 5 g/L Difco yeast extract, 10 g/L NaCl, 3 g/L K_2HPO_4 and 1 g/L KH_2PO_4, pH 7.0) and M9 medium (6 g/L Na_2HPO_4, 3 g/L KH_2PO_4, 1 g/L NH_4Cl, 2.5 g/L NaCl, 3 mg/L $CaCl_2$, 0.1% thiamine, 0.2 mL of 1M $MgSO_4.7H_2O$ and 1 mL of 20% Casamino acids) were used. For all experiments, media were supplemented with 0.2% glucose as the carbon source. 50 mg/L ampicillin was added for selection. Shake flask experiments were carried out at 275 rpm in a New Brunswick INNOVA 4000 incubator shaker at 37^{0}C. Batch and fed-batch fermentations were carried out in a BiofloIII fermentor (New Brunswick Scientific) with a working volume of 2.5 L at 37^{0}C and pH 7.0. The inoculum (1:50) was grown in 100 mL of the same medium in a 250 mL flask for approximately 16 hr. For the induction of CAT production, 1 mM IPTG was added unless otherwise described. Details of feeding protocols are described in the captions to figures.

Chemicals, reagents, and DNA manipulations

All restriction endonucleases, modifying enzymes (T4 DNA polymerase, Klenow fragment, T4 DNA Ligase) and IPTG were purchased from either New England BioLabs or Boehringer Mannheim Biochemicals. ^{14}C-labeled butyryl coenzymeA was obtained form New England Nuclear and ^{32}P-dCTP was purchased from Amersham. All DNA

manipulations were done according to standard methods.[16] DNA fragments were eluted from agarose gels using a Geneclean Kit (Bio 101).

Protein and CAT assays

Cells were harvested by centrifugation and disrupted by sonication. The soluble fraction was used for protein and CAT analysis. Total protein concentration was determined using a Sigma kit (No. P5656). CAT activity was determined with ^{14}C-labeled butyryl coenzyme A (New England Nuclear) according to recommended protocols.[14] For this assay, 1 µl of cell extract was added to a 7 ml glass miniscintillation vial (Kimble) containing sufficient 100 mM Tris-HCl (pH 7.8) to give a total volume of 50 µl. 200 µl of freshly prepared 1.25 mM chloramphenicol in 100 mM Tris-HCl (pH 7.8) and 0.1 µCi of ^{14}C-Butyryl CoA were added. 5 mL of a scintillation fluor (Econofluor) was gently overlaid on top of the reaction mixture. The vial was counted in a liquid scintillation counter (Beckman model LS5801) and the CAT activity was calculated from the slope of the cpm versus incubation time data. CAT activity is expressed in units of CAT per milligram of total soluble protein. Total CAT content (g CAT) was measured by a CAT ELISA kit obtained from 5 Prime, 3 Prime, Inc.

RNA analysis

Total ribonucleic acid (RNA) form *E. coli* was obtained as described by Ausubel et al.[8] 1 mL of cells was harvested by centrifuging for 2 minutes at 4^{0}C. The cell pellet was then 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 (ATA), Sigma) and 8 µL of 50 mg/mL lysozyme. Protoplasts were collected by centrifuging at 6600 X g for 5 min and the pellet was resuspended and incubated at 37^{0}C

for 5 min in 250 μL of gram- lysing buffer (10 mM Tris-HCl, pH 8.0, 10 mM NaCl, 1 mM Na-citrate and 1.5% SDS) with 7.5 μL DEPC. 125 μL of saturated NaCl solution was added and the mixture was incubated on ice for 10 min. The insoluble fraction was removed by centrifuging for 10 min at 4°C and the supernatant was transferred to a new tube. 2.5 volumes of ethanol were added and the RNA was precipitated for 30 min at -20°C

A 736 bp *Eco*RI/*Pst*I fragment of the *cI* gene and a 1.2 kb *Eco*RI/*Pst*I *lacI* fragment isolated from plasmid pKC7 were used as probes to analyze total CAT and *lacI* mRNA levels, respectively. This particular *cI* probe should allow detection of any non-mature transcripts arising from the synthetic operon because the *cat* gene is co-transcribed with the *cI* gene. DNA probes were isolated from a 0.8% agarose gel and purified using a Geneclean kit (Bio101). Probes were radiolabeled by random primed labeling with [^{32}P]dCTP (Boehringer Mannheim). Label DNA was purified from unincorporated radioactive nucleotides using NENSORB 20 cartridges (Dupont).

Gel electrophoresis and hybridization of RNA

Approximately 20-30 μg of total RNA were dissolved in 50 μL H_2O. 5 μL of each sample was mixed with 2 μL 5 X formaldehyde gel-running buffer (0.1 M MOPS (pH 7.0), 40 mM sodium acetate and 5 mM EDTA (pH 8.0)), 3.5 μL formaldehyde and 10 μL formamide. Samples were then heated at 65°C for 15 min and immediately chilled on ice. 2 μL of formaldehyde gel-loading buffer (50% glycerol, 1 mM EDTA (pH 8.0) and 0.25% bromophenol blue) was added, and the samples were loaded on an 1% agarose gel containing 1 X formaldehyde running buffer and 2.2 M formaldehyde. The gel was run at 4 V/cm for 5 hr. RNA was then transferred from the gel onto a nitrocellulose membrane in 20 X SSC for 16 hr.[17] After transfer, the membrane was

dried at 80^{0}C for 2 hr. Then 20 mL of prehybridization solution (50% formamide, 5 X SSPE, 2 X Dehardt's reagent and 0.1% SDS) was added to the membrane and allowed to incubate at 48^{0}C for 2 hr. The DNA probe (5x10^6 cpm) was then added to the membrane and incubated overnight at 42^{0}C. After hybridization, the membrane was washed once with 1 X SSC, 0.1% SDS at 20^{0}C, followed by three washes in 0.2% SSC, 0.1%SDS at 68^{0}C for 20 min.

Quantitation of CAT and *lacI* mRNA levels

Quantitation of CAT and *lacI* mRNA levels was done as follows. Total RNA isolated were run on a 1% agarose gel. The gel was electrophoresed for 3 hr at 80V and was subsequently stained for 40 min with 1 mg/L ethidium bromide and destained 2 X 20 min with water. The gel was illuminated with a Chromato-vue transilluminator (Ultraviolet Products, INC.), photographed with Polaroid type 665 film, and the negatives were developed according to the manufacturer's instructions. The negatives were scanned with a Molecular Dynamics scanning densitometer. Autoradiograms obtained form the CAT and *lacI* mRNA analysis were quantified by scanning with a Molecular Dynamics scanning densitometer. Counts thus obtained were normalized with total RNA counts in order to obtain a measure of specific gene activity.

4.4 Results

Effect of induction on cell growth and CAT production

It is well known that high expression of recombinant protein can lead to a reduction in cell growth and protein synthesis capacity.[3] Clearly, the amount of product synthesized by the recombinant culture is maximized by choosing an optimum operating

strategy. Previous modeling[2,17] and experimental[1] results have indicated that one of the most important factors in determining recombinant protein yield is the time of induction. The effect of induction time on culture growth and CAT production using the cross-regulation system was investigated using both LB and M9 medium supplemented with 0.2% glucose as the carbon source. Investigation of response to different induction times is particularly important for the cross-regulation system since our previous results showed that full activation of this system occurs approximately 30 min after addition of inducer.[5]

In order to study the effect of induction time on CAT production, shake flask experiments were carried out in which cultures were induced with 1 mM IPTG at various stages of growth. Cell density and normalized CAT activity as a function of induction times are listed in Table I. These results clearly indicate that during the growth phase inducing CAT production early has a large detrimental effect on cell growth leading to a dramatic reduction in the cell density. However, induction at this time point also provides the highest specific CAT activity. Considering that the most important parameter in industrial recombinant protein production is to maximize the volumetric productivity (g protein/L), which depends on both the cell density and specific productivity, our results suggest that induction at approximately 2-3 hr after the start of exponential growth offers the best compromise between maintaining cell growth and CAT production. These results are in good agreement with simulation predictions by Seressiotis and Bailey[17] and Bentley and Kompala.[2] For an efficient expression system such as the cross-regulation system, in which the addition of an inducer quantify yields an immediate and severe growth rate reduction, the best induction time is more toward the end of the batch as indicated by our results.

Regulated CAT expression in batch fermentations

The above results combined with our previous studies[5] regarding the response of the cross-regulation system to different IPTG dosages provide a basis for the development of a convenient protocol to be used with the cross-regulation system in a batch fermentation process. Batch fermentation results with LB and M9 medium are depicted in Figure 2 and 3, respectively. In the absence of IPTG, cells grown in both media show a low level of specific CAT content (in the range of 0.01% of total cellular proteins). In both cases, the maximum specific CAT content following induction by IPTG is at least 60-fold greater than those obtained before induction. However, the specific CAT content eventually stops increasing and levels out 5 hr after induction. This coincides with the shifting of cell growth from the exponential to the stationary phase. It is reasonable to speculate that some key factors involved in the CAT production process may be rate-limiting after this transition. In order to investigate whether this limitation in CAT production occurs at the CAT mRNA level, we performed Northern blot analysis to quantify the CAT mRNA levels. As depicted in Figure 4, the CAT mRNA content gradually increases after induction until cells shift into the stationary phase after which it actually declines. This result suggests that transcription of the CAT mRNA may be rate-limiting, although the possibility of mRNA degradation cannot be ignored.

Transcriptional analysis of the cross-regulation system

Another useful information obtained from mRNA assays is the correlation between CAT mRNA level and CAT activity for the cross-regulation system. One of the major benefits of using the cross-regulation system is that it provides the necessary regulation of recombinant protein production at the transcription level. It is, therefore, of interest to investigate the response of the cross-regulation system to induction at the transcriptional level. As indicated in Figure 5, the CAT mRNA level is essentially zero

pre-induction. This confirms that a low basal expression of CAT post-induction is due to a tight control of transcription but not some artifacts (such as a deficiency in translation) of our system. Furthermore, an increase in CAT production matches well with a corresponding increase in the CAT mRNA level, showing that CAT production is the direct result of a release of transcriptional control. It should be noted that, as the level of CAT mRNA drops about 90% to almost the pre-induction level after transition into the stationary phase, no further increase in the CAT level is observed. This strongly suggests that by maintaining the CAT mRNA level, we should be able to further increase the CAT production.

One of the major reason why the cross-regulation system should offer better induced expression over commonly used system is that induction provides a signal (production of the *cI* repressor) which cross-regulates and turns off transcription of the *lacI* gene. This behavior has been previously demonstrated at the protein level.[5] Here we explore this change at the mRNA level. This is illustrated in Figure 6. As indicated, a major band corresponding to the *lacI* mRNA is visible for the pre-induction samples (lanes 1, 2, and 3). The intensity of this band gradually decreases with time post-induction and eventually declines to almost zero at the end of cultivation. This is exactly the behavior that we are expecting since a single copy of the *lacI* gene (under control of its own promoter) is still present in the chromosome.

Regulated CAT expression in fed-batch fermentations

Our batch fermentation results suggest that the production of CAT can be increased by extending cell growth. This was accomplished by employing a fed-batch fermentation mode. Cells were grown under regular batch mode for the first 5 hr. Thereafter, a feed medium consisting of 5 X M9 and 20% glucose was fed in at specific

feeding rate (Figure 7a). The continued feeding of nutrients enables the culture to reach a much higher cell density compared to that previously attained from batch fermentation (final O.D. of 2.7 to 1.5). More importantly, cells are able to maintain transcription of the CAT mRNA throughout the course of the fermentation (Figure 7b). Consequently, there is also a corresponding increase in the CAT production (two-fold) compared to results obtained from the batch fermentation (Figure 7a). A very similar behavior was also observed using LB medium with the same feed strategy (data not shown).

Another approach that can be used to improve cell growth is to employ an unmutagenized *E. coli* strain as a production host. Although the strain MG1655 does not contain a *recA* mutation which could lead to plasmid instability problems, current advance in genetic technique such as the introduction of a *par* stabilizing locus has been shown to be effective in increasing plasmid stability.[20] As indicated in Figure 8 wild-type strain MG1655 carrying the plasmid pCRR248 reaches a final O.D. of 8.17 using M9 medium and the same feeding strategy mentioned earlier. This is approximately three fold higher than what we obtained using the mutagenized *E. coli* strain HB101. Surprisingly, the specific CAT content reaches only about 7% of the total cellular proteins which is approximately 2/3 that from strain HB101. However, this is compensated by a 3-fold increase in the final cell density; thus, the final volumetric CAT productivity is approximately 2-fold higher.

4.5 Discussion

Induction strategy in batch cultivations of recombinant cells with a strong regulated promoter has a large effect on process performance. Early induction offers the benefit of allowing product synthesis for a longer period of time with the tradeoff of hampering the cell's own metabolism. This is clearly illustrated by a major reduction in the cell density from our early induced culture (Table I). On the other hand, since the induction occurs very early in the growth phase, cells have more time to reach a higher specific CAT content. Examination of results from Table I indicates that the maximum CAT productivity is obtained by inducing cells in the mid-exponential phase which allows sufficient opportunity for CAT synthesis and also greater overall cell growth.

From previous studies with α-amylase production in *Bacillus subtilis*, it appears that the main reason for the ceased α-amylase production at the end of the growth phase is a 10 fold decrease in the mRNA levels.[11] A very similar connection is obtained in our batch fermentation results. It is well known that upon nutrients limitation, most organisms have the ability to alter their metabolism in order to adapt to this new environment. To accomplish that, new enzymes are synthesized in order to carry out different types of metabolic reaction. Examples such as the responses to carbon[12] and nitrogen[13] starvation in *E. coli* are well documented. Many proteins that are synthesized under normal growth condition will be turned off and only those presumably necessary for cell maintenance will be produced. This may explain why further transcription of the CAT gene is greatly reduced once cells enter the stationary phase.

Our batch fermentation results suggest that a decrease in the CAT mRNA level is rate limiting in the CAT production process. Since this decrease in the CAT mRNA level correlates with the shifting of culture growth into the stationary phase, one possible solution is to extend cell growth such that cells can maintain their CAT mRNA level.

Two of such methods have been applied in this work. First, growing cells under the fed-batch mode improves not only cell growth but also the CAT production level. A CAT mRNA analysis reveals that the CAT mRNA level remains at a much higher level than pre-induction throughout. Consequently, the final specific CAT content reaches a level of 9% of the total cell proteins (a two fold improvement over batch fermentation).

An alterative method to extend cell growth is by employing the unmutagenized *E. coli* strain MG1655. Since this strain does not contain all the mutations necessary for DNA cloning, it is not unreasonable to expect that it should be a more suitable production host. Our results support this hypothesis showing that this strain enjoys a better growth advantage than strain HB101 with an approximately 3-fold increase in the final cell density. Furthermore, using this type of unmutagenized strain has been demonstrated here to provide a two-fold increase in the volumetric CAT production.

Most of the useful regulated promoters for cloned-gene expression in *E. coli* provide regulation at the transcription level. Before one could apply these systems for practical recombinant protein production, it is of great importance to fully understand each system at the transcription level. This is particularly true for the cross-regulation system which involves complex transcriptional controls. An analysis of the CAT mRNA levels confirms the predicted transcriptional behavior of the cross-regulation system with a low message level before induction and a high message level post-induction. Moreover, the CAT protein activity is roughly proportional to the CAT mRNA level, suggesting that transcription but not translation is rate limiting. In addition, we have also demonstrated that the amount of *lacI* mRNA decreases by more than 3-fold after induction providing evidence that a shift in the transcriptional steady state from the λP_L to the *tac* promoter has indeed been accomplished. Validation of this result provides the basis for another potential application of the cross-regulation system as a transcriptional switch (metabolic

switch) such that different proteins (pathways) can be turned on under different conditions. We are currently pursuing this direction in our laboratory. Results presented here for the cross-regulation system pertain specifically to the production of the CAT protein. Trends may well differ for other constructs expressing other proteins.

4.6 Acknowledgements

This work was supported by the Advancement Industrial Concepts Division of the U.S. Department of Energy. W. Chen was supported in part by a Predoctoral Training Grant in Biotechnology from the National Institute of General Medical Sciences (National Research Service Award 5 T32 GM 08346-04, Pharmacology Science Program).

4.7 References

1. **Bentley, W. E., Davis, R.H. & Kompala, D.S.** (1991). Dynamics of induced CAT expression in *E. coli*. *Biotechnol. Bioeng.* **38**, 749-760.
2. **Bentley, W.E. & Kompala, D.S.** (1991). Optimal induciton of protein synthesis in recombinant bacterial cultures, *Ann. N.Y. Acad. Sci.* 121-138.
3. **Bentley, W.E., Morkajalili, N., Anderson, D.C., Davis, R.H. & Kompala, D.S.** (1990). Plasmid-encoded protein: The principal factor in the "Metabolic burden" associated with recombinant bacteria. *Biotechnol. Bioeng.* **35**, 668-681.
4. **Chen, W., Bailey, J.E. & Lee, S.B.** (1991). Molecular design of expression systems: Comparison of different repressor control configurations using molecular mechanism models. *Biotechnol. Bioeng.* **38**, 679-687.
5. **Chen, W., Kallio, P.T. & Bailey, J.E.** (1992). Construction and characterization of a novel cross-regulation system for regulating cloned gene expression in Escherichia coli. *Gene*, **Submitted.**
6. **de Boer, H.A., Comstock, L.J. & Vasser, M.** (1983). The *tac* promoter: a functional hybrid derived form the *trp* and *lac* promoters. *Proc. Natl. Acad. Sci. USA*. **80**, 21-25.
7. **Georgiou, G.** (1988). Optimizing the production of recombinant proteins in microorganisms. *AICHE J.* **34**, 1233-1247.
8. **Gilman, M.** (1989). Isolation of RNA from gram-negative bacterium. p. 146-148. In: F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, K. Struhl (eds.)., Short protocols in molecular biology, Greene Publishing Associates and Wiley-Interscience, NY.

9. **Goeddel, D., Heyneker, H., Hozumi, T., Arentzen, R., Itakura, K., Tansura, D. Ross, M., Miozzari, G., Crea, R. & Seeburg, P.** (1979). Direct expression in *Escherichia coli* of a DNA sequence coding for human growth hormone, *Nature*, **281**, 544-548.

10. **Johnson, I.S.** (1983). Human insulin from recombinant DNA technology, *Science*, **219**, 632-637.

11. **Kallio, P., Palva, A. & Palva, I.** (1987). Enchancement of α-amylase production by integrating and amplifying the α-amylase gene of *Bacillus amyloliquefaciens* in the genome of *Bacillus subtilis*. *Appl. Microbiol. Biotechnol.* **27**, 64-71.

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

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

14. **Newman, B.M., Morency, C.A. & Russian, K.O.** (1987). A novel rapid assay for chloramphenicol acetyltransferase gene expression. *Biotechniques*, **7**, 444-447.

15. **Remaut, E., Stanssens, P. & Fiers, W.** (1981). Plasmid vectors for high-efficiency expression controlled by the P_L promoter of coliphage lambda. *Gene*, **15**, 81-89.

16. **Sambrook, J., Fritsch, E.F. & Maniatis, F.E.** (1989). Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

17. **Seressiotis, A. & Bailey, J.E.** (1987). Optimal gene expression and amplification strategies for batch and continuous recombinant cultures. *Biotechnol.*

Bioeng. **29**, 392-398.

