# Expression of Human Cystatin C by the Methylotrophic Yeast *Pichia pastoris*

by

David Andrew Files

Bachelor of Engineering Science, The University of Western Ontario, 1997

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Applied Science

in

The Faculty of Graduate Studies

Department of Bio-Resource Engineering

We accept this thesis as conforming
to the required standard

The University of British Columbia

April 2000

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Bio-Resource Engineering

The University of British Columbia
Vancouver, Canada

Date April 13/2000

DE-6 (2/88)

# Abstract

Human cystatin C is a cysteine-proteinase inhibitor with great potential as a therapeutic protein. The methylotrophic yeast, *Pichia pastoris*, was selected as a host to express a human cystatin C variant in which site directed mutagenesis was used to introduce two consensus sequences for N-linked glycosylation. Under the regulation of the methanol-inducible *AOX1* gene promoter, both glycosylated and nonglycosylated cystatin were expressed simultaneously.

Shake-flask experiments were performed to investigate the effects on cell density of glycerol and glucose as carbon sources and yeast extract, yeast nitrogen base, and ammonium sulfate as nitrogen sources. Glycerol was found to be superior to glucose, and the addition of yeast extract or ammonium sulfate (up to 4.3 $g \cdot l^{-1}$) or yeast nitrogen base (up to 6.8 $g \cdot l^{-1}$) had positive effects on the cell densities.

The effects of pH and of feeding both methanol and glycerol during induction were investigated in a 2-liter bioreactor. The maximum expression was achieved at pH 6 and when only methanol was fed during induction. Glycerol feeding during induction showed potential to increase the concentration and yield of biologically active cystatin. The maximum concentration of biologically active cystatin in the culture supernatant was 54 $\mu moles \cdot l^{-1}$, equivalent to 0.72 $g \cdot l^{-1}$ of human cystatin C. The maximum yield of biologically active cystatin was 1.0 $\mu mole \cdot g^{-1}$ dry cell weight.

The effects of gene copy number as well as methanol utilization were considered. Resistance to the Zeocin™ antibiotic was used as a base for isolating transformants that were suspected to have a range of copy numbers. An inverse relationship between Zeocin™ resistance and active cystatin expression was observed. It is hypothesized that transformants with a higher resistance to Zeocin™ encountered inefficiencies in their secretion pathways. In terms of methanol utilization phenotype, the slow-growing transformant ($Mut^S$) expressed higher concentrations and higher yields of active cystatin. The fast-growing transformant ($Mut^+$) experienced oxygen mass transfer limitations within the 2-liter bioreactor.

# Table of Contents

# List of Tables

# List of Figures

# Preface

This thesis summarizes the research that was performed by David Files under the supervision of Dr. Susan Baldwin in the Department of Bio-Resource Engineering at the University of British Columbia. The project was a collaborative effort with Dr. Masahiro Ogawa and Dr. Christine Scaman of Food, Nutrition, and Health at the University of British Columbia.

Achieving the objectives of this project involved many stages. For those of us in Bio-Resource Engineering, the collaboration with Food, Nutrition and Health was the first stepping-stone. The wealth of information that they had gathered over the years, and the success of their previous experiments were a valuable asset that we were made privy to. The information from Food, Nutrition and Health was then complemented with an extensive literature review of the topic. From there, materials and methods were shared between the departments and the expression of cystatin C had expanded into two departments on the UBC campus. Preliminary experiments were performed on a small scale and were followed with bench-scale experiments in a bioreactor.

Troubleshooting was a group effort throughout the project. Gary Lesnicki and Marta Guarna provided expert opinion and peer feedback was enthusiastically received at conferences (Current Topics in Gene Expression Systems, San Diego, CA, March 28 – 31, 1999 as well as The 49$^{th}$ Canadian Chemical Engineering Conference, Saskatoon, SA, October 3 – 6, 1999).

As partial fulfillment of the degree of Masters of Applied Science, this thesis is a culmination of research efforts during the past two years.

# Acknowledgments

Dr. Susan Baldwin was the first to inform me of this exciting project. I would like to thank her for her initial efforts and her continued support and enthusiasm. As my supervisor, she has been an excellent source of knowledge and feedback.

A special thanks goes out to our collaborators in the Food Science Department. Dr. Christine Scaman's research efforts have made this project possible. We in the Department of Bio-Resource Engineering appreciate her invitation for us to participate in the project. Her expertise in recombinant DNA technology was able to shed light on many of the challenging issues that arose. It has also been a pleasure to work with Dr. Masahiro Ogawa. His laboratory door was always open and he showed patience and skill as he taught me the methods of genetic modification.

During the initial stages of this project, Dr. Jamie Piret guided me through the process of gathering and analyzing the masses of information relating to the topic. As I have a great respect for his opinion, I would like to thank him for being a member of my evaluation committee. I would also like to thank Dr. Marta Guarna and Gary Lesnicki for their expert opinions. Their years of experience with the engineering aspects of *Pichia pastoris* were an asset to the project.

From before this project began for me, and right through to its completion, my family and friends have been a tremendous source of encouragement. On top of this, they have shown interest and understanding in the work that I have done.

Funding for this work was provided for by a scholarship from NSERC (The Natural Sciences and Engineering Research Council of Canada) and an NSERC operating grant.

# Chapter 1

# Introduction

## 1.1 The Evolution of Biotechnology

The fascinating field of pharmaceutical biotechnology evolved from a moldy Petri dish. When cleaning his lab one day in 1928, Alexander Fleming noticed that some mold was growing on one of his discarded bacterial cultures. To his amazement, the bacteria all around the mold had been killed. This mold was found to be of the *Penicillium* genus. When asked about his contribution to science, Fleming replied, "Nature makes penicillin, I just discovered it." In his time, that is exactly what biotechnologists did, they looked for ways to isolate, harvest, and administer pharmaceuticals that were available in the natural environment. Biotechnologists today operate under a similar principle, but they are armed with the knowledge of DNA and the tools that allow them to recombine genetic material.

After the discovery of the three dimensional structure of DNA, geneticists learned how genetic information is stored and expressed. With this new understanding, it soon became evident that some genetic diseases occurred because a certain protein is not produced, or not produced properly, by the organism. For example, people with diabetes either do not produce enough insulin to properly metabolize glucose, or the insulin they have is ineffective. This discovery led to attempts to treat these types of genetic diseases by replacing the missing protein. As was the case for penicillin, biotechnologists isolated human-like insulin from natural sources. In its first application, insulin from pigs and cows was harvested and used it to treat the human condition of diabetes. Advances in recombinant DNA technology now allow for mass-production of authentic human insulin that is safer and more economical than its isolation from nature.

Today, geneticists have focused their attention to mapping the entire genome of human DNA. This ambitious task has been made feasible by advances in gene cloning and automated DNA sequencing technologies. Of the estimated 100,000 genes in the genome, many of them encode for proteins whose functions are unknown (Kost 1997).

To aid in the effort of identifying novel proteins, bioinformatics (the powerful combination of mathematics, computer science, and biology), is being used to classify and even predict protein function based on known protein sequences and structures. However, in order to validate predictions based on bioinformatic analysis, systems for the production of sufficient quantities of authentic proteins are required. As a result, recombinant DNA technology will be called upon to play a key role in humankind's genetic efforts in the 21$^{st}$ century.

## 1.2 Objectives of this Project

The overall objective of this project was to maximize the expression of a recombinant variant of human cystatin C using the methylotrophic yeast *Pichia pastoris* as a microbial host. To achieve this goal, experiments were designed to identify and quantify the most important factors for gene expression in *P. pastoris.* The factors considered for these experiments were 1) the growth medium, 2) the cystatin gene copy number, 3) the pH for expression, 4) the substrate feeding strategy used in the bioreactor, and 5) the *P. pastoris* phenotype.

## 1.3 Motivation for this Project

This project was undertaken in order to make a contribution to the field of pharmaceutical biotechnology. The main impetus for this project is cystatin C's potential as a therapeutic protein. In the future, recombinant cystatin C may save lives. Not only this, the knowledge gathered by studying cystatin C would be applicable to many other potential pharmaceuticals. The same is true for the yeast *Pichia pastoris.* As an expression system for recombinant proteins, *P. pastoris*, shows great potential. In fact, *P. pastoris* has become the expression system of choice for many recombinant proteins. With the plethora of novel genes being generated by the genome project, an efficient and safe expression system will be vital to express these genes as recombinant proteins. The experiments that are described in this thesis do in fact make a contribution to the field of pharmaceutical biotechnology.

## 1.4 Thesis Layout

This thesis is divided into chapters, each describing a specific aspect of the optimization process. The current chapter introduces the reader to the topic and defines the objectives of the project and the layout of the thesis. Chapter 2 provides background information to brief the reader on the applicable technology and terminology. Chapters 3, 4, and 5 describe the experiments that were performed and discuss the implications of the findings. Each of these chapters includes an introduction that reviews the relevant literature. Chapter 6 provides overall conclusions based on the experiments and makes recommendations for future experiments. A list of nomenclature and the cited references follow Chapter 6. Finally, the Appendices have lists of recombinant therapeutic proteins approved by the FDA, proteins that have been expressed by *P. pastoris,* medium formulations and tables of raw data.

# Chapter 2

## Background Information

This chapter provides background information on recombinant DNA technology, protein expression systems, cystatin C and *Pichia pastoris*. A review of these topics is important to achieve an understanding of the experiments that are described in Chapters 3, 4, and 5.

### 2.1 Therapeutic Recombinant Proteins

In 1982, insulin became the first therapeutic recombinant protein to gain approval by the FDA (Steinberg and Raso 1998). Since this time, there has been an overwhelming increase in the numbers and varieties of recombinant proteins that have been approved for therapeutic applications. The list of human health conditions for which the FDA has approved recombinant proteins includes AIDS, anemia, certain cancers (Kaposi's sarcoma, leukemia, and colorectal, kidney, and ovarian cancers), certain circulatory problems, certain hereditary disorders (cystic fibrosis, familial hypercholesterolemia, Gaucher's disease, hemophilia A, severe combined immunodeficiency disease, and Turner's syndrome), diabetic foot ulcers, diphtheria, genital warts, hepatitis B, hepatitis C, human growth hormone deficiency, and multiple sclerosis. A list of various FDA-approved recombinant proteins is provided in Appendix I.

Therapeutic recombinant proteins are derived from genetically modified organisms. Multi-cellular organisms such as cows, goats, sheep, and rats, and corn, potato, and tobacco plants, have been genetically engineered to produce proteins that are medically useful to humans. Single-celled microorganisms such as bacteria, yeast, and mammalian cells have also been modified to produce recombinant proteins. Genetically modified yeast cells were used in this work for the production of a recombinant variant of cystatin C.

### 2.1.1 Recombinant DNA Technology

Recombinant DNA technology, also known as genetic engineering, bioengineering, and gene splicing, refers to the transfer of genetic material from one organism into another organism. The result is a recombination of DNA from different sources. This process usually involves the isolation of a human gene that has therapeutic potential and its introduction into a strain of bacteria, yeast, or mammalian cell. These cells grow and multiply and are induced to produce the protein in high quantities under controlled conditions. As a result, large quantities of a highly purified protein can be made available for studies and clinical use.

The DNA from humans can be introduced to other organisms as follows: Using proteins called restriction enzymes, individual genes from human DNA can be isolated and separated from the genome. These genes can then be introduced into small circular pieces of DNA called plasmid vectors. Vectors are used to carry genetic information that can be shared between different types of organisms. Prior to insertion of the human DNA, the vector is cut with the same restriction enzymes as those used to isolate the human gene. The human gene is then fixed in place within the vector using an enzyme called DNA ligase. Essentially, restriction enzymes and DNA ligase are the scissors and glue of recombinant DNA technology. The recombinant plasmid is then inserted into a bacterial, yeast, or mammalian cell by the process of transformation. Methods of transformation are presented in section 2.4.3.

Transformed cells can be separated from non-transformed cells by a selection procedure that takes advantage of an antibiotic-resistance gene that is included in the vector. The process of cloning can then establish a homogenous population of transformed cells. In cloning, a single cell is selected that gives rise to a whole population of identical cells, or clones, by normal cell division. All of the resultant cells would be expected to contain a copy of the vector carrying the inserted human gene.

Once the cell line carrying the vector has been cloned and cultured, the cells can be induced to express the human gene. Depending upon the protein, the cell expression system, and the vector that was selected, the recombinant protein may be retained inside the cells (intracellular), or secreted out of the cell and into the culture medium (extracellular).

## 2.2 Cystatin C as a Therapeutic Recombinant Protein

### 2.2.1 Cystatin C in Nature

In humans, cystatin C serves a protective function to regulate the activities of cysteine proteinase enzymes. Without cystatin C, these proteinase enzymes would cause uncontrolled protein degradation and tissue damage (Abrahamson et al. 1986). Cysteine proteinase enzymes have a cysteine amino acid in their active site and they degrade proteins by the process of hydrolysis. Cystatin C binds to these proteinase enzymes and renders them inactive.

Cystatin C is a member of the cystatin superfamily of cysteine proteinase inhibitors. Human cystatin C is part of the Family 2 cystatins whose members have approximately 120 amino acid residues and two disulfide bonds. These cystatins are present in tissues and secreted into bodily fluids. Cystatin C expressed by other mammals, such as chickens and mice, share a high degree of homology to human cystatin C and perform similar functions. Chicken cystatin C shares 42% homology with human cystatin C (Barrett et al. 1984).

Human cystatin C shows promise as a therapeutic protein. Apart from its regulation of cysteine proteinases, cystatin C may also be involved in the defense against bacteria and viruses. Various cell culture experiments have indicated that human cystatin C inhibits the growth of bacteria such as group A streptococci (Björck et al. 1989) and pathogenic *Porphyromonas gingivalis* (Blankenvoorde et al. 1998) as well as viruses such as herpes simplex type 1 (Björck et al. 1990). The key to cystatin C's inhibitory activity is its strong and site-specific binding to proteinase enzymes. Equimolar binding occurs for ficin, papain, and cathespin B. Cystatin C also binds dipeptidly peptidase I and scarcely binds bromelain, trypsin, chymotrypsin, as well as various bacterial proteinases (Barrett 1981).

In plant cells, the defense role for cystatin has been verified. Cystatin in chestnut seeds (called cystatin CsC) has antifungal properties and is expressed in plants that are subjected to chemical and climatic stresses (Pernas et al. 2000). Cystatin CsC is also active against digestive proteinases from the flour beetle (*Tribolium castaneum*) and the dust mite (*Dermatophagoides farinae*), two significant agricultural pests. Cystatin CsC

is similar to cystatin C in that they both inhibit papain, ficin, chymopapain, and cathepsin B (Parnas et al. 1998).

Cystatin C is produced by all nucleated cells at a constant rate and is often secreted into the extracellular environment (Randers et al. 1998). In humans, cystatin C is most abundant in the cerebrospinal fluid and in seminal fluid (Abrahamson et al. 1986). Its concentrations in blood serum, cerebrospinal fluid, and seminal fluid are approximately 1, 5, and 49 $mg \cdot l^{-1}$ respectively (Barrett et al. 1984). In chickens, cystatin C concentrations in blood serum and egg whites are 1 and 60 $mg \cdot l^{-1}$ (Barrett 1981). Despite its relatively high concentration in egg whites, there have been no transgenic attempts to express human cystatin C in chickens.

### 2.2.2 Cystatin C in the Human Body

Various medical conditions arise from too much or too little cystatin C in the human body. In most cases, the problem arises due to insufficient concentrations or a mutation in the cystatin C gene that produces a nonfunctional protein.

A cystatin C deficiency has been linked to vascular disease (Shi et al. 1999). In locations of aortic lesions, it has been observed that the cystatin C expression is severely reduced. In its absence, cysteine proteinases such as cathepsins S and K are free to break down the elastic tissue of the vascular wall, resulting in either swelling (aneurysm) or hardening (atherosclerotic) of the arteries. It has been confirmed by in vitro studies that cystatin C does in fact inhibit cathepsins secreted from vascular muscle cells and would likely prevent such lesions from occurring.

A mutated variant of cystatin C causes the condition known as cerebral amyloid angiopathy (CAA) (Shimode et al. 1996). This condition is increasingly recognized as a major cause of strokes in the elderly. The term cerebral amyloid angiopathy refers to protein deposits in blood vessels of the brain that can allow blood to leak out and cause hemorrhagic strokes. The protein deposits are aggregates of mutated cystatin C. The mutant is missing 10 amino acids from its N-terminal and has a L68Q substitution (leucine (L) at position number 68 is replaced by glutamine (Q))(Wei et al. 1998). Its activity is similar to wild-type cystatin C, but it forms dimers (self-aggregates) at lower concentrations and shows an increased susceptibility to a serine proteinase in human

cerebrospinal fluid. As a result, the mutated cystatin C is depleted in the cerebrospinal fluid and is deposited in the brain of CAA patients. Therefore, measurement of cystatin C in the cerebrospinal fluid is a useful method of diagnosing conditions relating to cerebral amyloid angiopathy.

Cystatin C monitoring can also be useful for the recognition of abnormal kidney function. Blood serum cystatin C can facilitate the recognition of diabetic nephropathy, the most serious secondary complication of diabetes (Bokenkamp et al. 1998a). Patients suffer from a progressive loss of kidney function, which ultimately requires dialysis and frequently results in the need for a kidney transplant. Cystatin C has been proposed as a marker of kidney function by measuring its glomerular filtration rate (GFR). Cystatin C is produced at a constant rate in humans and is eliminated in the kidneys by glomerular filtration (Randers et al. 1998). Concentrations of cystatin C in blood serum have shown better correlations with GFR than concentrations of creatinine. Unlike creatinine, cystatin C has the advantage of being independent of gender and muscle mass.

### 2.2.3 Therapeutic Applications for Cystatin C

Cystatin C has been suggested as an aid for the treatment of the neurological condition known as sciatica (Grubb and Löfberg 1988). The back pain associated with sciatica results from compression of a spinal nerve root. A non-surgical option for treatment is an intradiscal injection of chymopapain (LeBlanc 1991). Although the actual mechanism of action is not clear, chymopapain acts to lyse the nucleus of cells associated with compression of the nerve and may possess anti-inflammatory properties that provide rapid relief from compression (Sawin et al. 1997). However, accidental chymopapain injection into the cerebrospinal fluid results in massive cerebral hemorrhage due to saturation of the cysteine proteinase-inhibiting capacity of the cerebrospinal fluid (Grubb and Löfberg 1988). Since cystatin C has been shown to inhibit chymopapain activity, an injection of cystatin C might prevent this side effect (Buttle et al. 1986).

Cystatin C may also be beneficial for inhibiting tumour metastasis (Cox et al. 1999). Experiments indicated that the metastic ability of tumour cells was greatly

reduced for a melanoma cell line that was altered to overexpress cystatin C. Cystatin C was shown to inhibit the cysteine proteinase enzymes that were correlated with tumour metastasis.

Considering its many functions and applications, cystatin C has the potential to be an important therapeutic protein. An efficient expression system for human cystatin C and its mutated variants would allow for clinical trials for human cystatin C and studies of structure/function relationships for the mutants. Fortunately, cystatin C has physical and chemical properties that are suitable for recombinant expression.

### 2.2.4 Physical and Chemical Properties of Cystatin C

Human cystatin C is composed of a single polypeptide chain of 120 amino acid residues and has a calculated molecular weight of 13.4 kDa (Grubb and Löfberg 1982). A schematic diagram of human cystatin C is presented in Figure 2.1. The branched chains that are attached to the protein in this schematic diagram represent carbohydrate polymers that can be attached by the process of glycosylation. Information about glycosylation is provided in section 2.4.5.

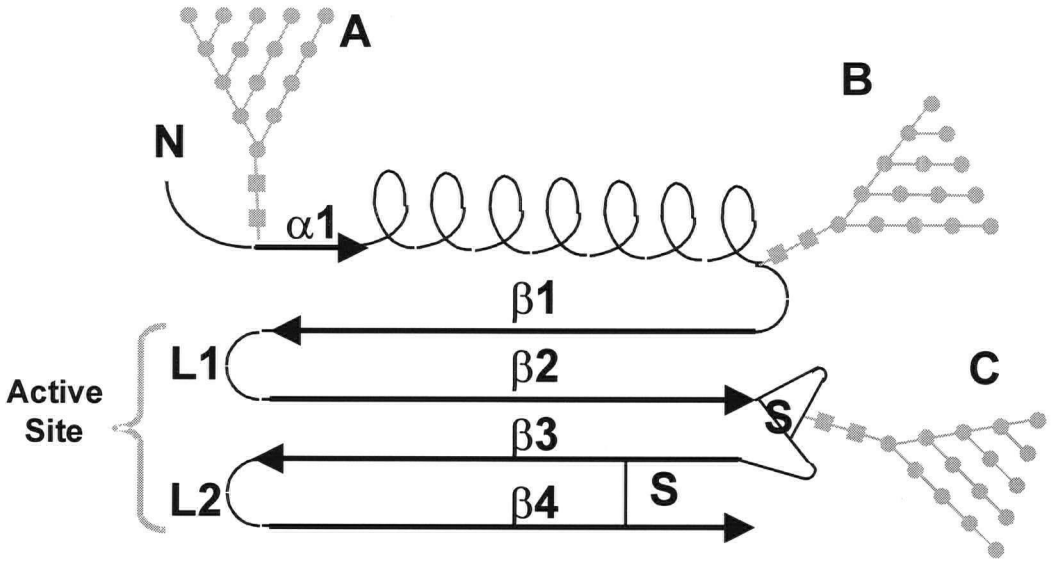


Figure 2.1. Schematic diagram of human cystatin C. The two β-hairpin loops (**L1** and **L2**) and the protein's N-terminal (**N**) form a wedge-shaped edge (indicated by a star) that binds to cysteine proteinase enzymes and renders them inactive. The numbering starts with the first amino acid (serine) and proceeds to the 120$^{th}$ amino acid (alanine). Cystatin C is a tightly folded protein, having regions of α-helix (α) and β-sheet (**β**) as well as two disulfide bridges (**S**). The attached branches (**A**, **B**, and **C**) represent carbohydrate chains that can be attached to genetic variants of cystatin C.