18. **Tabor, S. & 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.

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

20. **Wood, T.K., Kuhn, R.H. & Peretti, S.W.** (1990). Enhanced plasmid stability through post-segregational killing of plasmid-free cells. *Biotechnol. Technol.* **4**, 36-41.

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

4.8 Tables

Table I. Effect of induction time on CAT production and cell growth.

Induction time, hr	O.D. after 7 hr of growth	Normalized CAT activity[1]
LB		
No induction	8.16	0
2	6.30	1
3	7.27	0.625
4	7.32	0.380
5	7.33	0.320
6	7.01	0.325
M9		
No induction	1.45	0
3	0.98	1
4.5	1.33	1.013
5.5	1.32	0.771
6.5	1.17	0.399

[1]Normalized CAT activity was determined by using the specific CAT activity at the earliest induction time point as the basis.

4.9 Figures

Figure 1. Plasmid map of pCRR248 carrying the cross-regulation system for regulated CAT expression used in this work.

Figure 2. Specific CAT content (●) and optical density (O) for strain HB101/pCRR248 grown in LB medium supplemented with 0.2 % glucose. Batch fermentation were carried out in a New Brunswick BiofloIII fermentor at 37^{0}C and pH 7.0. A constant air flow rate of 1 L/min and an agitator speed of 300 rpm were maintained throughout.

Figure 3. Specific CAT content (●) and optical density (O) for strain HB101/pCRR248 grown in M9 medium supplemented with 0.2 % glucose. Batch fermentation were carried out in a New Brunswick BiofloIII fermentor at 37^{0}C and pH 7.0. A constant air flow rate of 1 L/min and an agitator speed of 300 rpm were maintained throughout.

Figure 4. Northern blot analysis of the CAT mRNA levels during the batch fermentation with LB medium + 0.2% glucose. Lanes are loaded with samples obtained at different fermentation time points (Lanes 1-7 correspond to samples from 1, 2, 3, 4, 5, 6, and 9 hr, respectively).

Figure 5. Correlation between the CAT mRNA (●) and protein (O) levels for the batch fermentation with LB medium + 0.2% glucose. Normalized RNA content was obtained by using the highest RNA content as the basis.

Figure 6. Transcriptional analysis of the *lacI* mRNA levels. This is an autoradiogram obtained from the LB + 0.2% glucose batch fermentation. Lanes 1-3 correspond to time points 2 hr, 1 hr, and 0 hr before induction. Lanes 4-7 correspond to 1, 2 , 3, and 4 hr post-induction.

Figure 7. a. Specific CAT content (●) and optical density (O) for strain HB101/pCRR248 grown in M9 medium supplemented with 0.2 % glucose in a fed-batch mode. Fermentation were carried out in a New Brunswick BiofloIII fermentor at 37^{0}C and pH 7.0. A constant air flow rate of 1L/min and an agitator speed of 300 rpm were maintained throughout. Feeding was commenced at 2.95 mL/hr.

b. Correlation of the CAT mRNA and protein levels for the fed-batch fermentation. Normalized RNA content was obtained as before (above). Northern blot analysis of the CAT mRNA transcription shown here for different time points post-induction (below).

Figure 8. Specific CAT content (●) and optical density (O) for strain MG1655/pCRR248 grown in M9 medium supplemented with 0.2 % glucose in a fed-batch mode. Fermentation were carried out in a New Brunswick BiofloIII fermentor at 37^{0}C and pH 7.0. A constant air flow rate of 1L/min and agitator speed of 300 rpm were maintained throughout. Feeding was commenced at 2.95 mL/hr.

Figure 1.

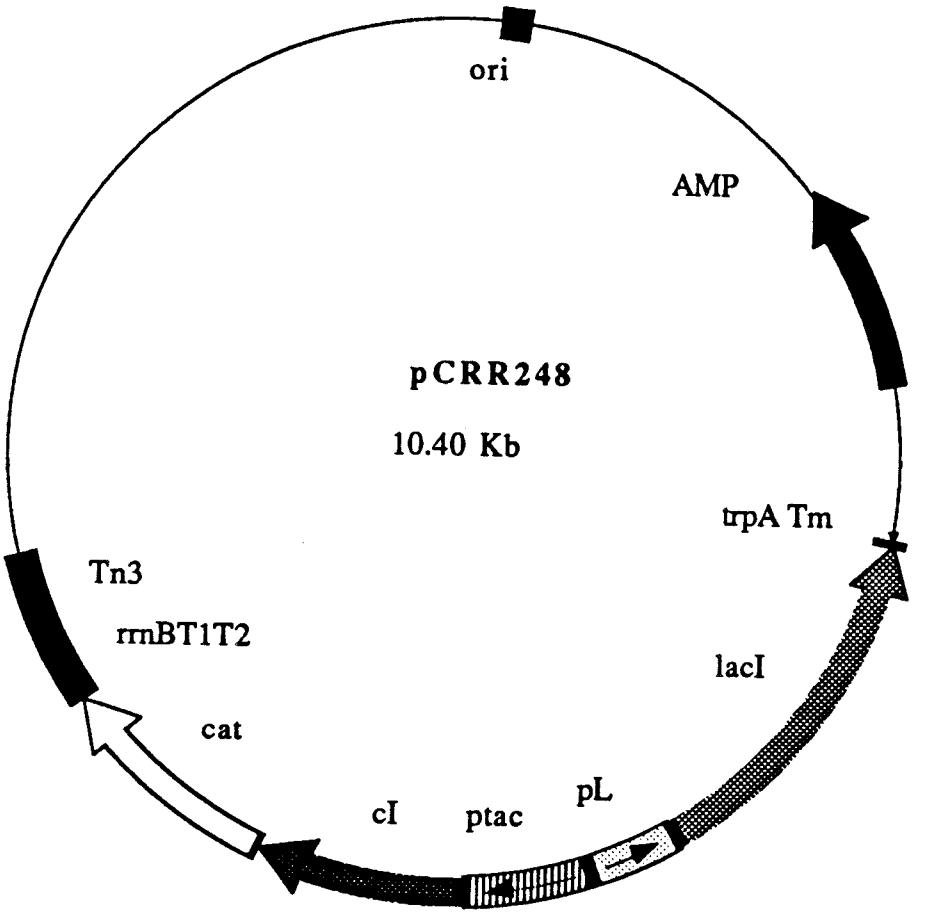

ori
AMP
pCRR248
10.40 Kb
trpA Tm
Tn3
rrnBT1T2
lacI
cat
pL
cI
ptac

Figure 2.

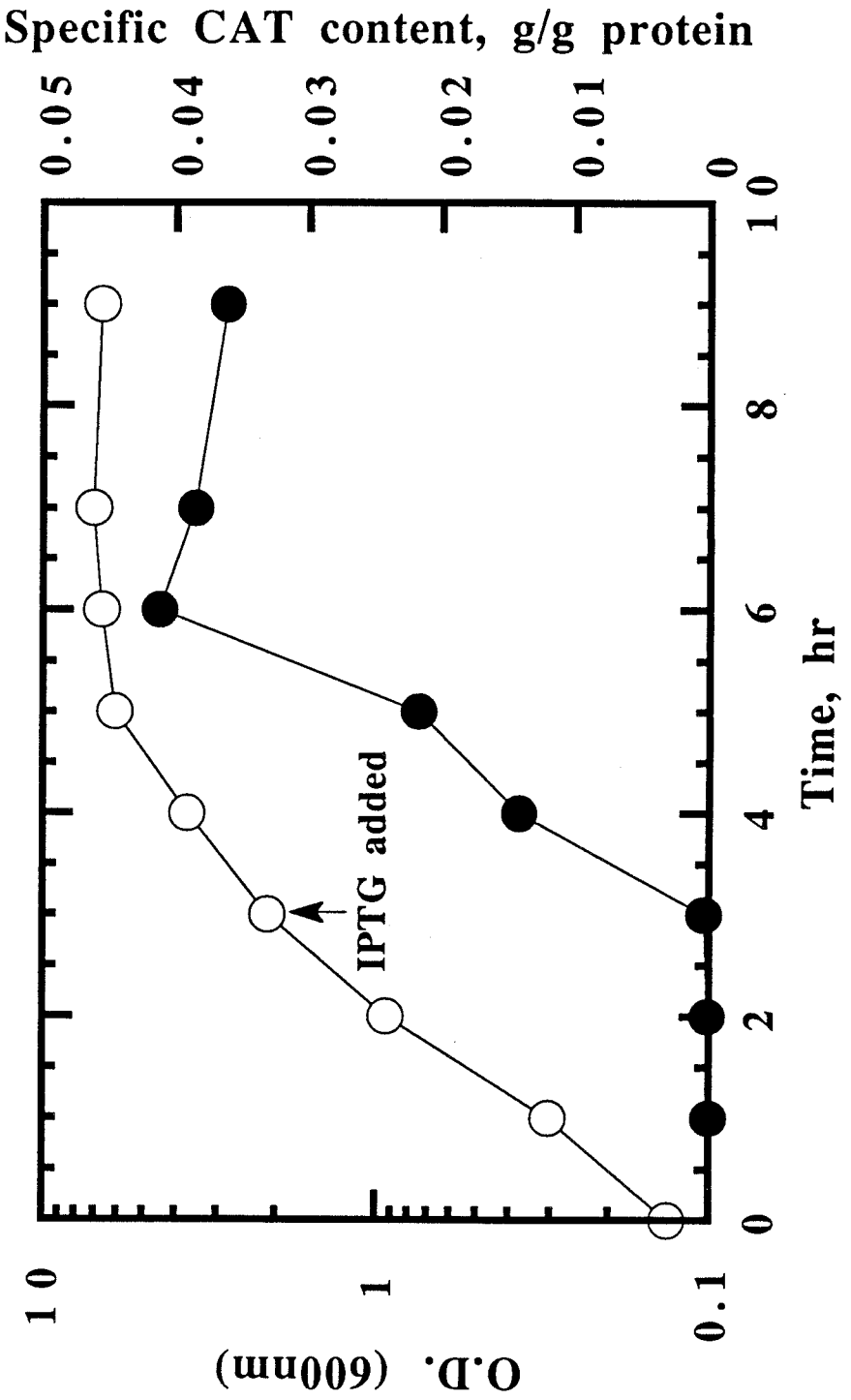

Specific CAT content, g/g protein
0.05
0.04
0.03
0.02
0.01
0
10
8
6
4
2
0
Time, hr
IPTG added
10
1
0.1
O.D. (600nm)

Figure 3.

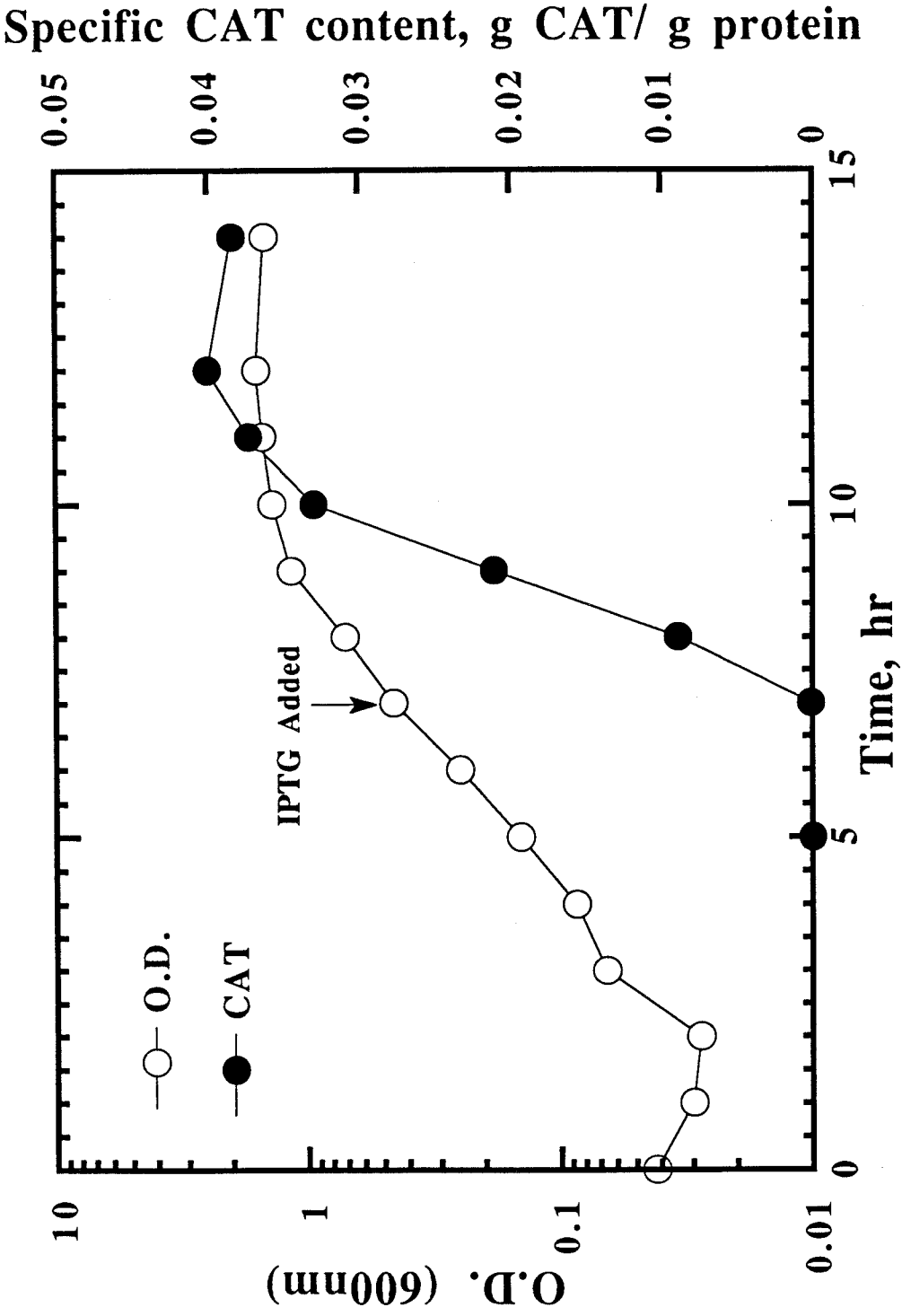

Specific CAT content, g CAT/ g protein
0.05
0.04
0.03
0.02
0.01
0
15
10
5
0
Time, hr
IPTG Added
O.D.
CAT
10
1
0.1
0.01
O.D. (600nm)

Figure 4.

Figure 5.

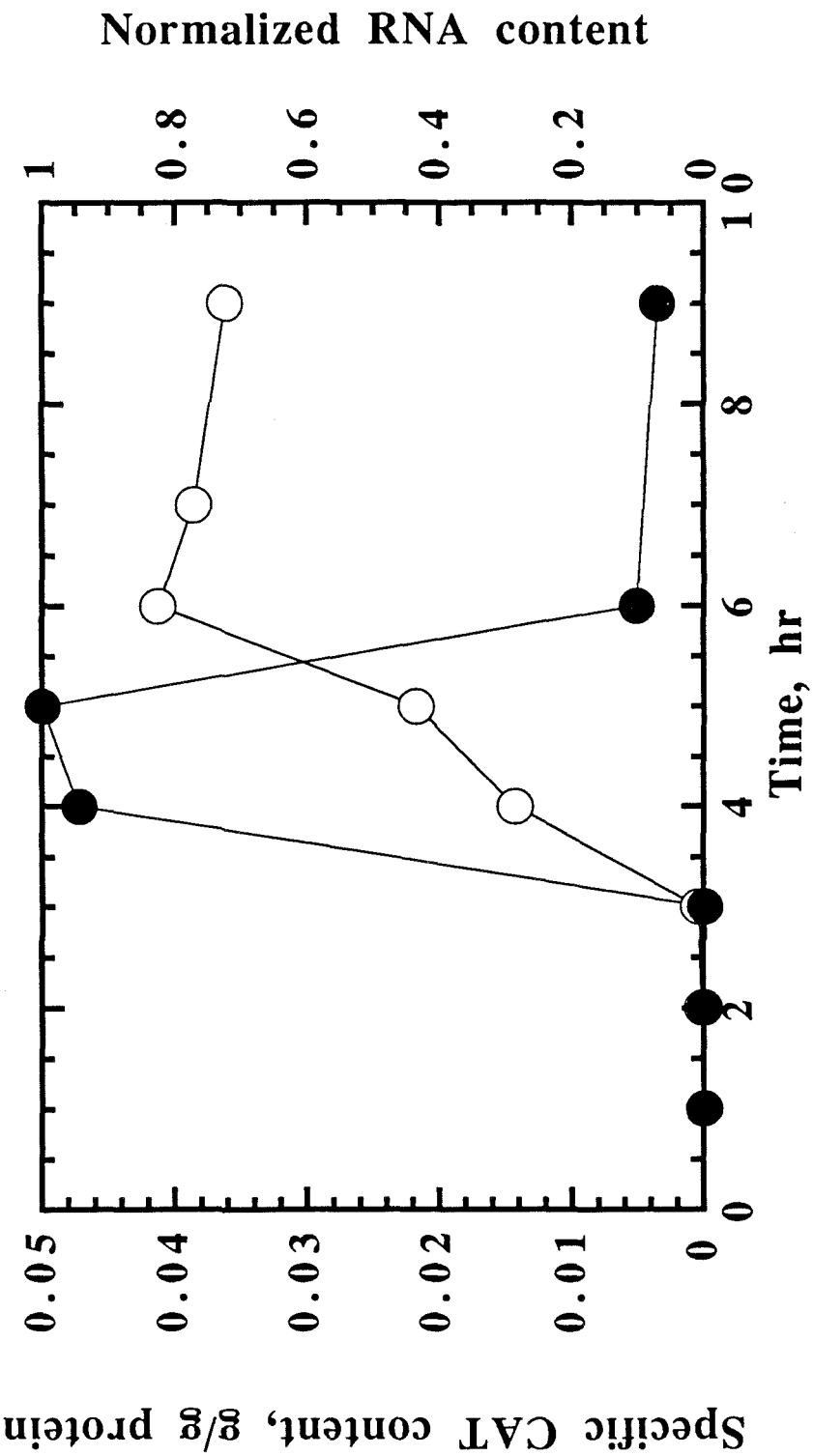

Normalized RNA content
1
0.8
0.6
0.4
0.2
0
10
8
6
4
2
0
Time, hr
0.05
0.04
0.03
0.02
0.01
0
Specific CAT content, g/g protein

Figure 6.