The active site of cystatin C is indicated in Figure 2.1 with a star. The two β-hairpin loops as well as the 10 amino acids that make up the N-terminal are jointly involved in cystatin C's inhibitory properties. The three-dimensional shape of this region compliments the shape of the active site of cysteine proteinase enzymes (Bode et al. 1988). X-ray crystal structures and mathematical models have been used to study the shapes of active sites for cystatin and its mutants and compare their binding compatibility with various cysteine proteinases.

Cystatin C's compact structure and disulfide bridges impart excellent physical and chemical stability. Cystatin C is completely stable to 80°C for 10 minutes (pH 6.5) and is stable down to pH 2.0 at 25°C for 10 minutes in a 0.10 M glycine/HCl buffer (Barrett et al. 1984). Many biopharmaceutical substances lack stability and can become denatured upon passage through the gastrointestinal tract, the lungs, or the skin. Therefore, cystatin C's stability at low pH make ingestion a feasible option for its delivery. Ingestion has been shown to be effective for a biopharmaceutical that elicits

an immune response to cholera. A strain of potato has been genetically modified to produce this heat- and pH-resistant protein (Arakawa et al. 1998).

However, human cystatin C is vulnerable to oxidation (Berti et al. 1997), dimerization (Abrahamson and Grubb 1994b; Ekiel and Abrahamson 1996), and proteolysis (Lenarcic et al. 1991), all of which reduce its stability. These factors have to be considered when cystatin C is expressed as a recombinant protein in a bioreactor. For example, the oxygen rich environment required by aerobic expression systems has been shown to oxidize methionine residues in cystatin C. In addition, acidic byproducts produced by the cells can lower the bioreactor pH and induce cystatin dimerization. These self-aggregates of cystatin, which have been observed below pH 5, are difficult to break apart and they exhibit no inhibitory activity. Finally, protease enzymes that are produced by the host cell in order to protect it from foreign proteins can degrade cystatin C. A detailed discussion concerning each of these issues is presented in Chapter 5.

### 2.2.5 Measurement of Cystatin C

Human cystatin C can be identified, quantified, isolated, and purified with basic chemical and physical processes. Some of these techniques will be described in detail in Chapter 4.

The papain inhibitory assay is a convenient method for identifying and quantifying biologically active cystatin C (Barrett 1981). The assay makes use of the proteinase enzyme papain and its activity on the substrate benzoyl-DL-arginine p-nitroanilide hydrochloride (BAPNA). The assay measures the relative reduction in activity of papain due to irreversible binding by cystatin C. Since cystatin binds to papain equimolarly, a reduction in the papain activity is correlated with the presence of biologically active cystatin. The nitroaniline product of papain's residual proteinase activity can then be measured using a spectrophotometer. With minor modifications, the assay can also be performed with cathepsin B instead of papain.

The inhibitory assay, in combination with a verification of its molecular weight can give a good indication of the activity and authenticity of cystatin C. The molecular

weight can be approximated by SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) (Laemmli 1970).

The authenticity of cystatin C can be verified by radial immunodiffusion or western blot analysis (Mancini et al. 1965). For western blots, cystatin C is incubated with rabbit antiserum raised against human cystatin C. The antigen is then reacted with goat anti-(rabbit IgG).

Purification of cystatin C can be performed by affinity chromatography on carboxymethyl-papain (Cm-papain) linked to Sepharose (Barrett 1981). The final purification stages make use of gel chromatography on Sephadex G-75 or ultrafiltration cycles.

### 2.2.6 Expression of Recombinant Cystatin C

The efficient production of recombinant cystatin C would serve several purposes. As a potential therapeutic protein, in vivo trials aimed at increasing the cysteine proteinase inhibitory capacity of biological fluids could be performed. In addition, a recombinant system could be used for site-directed mutagenesis to produce variants of cystatin C for studies of its structure-function relationships.

To date, cystatin C has only been expressed in bacteria. Intracellular expression of human cystatin C has been achieved in recombinant *E. coli* at levels of 1000 $mg \cdot l^{-1}$ (Dalboge et al. 1989) and 30 $mg \cdot l^{-1}$ (Berti et al. 1997). In recombinant *E. coli*, cystatin C is expressed intracellularly and comprises less than 10% of the total soluble cell protein (Abrahamson et al. 1988). A product loss of 50% could be expected as a result of purification processes to separate cystatin C from *E. coli*'s endogenous proteins.

Extracellular expression of cystatin C has been achieved by a protease deficient strain of *Bacillus subtilis* (C. Scaman, per. commun. 1997). Recombinant cystatin C was produced at 2 $mg \cdot l^{-1}$.

### 2.2.7 Expression of a Recombinant Cystatin C Variants

As previously mentioned, human cystatin C is vulnerable to oxidation (Berti et al. 1997), proteolysis (Lenarcic et al. 1991), and dimerization (Abrahamson and Grubb 1994b; Ekiel and Abrahamson 1996), all of which reduce its stability. In the case of

mouse cystatin C, the presence of carbohydrate polymers attached by the process of glycosylation conferred resistance to heat deactivation and additional stability to proteolytic degradation compared to the nonglycosylated protein (Nakamura et al. 1998a). These results are shown in Figure 2.2.

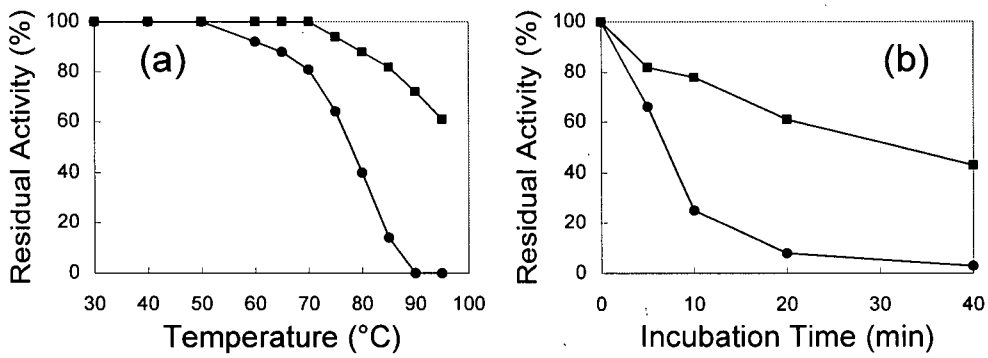


Figure 2.2. Heat stability **(a)** and susceptibility to proteolysis by cathepsin D **(b)** for recombinant cystatins. Shown here is glycosylated cystatin (■) compared to nonglycosylated cystatin (●). The residual activity of each incubated sample was determined by measuring the papain-inhibiting activity and comparing it with that of the corresponding sample without heating **(a)** or exposure to cathepsin D **(b)**. Reproduced from Nakamura et al. (1998a).

Other variants of human cystatin C have been expressed to study its functional properties. Expression of a cystatin C variant that lacked the N-terminal decapeptide (as is the case for the cystatin responsible for cerebral amyloid angiopathy (CAA)) was performed in *E. coli* (Abrahamson et al. 1991). Its proteinase inhibitory activity was 5-fold less than native human cystatin C. Subsequent studies supported the fact that this variant of cystatin C forms an inactive dimer much more readily than wild-type cystatin C, especially at elevated temperatures (Abrahamson and Grubb 1994b). These findings suggested that sufferers of CAA should seek treatment for fevers immediately, to avoid increases in body temperature that could trigger the amyloid formation that leads to cerebral hemorrhage.

In similar studies, cystatin C variants that had amino acid substitutions in the N-terminal region and in the active binding site were expressed in *E. coli* (Hall et al. 1995). A total of 10 variants were isolated, structurally verified, and compared to wild-type cystatin C. These studies confirmed the importance of the authenticity of the

N-terminal decapeptide. Interestingly, one of the variants showed increased inhibition of cathepsin S compared to wild-type cystatin C.

## 2.3 Hosts for Recombinant Proteins

As mentioned before, insulin is one of the classic therapeutic proteins. Initially, insulin was harvested from organisms that naturally produce a protein that was very similar to human insulin. With the establishment of recombinant DNA techniques, the human gene that encodes insulin has been transformed into organisms that do not normally produce insulin. As a result, large quantities of biologically authentic human insulin could be produced and easily harvested. For many years, the organism of choice as a host for recombinant proteins has been the bacterium *Escherichia coli*.

To date, a wide variety of biological systems have been developed for producing recombinant proteins. These include bacterial, yeast, fungal, viral, mammalian and insect cell systems as well as plants and transgenic animals (Kost 1997). These expression systems are useful for the production of large quantities of proteins for structural analysis, use as therapeutic agents, and various other applications. A wide variety of expression systems are required due to the vast differences in the types of proteins that are being expressed. Appendix I includes lists of some of the various types of recombinant proteins. No one host can provide the optimum conditions for every recombinant protein. The following sections describe some of the expression systems that are commonly used.

### 2.3.1 Non-Yeast Expression Systems

#### *Bacterial Expression Systems*

Bacterial host systems, such as those based on *Escherichia coli*, serve an important role for the production of recombinant proteins. The simplicity of bacterial expression systems allows for rapid genetic manipulation and ease of scale-up. The culture medium is inexpensive, and the cells do not require elaborate facilities for growth and maintenance.

Bacteria cells also have fast growth rates compared to other hosts. Many bacteria have segments of DNA in their cytoplasm called plasmids. Plasmids are easy

to isolate and manipulate. The plasmid can be removed, a desired gene can be inserted into it, and the plasmid can be reintroduced into the bacterium. The transformed bacterium will then produce the foreign protein as if the protein were native to the bacterium.

The primary disadvantage of these prokaryotes is that they cannot perform certain post-translational modifications that may be required for producing authentic and bioactive eukaryotic proteins (Kost 1997). These modifications may include processing of signal sequences, three dimensional folding, disulfide bridge formation, as well as O- and N-linked glycosylation. Bacteria do not have any of the membrane bound organelles found in eukaryotes. In eukaryotes, a protein is often modified after it is initially produced in order to create a functional protein. Various organelles such as the endoplasmic reticulum or the Golgi apparatus are involved in the modifications. Because of the limitations of bacterial expression systems, higher order organisms such as mammalian, insect, and yeast cells have been developed as hosts.

### *Mammalian, Insect and Virus Expression Systems*

Mammalian expression systems can process the most complex of the recombinant proteins. However, their disadvantages include an often expensive growth medium, slow growth rates, low cells densities, and sensitivities to shear, contamination, and environmental conditions within the bioreactor (Dr. J. Piret, per. commun., 1998).

Insect cells are gaining popularity as hosts for recombinant protein expression. Insect cells carry out many of the post-translational modifications of proteins comparable to those that occur in mammalian cells. In addition, insect cells are easy to maintain in suspension and gene expression levels can be tightly regulated. The production of functional recombinant proteins by either mammalian or insect cells requires that the synthesized proteins be targeted to the correct cellular location for appropriate processing.

Recombinant virus systems represent powerful tools for the expression of recombinant proteins in cultured cells, animals, and man. Hundreds of proteins have been expressed by this versatile system that is commonly based on recombinant baculovirus (Smith et al. 1983). In development is the use of recombinant baculovirus

to be used for the introduction and expression of heterologous genes in certain types of mammalian cells (Kost 1997).