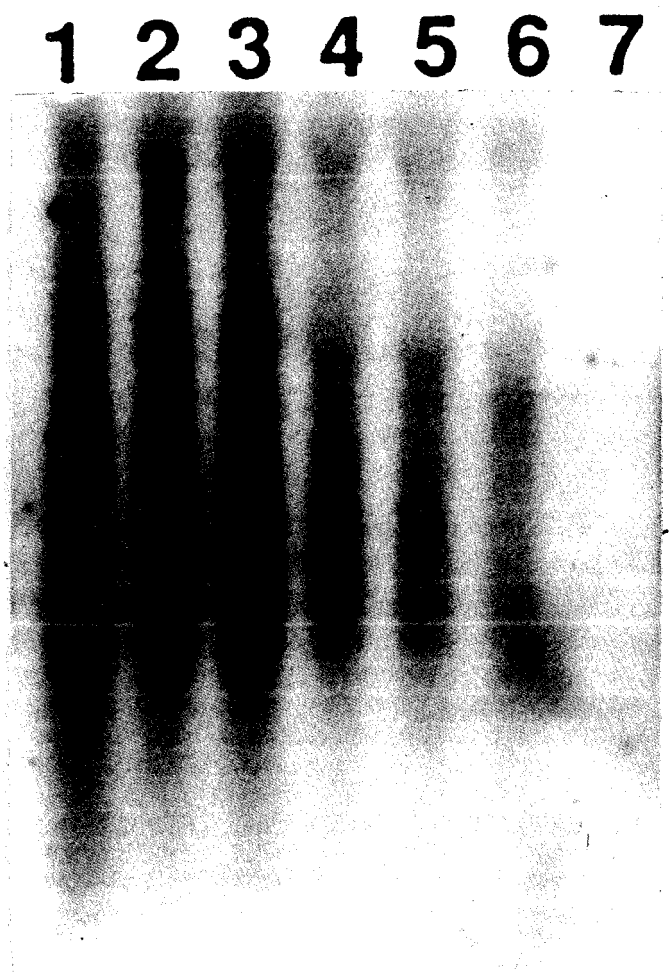


Figure 7a.

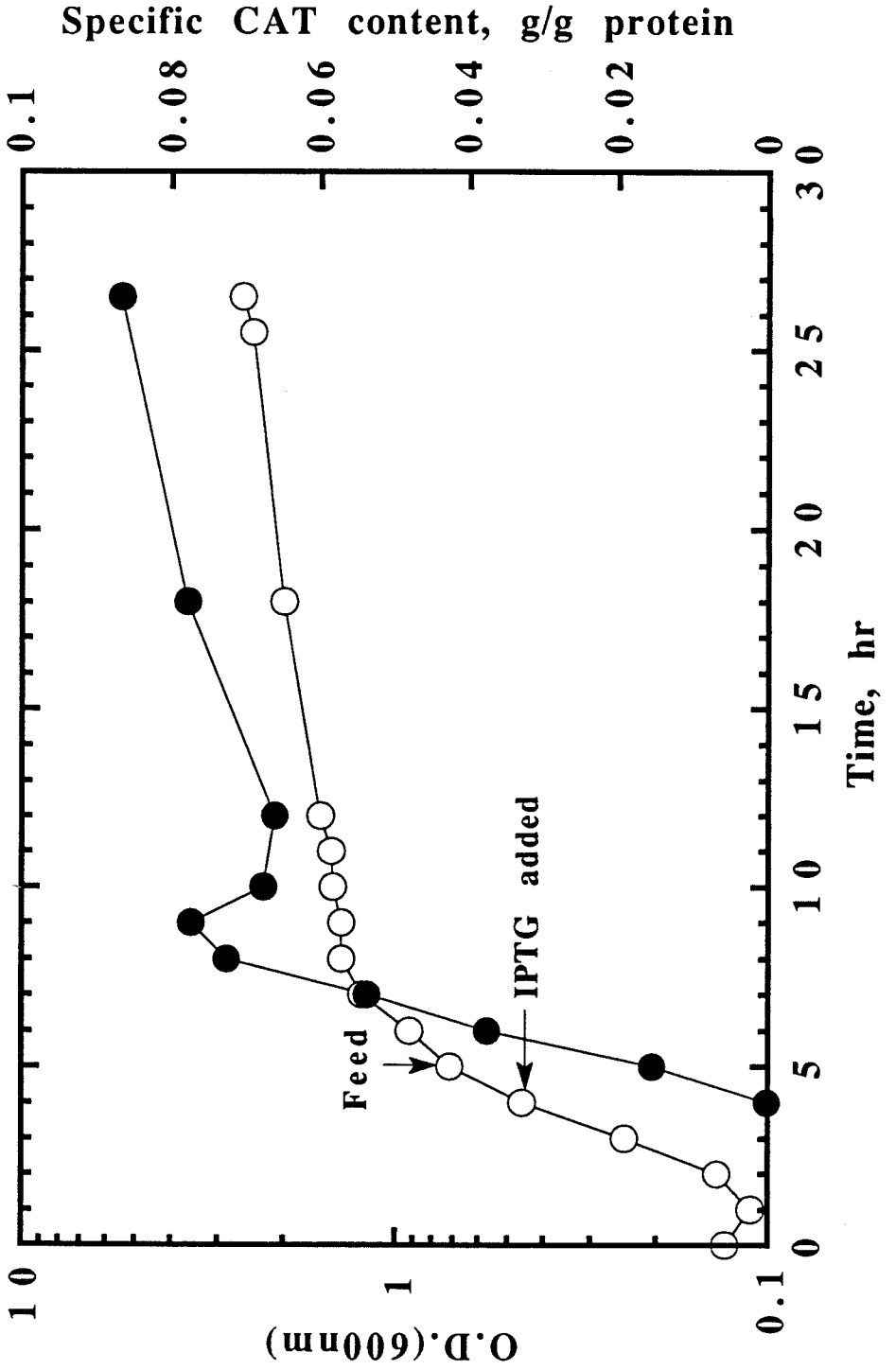
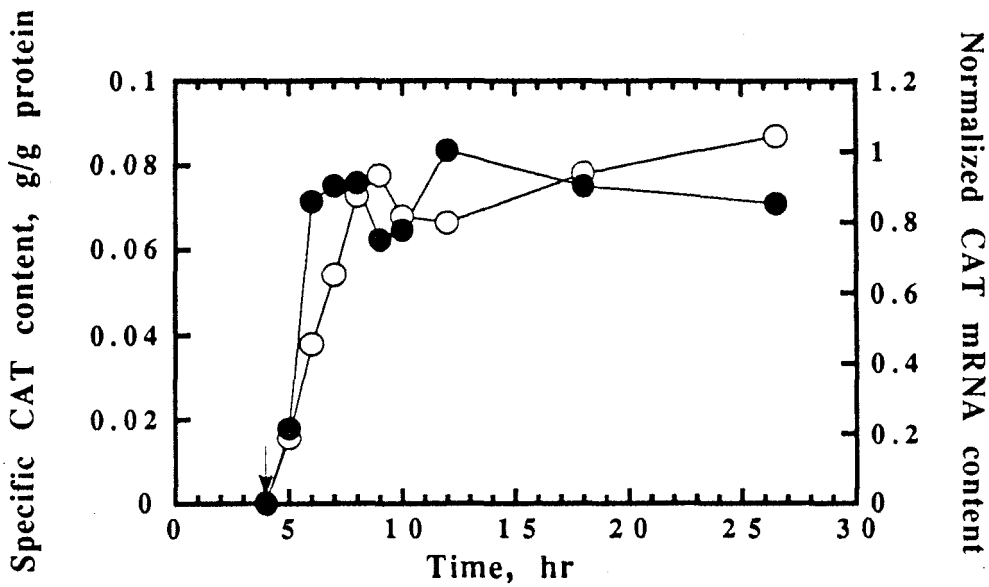

Specific CAT content, g/g protein
0.1
0.08
0.06
0.04
0.02
0
30
25
20
15
10
5
0
Time, hr
IPTG added
Feed
10
1
0.1
O.D.(600nm)

Figure 7b.


Specific CAT content, g/g protein
0.1
0.08
0.06
0.04
0.02
0
0
5
10
15
20
25
30
Time, hr
1.2
1
0.8
0.6
0.4
0.2
0
Normalized CAT mRNA content

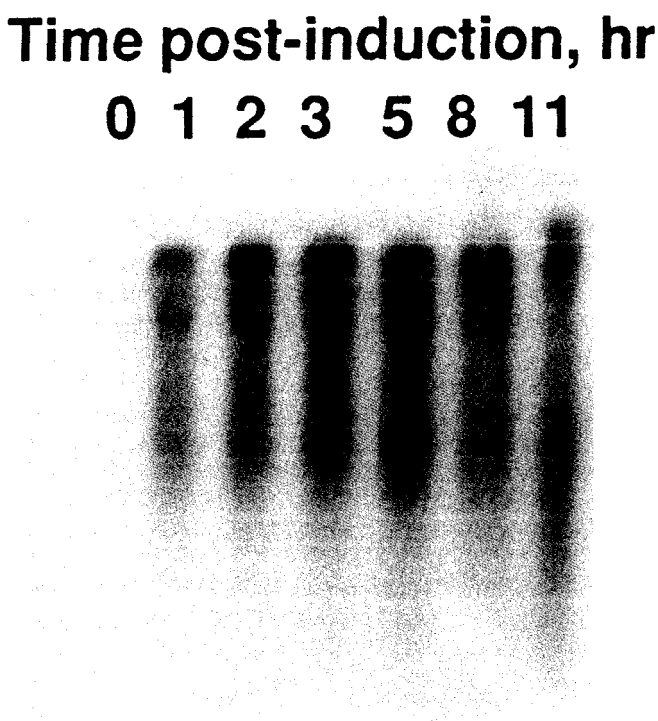
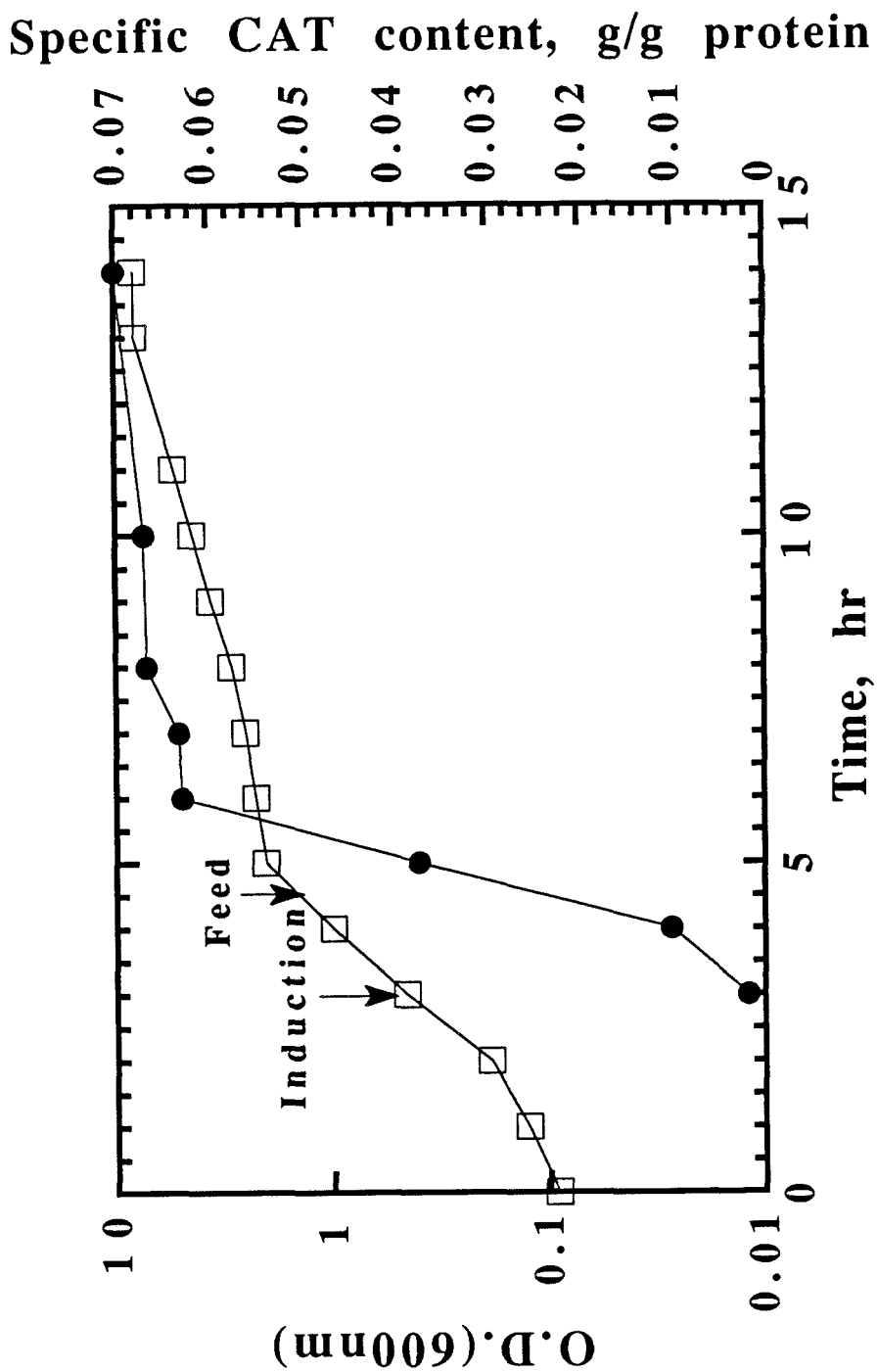

Time post-induction, hr
0 1 2 3 5 8 11

Figure 8.


Specific CAT content, g/g protein
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
15
10
5
0
Time, hr
Feed
Induction
10
1
0.1
0.01
O.D.(600nm)

CHAPTER 5

APPLICATION OF THE CROSS-REGULATION SYSTEM AS A METABOLIC SWITCH: A NOVEL WAY OF REDIRECTING METABOLIC FLUX

5.1 Summary

A novel way of redirecting metabolic flux is described by applying the cross-regulation system as a metabolic switch. This system carries a special property in which a shift in the transcriptional steady state from one promoter to another occurs upon induction. In particular, two model systems were constructed to illustrate this concept. In the first example, switching between VHb and CAT expression were demonstrated showing the validity of the metabolic switch. As a practical example, this system was applied to switch from accelerated glycogen synthesis to glycogen degradation in an *E. coli* strain carrying mutations of the genes responsible for the acetate synthesis pathway. Results indicated a 5 fold increase in glycogen synthesis preinduction and a 30-40% increase in glycogen degradation postinduction. It is expected that this concept and its extensions can be used with different combinations of promoter system and synthetic operon constructs to achieve complicated metabolic flux regulation in diverse hosts.

5.2 Introduction

The metabolic activities of living organisms are composed of highly regulated, coupled networks of reactions catalyzed by specific enzymes. Unfortunately, these naturally occurring networks are not generally optimized for practical applications. However, the performance of these organisms can be artificially improved by genetic manipulations of their metabolic networks. The ability to control flux distribution is a generally important objective. Current approaches usually involve transforming the host with the genes that encode for synthesis of the desired products. This strategy, however, does not consider that such an abrupt flux alteration could be detrimental to the host. For this reason the ability to adjust flux distribution from one configuration to another is important.

In order to achieve this goal, one must control synthesis of the enzymes catalyzing the different pathways. In this paper, our aim is to construct a new expression system to provide a metabolic switch. The term metabolic switch is defined here as the switching of metabolic flux from one pathway to another (Figure 1). A naturally occurring switch can be found in *Salmonella* in which a change from flagellin H1 to flagellin H2 takes place in order to evade the immune response of its host.[16] In order to achieve the same kind of switching using a synthetic genetic construct, the novel cross-regulation system can be applied.[2] Making use of the fact that this system can change from one transcriptional steady state to another upon induction,[4] transcription can be switched from one gene to another if theses genes are imbedded within the regulated operons. In principle any number of genes can be included in each operon, allowing the switch of even very complicated metabolic pathways.

Two examples are presented to illustrate the metabolic switch concept. First, the

Vitreoscilla hemoglobin (VHb) gene is inserted into the vector pKC7 which utilizes the cross-regulation system to regulate the production of chloramphenicol acetyltransferase (CAT).[3] Transcription of the VHb gene is under control of the λP_L promoter and cotranscribed with the *lacI* gene. There are two reasons in choosing the VHb gene: 1) It is a relatively small gene (approximate 500 bp) and restriction sites are available to make this construct easily; and 2) An activity assay and antibody are readily available. Our main goal for this first example is to demonstrate that the amount of VHb and CAT produced from the λP_L and the *tac* promoter can be altered upon induction, thus showing the validity of this concept.

The second example is a more practical application of the metabolic switch. Recently, it was shown that overproducing the glycogen synthesis enzymes in an *E. coli* mutant that is deficient in acetate production improved not only the amount of glycogen synthesized but also final density of the culture. It has been postulated that such an improvement in cell density can be attributed to greater energetic efficiency of using glycogen rather than pyruvate as the carbon source.[13] During this growth phase characterized by glycogen utilization, the rate of glycogen degradation may be rate-limiting. Growth of the culture ceases before all of the glycogen is consumed. In this example, our goal is to alter the rate of glycogen synthesis and degradation at different stages of growth. This can be accomplished by subcloning the glycogen synthesis (*glyA* and *glyC*) and degradation (*glyP*) pathway genes into the λP_L and *tac* operons, respectively. This construct should enable an increased rate of glycogen synthesis preinduction to accumulate glycogen and can be switched after glucose depletion by addition of IPTG to favor glycogen degradation.

5.3 Materials and Methods

Bacterial strains and plasmids

Escherichia coli strains DH5α (F-, *endA1, hsdR17*($r_k^-m_k^+$), *supE44, thi-1, λ-, recA1, gryA96, relA1, φ80dlacAm15*) and TA3476[10] (*his Δ(pta-ack-dhuA-hisJ-hisQ-hisP)*) were used in this work. Plasmids pPR2[14] and pGTC12[5] were used as templates for the PCR amplification of the *glgC-glgA* and *glgp* genes, respectively. Plasmid pINT1[9] was used as the source of the VHb gene. Plasmids pSL1180 (Pharmacia), pKC7[3], pTCI[3] and pλ-lacI[3] were used in constructing the two metabolic switch examples.

Media and growth conditions

For all experiments LB medium (10 g/L Difco tryptone, 5 g/L Difco yeast extract, 10 g/L NaCl, 3 g/L K_2HPO_4 and 1 g/L KH_2PO_4, pH 7.0) supplemented with 0.2% glucose as the carbon source was used. 50 mg/L ampicillin was added for selection purpose. Shake flask experiments were carried out at 275 rpm in a New Brunswick INNOVA 4000 incubator shaker at 37°C.

Chemicals, reagents, and DNA manipulations

All restriction endonucleases, modifying enzymes (T4 DNA polymerase, Klenow fragment, T4 DNA Ligase) and IPTG were purchased from either New England BioLabs or Boehringer Mannheim Biochemicals. All DNA manipulations were done according to standard methods.[1] DNA fragments were eluted from agarose gels using a Geneclean Kit (Bio 101).

Immunoblot analysis of VHb

Cell pellets were boiled for 5 min in a lysis buffer containing 10% glycerol, 5% 2-mercaptoethanol, 3.3% SDS and 0.5 M Tris, pH 6.8 and then electrophoresed on a 12.5% polyacrylamide gel according to the method of Laemmli.[10] The proteins were electrophoretically transferred to a nitrocellulose membrane as described elsewhere.[1] The proteins were screened with antiserum generated against *Vitreoscilla* hemoglobin as described elsewhere.[1] The hemoglobin standard was produced in recombinant *E. coli*.[9]

Protein and CAT assays

Cells were disrupted by sonication. The soluble fraction was used for protein and CAT analysis. Protein concentration was determined according to the method of Lowry et al.[12] using a Sigma kit (No. P5656). Total CAT content was measured by a CAT ELISA kit obtained from 5 Prime, 3 Prime, Inc.

Glycogen assay

Glycogen assay was performed according to Gunja-Smith et al.[7] Briefly, cell pellets were suspended in 2 mL of 20% potassium hydroxide solution and boiled at 100^{0}C for 1 hr. After cooling, 0.5 mL of 0.1 M sodium phosphate buffer solution (pH 6.8) was added and the suspensions were adjusted to pH 6 to 7 with 5 N hydrochloric acid. Two volumes of ethanol were added, and the resulting precipates were recovered by centrifugation, washed three times with 70% ethanol, and dried for 10 min. The resulting pellets were resuspended in 1.8 mL of water and 0.2 mL of 10X reaction buffer (5 mM calcium chloride, 50 mM sodium acetate buffer, pH 5.0, 1U α-amylase and 2U glucoamylase) and incubated at 37^{0}C for 1 hr and 65^{0}C for another 1 hr. Glycogen was then determined by the amount of glucose released using a Sigma kit (#510-A).

Polymerase chain reaction (PCR)

The PCR reaction was carried out in a 50 µl final reaction volume containing 2.5 ng of each respective template DNA, 5 µl of 10X reaction buffer (100 mM Tris-HCl pH. 8.3, 500 mM KCl, 15 mM $MgCl_2$, 0.01%(W/V) gelatin), 8 µl of dNTPS mix (1.25 mM each), 2 mM each of the primers, and 0.5 µl (1U/µl) of Taq DNA polymerase (Cetus). The amplification reaction was carried out for 36 cycles in a DNA thermal cycler (Perkin Elmer-Cetus). The DNA was denatured at 92^{0}C for 1 min, annealed at 42^{0}C for 2 min, and extended at 72^{0}C for 5 min.

5.4 Results

Example 1

Construction of the plasmid pMSW1

A 1.1 kb *Bam*HI/*Sph*I fragment containing the VHb gene was isolated from plasmid pINT1. After gel purification, this fragment was cleaved with *Hae*III and the resulting 460 bp *Bam*HI/*Hae*III fragment carrying the VHb structural gene without the transcriptional stop signal was subcloned into the *Bam*HI/*Eco*RV sites of pSL1180 to create pSLVHb. This plasmid was subsequently cut with *Bam*HI and filled in with Klenow fragment to generate a blunt end. A 600 bp fragment was isolated by cleaving with *Kpn*I. A *BspE*I partial digestion was performed on plasmid pKC7 and the fragment with only one site cut was isolated. This fragment was then cleaved with *Kpn*I and the 600 bp fragment from pSLVHb was inserted to give plasmid pMSW1 (Figure 2). The resulting construct contains the λP_L-*vhb-lacI* and *tac-cI-cat* operons.

Effect of induction on VHb and CAT expression

To investigate the response of VHb and CAT expression to induction, a shake-flask experiment was carried out with strain DH5α/pMSW1. Cultivation was started with a 5% inoculum in order to increase sample concentration. Production of CAT and VHb were followed closely throughout the cultivation. As depicted in Figure 3, a band corresponding to the VHb protein is clearly visible for the sample right before induction (see Figure 3a, 4.5 hr sample). This band immediately disappears for all subsequent samples after IPTG addition (Figure 3a, 6.5-9.5 hr samples). From previous experiment, we do not anticipate that VHb degradation is a major factor, therefore, the disappearance of the VHb can be attributed mainly to a combination of a cease in the VHb production and the effect of dilution due to cell growth. In contrast, expression of CAT is not observed from any sample before induction. However, after IPTG addition the level of CAT continues to increase and eventually reaching 4% of total soluble proteins. These results are indicative of a complete change in the CAT and VHb expression patterns before and after induction. In fact, the validity of the metabolic switch is demonstrated.

Example 2

PCR amplification of *glgC-A* and *glgP* gene fragments

The polymerase chain reaction was used to synthesize both the *glgC-A* and *glgP* structural gene with plasmids pPR2 and pGTC12 as templates, respectively. Primers I and II (Figure 4a) were used to amplify the *glgC-A* gene. These two primers contain only the S-D sequence of the *glgC* gene but not the trancriptional stop sequence such that the amplified fragment can be used to create a λP_L-*glgC-A-lacI* operon. Restriction sites *Sac*I were created at both ends for subcloning. To synthesize the structural gene of the *glgP* gene, primers III and IV were used (Figure 4b). Similarly, only the S-D sequence of the

glgP gene is included. *Nsi* I sites were created to facilitate subcloning.

Construction of the glycogen switch plasmid pMSW2

The 2.76 kb *glgC-A* fragment generated from PCR amplification was digested with *Sac*I overnight. After gel purification, the resulting fragment was subcloned into pUC18 to create pUC18-glgCA. The functionality of the *glgC-A* fragment was confirmed by comparing the glycogen content of *E. coli* TA3476 with and without this plasmid. The results indicate a ten-fold increase in the glycogen content for cells carrying this plasmid (Table I). This plasmid was then cleaved with *Sac*I and the *glgC-A* fragment was inserted into the same site of plasmid pλlacI to give pλCAI.

Similarly, the 2.36 kb *glgP* PCR fragment was cleaved with *Nsi*I and subcloned into pSL1180 to yield pSLglgP. Enzymatic assays were carried out to confirm the functionality of the *glgP* fragment with results as shown in Table I. To construct plasmid pTCIP, the *glgP* fragment was cleaved from pSLglgP and inserted into pTCI. Finally, the glycogen switch construct pMSW2 (Figure 5) is obtained by transferring a 5.8 kb *Nde*I/*Bam*HI fragment from pλCAI into the *Nde*I/*Sma*I sites of pTCIP with blunt end ligation at the *BamHI*/*Sma*I sites.

Alteration of glycogen synthesis and degradation

As mentioned earlier in the Introduction, our main goal in this example is to alter the glycogen synthesis and degradation rate at different stages of growth. Shake flask experiments were carried to investigate this effect. In the first set of experiment, the glycogen contents between strain TA3476 and TA3476/pMSW2 were compared to determine whether glycogen synthesis is augmented by the presence of pMSW2. Results from this experiment are shown in Figure 6. IPTG was added to induce glycogen

glycogen degradation right after glucose exhaustion at the 7 hr. As expected, the specific glycogen content is approximately 5 fold higher for cells carrying this plasmid. Interestingly, the strain carrying pMSW2 has a noticeable lag at the start of the cultivation. We do not know at present whether this is the result of an overproduction of glycogen during the start of the culture. Nevertheless, at the end of the cultivation, the specific glycogen content decreases to a compatible level to the wild-type strain. Unfortunately, without doing a control experiment in which we compare the degradation rate with and without IPTG addition, it is impossible to tell whether the rate of glycogen degradation has been increased.

In order to determine whether the glycogen degradation rate has been increased by overproducing *glgP*, we performed a series of shake flask cultures using strains TA3476, TA3476/pBR322 and TA3476/pMSW2 with and without induction. The pBR322 carrying strain was used to separate any effect due to the presence of the plasmid alone. As depicted in Figure 6, the specific glycogen content is again increased by about 5 fold over the wild-type strain and the pBR322 carrying strain in which there is no observable differences. Moreover, the amount of glycogen accumulated in strains TA3476/pMSW2 with and without induction is the same. In contrast, the rate of glycogen degradation is increased by 30-40% when IPTG was added. This is indicative of an improvement in the glycogen depolymerization process due to an increase in the *glgP* expression. Results from this example indeed solidify the use of the metabolic switch as a novel mean to redirect metabolic flux at our desired time and level.

5.5 Discussion

It is quite clear that with all the currently available genetic techniques, we now have the capability to gather genetic elements from different organisms into a single host

in order to enhance existing pathways or to construct novel pathways. Many theories have been formulated to predict what genetic manipulations are necessary to achieve the optimum results.[6,15] Unfortunately, no genetic tools are available to carry out the precise pathway regulation. Such a novel system is described here by applying the cross-regulation system as a metabolic switch in which this system provides the possibility of manipulating different pathways at our desired time and level.

The validity of this novel system has been shown by using a model system in which the expression of VHb and CAT proteins were alternated under non-induced and induced conditions. We have shown that only the VHb protein is expressed during the pre-induction period. However, the opposite is true post-induction in which VHb expression is turned off and CAT expression is turned on. This type of alternation in the protein production pattern is exactly the expected behavior for the metabolic switch.

However, we still do not have the proof that such changes in the protein level are enough to ensure corresponding changes in the metabolic pathways. To provide this proof, we must apply the metabolic switch to alternate the production of proteins that catalyze different metabolic pathways. Such experiments were performed using the glycogen synthesis and degradation pathways. Using this system, cells are able to synthesize at least 5 fold more glycogen and degrade glycogen 30-40% faster than the wild-type strain. Surprisingly, the augmentation in glycogen degradation is not as good as glycogen synthesis leading to the question whether glycogen degradation is actually rate limiting. Alteratively, there may be other factors that can contribute to this behavior such as protein level inhibition. Moreover, it appears that augmenting glycogen synthesis at the start of the cultivation has a big effect as indicated by a long lag for about 2 hr. This lag is consistently observed for the two different sets of cultivation. However, it does not appear to have any impact on the final cell density achieved. Nevertheless, increasing glycogen degradation does has a small positive effect on cell growth both in

the rate and final density. This is contradictory to previously observed behavior in which overproducing glycogen in an *E. coli* acetate deficient mutant drastically improved cell growth. We believe that this is because of a lower amount of glycogen produced from this system (about 6 times less) as compared to the system used in the other study. This is probably due to a lower plasmid copy number (5 times).

In order to totally understand the metabolic switch system, further and more complete characterization of the glycogen switch system must be carried out. For example, adding a small amount of IPTG at the start of the culture to curtail glycogen synthesis seems very logical. Using a higher copy number construct or different promoter systems are also feasible alternatives.

In conclusion, we have demonstrated in this work the successful application of the metabolic switch as a novel mean of redirecting metabolic flux and anticipate that this should open up new opportunities for more complicated pathway manipulations.

5.6 Acknownledgements

This research was supported by the National Science Foundation (Grant # NSFBCS-891284) and the Advancement Industrial Concepts Division of the U.S. Department of Energy. W. Chen was supported in part by a Predoctoral Training Grant in Biotechnology from the National Institute of General Medical Sciences (National Research Service Award 5 T32 GM 08346-04, Pharmacology Science Program). We wish to thank Dr. Keqin Chen for technical assistance in performing the PCR reaction. We would also like to thank Dr. Romeo and Dr. Komano for providing plasmids pPR2 and pGTC12, respectively.

5.7 References

1. **Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K.** (1989). Short protocols in molecular biology, Greene Publishing Associates and Wiley-Interscience, NY.
2. **Chen, W., Bailey, J.E. & Lee, S.B.** (1991). Molecular design of expression systems: Comparison of different repressor control configurations using molecular mechanism models. *Biotechnol. Bioeng.* **38**, 679-687.
3. **Chen, W., Kallio, P.T. & Bailey, J.E.** (1992). Construction and characterization of a novel cross-regulation system for regulating cloned gene expression in *Escherichia coli*. *Gene*, **Submitted**.
4. **Chen, W., Kallio, P.T. & Bailey, J.E.** (1992). Factors influencing recombinant protein yield in *E. coli* using a novel cross-regulation system. **In preparation**.
5. **Choi, Y.L., Kawamuka, M., Utsumo, R., Sakai, H. & Komano, T.** (1989). Molecular cloning and sequencing of the glycogen phosphorylase gene from *Escherichia coli*. *FEBS Letters*, **243**, 193-198.
6. **Fell, D.A. & Sauro, H.M.** (1985). Metabolic control and its analysis. Additional relationships between elasticities and control coefficients. *Eur. J. Biochem.* **148**, 555-561.
7. **Gunja-Smith, Z., Patil. N.D. & Smith, E.E.** (1980). Two pools of glycogen in *Saccharomyces*. *J. Bacteriol.* **130**, 818-825.
8. **Helmreich, E. & Cori, C.F.** (1964). The role of adenylic acid in the activation of phosphosylase. *Proc. Natl. Acad. Sci. USA* **51**, 131-138.
9. **Khosla, C. & Bailey, J.E.** (1989). Evidence for partial export of Vitreoscilla

hemoglobin into the periplasmic space in *Escherichia coli*. *J. Mol. Biol.* **210**, 79-89.

10. **Laemmili, U.K.** (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-685.
11. **LeVine, S.M., Ardeshir, F. & Ferro-Luzzi Ames, G.** (1980). Isolation and characterization of acetate kinase and phosphotransacetylase mutants of *Escherichia coli* and *Salmonella typhimurium*. *J. Bacteriol.* **143**, 1081-1085.
12. **Lowery, O.H., Rosebrough, N.J., Farr, A.L. & Randall, R.J.** (1951). Protein measurement with the folin phenol reagent. *J. Biol. Chem.* **193**, 265-275.
13. **Neilay D.H. & Bailey, J.E.** Personal communication.
14. **Romeo, T. & Preiss, J.** (1989). Genetic regulation of glycogen biosynthesis in *Escherichia coli*: In vitro effects of cyclic AMP and guanosine 5'-diphosphate 3'-diphosphate and analysis of in vivo transcripts. *J. Bacteriol.* **171**, 2773-2783.
15. **Savageau, M.A., Voit, E.O. & Irvine, D. H.** (1987). Biochemical systems theory and metabolic control theory: 1. Fundamental similarities and differences. *Mathematical Biosciences*, **86**, 127-145.
16. **Zeig, J. Silverman, M. Hilmem, M. & Simon, M.** (1978). Recombinational switch for gene expression. *Science*, **196**, 170-172.

5.8 Tables

Table I. Functional assay for the PCR *glgC-A* and *glgP* fragments.

Strain	Glycogen content,mg/L	Normalized *glgP* activity[1]
TA3476	4	1
TA3476/pUC18-glgCA	52	----
TA3476/pTCIP	----	22

1. *glgP* activity assay was carried out according to Helmreich and Cori.[8]

5.9 Figures

Figure 1. Basic concept of the metabolic switch. Gene 1 and gene 2 are enzymes catalyzing the reactions from B to C and from B to F. At the preinduction state, expression of gene 1 is turned on to favor the pathway from B to C. Whereas, after induction the pathway from B to F is favored.

Figure 2. Plasmid map of pMSW1. This is a derivative of pKC7 containing a VHb gene inserted between the λP_L promoter and the *lacI* gene. This plasmid contains two separate operons with co-transcription of the *cI* and *cat* genes under control of the *tac* promoter and the *vhb* and *lacI* genes under control of the λP_L promoter.

Figure 3. Expression of VHb and CAT from strain DH5α/pMSW1 under non-induced and induced conditions. a) Western blot analysis of VHb expression. Samples were harvested at different cultivation times indicated. b) Specific CAT content (●) and growth (O) for DH5α/pMSW1.

Figure 4. a) Primer I and II used for PCR amplification of *glgC-A* fragment. Sequences homologous to the template pPR2 are underlined. Start codon and stop codon are indicated in italics. b) Primer III and IV used for PCR amplification of *glgP* fragment. Sequences homologous to the template pGTC12 are underlined. Start codon and stop codon are indicated in italics.

Figure 5. Plasmid map of pMSW2. This plasmid contains the *tac-cI-glgP* and λP_L-

glgCA-lacI operons. Under this arrangement, expression of the *glgCA* genes is turned on at the preinduction state. Whereas, a shift in steady state is achieved after induction to favor *glgP* expression.

Figure 6. Comparison of growth properties between stains TA3476 (close symbol) and TA3476/pMSW2 (open symbol). a) Glucose profile (O) and growth curve (Δ) for the two strains. b) Specific glycogen content (O) and glucose profile (Δ) for the two stains.

Figure 7. Comparison of growth properties between stains TA3476 (❑) and TA3476/pBR322 (Δ), TA3476/pMSW2, no induction (O) and TA3476/pMSW2, with induction (◊). a) The growth curve for the four strains. b) Specific glycogen content (open symbol) and glucose profile (close symbol) for the four stains.

Figure 1.

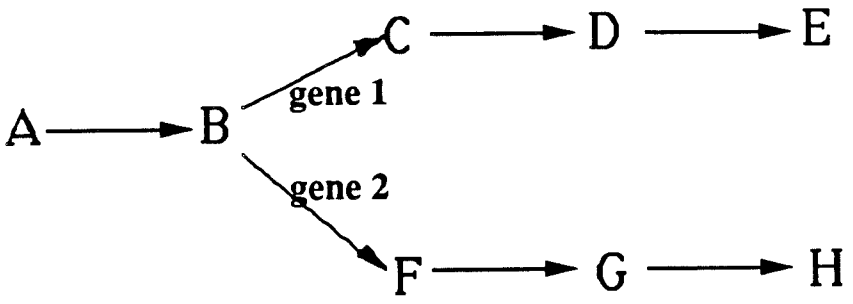


Figure 2.

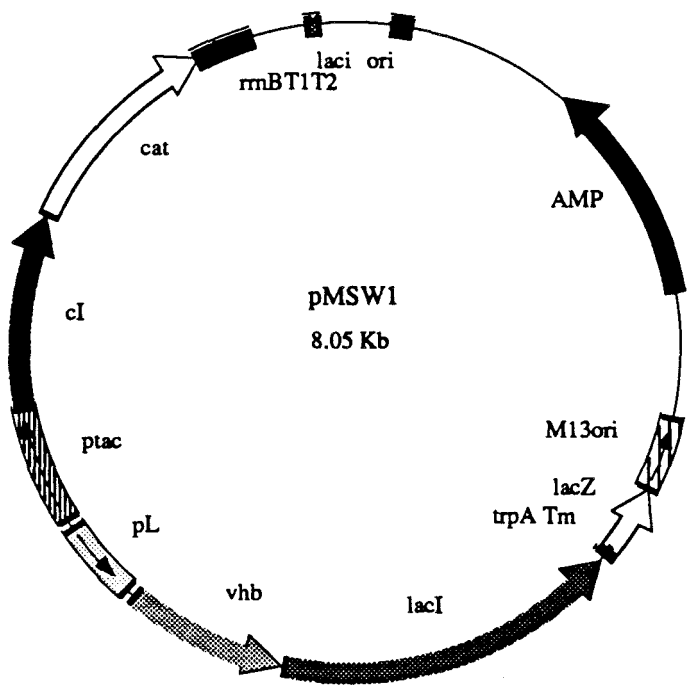
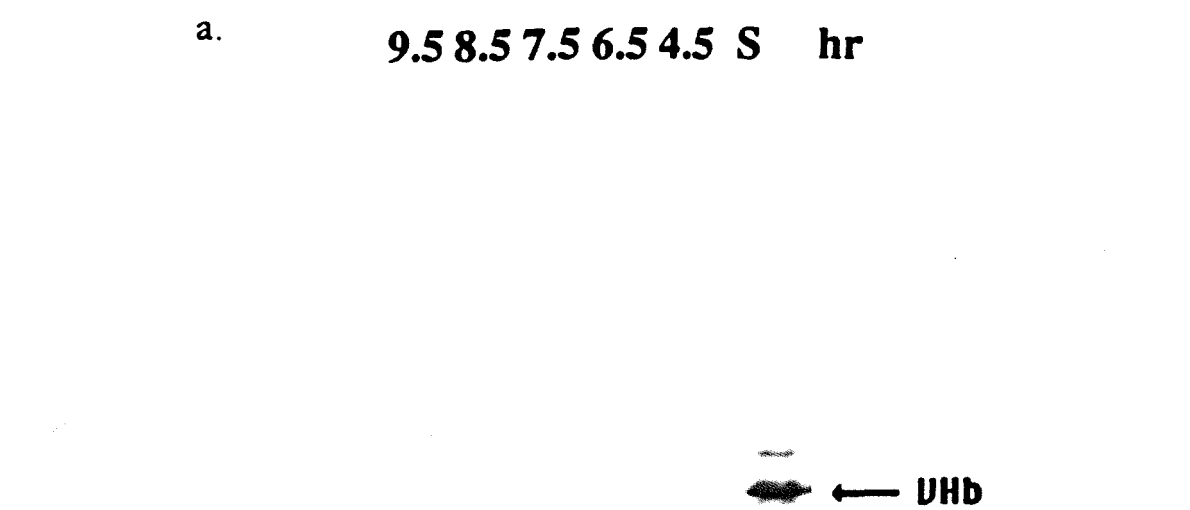

laci ori
rrnBT1T2
cat
AMP
pMSW1
8.05 Kb
cI
ptac
M13ori
lacZ
trpA Tm
pL
vhb
lacI

Figure 3.

a.