### 2.3.2 Yeast Expression Systems

Yeast cells combine the ease of genetic manipulation and rapid growth characteristics of a prokaryotic organism with the subcellular machinery for performing post-translational protein modifications of eukaryotic cells (Cregg et al. 1993). The basic methodology of recombinant protein production in yeast was established by the mid 1980's and led to the successful production of a hepatitis B vaccine based on recombinant HBsAg (hepatitis B surface antigen) (Sudbery 1996).

Many foreign proteins have been expressed in the yeast *Saccharomyces cerevisiae*. This species is popular because of the wealth of knowledge that has been accumulated about its genetics and physiology. However, despite its capabilities and success, there are several limitations to this expression system. The product yields are low, the presence of foreign gene products puts additional stress on the cells, the production of the protein during the growth phase hinders growth, and plasmids tend to be unstable (Buckholz and Gleeson 1991). All of these factors combine to make scale-up a difficult and unpredictable task. Also, hyperglycosylation of secreted glycoproteins is a common problem. Hyperglycosylation can cause differences in immunogenicity, diminished activity, and decreased serum retention of the foreign protein (Romanos et al. 1992). The difficulties encountered with *Saccharomyces cereviciae* have been part of the impetus to investigate other species of yeast for recombinant protein production.

#### *Methylotrophic Yeasts*

Methylotrophic yeasts are able to metabolize methanol as the sole carbon and energy source (Ogata et al. 1969). In the early 1970's, when methanol was inexpensive, interest in the production of single-cell protein from methanol led to the study of methylotrophic yeasts (Wegner 1990). The methylotrophic yeasts *Hansenula polymorpha* and *Pichia pastoris* have now become mainstream host systems for the production of recombinant proteins (Faber et al. 1995). For reasons that will be presented in the following sections, these yeasts are rapidly becoming the preferred

hosts for high-level protein production. However, the relative newness of methylotrophic systems may make the acceptance of recombinant products by regulatory authorities more difficult.

## 2.4 *Pichia pastoris* as a Host for Recombinant Proteins

More than 300 recombinant proteins have been produced by *Pichia pastoris* expression systems (Cereghino and Cregg 1999). These include enzymes, proteases, protease inhibitors, receptors, single-chain antibodies, and regulatory proteins (Higgins and Cregg 1998) (see Appendix I). Some of these proteins have been expressed to levels as high as grams per liter. Its past success and future potential have been the motivation for ongoing efforts to improve vectors, strains, and protocols for the *P. pastoris* expression system.

The work presented in this thesis is based on the *Pichia pastoris* expression system. The selection of *P. pastoris* for the production of human cystatin C was based on *P. pastoris'* history of high-level expression of recombinant proteins (Clare et al. 1991a; Tschopp et al. 1987a), its tightly-regulated alcohol oxidase 1 (*AOX1*) gene promoter (Tschopp et al. 1987b), its sub-cellular organelles such as the endoplasmic reticulum and Golgi apparatus that allow for post-translational modifications, and its ability to secrete proteins into the culture medium.

### 2.4.1 Metabolism in *Pichia pastoris*

*P. pastoris* can metabolize a wide variety of carbon sources. These include glucose, glycerol, fructose, sorbitol, ethanol, methanol, alanine, lysine, succinate, ethyl amine, cadaverine, glucitol, mannitol, L-rhamnose, and trehalose (Sreekrishna and Kropp 1996). Methanol metabolism is of particular importance for recombinant protein production in *P. pastoris*.

The initial reactions for methanol metabolism take place in specialized microbodies known as peroxisomes (Faber et al. 1995). The presence of peroxisomes is related to environmental conditions. When the cells metabolize non-methanol carbon sources, very few peroxisomes are present. However, when grown on methanol, peroxisomes may occupy 80% of the total cell volume (Veenhuis et al. 1983).

In the first step of methanol metabolism, methanol enters the peroxisomes and is oxidized by alcohol oxidase (AOX) to generate formaldehyde and hydrogen peroxide. Since alcohol oxidase has a low affinity for oxygen, *P. pastoris* compensates by producing excessive amounts of the enzyme. AOX can comprise up to 30% of the total soluble cell protein in *P. pastoris* cells (Couderc and Barratti 1980). Hydrogen peroxide is a toxic byproduct of the oxidase reaction that must be decomposed to water and molecular oxygen by a peroxisomal catalase. This in fact is the function of the peroxisomes, to protect the rest of the cell from the hydrogen peroxide.

The formaldehyde generated by the oxidase reaction enters either 1) a dissimilatory pathway to yield energy or 2) the assimilatory pathway to generate biomass. In the assimilatory pathway, glyceraldehyde-3-phosphate (GAP) is generated and is further assimilated in the cytosol. One-third of the GAP generated from formaldehyde is converted into biomass by standard yeast reactions. Formaldehyde used for energy generation enters the cytosol where it is catabolized by two subsequent dehydrogenase reactions. The final product of the dissimilatory pathway is carbon dioxide.

Alcohol oxidase and peroxisomes are generally not present in the absence of methanol. In fact, when the carbon source is changed from methanol to a repressive carbon source such as glucose, a rapid degradation of peroxisomal enzymes and peroxisomes is observed (Gellissen et al. 1995). This process is initiated by the formation of a membrane that encloses the organelle. Hydrolytic enzymes are then released that degrade alcohol oxidase.

The expression of the alcohol oxidase gene is tightly regulated at the transcription level (Cregg et al. 1985) and is under the control of a repression/derepression mechanism as well as an induction mechanism (Higgins and Cregg 1998). Growth on certain non-methanol carbon sources, such as glucose, represses the *AOX* gene promoter. In this case, no alcohol oxidase is synthesized even in the presence of methanol. Some carbon sources have a derepressing effect, meaning that some alcohol oxidase is synthesized, even in the absence of methanol. In the case of glycerol as the sole carbon source, alcohol oxidase amounting to 2 to 5% of that in methanol-grown cells has been observed (Brierly et al. 1990; Gellissen et al.

1995). However, some researchers insist that growth on glycerol alone (derepression) is not sufficient to generate even minute quantities of AOX1 (Ellis et al. 1985; Tschopp et al. 1987b). In the absence of a repressive carbon source, and upon addition of methanol, the gene promoter is heavily induced. However, it is not known if methanol itself, or another metabolic intermediate acts as the inducing agent (Gellissen et al. 1995).

There are actually two alcohol oxidase genes, *AOX1* and *AOX2*, which encode for alcohol oxidase 1 (AOX1) and alcohol oxidase 2 (AOX2) respectively (note that italicized acronyms refer to genes and non-italicized acronyms refer to the proteins that they encode for). AOX2 is 97% identical to, and has the same specific activity, as AOX1 (Cregg et al. 1989). However, the *AOX1* gene promoter is much more active than the *AOX2* gene promoter. As a result, AOX1 comprises the majority of the total alcohol oxidase in methanol-grown cells.

### 2.4.2 Regulation of Protein Expression in *Pichia pastoris*

Protein expression is regulated by gene promoters. A promoter is the segment of DNA located immediately in front of each gene (according to the direction of transcription). The promoter regulates when, how much and how often the gene is transcribed (Kost 1997). The promoter used for gene expression must originate from the host organism.

The alcohol oxidase 1 gene promoter is commonly exploited as the gene promoter for the recombinant protein expression in *P. pastoris*. Under the control of the *AOX1* promoter, foreign gene expression can be "turned off" (repressed) when the cells are grown on a non-methanol carbon source and "turned on" (induced) when the carbon source is changed to methanol.

Typically, *P. pastoris* cells are grown to high cell densities utilizing glucose or glycerol as the sole carbon source. There are two reasons for generating biomass using a repressing carbon source. Primarily, the doubling time on glycerol or glucose is less than 2 hours compared to 6 hours for growth on methanol (Sreekrishna and Kropp 1996). Also, non-methanol carbon sources repress the expression of the recombinant protein which prevents the selection of wild-type cells that do not contain the

recombinant gene (Cereghino and Cregg 1999). If methanol were to be used to generate high cell densities, wild-type cells may grow faster than cells that are burdened with the task of expressing the recombinant protein. Upon depletion of the repressing carbon source, methanol is fed to the culture to induce the *AOX1* promoter. This overall process is known as step-wise induction.

### 2.4.3 Recombinant Gene Incorporation into *Pichia pastoris*

Now that an overview of how proteins are expressed by *P. pastoris* has been provided, it is necessary to investigate how their genes are incorporated into the host cells. Incorporating the recombinant gene into *P. pastoris* requires two steps. The first step is to insert the recombinant gene into a vector that contains a *P. pastoris* gene promoter and a transcriptional termination sequence. The second step is to incorporate the vector into the host's genome by the process called transformation.

The vectors commonly used to generate recombinant *P. pastoris* strains are integrative plasmids (Gellissen et al. 1995). They are also shuttle-type vectors meaning that they are composed of sequences necessary to be maintained by either *E. coli* or *P. pastoris* (Higgins and Cregg 1998). Two of these vectors are pPICZ and pPICZα (Invitrogen 1998). The pPICZ vector is used for intracellular expression of a recombinant protein and the pPICZα vector is selected for extracellular expression. A schematic diagram of the pPICZα vector is shown in Figure 2.3.

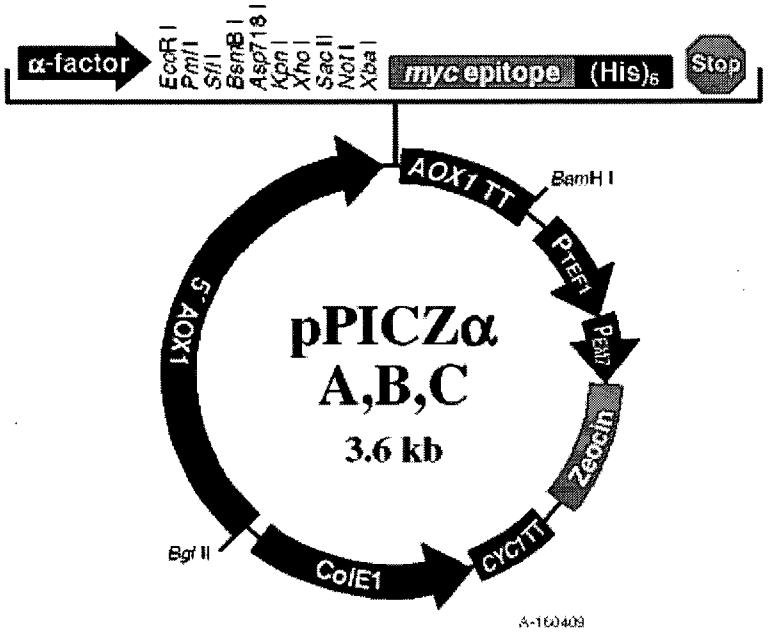


Figure 2.3. Graphic map of the pPICZα A, B, C vector series. The vector is composed of 3330 nucleotides and its incorporation into the host cell (*P. pastoris* or *E. coli*) is verified by conferred resistance to the antibiotic Zeocin™. Starting from the left, and moving clockwise, is the 5' *AOX1* gene promoter sequence that allows for methanol-driven induction. Next is the α-factor signal sequence that allows for secretion followed by 10 unique restriction sites for inserting the recombinant gene. The *myc* epitope tag and the $(HIS)_6$ polyhistidine tag allow for antibody detection and metal-binding affinity purification respectively. The *AOX1* TT is the transcription termination sequence for alcohol oxidase 1. For *P. pastoris*, expression of the Zeocin™ resistance gene is driven by the promoter $P_{TEF1}$ and is terminated by the *CYC1* transcription termination sequence. For *E. coli*, the recognized promoter is $P_{EM7}$. Finally, the ColE1 sequence allows for replication and maintenance of the vector in *E. coli*.
Source: www.invitrogen.com

The pPICZα vector contains the *S. cerevisiae* α-factor signal sequence that directs the protein to the appropriate secretion pathway. These vectors also contain a gene to express a protein that binds and inactivates the antibiotic Zeocin™. The conferred resistance to Zeocin™ allows only those cells that have incorporated the vector to grow in media that have a Zeocin™ concentration of at least 100 $\mu g \cdot ml^{-1}$. The pPICZα vectors also have more than 10 unique restriction sites to allow for insertion of the recombinant gene using a variety of restriction enzymes.