9.5 8.5 7.5 6.5 4.5 S hr
VHb

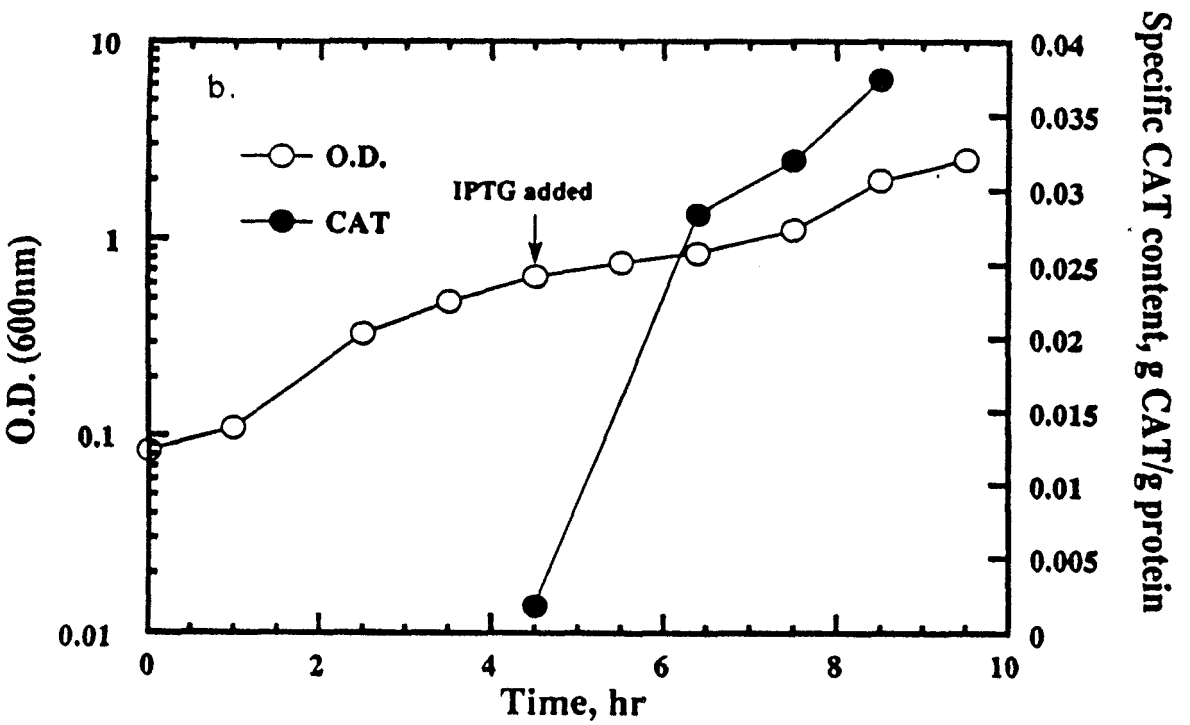

b.
O.D.
CAT
IPTG added
O.D. (600nm)
Specific CAT content, g CAT/g protein
Time, hr
10
1
0.1
0.01
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
0
2
4
6
8
10

Figure 4.

a.

Primer I

5' 3'
GGGAGCTCGGAAGGAGTTAGTC*ATG*GTTAG

Primer II

5' 3'
CCGAGCTCCCAC*TATT*TCGAGCGATAGTAAAGCTC

b.

Primer III

5' 3'
GGATGCATGGCTATCGCTTACAAGCTG*ATG*

Primer IV

5' 3'
CCATGCATCC*CTTA*CAATCTCACCGGATCG

Figure 5.

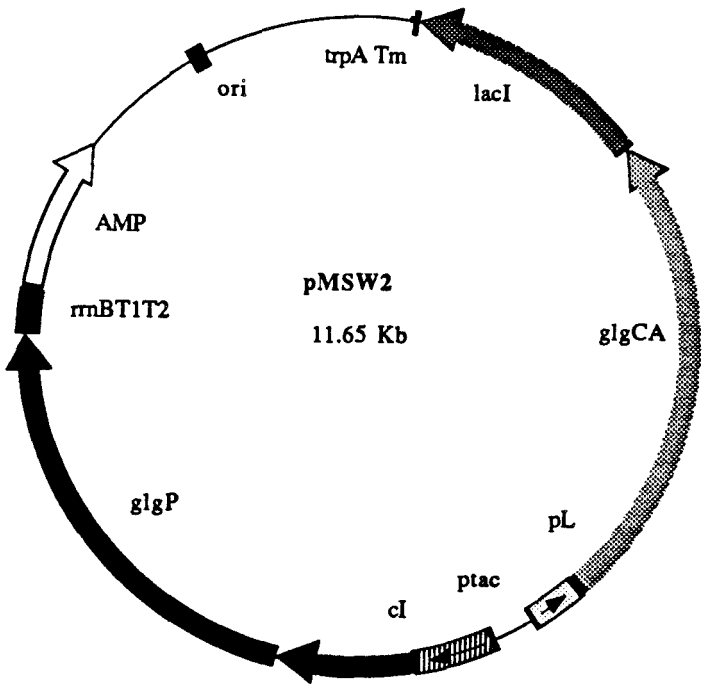

trpA Tm
lacI
ori
AMP
pMSW2
11.65 Kb
rrnBT1T2
glgCA
glgP
pL
ptac
cI

Figure 6.

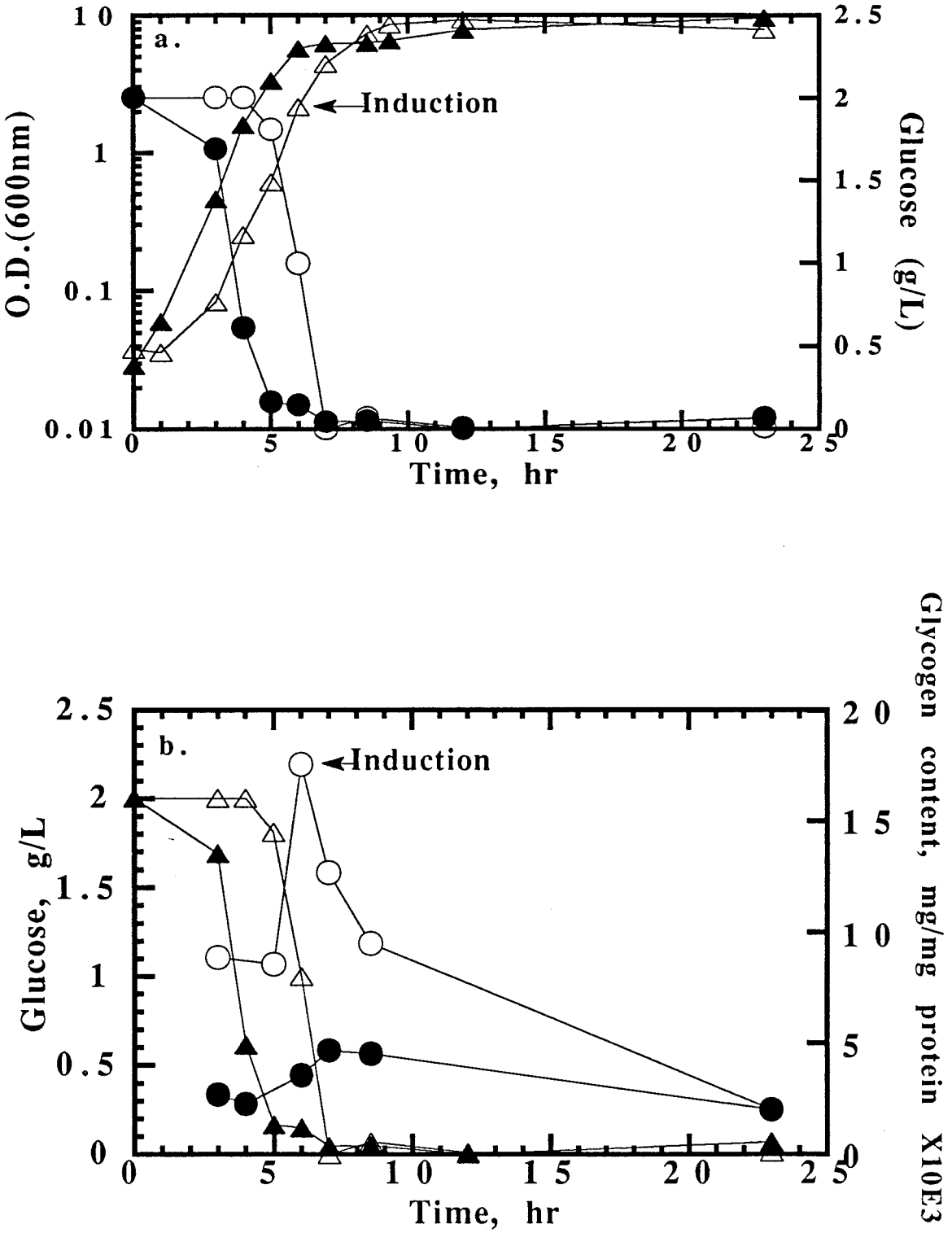
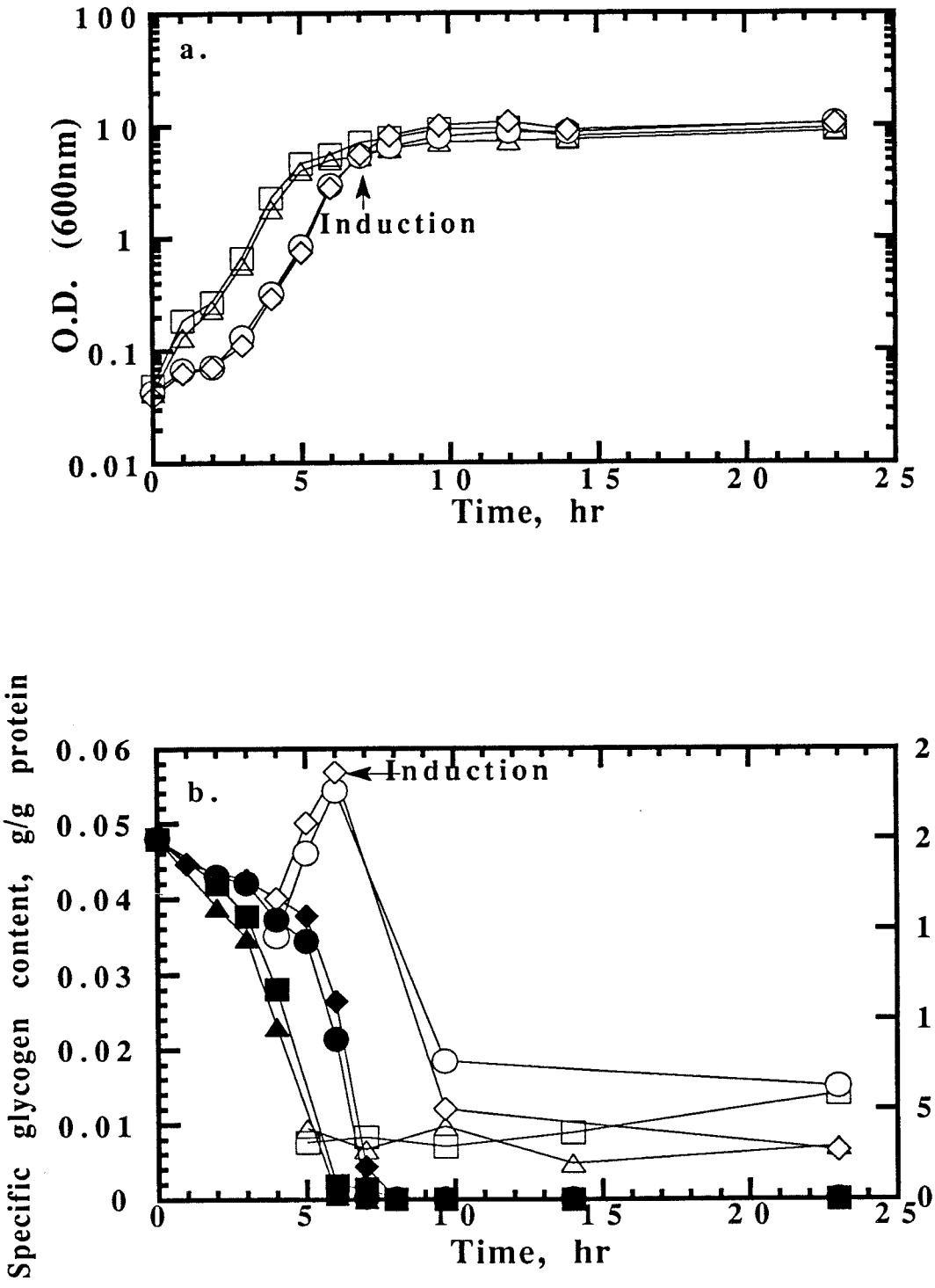

a.
Induction
O.D.(600nm)
Glucose (g/L)
Time, hr
b.
Induction
Glucose, g/L
Glycogen content, mg/mg protein X10E3
Time, hr

Figure 7.


a.
O.D. (600nm)
100
10
1
0.1
0.01
Induction
0
5
10
15
20
25
Time, hr
b.
Induction
Specific glycogen content, g/g protein
0.06
0.05
0.04
0.03
0.02
0.01
0
Glucose, g/L
25
20
15
10
5
0
Time, hr

CHAPTER 6

CONCLUSIONS

6.1 Summary of findings

Bearing in mind the dramatic impact of recombinant proteins on therapeutic, diagnostic and research purposes during the past decade, much effort has been made in the past to optimize their production. However, the process of recombinant protein production involves complicated host-vector interactions[1-2] and optimizing production requires very careful balance between host cell growth and protein synthesis. Many of the current regulated promoter systems suffer drawbacks in one or more ways.[3-5] In view of these deficiencies, we have focused our work on characterizing a novel cross-regulation system for regulated recombinant protein expression and its application as a metabolic switch.

First, using molecular-level mathematical models of the *lac* and λP_R promoter-repressor systems as examples, we were able to explore different repressor synthesis configurations (constitutive synthesis, autoregulation and cross-regulation) and to compare their effectiveness. Simulated trends for the uninduced transcription efficiency agreed well with experiment indicating a loss of control for all configurations if the repressor gene is only included in the chromosome. However, maximizing the induced expression is as important as maintaining tight control before induction. Our results showed that the cross-regulation system maintained the tightest control before induction and also provided the highest transcription rate post-induction.

Since this cross-regulation system offered several interesting characteristics, experimental verification of simulation results was carried out. A particular experimental realization of the cross-regulation system was constructed using the *tac-lacI* and λP_L-*cI* promoter-repressor systems. The polymerase chain reaction (PCR) technique was employed to obtain the precise forms of both the *cI* and *lacI* structural genes required for

this construct . For the purpose of comparison, a commonly used constitutive repressor synthesis system was also constructed. Induction results using the easily assayed reporter enzyme chloramphenicol acetyltransferase (CAT) as a model product protein revealed that the specific CAT content reached at least a two-fold higher level from the cross-regulation system than from a system with constitutive repressor synthesis. SDS-PAGE analysis of protein samples from this experiment showed that the synthetic operon constructed is fully functional as indicated by the presence of the CAT and *cI* proteins post-induction. In addition, it also demonstrated that the amount of *lac* repressor produced from the cross-regulation system is much less after induction which is the likely reason for higher expression. The applicability of this system over a broad range of copy number was also demonstrated by transferring the expression cassette into a series of different copy number plasmids. These experiments showed that CAT expression from the cross-regulation system was at least two-fold higher independent of copy number.

Further characterization of the cross-regulation system was carried out in order to identify the various factors affecting the recombinant protein yield. We have investigated the response of the cross-regulation system to different IPTG dosages. Our results indicated an IPTG concentration of at least 0.5 mM must be used to achieve the optimum induction. The effect of induction on CAT production and cell growth were determined by inducing CAT expressing with IPTG at different times during a batch cultivation. The results suggested that the best time for induction is about 2-3 hr into the exponential phase which provided the best compromise between maintaining cell growth and CAT production. Furthermore, using Northern blot analysis we were able to identify that transcription of the CAT mRNA correlated with CAT expression upon transition into the stationary phase. Subsequently, no further increase in the CAT level was observed. This limitation can be resolved by extending cell growth either by using a fed-batch

fermentation or by using an unmutagenized *E. coli* strain. Moreover, we have investigated the *lacI* mRNA level in response to induction and observed that its level has decreased by at least three fold. This illustrated the special property of the cross-regulation system with a shift in the transcriptional steady state from one promoter to another upon induction.

We have demonstrated the feasibility of a novel way to redirect metabolic flux by applying the cross-regulation system as a metabolic switch. The validity of this system was showed by alternating VHb and CAT expression under uninduced and induced conditions. Furthermore, the metabolic switch was applied to alter glycogen synthesis and degradation. Results indicated a five-fold increase in glycogen synthesis and a 40% augmentation in glycogen degradation.

Much of the work described here can be readily applied to other promoter systems and is not just restricted to the particular examples illustrated. It is our desire to demonstrate the feasibility of the cross-regulation system with these few examples. We believe that using the same kind of principle, it is possible to extend our cross-regulation results into diverse hosts using different types of promoter.

In another project, we have constructed new vectors to functionally express *Vitreoscilla* hemoglobin (VHb) in *Saccharomyces cerevisiae*, and the influence of VHb expression on yeast aerobic metabolism was investigated. Growth experiments indicated that strain expressing VHb grew to lower final density but produced a greater amount of ethanol. A subsequent respiration inhibition experiment revealed that this effect was likely coupled with respiration. With the use of an adh^0 mutant, which produces ethanol only by a novel mitochondrial acetaldehyde dismution pathway, we were able to demonstrate that the presence of VHb reduced the production of ethanol through this pathway. Since this pathway is highly coupled with electron transport, we have

established the strong link between the VHb effect and respiration. Unlike what was observed with the strain SEY2101, the excess carbon flux that did not flow into ethanol was diverted into cell mass in this strain.