The recombinant gene is cloned into the vector by ligation techniques. Prior to incorporation into *P. pastoris*, multiple copies of the vector are created. *E. coli* is the common host for vector amplification. Once an ample quantity of the vector has been harvested from the *E. coli* cells, the vector is prepared for transformation into *P. pastoris*.

The vector can establish itself in two ways: integration into chromosomal DNA by homologous recombination or autonomous replication as a circular plasmid vector (Sreekrishna and Kropp 1996). To maximize the stability of foreign protein production, expression vectors should be integrated into the *P. pastoris* genome (Cregg et al. 1985; Cregg et al. 1989). Chromosomal integration requires homology of the introduced DNA with a chromosomal locus. Prior to incorporation, the vector is linearized within a gene sequence that is homologous to the *P. pastoris* locus. Homologous recombination usually results in the integration of a single copy of the recombinant gene. Multiple insertion events occur spontaneously at about 1 – 10% of the single insertion events and can be detected by quantitative dot blot analysis, Southern blot analysis, and differential hybridization (Invitrogen 1998). Since the yield of recombinant protein is often gene dose-dependent, protocols have been developed to obtain multi-copy integrants. They are obtained either by using vectors containing multiple gene copies or by screening for multiple integration events. Chromosomal integration followed by screening for multiple insertion events were the methods that were used in this work.

The vector can be inserted upstream or downstream of the *AOX1* gene. In these cases, the transformant retains its ability to express the alcohol oxidase 1 gene. The phenotype of such a transformant is $Mut^+$ (methanol utilization plus). If the vector is inserted within the locus for the *AOX1* gene, then the transformant can no longer express AOX1. Instead, the cell must rely upon the transcriptionally weak *AOX2* gene. These transformants, known as $Mut^s$ (methanol utilization slow) exhibit much slower growth on methanol as a sole carbon source compared to wild-type *P. pastoris* cells. The advantages and disadvantages of $Mut^+$ and $Mut^s$ phenotypes are described in more detail in Chapter 5.

## *Transformation Protocols*

Introduction of recombinant DNA into *P. pastoris* can be accomplished by the lithium chloride method (Ito et al. 1983), the spheroplast method (Cregg et al. 1985), and by electroporation (Becker and Guarente 1992).

The lithium chloride method provides the simplest means to transform *P. pastoris* cells. This method is based on the observation that alkali cations (lithium chloride or lithium acetate) make yeast competent to take up DNA by making their cell walls more permeable. After incubation of whole yeast cells with lithium chloride, the linearized vector is added along with a carrier DNA. The uptake of the vector by the cells is triggered by the addition of polyethylene glycol (PEG) and a heat shock treatment. Despite its simplicity, the lithium chloride method does not give rise to high frequencies of multi-copy transformants (Invitrogen 1998) and has a low efficiency of transformation (~$10^2$ transformants/μg DNA) (Higgins and Cregg 1998).

Transformation by the spheroplast method tends to produce transformants with a wide range of copy numbers (Sreekrishna et al., 1997) and has a higher efficiency of transformation (~$10^5$ transformants/μg DNA) (Higgins and Cregg 1998). However, this method is laborious and time-consuming. Cells are prepared for transformation by enzymatic digestion of the cell wall that is time-dependent and requires careful monitoring. Several isotonic washes are then required to remove the digestive enzyme. In a similar manner to the lithium chloride method, the linearized vector is added to the spheroplasts suspension and treated with PEG to promote vector uptake. After removal of the PEG and incubation in a regeneration medium, the spheroplasts are transferred to molten regeneration agar (45°C) and poured into Petri dishes. After incubation for several days, the transformant colonies will be embedded within the agar and must be removed with an inoculation loop.

The spheroplast method can randomly generate multi-copy transformants. Up to 10 copies of the linearized vector can randomly integrate into the spheroplast cells (Romanos, 1995). However, the frequency of occurrence is unpredictable and is often inversely related to the copy number. For example, when the goal of transformation is to generate $Mut^S$ transformants by transplacement of the *AOX1* gene, the actual outcome is variable. For example, only 5 – 30% of the transformants will be of the

desired $Mut^s$ phenotype. A high proportion of the transformants will contain no vector, and the remainder will have retained their $Mut^+$ phenotype. Of the $Mut^s$ transformants, 1 – 10% will have up to 30 integrated copies (often in tandem head to tail repeats). Traditionally, spheroplasting has been used to transform *P. pastoris*, but this method of transformation does not allow for direct selection on Zeocin™. Damage to the cell wall leads to increased sensitivity to Zeocin™, causing transformants to die before they express the Zeocin™ resistance gene (Invitrogen 1998).

Instead of using chemicals or enzymes to prepare the cells for transformation, electroporation uses high intensity electric pulses to create transient pores in the cell membrane (Prasanna and Panda 1996). The vector DNA enters the cells without the assistance of a carrier and the electropores reseal upon removal of the electric field. Electroporation is strongly recommended for isolating transformants with multiple copies of the vector (Invitrogen 1998). The transformation efficiency is comparable to spheroplasting (~$10^5$ transformants/μg DNA), but the frequency of multi-copy insertions is only from 1 - 10% (Higgins and Cregg 1998). Electroporation was selected for this project because of its convenience and potential to generate multi-copy transformants. The electroporation protocol is described in more detail in Chapter 4.

### 2.4.4 *Pichia pastoris* Transformant Selection

Depending on the host strain of *P. pastoris*, the vector selected, and the method and location of gene recombination, various types of transformants can be generated. The phenotype can be $Mut^+$ or $Mut^s$, the gene can have single or multiple insertions, and the protein can be expressed in the intracellular or extracellular environment. The selection of gene copy number and a comparison of methanol utilization phenotypes are presented in Chapter 4 and Chapter 5 respectively. The choice of intracellular or extracellular expression is considered in section 2.4.5 that follows.

### 2.4.5 Protein Processing by *Pichia pastoris*

#### *Extracellular versus Intracellular Expression*

As discussed previously, vectors are available for both intracellular and extracellular expression of recombinant proteins. Because *P. pastoris* secretes only low

levels of native proteins and because its commonly used culture medium contains no added proteins, a secreted recombinant protein comprises the majority of the total protein in the medium (Barr et al. 1992; Tschopp et al. 1987a). As a result, the purification process is greatly simplified. Purification is necessary because foreign proteins and impurities can cause allergic reactions or make the therapeutic effects of the recombinant protein different from the intended therapeutic effects.

Generally, secretion into the extracellular environment is the preferred mode of expression for naturally secreted proteins (Sreekrishna and Kropp 1996). These naturally secreted proteins usually require post-translational modifications that are associated with secretion pathways. These proteins also require a secretion signal to direct them to the appropriate pathways. It is possible to use the native secretion signal for the recombinant protein, but the signal may not be recognized by *P. pastoris* (Higgins and Cregg 1998). The α-factor secretion signal from *S. cerevisiae* has been used with the most success. This signal works very efficiently and is particularly effective in secreting smaller-sized proteins (<10 kDa). The α-factor signal is included in the pPICZα vector available from Invitrogen (San Diego, CA).

Intracellular expression may be more appropriate than extracellular expression. It is possible that the protein may be toxic to *P. pastoris* in the extracellular environment. In this case, targeting the protein to the intracellular peroxisomes may avoid toxicity problems. This type of approach has been used for insulin-like growth factor-II in the methylotrophic yeast *H. polymorpha* (Faber et al. 1996). Another scenario may be that the protein may form a gel in the extracellular environment. An increased viscosity may hinder mass transfer in the reactor that may cause cell death or even damage to the bioreactor. And in some cases, a mysterious inability to secrete a recombinant protein has been observed for *P. pastoris* (Romanos 1995). Reasons for this observation remain unknown.

In this work, secretion was selected for the expression of a glycosylated variant of cystatin C. The choice was based on the increased purity and ease of isolation of the product in the extracellular environment.

## *Glycosylation in Pichia pastoris*

Glycosylation is the addition of carbohydrate chains to a protein molecule. In yeast expression systems, glycosylation often occurs at the asparagine amino acid when the consensus sequence N-X-S/T is recognized in the secretion pathway (Nakamura et al. 1993). N, S, and T are the one-letter codes for asparagine, serine, and threonine respectively, and X can be any amino acid. This type of glycosylation is known as N-linked glycosylation . This sequence is recognized in the endoplasmic reticulum where the core structure of the oligosaccharide is attached to the asparagine residue (Gellissen et al. 1995).

Cystatin C found in humans is not glycosylated. However, glycosylated cystatin C has been observed in the seminal vesicles of rats (Esnard et al. 1988). Also, within the superfamily of the cystatins, the high molecular weight kininogens are highly glycosylated and have hydrophobic leader sequences (Abrahamson et al. 1987). Cystatin C also has a hydrophobic leader sequence. The kininogens are also proteinase inhibitors that are found in extracellular bodily fluids.

Recently, two N-linked glycosylated cystatins have been discovered in the human genome. These are cystatin F (Ni et al. 1998) and cystatin E/M (Ni et al. 1997; Sotiropoulou et al. 1997). Cystatins F and E/M are only 29% homologous to each other, but they each share 30 and 35% homology to human cystatin C (Ni et al. 1998).

Each of these glycosylated cystatins (F, E/M, and mouse cystatin C) exhibited N-linked glycosylation in positions that are opposite to the active site of the protein molecule. These positions would approximately correspond to glycosylation at location B or C as shown in Figure 2.1. In this location of the cystatin molecule, the inhibitory activity is not impaired. In addition, site-specific N-linked glycosylation of mouse cystatin C conferred additional stability to proteolytic degradation and thermal inactivation compared to the nonglycosylated protein (Nakamura et al. 1998a).

For these reasons, this work used site-directed mutagenesis to introduce a consensus sequence for N-linked glycosylation at two sites in the human cystatin C gene (labeled B and C in Figure 2.1). The appropriate amino acid substitutions were made in order to create the N-X-S/T consensus sequences. Dr. Masahiro Ogawa of

Food, Nutrition, and Health at the University of British Columbia performed this work. The methods of site-directed mutagenesis will be discussed in Chapter 4.