6.2 References

1. **Latta, M., Philit, M., Maury, I., Soubrier, F., Denefle, P. & Mayaux, J-F.** (1990). Tryptophan promoter derivatives on multicopy plasmids: A comparative analysis of expression potentials in *Escherichia coli*. *DNA and Cell Biology*, **9**, 129-137.
2. **Mieschendahl, M. Petri, T. & Hanggi, U.** (1986). A novel prophage independent *trp* regulated lambda P_L expression system. *Bio/Technology*, **4**, 802-808.
3. **Peretti, S.W & Bailey, J.E.** (1987). Simulations of host-plasmid interactions in *Escherichia coli*: Copy number, promoter strength, and the ribsome binding site strength effects on metabolic activity and plasmid gene expression. *Biotechnol. Bioeng*. **24**, 316-328.
4. **Seo, J-H. & Bailey, J.E.** (1985). Effect of recombinant plasmid content on growth properties and cloned gene product formation in *Escherichia coli*. *Biotechnol. Bioeng*. **27**, 1668-1674.
5. **Stark, M.J.R.** (1987). Multicopy expression vectors carrying the *lac* repressor gene for regulated high level expression of genes in *Escherichia coli*. *Gene*, **51**, 255-267.

APPENDIX A

INTRACELLULAR EXPRESSION OF BACTERIAL HEMOGLOBIN ALTERS AEROBIC METABOLISM OF *SACCHAROMYCES CEREVISIAE*

Abstract

Vitreoscilla hemoglobin (VHb) has been functionally expressed in *Saccharomyces cerevisiae* and its influence on yeast aerobic metabolism was investigated. New expression vectors were constructed to express VHb constitutively under control of the ADH promoter. The presence of VHb was shown by Western blot analysis. The functional activity of this protein was also demonstrated using a CO binding assay. Localization of this protein was shown to be predominately in the cytoplasm. Batch fermentation results indicated that the wild-type strain expressing VHb redirected carbon flux away from cell mass into ethanol with no significant alteration in the specific growth rate. This effect was not observed if cells were grown under respiration inhibition, indicating the metabolic effect of VHb is likely linked to respiration. Expression of VHb in the *adh*^{0} strain MC65-2A, which produces ethanol only by a respiration-coupled pathway, revealed that ethanol production was decreased and cells reached a higher final cell density. Growth enhancement due to VHb expression was observed only during the final stage of culture growth when the acetaldehyde produced during the glucose growth phase was used as a substrate. This metabolic effect of intracellular VHb was more clearly seen in an acetaldehyde fed-batch fermentation in which VHb-expressing cells grew to at least 3-fold higher final cell density. These results suggested that the action of VHb is likely linked to electron transfer.

A.2 Introduction

The recent advance of recombinant DNA technology has enabled us to genetically engineer industrial organisms in order to achieve some special tasks. Much of this work has been focused on the massive production of desirable foreign proteins or the redirection of metabolic flux toward a particular end product (for example, amino acids or antibiotics). However, in recent years, researchers have moved toward the goal of genetically engineering cells to function as "designer biocatalysts", in which certain desirable heritable properties from different organisms are brought together in one single host with the aim of performing some specific task(s).[2,19,14]

Recently, Wakabayashi et al.[18] reported the primary amino acid sequence of a soluble dimeric hemoglobin-like molecule found in the obligately aerobic bacterium, *Vitreoscilla*, and demonstrated significant similarity of this protein with known eukaryotic globin sequences. As an obligate aerobe, *Vitreoscilla* is commonly found in oxygen-poor habitats[11]; therefore, it is possible that the hemoglobin-like molecule, denoted*Vitreoscilla* hemoglobin (VHb), is playing some role in supporting growth under such conditions.

Khosla and coworker first demonstrated that the expression of VHb enhanced growth properties[5] and protein synthesis[7] in recombinant *Escherichia coli*. It was proposed that these effects are due to an improved overall efficiency of oxygen-limited ATP production. Enhancement in actinorhodin production was also observed for the VHb-expressing *Streptomyces*.[9]

Although not as well characterized as *E. coli*, *Saccharomyces cerevisiae* (baker's yeast) is one of the best understood eucaryotes. It is extensively used as an industrial microorganism for the production of ethanol, cloned proteins, and food-related

products, and the technology required to grow yeast under both aerobic and anaerobic conditions to high cell densities in large, process-scale bioreactor is well established. In order to investigate whether significant metabolic effects can be obtained in *Saccharomyces cerevisae*, VHb has been functionally expressed in this yeast and its effect on yeast metabolism examined.

A.3 Materials and Methods

Strains and media

Saccharomyces cerevisiae SEY2101 (*ura 3-52, leu 2-3 112, suc 2-Δ9, ade 2-7, gal 2*) was utilized as the host strain in most experiments. Strain MC65-2A (*Matα, adh 1-Δ, adh 2, adh 3, adh 4:URA3, trp 1 (adh°)*) is an *adh°* strain provided by Dr. M. Ciriacy which carried mutations in all four known cytoplasmic alcohol dehydrogenases.[4] *E. coli* DH5α (F-, *endA1*, *hsdR17*(r_k-m_k+), *supE44*, *thi-1*, λ-, *recA1*, *gryA96*, *relA1*, *φ80dlacAm15*) was used for all intermediate DNA cloning steps.

SD medium was used for all routine cultivation of strain SEY2101 carrying plasmids pEX-2 and pAAH5. This medium contains 6.7 g/L of Difco yeast nitrogen base without amino acids, 2 g/L of amino acid powder mix (Table I), and the desired amount of glucose (1-4 g/L). Leucine was added to the culture medium for the plasmid-free host at 25-100 mg/L. Citrate buffer (4.93 g/L citric acid and 7.8 g/L sodium citrate, pH. 4.5) was used to maintain pH at 5.0 during fermentation. For the cultivation of strain MC65-2A, SD medium was used except that tryptophan (50-100 mg/L) instead of leucine was eliminated for plasmid-containing culture for selection.

Chemicals, reagents, and DNA manipulations

All restriction endonucleases, modifying enzymes (T4 DNA polymerase, Klenow fragment, T4 DNA Ligase) were purchased form either New England BioLabs or Boehringer Mannheim Biochemicals. All DNA manipulations were done according to standard methods.[12] DNA fragments were eluted from agarose gels using a Geneclean Kit (Bio 101). Yeast transformation was performed as described by Sherman et al.[13]

Preparation of cells extract and immunoblot analysis

Yeast extracts were prepared as follows. Cell pellets were incubated for at least 2 hr in a solution containing 1.0 M sorbitol, 0.05 M potassium phosphate pH 7.5, 14 mM β-mercaptoethanol, and 20 unit/mL of lyticase (Sigma L-8012). After 2 hr the spheroplasts were microfuged out and resuspended in a solution containing 100 mM Tris-HCl, pH 8.0, 50 mM NaCl and 1 mM EDTA. The protoplasts were then sonicated using a microtip at 20% power for 10 X 30 seconds bursts interspersed with 30 seconds cooling. The lysed cell debris was then removed by microfuging for 15 minutes. For the immunoblot analysis, an aliquot (3-5 mg protein) was electrophoresed on a 15% polyacrylamide-SDS gel according to the method of Laemmli.[8] The proteins were then electrophoretically transferred to a nitrocellulose membrane as described elsewhere.[1] The proteins were screened with antiserum generated against *Vitreoscilla* hemoglobin as described elsewhere.[1] The hemoglobin standard was produced in recombinant *E. coli*.[6]

CO binding assay

Cell lysates were made up to 3 mL and separated into two epifuge tubes. A few grains of sodium dithionite were added to each. One tube was allowed to stand exposed to air for 10 minutes while the other was left in a pressurized chamber filled with carbon

monoxide (CO) gas for 20 minutes. The CO difference spectrum (400-500 nm) was then taken in a quartz cuvette using a Shimadzu UV-260 recording spectrophotometer (Kyoto, Japan).

Analytical methods

Glucose was analyzed by an enzymatic kit available from Sigma (Cat. #510-A). Glycerol, acetaldehyde and acetate were analyzed using assay kits from Boehringer Mannheim (Cat. #148270, 668613, and 148261). Ethanol was analyzed by gas chromatography with a Supelco Inc. Chromosorb. 101 (80/100) column (2m X 0.5cm); *col. temp. 150°C; carrier flow (50mL/min). Glucose-6-phosphate* dehydrogenase and isocitrate dehydrogenase activities were assayed with Sigma kits (Cat. 400K-25 and 153-A).

Fermentation and growth conditions

The batch and fed-batch fermentations were performed at 30°C, pH 5.0 and 250 rpm in a BiofloIII fermentor with a 2.5 liter working volume. Cultures were started with a *1% (v/v) seed inoculum. Optical density was determined by a Spectronic 21* spectrophotometer (Milton Roy). Shake flask experiments were performed in a New Brunswick G1000 Incubator Shaker at 30°C and 300 rpm in 1 L Erlenmeyer flasks containing 400 mL medium.

Preparation of cell fractions

Cell fractionation was performed as described by Oeda et al.[10] Cell lysates were centrifuged at 3000 X g for 10 *minutes to precipitate cell debris. The mitochondrial* fraction was recovered by further centrifuging at 10000 X g for 20 minutes. The

supernatant obtained was assigned as the cytoplasmic fraction. Enzymatic activities of mitochondrial enzymes were determined by disruption of the membrane by lysis with 1% Triton X-100 according to Campbell et al.[3]

A.4 Results

Plasmids for expressing *Vitreoscilla* hemoglobin (VHb) in *Saccharomyces cerevisiae*

To express VHb in strain SEY2101, plasmid pEX-2 was constructed from plasmid pAAH5 by inserting the *vhb* gene after the ADH promoter (Figure 1). A *Bam*HI-*Sph*I fragment from pRED302 (provided by Chaitian Khosla) was subcloned into the *Bam*HI/*Sph*I site of the vector pUC18. From this new construct, a *Hin*dIII-*Bam*HI fragment was isolated and cloned into the *Hin*dIII site of pAAH5 with blunt-end ligation at the *Bam*HI end. The resulting construct expresses VHb constitutively under control of the ADH promoter.

Plasmid pVHb-trp1 was constructed to express VHb in strain MC65-2A with *trp1* as the selection marker (Figure 2). First, a 3 kb *Bam*HI fragment containing the entire VHb-expressing cassette from pEX-2 was inserted into plasmid pYE[17] to yield plasmid pVHb-tm1. To construct plasmid pVHb-trp1, a 4.7 kb *Nhe*I/*Eag*I fragment containing the *trp1* marker and the VHb expression cassette was inserted into pEX-2.

Transformed cells expressed the VHb protein

Immunoblot analysis was performed to demonstrate the expression of VHb in *S. cerevisiae* transformed with plasmids pEX-2 and pVHb-trp1. Whole cell extracts were electrophoresed on a 15% SDS-PAGE gel. The proteins were then transferred to a

nitrocellulose membrane and incubated with VHb antiserum. As shown in Figure 3, a major band that comigrates with the VHb standard is present in SEY2101/pEX-2 (3a, lane 2) and MC65-2A/pVHb-trp1(3b, lane 2). This band however does not appear for cells carrying control plasmids pAAH5 and pYE. These results indicate VHb protein of the correct molecular weight is being produced by cells harboring VHb expression plasmids.

VHb produced in *Saccharomyces cerevisiae* is biologically active

In order to verify that the VHb protein produced is functionally active, a carbon monoxide (CO) binding assay was performed. Initial analyses using lysates obtained using sonication method failed repeatedly. It is postulated that VHb degradation may be occurring during prolonged incubation for the spheroplast formation. To minimize this problem, fresh cell extract was prepared using a cell homogenizer. This sample was then used for the CO binding assay. Results using lysate so obtained are presented in Figure 4. As can be seen, the difference spectrum of SEY2101:pEX-2 cells (Figure 4a) contains a peak at 419 nm followed by a depression at 440 nm, characteristic of functional VHb. This profile is not observed in the difference spectrum from the control SEY2101:pAAH5 (Figure 4b).

Subcellular localization of VHb synthesized in yeast

Previous results with recombinant *E. coli* indicated that the VHb protein was equally distributed between the cytoplasm and the periplasm;[6] thus suggesting that this protein can translocate across the periplasmic membrane. Since the location of the VHb may be important in determining its function in yeast, we examined the subcellular localization of VHb. In this experiment, the cytoplasmic fraction was separated from the mitochondrial fraction as described in the Materials and Methods. These samples were

analyzed by Western blotting and results reveal that most of the VHb is present in the cytoplasmic fraction as indicated by a much higher intensity of the VHb band (see Figure 5). Only a small amount of VHb is observed in the mitochondrial fraction.

In order to verify the fractions obtained are actually from the cytoplasm and the mitochondria, glucose-6-phosphate dehydrogenase (G-6-PDH) and isocitrate dehydrogenase (IDC) were assayed as cytoplasmic and mitochondrial control markers, respectively. Table II shows the percentage of control markers in each fraction. It can be seen that 96% of the cytoplasmic marker G-6-PDH activity is present in the cytoplasmic fraction. In contrast, about 30% of the IDC activity is found in the cytoplasmic fraction. The data indicate some breakage of the mitochondria during the fractionation procedure. However, the great majority of VHb is found in the cytoplasmic fraction with almost none in the mitochondrial fraction, thus indicating that VHb in yeast is localized primarily in the cytoplasm.

Effect of VHb presence on cell growth and by-product formation

Previous fermentation results with recombinant *E. coli* expressing VHb illustrated that VHb-expressing *E. coli* have higher specific growth rate, higher final cell density, and higher oxygen consumption rate compared to a control strain without VHb when both are cultivated under mircoaerobic conditions.[5,7] In this work, batch fermentations were performed under poorly aerated conditions in order to explore the effect of VHb on yeast cell growth and carbon metabolism. Yeast strains SEY2101, SEY2101:pAAH5 and SEY2101:pEX-2 were grown at 37^{0}C in SD medium supplemented with 1 g/L of glucose as the carbon source. Figure 6 shows results of the experiment under mircoaerobic conditions in which the agitation rate was manually adjusted such that the dissolved oxygen (DO) was lower than 10%. Strain SEY2101:pEX-2 grows to a much lower cell

density than the other two with no significant difference in the exponential-phase specific growth rate (Figure 6a). Glucose consumption and ethanol production were monitored as shown in Figure 6b. Results indicate that more ethanol is produced from SEY2101:pEX-2 than the other two strains. Since the same amount of glucose was used for the different strains in this experiment, we postulate that the decrease in final cell density for the VHb-expressing strain SEY2101:pEX-2 is due to a redirection of the carbon flux away from cell mass into other by-products (for example, ethanol). To ensure that VHb is present during the course of the SET2101:pEX-2 fermentation, Western blot analysis was performed. Those results indicates that VHb is present at all time during the batch fermentation for SEY2101:pEX-2 (data not shown).

One of the main hypothesis for VHb action in recombinant *E. coli* is the interaction of VHb (or oxygenated VHb) with the respiratory apparatus leading to an improved overall proton translocation efficiency[7]. To test whether the same type of interaction exists in yeast, we carried out experiments in which cells were grown under respiration inhibition which was accomplished by the addition of respiration inhibitor antimycin A. In this growth study, cells were grown in SD medium for 4 hours before the addition of antimycin A. Results from this experiment are shown in Figure 7. Under this condition, the previously observed VHb effect no longer exists. As indicated, no difference in growth is observed between stains SEY2101:pEX-2 and SEY2101:pAAH5 both in specific growth rate and finally cell density (Figure 7a). Furthermore, no difference in ethanol production is observed between the two strains under this condition (Figure 7b). These results indicate a strong link between respiration and the effect of VHb in yeast.

Investigate the interaction of VHb with respiration using a cytoplasmic *adh*^{o} strain MC65-2A

Our fermentation results suggest that the action of VHb in yeast is linked to respiration. How does this propagate into a redirection of carbon flux as indicated by an increase in ethanol production? Recently, a novel mitochondrial acetaldehyde-reducing pathway was identified to be responsible for the production of approximately one third of the ethanol under aerobic conditions.[4] This pathway is highly coupled with the respiratory chain. Therefore, if the presence of VHb alters the production of ethanol, this pathway is a potential candidate to be investigated.

To carry out this experiment, VHb was functionally expressed into strain MC65-2A. This strain carries irreversible mutations in the genes ADH1, ADH3, and ADH4 and a point mutation in ADH2. The only way to produce ethanol in this mutant is via the mitochondrial acetaldehyde-reducing pathway. Construction of the plasmid pVHb-trp1 to express VHb was described earlier. Production of the VHb protein in MC65-2A:pVHb-trp1 has been demonstrated by Western blotting (see Figure 3b). Shake flask experiments were carried out using SD medium supplemented with 1 g/L glucose as the carbon source. No difference in cell growth between strains with and without VHb is observed for the glucose growth phase (Figure 8a). After glucose exhaustion, the strain expressing VHb grows to a higher cell density. Analyses of ethanol and acetaldehyde reveal that the same amount of acetaldehyde is produced during the glucose growth phase in which no growth difference was observed (Figure 8b). Only when the acetaldehyde was used as the carbon source did we observe an advantage in growth for the VHb-expressing cells. This is consistent with lower ethanol production from this strain during this growth phase. Moreover, Thielen et al. reported that the same amount of acetate is also produced from this pathway,[16] it will be of interest to investigate whether an alteration in acetate

production has also occurred here. Results from this experiment confirm that the effect of VHb is indeed respiration-linked.

Fed-batch fermentations were then carried out using acetaldehyde as the feeding substrate. Our goal here is to define the conditions for which the effect of VHb is more apparent. In this experiment, fermentation was started with cells using glucose as the main carbon source. After glucose exhaustion, acetaldehyde was fed in to maintain cell growth. As depicted in Figure 9, cells expressing VHb grow to a much higher cell density. This difference is much larger than that observed from the shake-flask culture emphasizing the effect of VHb with acetaldehyde as the carbon source. Similarly, more ethanol is being produced from the strain not expressing VHb (data not shown). Possible mechanisms will be discussed later.

A.5 Discussion

We have investigated in this study the effect of VHb on yeast aerobic metabolism. New expression vectors were constructed in order to express VHb in different yeast strains. Production of the VHb protein is shown by Western blot analysis. Furthermore, using the CO binding assay, we are able to demonstrate that the VHb protein produced is functional. In contrast to previous studies with *E. coli* and *Streptomyces* in which the presence of VHb improves cell growth, our results with yeast show otherwise. When VHb is expressed in yeast, cell growth is decreased, and the remaining carbon flux is redirected to other by-products such as ethanol. However, similar to VHb effects observed in *E. coli*, the metabolic changes caused by VHb in yeast appears to be closely linked to respiration. How could such a respiration-linked interaction influence the carbon flux distribution? There are two possible explanations to this question. First, previous reports have indicated that an increase in the glycolytic flux can lead to an

increase in ethanol production and a decrease in cell yield very similar to what we have observed here[17]. It is highly unlikely that the action of VHb will be to alter the glycolytic flux since we have shown that it is linked to respiration. Furthermore, all VHb effects occur after glucose exhaustion, therefore, it does not appear that glycolysis is playing another role in this alteration.

A much more reasonable explanation can be found by the recent discovery of a novel mitochondrial acetaldehyde dismutation pathway. Production of ethanol and acetate from this pathway inside the mitochondria membrane is dependent on electron transport.[16] Therefore, if the effect of VHb is to influence some key steps involved in the electron transport (respiration), the net result should be an alteration in the ethanol production. Such a result is observed here as indicated by a decrease in the ethanol production from the strain expressing VHb. Since none of the cytoplasmic ADHs still exists, the excess acetaldehyde is probably processed by other pathways (VHb enhanced?) to generate the extra cell mass. A more direct demonstration of this hypothesis is illustrated by using an acetaldehyde fed-batch fermentation in which the difference in cell growth and ethanol production is much more profound. This is indicative of the important link between this mitochondrial acetaldehyde dismutation pathway and the effect of VHb.

How does our *adh*^{0} mutant results relate to our results using the strain SEY2101? We can speculate that since the amount of ethanol produced through this dismutation pathway is less for the VHb-expressing strain, the excess acetaldehyde can be processed by the cytoplasmic ADHs and converted into ethanol. The amount of ethanol produced is more because, unlike the dismutation pathway, only ethanol will be produced. Moreover, since this dismutation pathway is highly coupled with the respiratory chain, it is reasonable to expect that such a change can also affect the amount of energy produced

through this pathway. This hypothesis may help explain why the VHb-expressing strain has lower cell yield but higher ethanol production. Further evidence for this hypothesis could perhaps be obtained by studying the effect of an acetaldehyde fed-batch fermentation on cell growth and ethanol production from the strain SEY2101.

A.6 Acknowledgements

This work was supported by the Advancement Industrial Concepts Division of the U.S. Department of Energy. W. Chen was supported in part by a Predoctoral Training Grant in Biotechnology from the National Institute of General Medical Sciences (National Research Service Award 5 T32 GM 08346-04, Pharmacology Science Program). Plasmid pEX-2 was constructed by Dr. Dallas E. Hughes at Exogene. We wish to thank Dr. M. Ciriacy and Dr. G. Bitter for providing the *adh*^{0} stain MC65-2A and the plasmid pYE, respectively.

A.7 References

1. **Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K.** (1989). Short protocols in molecular biology, Greene Publishing Associates and Wiley-Interscience, NY.
2. **Brennand, J. & Margison G.P.** (1986). Reduction of the toxicity and mutagenicity of alkylating agents in mammalian cells harboring the *Eshcerichia coli* alkylatransferase gene. *Proc. Natl. Acad. Sci.* **83**, 6292-6296.
3. **Campbell, I. & Duffus, J.H.** (1988). Yeast, a practical approach. IRL Press, Oxford.
4. **Drewke, C., Thielen, J. & Ciriacy, M.** (1990). Ethanol formation in *adh*^{0} mutants reveals the existence of a novel acetaldehyde-reducing activity in *Saccharomyces cerevisiae*. *J.Bacteriol.* **172**, 3909-3917.
5. **Khosla, C. & Bailey J.E.** (1988). Heterologous expression of a bacterial haemoglobin improves properties of recombinant *Escherichia coli*. *Nature*, **332**, 633- 635
6. **Khosla, C. & Bailey, J.E.** (1989). Evidence for partial export of Vitreoscilla hemoglobin into the periplasmic space in *Escherichia coli*. *J. Mol. Biol.* **210**, 79-89.
7. **Khosla, C., Curtis, J.E., Demondena, J., Rinas U. & Bailey, J.E.** (1990). Expression of intracellular hemoglobin improves protein synthesis in oxygen-limited *Escherichia coli*. *Bio/Technology*, **8**, 849-853.
8. **Laemmili, U.K.** (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-685.
9. **Magnolo, S.K., Leenutaphang, D.L., DeMondena, J.A., Curtis, J.E.,**

Bailey, J.E., Galazzo, J.L. & Hughes, D.A. (1991). Actinorhodin producing by streptomyces coelicolor and growth of streptomyces lividans are improved by the expression of a bacterial hemoglobin. *Bio/Technology*, **9**, 473-476.

10. **Oeda, K., Sakaki, T. & Ohkawa, H.** (1985). Expression of rat liver cytochrome P-450 cDNA in *Saccharomyces cerevisiae. DNA*, **4**, 203-210.
11. **Pringsheim, E.C.** (1951). The *Vitreoscillaceae*: a family of colorless, gliding, filamentous organism. *J. Gen. Microbiol.* **5**, 124-149.
12. **Sambrook, J., Fritsch, E.F. & Maniatis, F.E.** (1989). Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
13. **Sherman, F., Fink, G.R. & Hicks, J.** (1986). Lacboratory course manuel for method in yeast genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
14. **Stanzak. R. Matsushima, P. Baltz, D.H. & Roa, R.N.** (1986). Cloning and expression in *Streptomyces lividans* of clustered Erythromycin biosynthesis genes from *Streptomyces erythreus. Bio/Technology*, **4**, 229-232.
15. **Tschumper, G. & Carbon J.** (1980). Sequence of a yeast DNA fragment containing a chromosomal replictor and the *Trp1* gene. *Gene*, **10**, 157-166.
16. **Thielen, J. & Ciriacy, M.** (1991). Biochemical basis of mitochondrial acetaldehyde dismutation in *Saccharomyces cerevisiae. J. Bacteriol.* **173**, 7012-7017.
17. **Van der Aar, P.C., Lopes, T.S., Klootwijk, J., Groneveld, Ph., van Verseveld, H.W. & Stouthamer, A.H.** (1990). Consequence of phosphoglycerate kinase overproduction for the growth and physiology of *Saccharomyces cerevisiae. Appl. Microbiol. Biotechnol.* **32**, 577-587.
18. **Wakaboyashi, S., Matsubara, H. & Webster, D.A.** (1986). Primary

sequence of a dimeric bacterial hemoglobin from *Viteroscilla*. *Nature*, **332**, 481-483.

19. **Windass, J.D., Worsey, M.J., Pioli, Z.M., Barth, P.T., Atherton, K.T., Dart, E.C., Byrom, D., Powell, E. & Senior R.J.** (1980). Improved conversion of methanol to single-cell protein by *Methylophilus methylotrophus*. *Nature*, **287**, 396-401.

A.8 Tables

Table I. Composition of amino acid powder mix.

4 g adenine sulfate	4 g Gly
4 g Phe	4 g Val
4 g Trp	4 g Asn
4 g His	4 g Thr
4 g Met	4 g Uracil
4 g Ile	4 g Inositol
4 g Lys	0.4 g p-aminobenzoic acid
4 g Glu	4 g Ala
4 g Tyr	4 g Cys
4 g Pro	4 g Gln
4 g Asp	

Table II. Percentage of cytoplasmic (G-6-PDH) and mitochrondrial (IDC) marker enzymes in the two different fractions.

Fraction	G-6-PDH	IDC
Cytoplasm	96%	27%
Mitochonrdria	4%	73%

A.9 Figures

Figure 1. Plasmids for expressing *Vitreoscilla* hemoglobin (VHb) in yeast strain SEY2101. Plasmid pAAH5 is the parental shuttle vector containing the 2μ origin of replication and the *LEU2* selection marker. Pasmid pEX-2 is derived from pAAH5 expressing VHb constitutively under control of the ADH promoter.

Figure 2. Plasmid pVHb-trp1 for expressing VHb in strain MC65-2A. This plasmid is essentially the same as pEX-2 except it also contains a *trp1* selection marker isolated from plasmid pYE[17].

Figure 3. Western blot analysis of whole cell extracts from *S. cerevisiae*. (A) lane 1: SEY2101:pEX-2; Lane 2: SEY2101:pAAH5; Lane 3: VHb standard. (B) Lane 1: VHb standard; Lane 2: MC65-2A:pVHb-trp1; Lane 3: MC65-2A:pYE.

Figure 4. The carbon monoxide (CO) difference spectra of cell extracts. (A) SEY2101:pEX-2; (B) SEY2101:pAAH5. The difference in absorbance between CO-reduced and O_2-reduced samples is plotted.

Figure 5. Western blot analysis of extract from the cytoplasmic and mitochondrial fractions. Lane 1: Cytoplasmic fraction; Lane 2: Mitochondrial fraction; Lane 3: Vhb standard.

Figure 6. Comparison of growth properties of SEY2101(❑), SEY2101:pAAH5(O), and SEY2101:pEX-2 (Δ) under poorly aerated conditions. The batch fermentations were carried out at 30^{0}C, pH 5.0 and 250 rpm. Dissolved oxygen (DO) was manually adjusted to be lower than 10%. (A) The growth curves as measured in term of OD_{600} are shown here for these three strains. (B) Glucose (opened symbol) and ethanol (filled symbol) profiles for these three stains.

Figure 7. Comparison of growth properties of SEY2101:pAAH5 (Δ), and SEY2101:pEX-2 (O) under respiration inhibition condition. Cells were grown under same conditions as described in Figure 6. (A) The growth curves as measured in term of OD_{600} are shown here for these three strains. Arrow indicates time of antimycin A addition. (B) Glucose (opened symbol) and ethanol (filled symbol) profiles for these two stains.

Figure 8. Comparison of growth properties of MC65-2A (O), and MC65-2A:pVHb-trp1 (Δ). Experiments were carried out in shake flasks at 30^{0}C, pH 5.0 and 300 rpm. (A) The growth curves as measured in term of OD_{600} (opened symbol) and the glucose profile (filled symbol) are shown here for these two strains. (B) Acetaldehyde (opened symbol) and ethanol (filled symbol) profiles for these two stains.

Figure 9. Comparison of growth properties of MC65-2A (O), and MC65-2A:pVHb-trp1 (Δ). The batch phase of the fermentation were carried out at 30^{0}C, pH 5.0, 250 rpm and a constant air flow rate of 0.4 L/min. Arrows

indicate time of acetaldehyde feeding. 50% acetaldehyde is commenced at 2.5 mL/hr. Growth was measured in term of OD_{590}.

Figure 1.

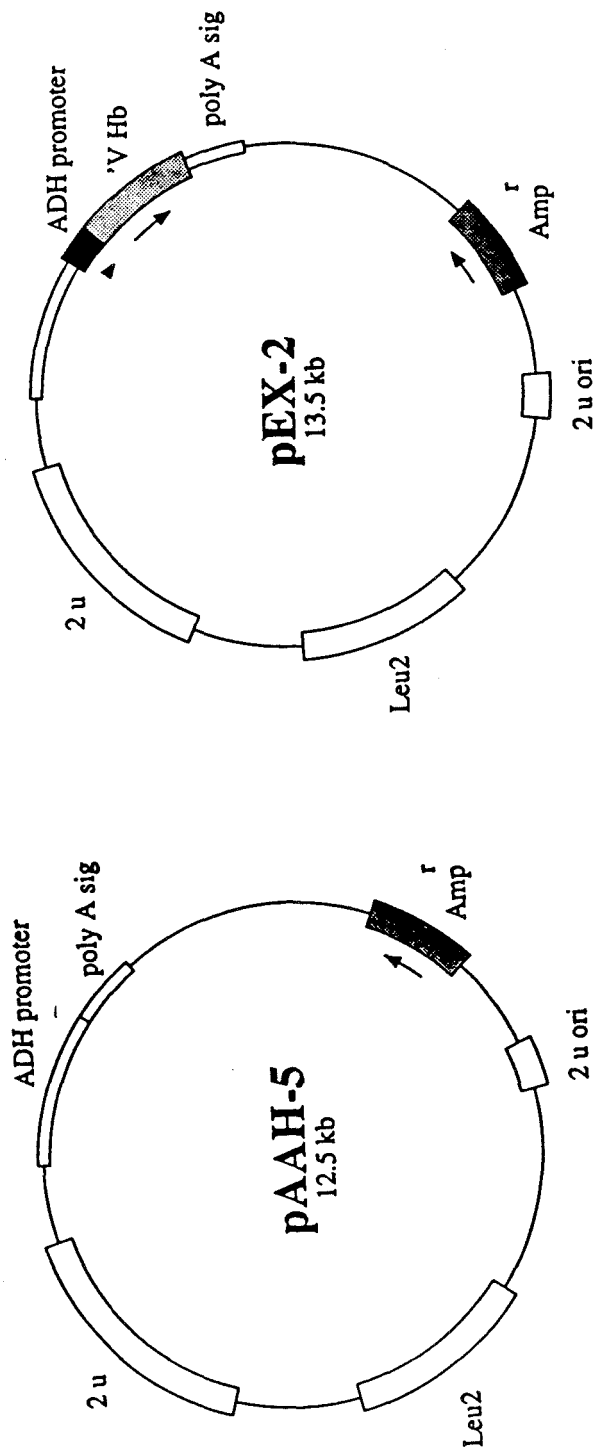

ADH promoter
'V Hb
poly A sig
r
Amp
2 u ori
2 u
Leu2
pEX-2
13.5 kb
ADH promoter
poly A sig
r
Amp
2 u ori
2 u
Leu2
pAAH-5
12.5 kb

Figure 2.

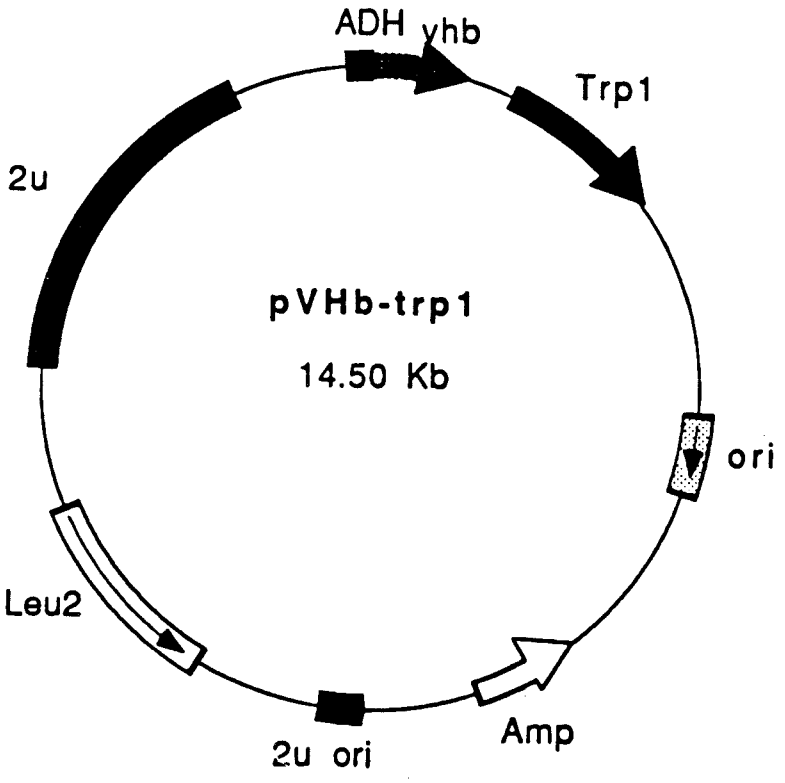


Figure 3.

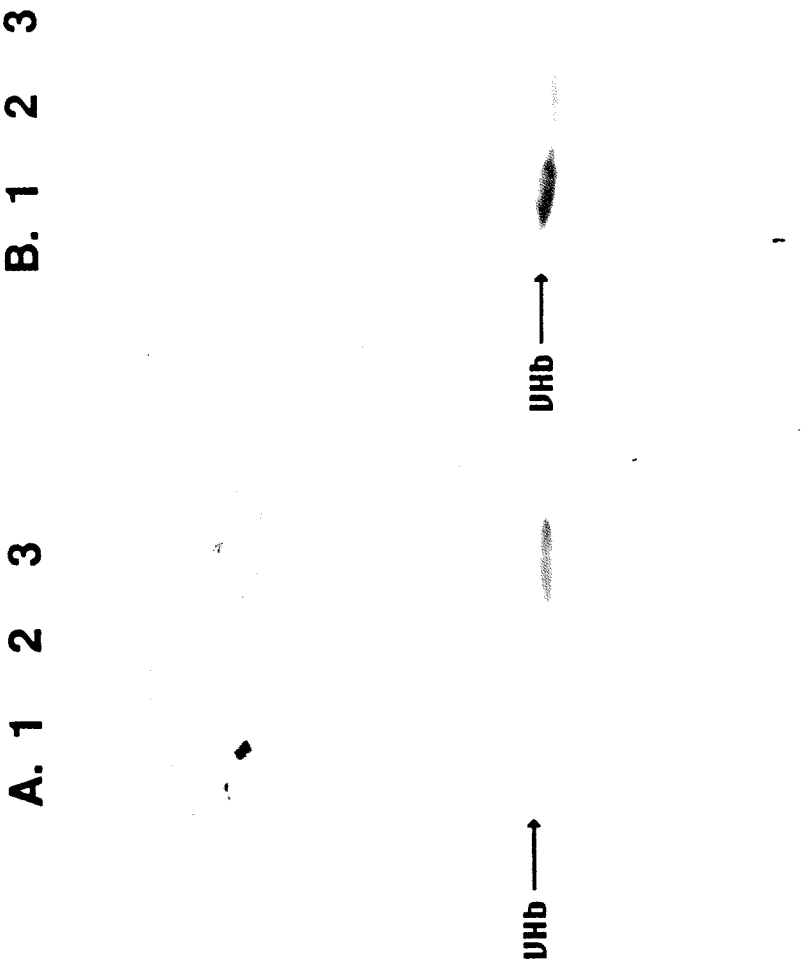


Figure 4.