The structure of the carbohydrate chains of glycoproteins is organism specific. Many proteins secreted from *S. cerevisiae* have been demonstrated to be antigenic when introduced into mammals (stimulate the production of antibodies). These therapeutic proteins are quickly cleared from the body. As a result, the use of glycoprotein products synthesized by *S. cerevisiae* for therapeutic purposes has been avoided. The antigenic nature of these glycoproteins is a result of hyperglycosylation (50 to 150 mannose residues) and the presences of $\alpha$-1,3 linkages (Cregg et al. 1993; Romanos et al. 1992). The length and structure of these carbohydrate chains is significantly different than those of mammalian cells.

It has been shown that the post-translational modifications made by *Pichia pastoris* are more suitable for proteins intended for therapeutic use in humans. The length of the glycosylation saccharides are often much shorter in proteins expressed in *P. pastoris* (<30 mannose residues) compared to *S. cerevisiae* (Grinna and Tschopp 1989). The small and uniformly sized oligosaccharides (oligo = a few or little) synthesized by *P. pastoris* are similar in size and structure to oligosaccharides synthesized by mammalian cells (Kornfeld and Kornfeld 1985). Recently, six glycoproteins were expressed in *P. pastoris* and their oligossacharides were investigated (Montesino et al. 1998). No $\alpha$-1,3 linked mannosylation was observed in any of the secreted proteins. In fact, the enzyme that makes $\alpha$-1,3 linkages is $\alpha$-1,3 mannosyl transferase and it is undetectable in *P. pastoris*. The oligosaccharides of glycoproteins expressed by *P. pastoris* contain $\alpha$-1,2 linked mannose residues that have a close resemblance to oligosaccharides observed in glycoproteins of higher eukaryotes (Duman et al. 1998; Tschopp et al. 1987a). For the production of vaccines, totally authentic glycosylation may be unnecessary, provided the proteins elicit a protective immune response (Romanos 1995).

More clinical trials are required to study the potential antigenic nature of glycoproteins expressed by *P. pastoris* and how the pharmacological properties of these products might be affected (Romanos 1995). This project will contribute to these studies. By site-directed mutagenesis, Dr. Masahiro Ogawa mutated the gene encoding

for human cystatin C by introducing two consensus sequences for N-linked glycosylation. As previously observed for cystatin F, cystatin E/M, and mouse cystatin C, the location of the consensus sequences are opposite to the active site of the molecule.

## 2.5 Conclusion

In the 21st century, the intensive application of gene expression systems will be necessary to make effective use of the vast quantity of information currently being generated by genomics efforts. In turn, the knowledge gained from genomics programs will provide an increased understanding of the genetics of host organisms. Creative use of this knowledge will lead to improvements in current expression systems and the development of innovative expression approaches.

*P. pastoris* has already been widely accepted as an important biotechnological host organism. Many *P. pastoris*-expressed recombinant proteins are moving through to clinical trials and beyond. IGF-1 and HSA are marketed products for the treatment of amylotrophic lateral sclerosis and as a serum replacement respectively (Romanos 1995). Numerous cytokines, vaccines, and other biological products are under development and a Cuban group has developed a hepatitis B vaccine that is currently being sold in South America (Romanos 1995). The experiments presented in this thesis will contribute to the surge of information regarding the *P. pastoris* expression system.

# Chapter 3

## Growth Medium Optimization for Batch Cultures of *Pichia pastoris*

Previous experiments in Food, Nutrition, and Health at the University of British Columbia had compared *S. cerevisiae* and *P. pastoris* for the expression of glycosylated mouse cystatin C (Nakamura et al. 1998a). A buffered minimal medium (BMM) designed for *S. cerevisiae* was selected as the medium for both yeasts. This recipe is listed in Appendix II. Despite selecting *P. pastoris* as the preferred host, for reasons of convenience and familiarity, *P. pastoris* cultures continued to be grown in the BMM medium. It soon became evident that the cell densities observed in the BMM medium were significantly lower than those achievable with media recipes designed specifically for *P. pastoris* (data not shown) (Invitrogen 1998). Therefore, the first objective of this project was to identify and rectify the limiting factors in the BMM medium. It was hypothesized that a nitrogen limitation was the most significant factor.

This chapter presents the factorial experiments that were performed to investigate the effects of various carbon and nitrogen sources on the cell density in shake-flask cultures of *P. pastoris*. The results verify the hypothesis of a nitrogen limitation and identify histidine concentration as an important *P. pastoris* strain-dependent factor.

This chapter is separated into two sections. Section 3.1 looks at the effects of carbon sources and nitrogen sources. Section 3.2 considers in more detail the effects of nitrogen sources and their concentrations for *P. pastoris* cultures.

# 3.1 Factorial Experiments to Investigate Carbon and Nitrogen Sources

## 3.1.1 Introduction

*P. pastoris* cultures have the potential to be grown to cell densities as high as 130 $g \cdot l^{-1}$ based on dry cell weight (dcw) (Wegner 1983). This is a great asset because the production of recombinant proteins by *P. pastoris* is often dependent upon the cell density in the bioreactor (Romanos et al. 1992). Therefore, optimizing the growth medium to achieve high cell densities is an important step in expressing recombinant human cystatin C.

Many growth media recipes exist for *P. pastoris.* The Invitrogen Corporation (San Diego, CA) sells a *P. pastoris* expression kit that describes several commonly used recipes (Invitrogen 1998). For fed-batch cultures for *P. pastoris*, two similar base media formulations, Basal Salts and FM22, have given the best results (Higgins and Cregg 1998). The carbon and nitrogen sources are glycerol and ammonium hydroxide for the Basal Salts medium and glycerol and ammonium sulfate for FM22. After autoclaving, both of these media are supplemented with a heat-sensitive trace element solution. The PMT1 solution is added to the Basal Salts medium and PMT4 is added to the FM22 medium. The recipes for these media are included in Appendix II.

When grown in a bioreactor, these defined media allow for cell densities that are equivalent to complex media that contain yeast extract. However, for small-scale batch cultures of *P. pastoris* (shake-flasks), the defined media recipes do not achieve cell yields like those of cultures grown in complex media. This is part of the reason why the concentrations of recombinant proteins are typically lower in shake-flasks compared to bioreactor cultures (Higgins and Cregg 1998). Therefore, a medium formulation that achieves high cell densities in shake-flask cultures is important for preliminary expression experiments.

The experiments described in this chapter investigate some of the differences in cell yields due to different medium compositions. The factors considered here are the growth substrate and nitrogen source. The substrate supplies the carbon and hydrogen

atoms that form the 'backbone' of all organic molecules. The nitrogen source is essential for protein formation.

An effective substrate is transported rapidly into the host cell and metabolized efficiently. Transport across the plasma membrane and across the mitochondrial membrane is a significant rate-controlling process in yeast metabolism (Barford 1990). In addition, inefficient metabolic pathways can result in the formation of unwanted byproducts or excessive oxygen consumption. Therefore, an investigation into the type of substrate is an important aspect of an optimization effort.

*P. pastoris* can metabolize a wide variety of carbon sources. These include glucose, glycerol, fructose, sorbitol, ethanol, methanol, alanine, lysine, succinate, ethyl amine, cadaverine, glucitol, mannitol, L-rhamnose, and trehalose (Sreekrishna and Kropp 1996). Autoclaved sucrose can be used as a sole carbon source due to the resulting breakdown of sucrose into glucose and fructose (Romanos et al. 1992). The advantage of sorbitol (10 $g \cdot l^{-1}$) or alanine (100 mM) is that they do not repress the *AOX1* gene promoter (Thorpe et al. 1999). However, sorbitol is growth-limiting at concentrations required for large-scale batch growth (50 $g \cdot l^{-1}$) and the cell yield (g cells$\cdot g^{-1}$ sorbitol) is less that that of glycerol. *P. pastoris* cannot metabolize galactose, arabinose, ribose, maltose, lactose, raffinose, melibiose, cellulose, or starch.

The two most common carbon sources for *P. pastoris* are glucose and glycerol. These carbon sources are similar in that they are available at low cost and their stoichiometric oxygen requirements for oxidation are nearly identical. However, compared to glycerol, glucose metabolism generates more ethanol as a toxic byproduct of metabolism (Gellissen et al. 1995). Therefore, the use of glycerol is likely to reduce ethanol formation. In addition, glycerol-grown cells synthesize more alcohol oxidase than glucose-grown cells (Gellissen et al. 1995). Even without methanol, low concentrations of alcohol oxidase (2% of the induced status) are present during growth under glycerol-limited conditions. However, some debate exists in that several researchers insist that methanol is necessary for even detectable levels of alcohol oxidase (Ellis et al. 1985; Koutz et al. 1989; Tschopp et al. 1987b). In general, researchers prefer to use glycerol rather than glucose for recombinant protein expression in yeast systems (Sreekrishna and Kropp 1996).

Methanol can also be used as a sole carbon source, but growth on methanol is slower than on glycerol and glucose (Gellissen et al. 1995). Methanol-grown cells are also smaller and have more rigid cell walls compared to cells grown on glycerol or glucose. Therefore, it is recommended to grow the cells initially on an alcohol oxidase-repressing carbon source, such as glucose or glycerol, and then switch to methanol feeding to induce protein expression. This culture strategy also minimizes the selection of nonexpressing mutant strains that may grow faster on methanol (Cregg et al. 1993).

There is not a strong consensus as to the quantity and quality of a nitrogen source for *P. pastoris* cultures. In fact, some recipes, such as the Basal Salts medium, do not include a nitrogen source and rely on its supply from ammonium hydroxide ($NH_4OH$) that is used for pH control (Sreekrishna and Kropp 1996). However, many recipes include ammonium sulfate ($(NH_4)_2SO_4$) and or yeast nitrogen base (YNB). It has been reported that a supplemental nitrogen source such as ammonium sulfate or ammonium phosphate can increase cell density as well as inhibit protease enzymes (Tsujikawa et al. 1996). Concentrations for ammonium sulfate range from 4.3 $g \cdot l^{-1}$ (Guebel et al. 1991) to 10 $g \cdot l^{-1}$ (Invitrogen Expression Kit). The FM22 medium includes 5 $g \cdot l^{-1}$ $(NH_4)_2SO_4$ and uses KOH to adjust the pH. If YNB is used as the nitrogen source, the concentration is typically 3.4 $g \cdot l^{-1}$.

Yeast extract is a complex nitrogen source that is a common ingredient for media that feed many types of microorganisms. Complex media typically contain 10 $g \cdot l^{-1}$ of yeast extract (Invitrogen 1998). The addition of yeast extract in the culture medium typically increases the cell density relative to a defined medium (Barr et al. 1992). The increased cell density is due to the presence of amino acids and other complex components (vitamins, carbohydrates, etc.) that can be directly assimilated by the cells. These components may be absent from a defined medium. However these components may complicate the purification of secreted proteins. Several stages of protein purification may be required to isolate the protein from the complex supernatant. Therefore, the addition of yeast extract to culture media is most appropriate for intracellular expression of recombinant proteins in which the protein accumulates within the cells. Prior to cell lysis and purification of the protein, the complex culture medium is

separated from the cells. The cells are then resuspended in an appropriate buffer before cell lysis.