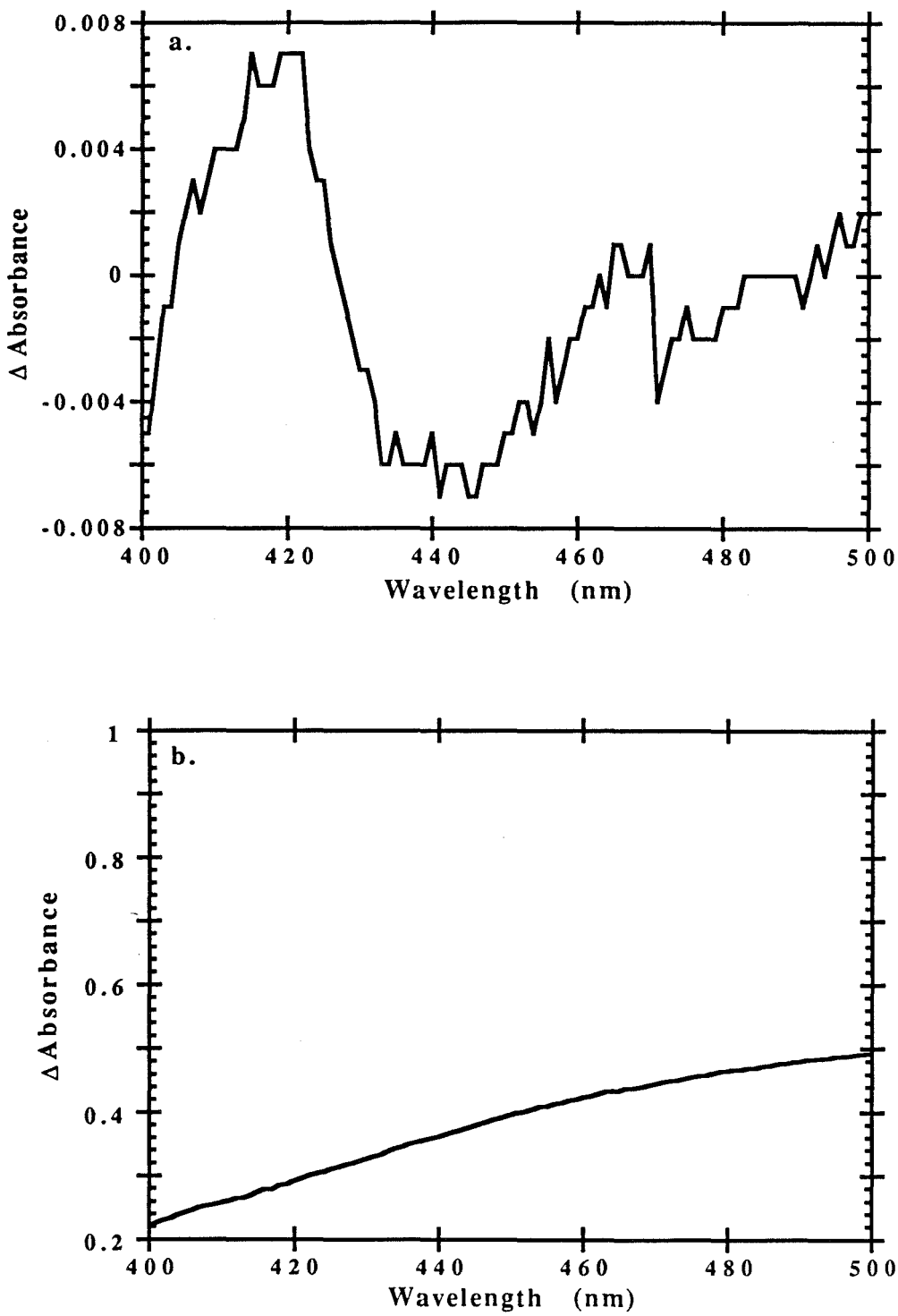
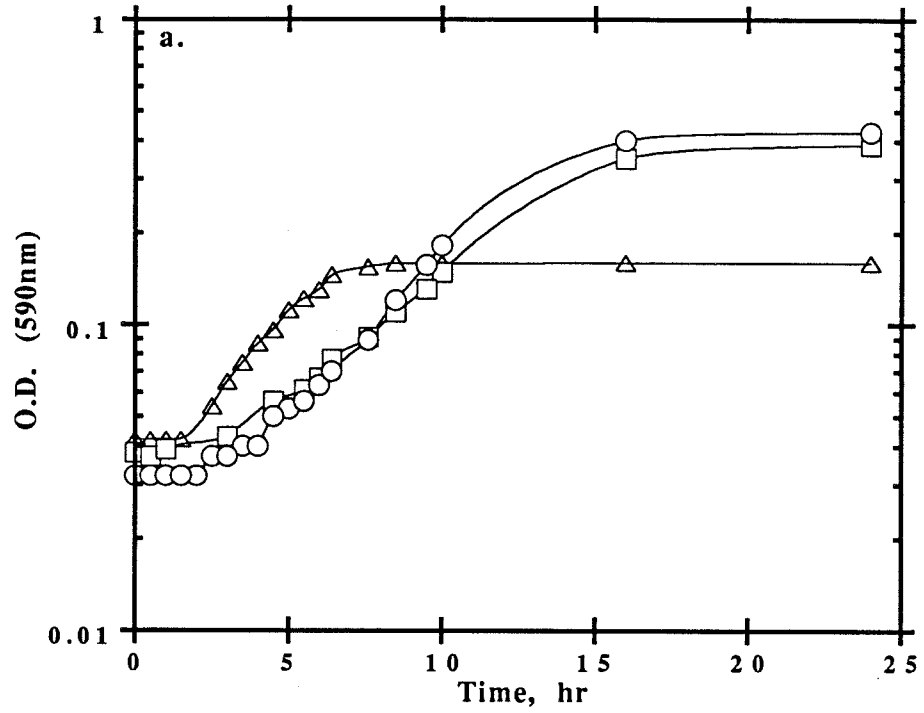

a.
0.008
0.004
0
-0.004
-0.008
Δ Absorbance
400
420
440
460
480
500
Wavelength (nm)
b.
1
0.8
0.6
0.4
0.2
Δ Absorbance
400
420
440
460
480
500
Wavelength (nm)

Figure 5.

Figure 6.


a.
1
0.1
0.01
O.D. (590nm)
0
5
10
15
20
25
Time, hr

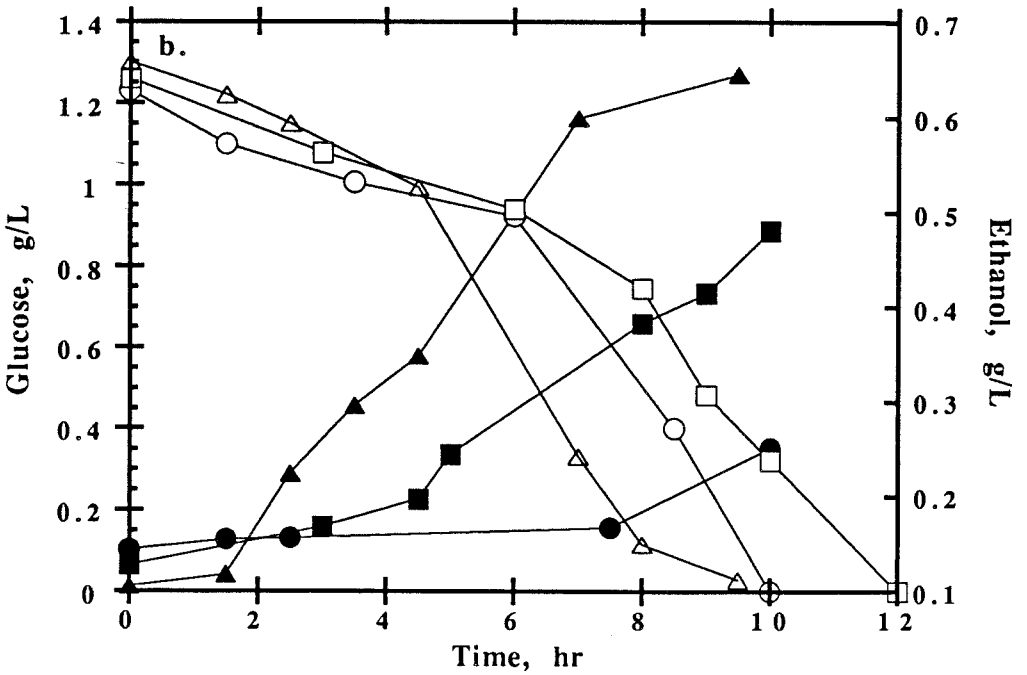
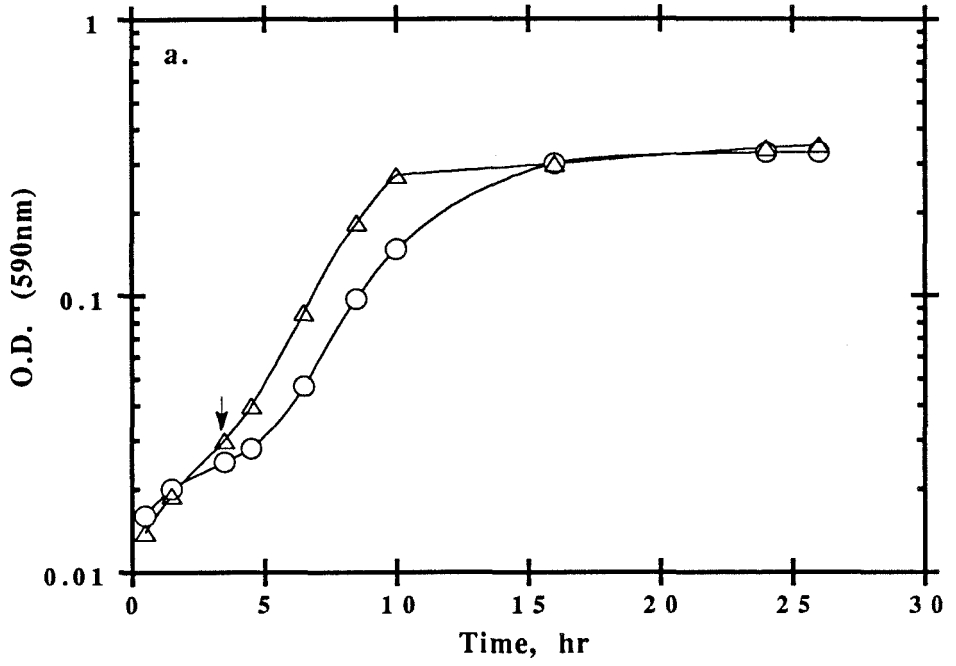

b.
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Glucose, g/L
0.7
0.6
0.5
0.4
0.3
0.2
0.1
Ethanol, g/L
0
2
4
6
8
10
12
Time, hr

Figure 7.


a.
O.D. (590nm)
1
0.1
0.01
0
5
10
15
20
25
30
Time, hr

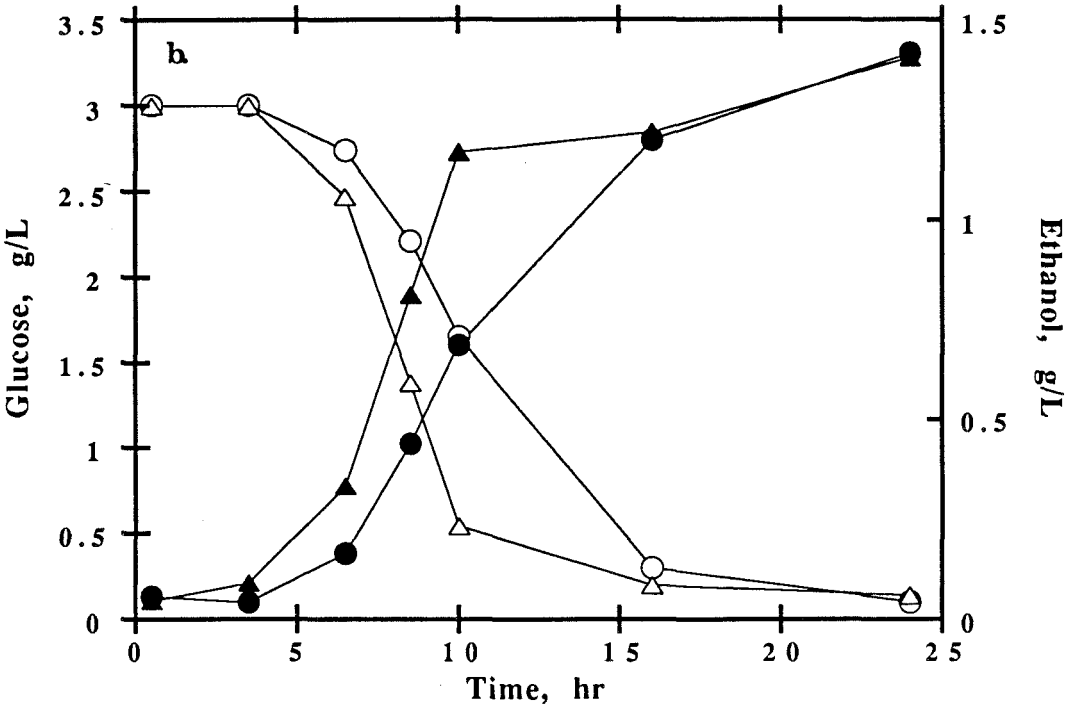

b.
Glucose, g/L
Ethanol, g/L
3.5
3
2.5
2
1.5
1
0.5
0
1.5
1
0.5
0
0
5
10
15
20
25
Time, hr

Figure 8a.

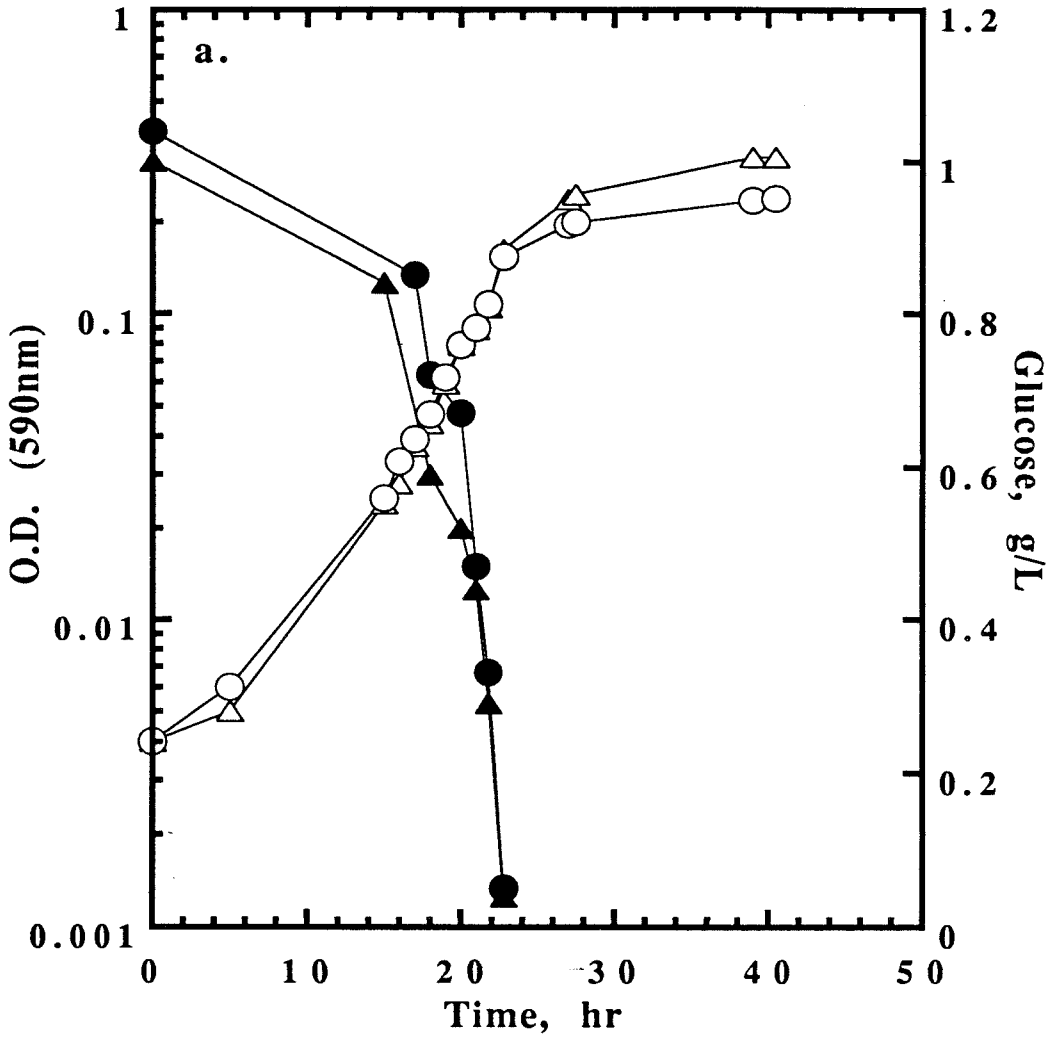

a.
1
0.1
0.01
0.001
O.D. (590nm)
1.2
1
0.8
0.6
0.4
0.2
0
Glucose, g/L
0
10
20
30
40
50
Time, hr

Figure 8b.

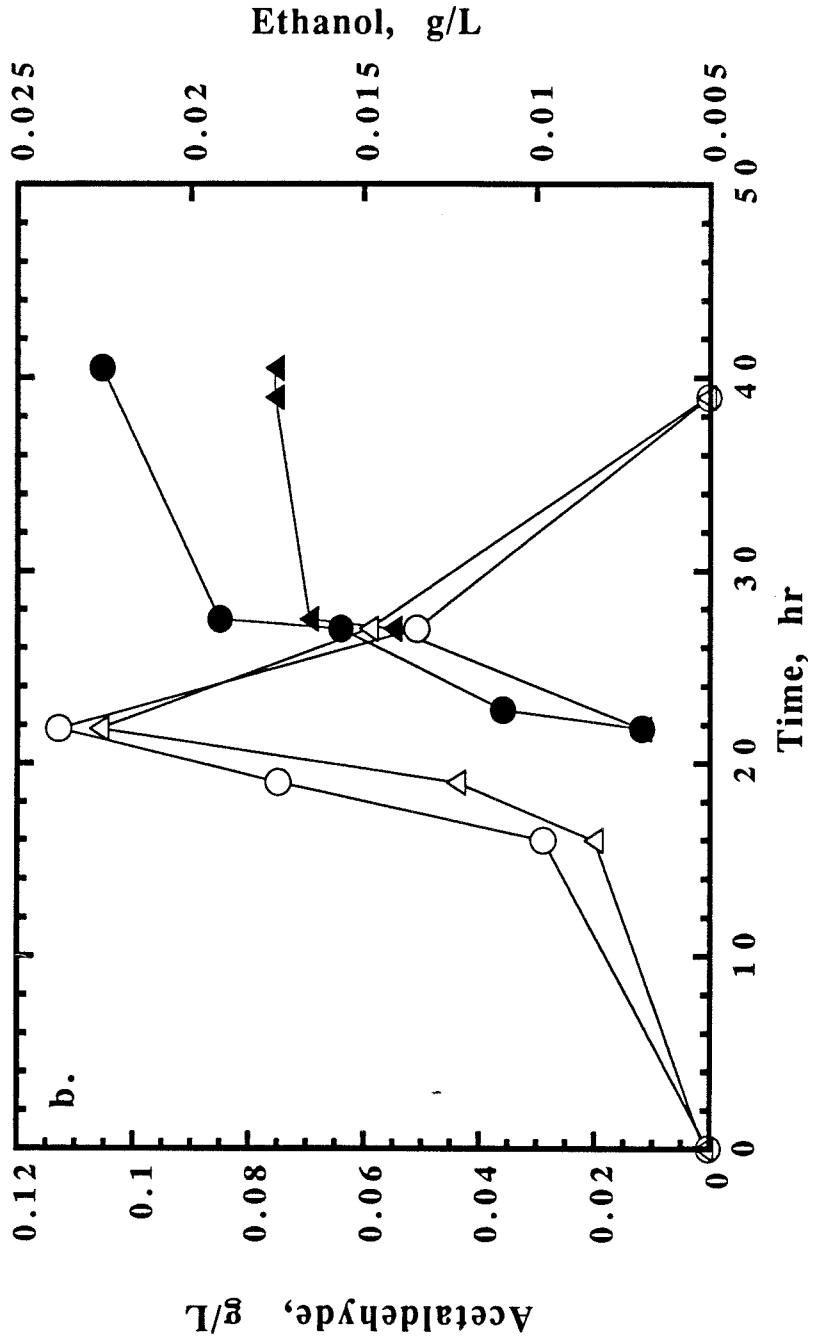

Ethanol, g/L
0.025
0.02
0.015
0.01
0.005
0
10
20
30
40
50
Time, hr
b.
0.12
0.1
0.08
0.06
0.04
0.02
Acetaldehyde, g/L

Figure 9.