A typical compromise for *P. pastoris* cultures exhibiting extracellular expression of a recombinant protein is to use a defined medium in the bioreactor (Basal Salts or FM22), and an inoculum that has been grown in a complex medium (Lesnicki 1999). This procedure reduces the inoculum growth time and may also serve to supply some trace nutrients that may be lacking in the defined medium.

An investigation into the effect of yeast extract in batch cultures can serve an important optimization function. A complex medium can serve as a benchmark from which to compare various defined media. Ideally, a defined medium should allow *P. pastoris* cells to grow to cell densities that are comparable to densities achieved in complex media. A large discrepancy would indicate the need to select a more effective defined medium or to modify the existing medium. It may be necessary to modify the carbon to nitrogen ratio or add amino acids in order to achieve higher cell yields in a defined medium.

In some cases, it is necessary to optimize or modify the medium in accordance to the recombinant protein that is being expressed. Growth medium optimization can be achieved by various methods. A "carbon source controlled shift technique" has been used to maximize the growth rate and growth-linked product formation of the methylotrophic yeast *Candia boidinii* (Beste et al. 1997). This method used computer-controlled feeding from 15 different stock solutions of individual medium components. The optimization of 7 variables in the medium required 2.5 months of experimentation, but the resulting medium formulation increased the specific growth rate by 19% compared to the commonly used growth medium. Another possibility is the "chemostat pulse and shift technique". In this case, a continuous culture receives a concentration pulse for each medium component and the change in cell density is observed. For both of these methods, the response time is slow, and there may be lingering effects from previous shifts or pulses. These methods are also not appropriate for fed-batch cultures that are typically used for *P. pastoris.*

Factorial experiments were selected for the growth medium optimization efforts that are presented in this chapter. Compared to the shift and pulse techniques, factorial

experiments can be performed on a small-scale, the process is rapid, and the interaction effects of different components can be investigated.

## 3.1.2 Materials and Methods

### *Microorganism*

The host microorganism was wild-type *Pichia pastoris* known as X-33 ($Mut^+$, $His^+$) (Invitrogen, San Diego, CA). The wild-type *P. pastoris* strain synthesizes the AOX1 enzyme for fast growth on methanol and also has the histidinol dehydrogenase gene (*HIS4*) for synthesizing histidine.

The cells were transformed according to the procedure described in Chapter 4. The pPICZα vector, containing the gene for a cystatin C variant, was inserted by transplacement of the *AOX1* gene resulting in a $Mut^s$ phenotype. It was later discovered that the transformant selected for these experiments was auxotrophic for histidine ($His^-$). The inability to synthesize histidine was likely due to a disruption in the histidinol dehydrogenase gene (*HIS4*) by transplacement during electroporation (Invitrogen 1998). Since the growth of this auxotrophic transformant was limited by the amount of histidine supplied in the defined medium, cell yields were less than expected. The implications of this limitation will be discussed in more detail section 3.1.4.

### *Culture Media*

Unless otherwise indicated, chemicals were from Difco, Detroit. The selected colony was plated and maintained on YPD agar plates (10 $g \cdot l^{-1}$ yeast extract, 20 $g \cdot l^{-1}$ peptone, 10 $g \cdot l^{-1}$ D-glucose (Fisher, New Jersey), 20 $g \cdot l^{-1}$ agar) and stored at 4°C.

Unless otherwise indicated, chemicals were from Sigma, St. Louis. The growth medium was based on a buffered minimal medium (BMM) for the yeast *Saccharomyces cerevisiae*. This recipe had been used in Food, Nutrition and Health at the University of British Columbia for the expression of recombinant mouse cystatin by *P. pastoris* and *S. cerevisiae* (Nakamura et al. 1998a). A 1-liter volume of the medium included: a carbon source (either 20 g of glucose or 20 g of glycerol (both from Fisher, New Jersey)), minerals (0.1 g $CaCl_2 \cdot 2H_2O$, 0.5 g $MgSO_4 \cdot H_2O$, 1.5 g $KH_2PO_4$), vitamins (0.2 mg thiamine, 0.2 mg pridoxine, 0.2 mg nicotinic acid, 10 mg inositol, 0.2 mg

pantothenic acid, 0.002 mg biotin), trace elements (0.1 mg KI, 0.06 mg $H_3BO_3$, 0.06 mg $MnSO_4 \cdot 5H_2O$, 0.3 mg $ZnSO_4 \cdot 7H_2O$, 0.04 mg $CuSO_4 \cdot 5H_2O$, 0.25 mg $FeCl_3 \cdot 6H_2O$, 0.025 mg Mo, 0.025 mg NaCl), histidine (20 mg), and biotin (0.4 mg biotin). Ammonium sulfate (0 or 4.3 $g \cdot l^{-1}$) (Fisher, New Jersey) and yeast extract (0 or 10 $g \cdot l^{-1}$) (Difco, Detroit) were added to the base medium in accordance with the experimental design.

### *Apparatus*

The experimental units were 250-ml baffled flasks that contained 50 ml of complete culture medium. All of the flasks were contained within an environmental shaker that was maintained at 30°C and shaken at 250rpm.

### *Cell Cultivation*

A flamed loop was used to select a colony from the agar plate and inoculate a 30-ml volume of YPD liquid medium in a 250-ml baffled flask. The inoculation was performed in triplicate in order to have a total of 90 ml of inoculum. The inoculum was grown overnight at 30°C and 250 pm. Meanwhile, the experimental units were prepared by aseptically transferring the appropriate amounts of the various stock media into autoclaved flasks. After 24 hours of incubation, the inoculum flasks were mixed together and 5-ml aliquots were added to each of the flasks. A 1-ml sample was taken to measure the initial optical density, and then the flasks were placed in the 30°C incubator and shaken at 250 rpm.

### *Optical Density Measurements*

The optical density (OD) of the culture was used to compare the cell densities in each of the flasks. Optical density values are commonly used to describe cell densities in shake-flask cultures of *P. pastoris* (Invitrogen 1998). Therefore, a calibration curve to relate optical density to dry cell weight was not determined.

The OD was measured with a Spectronic 2D spectrophotometer (Milton Roy Company) set to a wavelength of 600 nm. The samples were diluted with distilled water to ensure that the OD reading was within the linear range of the spectrophotometer.

## *Experimental Design*

Factorial experiments were selected to facilitate the investigation of several factors and to indicate trends and interactions. For these preliminary screening experiments, only two levels for each factor were investigated. Including a midpoint for each factor would have allowed for an indication of curvature, but this was not deemed necessary at this stage.

A completely randomized design was selected for the experiments. Each of the flasks was placed randomly within the environmental shaker. Blocking of the flask position within the apparatus was not necessary because the environmental conditions and base medium were considered homogeneous for each experimental unit.

Each of the three factors was investigated at two levels, denoted -1 and +1 (low and high). Therefore, the total number of treatments was $2^3 = 8$. The levels for the three factors are shown in Table 3.1.

Table 3.1. Factors and their levels for preliminary factorial experiments.

| Factor | Symbol | Coded Level | |
|---|---|---|---|
| | | -1 | +1 |
| Substrate | G | 20 $g \cdot l^{-1}$ glucose | 20 $g \cdot l^{-1}$ glycerol |
| Nitrogen $[(NH_4)_2SO_4]$ | N | 0 $g \cdot l^{-1}$ | 4.3 $g \cdot l^{-1}$ |
| Yeast Extract | Y | 0 $g \cdot l^{-1}$ | 10 $g \cdot l^{-1}$ |

Since either glycerol or glucose must be present in the culture medium, they were combined as one factor. Glucose was used as the low level for substrate and glycerol was used as the high level. This modification to a typical factorial design allowed for more factors to be studied with relatively few experiments.

It should also be noted that the typical concentration of glycerol was not selected. Usually, a concentration of 0.11 Molar is used for either glucose or glycerol (Invitrogen 1998). This is equivalent to 20 $g \cdot l^{-1}$ of glucose and 10 $g \cdot l^{-1}$ of glycerol. In order to provide the same amount of carbon in the medium (8 grams of carbon per liter), the concentration of glycerol was increased to 20 $g \cdot l^{-1}$ (as shown in Table 3.1). Since both

glycerol and glucose are composed of 40% (w/w) carbon, it is more reasonable to use equal masses instead of equimolar quantities.

For interpretation of the results, the level of significance, α, was selected to be 0.1. This is less strict than the typical value of 0.05. The p-values are presented to allow the reader some subjectivity when interpreting the significance of the results.

### *Factorial Arrangement*

In order to facilitate sampling, only two replications were used for each flask. Therefore, a total of $2 \times 2^3 = 16$ flasks were used. These flasks were randomly numbered from 1 to 16. The complete factorial arrangement is shown in Table 3.2.

Table 3.2. Factorial arrangement of treatments in the standard order.

| Flask Number | Substrate G | Nitrogen N | Yeast Extract Y | Treatment Code |
|---|---|---|---|---|
| 1 and 9 | -1 | -1 | -1 | 1 |
| 2 and 10 | 1 | -1 | -1 | G |
| 3 and 11 | -1 | 1 | -1 | N |
| 4 and 12 | 1 | 1 | -1 | GN |
| 5 and 13 | -1 | -1 | 1 | Y |
| 6 and 14 | 1 | -1 | 1 | GY |
| 7 and 15 | -1 | 1 | 1 | NY |
| 8 and 16 | 1 | 1 | 1 | GNY |

## 3.1.3 Results

### *Growth Curves and Experimental Error*

The statistical analyses for the factorial experiments are based on a one-time response for each of the treatments. In order to determine the best time to make a comparison, optical density measurements were taken for the entire culture period.

The average growth curve for each of the treatments is shown in Figure 3.1. Since the maximum densities were reached soon after 24 hours, these values were used to perform the analysis of variance.

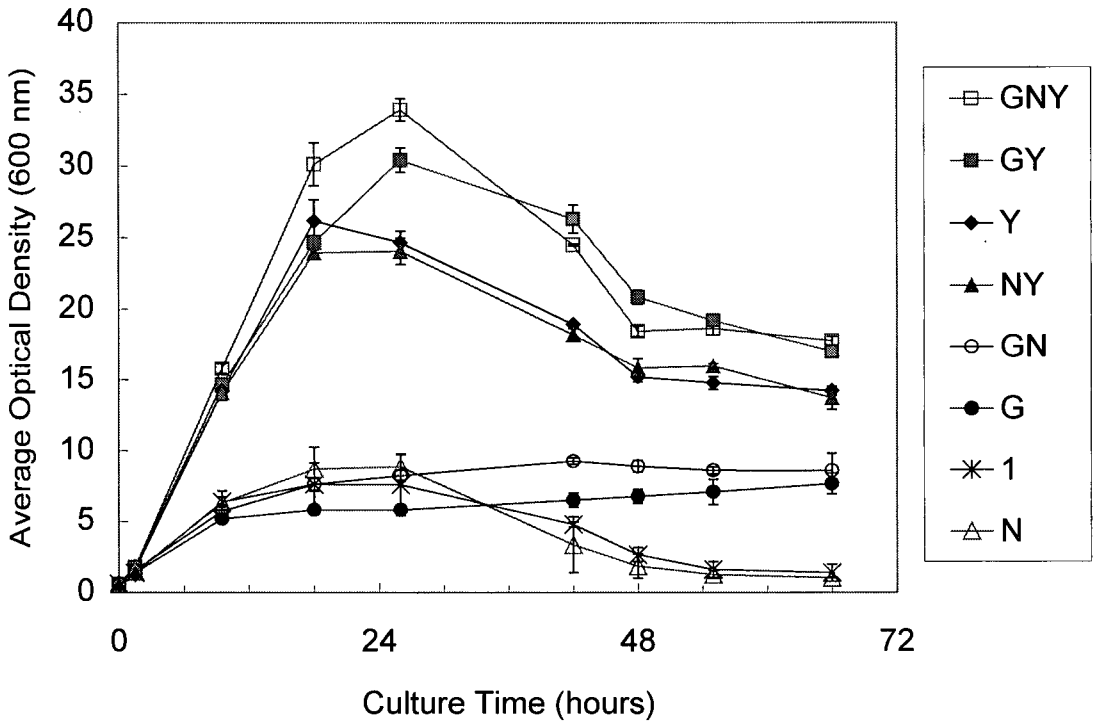


Figure 3.1. Growth curves for the eight treatment conditions of the two-level factorial experiment. The treatment codes are described in Table 3.2. The optical density was measured at 600 nm. The error bars are standard deviations based on two replicates.

The distribution of the experimental errors is shown in Figure 3.2. The error of an observation is equal to the difference between the observed OD and the mean OD for that particular treatment. For comparison purposes, a normal distribution having a mean (0.0) and standard deviation (0.34) equal to that of the experimental errors is included.

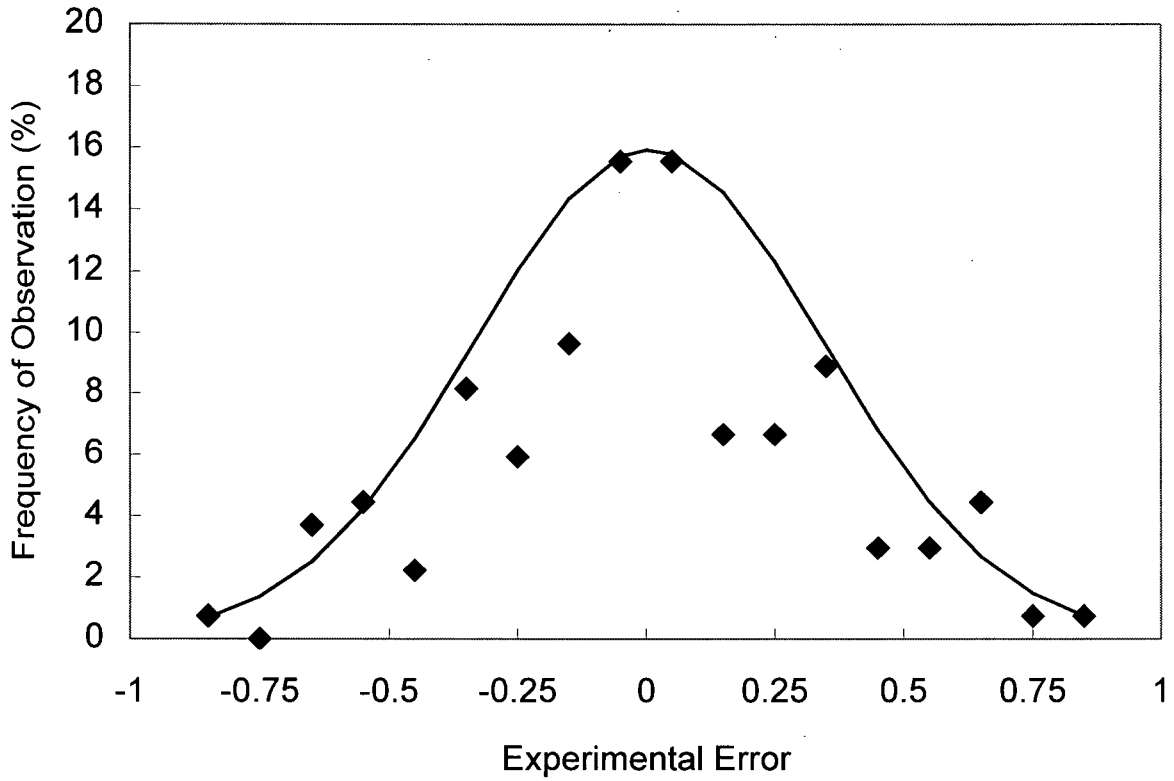


Figure 3.2. Distribution of the experimental errors of observations compared to a normal distribution having the same mean (0.0) and standard deviation (0.34) as the experimental errors. Each data point represents the percentage of observations within an experimental error interval having a magnitude of 0.1.

## *Optical Density Measurements at 26 Hours*

For most of the experimental units, the maximum culture OD was recorded at approximately 26 hours. Therefore, these values were selected for the analysis of variance. The OD readings at 26 hours are shown in Table 3.3.

Table 3.3. OD measurements at 600 nm and 26 hours after inoculation.

| **Optical Density Measurements** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **1** | **G** | **N** | **GN** | **Y** | **GY** | **NY** | **GNY** |
| 6.1 | 5.7 | 8.2 | 8.0 | 25.2 | 29.8 | 24.7 | 33.7 |
| 9.1 | 6.0 | 9.5 | 8.5 | 24.1 | 31.0 | 23.4 | 34.1 |

## *Box Plot*

A box plot is effective for summarizing data in three dimensions. The eight vertices of the box represent the eight treatment combinations. The response values that are shown are the average OD values for the specific treatments.

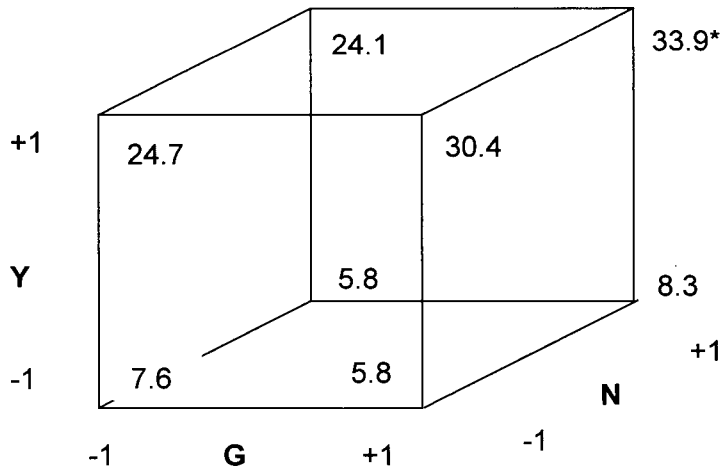


Figure 3.3. A box plot to indicate the OD response for all of the combinations of the three factors. The corner marked with 33.9* is the average OD for the GNY treatment (in this case, each of the factors was at the '+1' level).

## *Analysis of Variance*

The analysis of variance was performed on the data that are presented in Table 3.3. The analysis of variance and the effects are shown in Table 3.4. The effects are equal to the average response at the high level minus the average response at the low level. Higher order effects were calculated using Yate's Algorithm (procedure not shown) and confirmed using the SYSTAT 8.0 software package. As shown in Table 3.4, all of the main effects were significant as well as the GN and GY interactions.

Table 3.4. Analysis of variance and effects for the factorial experiments.

| Source | Effect | SS | df | MS | F-value | p-value |
|---|---|---|---|---|---|---|
| G* | 3.3 | 44 | 1 | 44 | 44 | 0.00 |
| N* | 1.7 | 11 | 1 | 11 | 11 | 0.01 |
| Y* | 20.6 | 1704 | 1 | 1704 | 1704 | 0.00 |
| GN* | 1.3 | 7 | 1 | 7 | 7 | 0.03 |
| GY* | 4.5 | 81 | 1 | 81 | 81 | 0.00 |
| NY | -0.2 | 0.2 | 1 | 0.2 | 0.2 | 0.71 |
| GNY | 0.7 | 2 | 1 | 2 | 2 | 0.17 |
| Error | | 8 | 8 | 1 | | |
| Total | | 1857 | 15 | | | |

SS = sum of squares, df = degrees of freedom, MS = mean squares
F-value = $MS_{TREATMENT}/MS_{ERROR}$
p-value is the probability that the effect was due to chance alone
* indicates that the effect was significant with at least 90% confidence

An effect was significant if its F-value is greater than the critical F-value that is taken from statistical tables. The critical F-value is based upon the degrees of freedom of the treatment, the degrees of freedom of the error, and the desired level of confidence. The confidence level, $\alpha$, is often selected as 0.05 or 0.10. For this set of experiments, the critical F-value is 5.32 for $\alpha = 0.05$. Increasing the number of replications or reducing the level of confidence will reduce the critical F-value. In addition to the comparison of F-values, p-values were calculated. The p-values are presented in the last column of Table 3.4. The p-value states the probability that the observed effect was due to chance alone (i.e. that the factor has no effect compared to the control and the observation was due to experimental error). P-values have an advantage over F-value comparisons in that they allow the experimenter and his/her colleagues, some flexibility when it comes to interpreting the significance of the results. Like the confidence level, $\alpha$, a p-value of 0.05 indicates a 5% chance that the observed effect is due to chance alone. Smaller p-values correspond to smaller percentages.

## *Response Surface*

A least squares regression of the data was performed using Microsoft Excel and verified with SYSTAT 8.0. SYSTAT 8.0 also calculated t-scores and two-tailed probability tests for each of the coefficients. A two-tailed probability test was used because it was not known if the levels of the various factors would have an increasing

or decreasing effect on the optical density of the culture. In agreement with the analysis of variance, the NY and GNY coefficients were not significant at the 90% confidence level. The results of this analysis are shown in Table 3.5.

Table 3.5. Statistical analysis of the three-factor response surface.

| Effect | Coefficient | Standard Error | t-score | p-value (2-tail) |
|---|---|---|---|---|
| Constant | 17.94* | 0.25 | 72.85 | 0.00 |
| G | 1.66* | 0.25 | 6.73 | 0.00 |
| N | 0.82* | 0.25 | 3.32 | 0.01 |
| Y | 10.31* | 0.25 | 41.84 | 0.00 |
| GN | 0.66* | 0.25 | 2.66 | 0.03 |
| GY | 2.24* | 0.25 | 0.11 | 0.00 |
| NY | -0.9 | 0.25 | -0.38 | 0.71 |
| GNY | 0.369 | 0.25 | 1.50 | 0.17 |

p-value is the probability that the effect was due to chance alone
* indicates that the coefficient was significant with at least 90% confidence

The resulting equation is a function of the five significant treatment combinations and can only be visualized in five dimensions. The equation estimates the optical density ($Y_i$) after 26 hours of culture time (based on a 10% (v/v) inoculum having an OD of 11).

$$Y_i = 17.9 + 1.66(G) + 0.82(N) + 10.31(Y) + 0.66(GN) + 2.24(GY)$$

$$R^2 = 0.995;\ SEE = 1.0$$

The $R^2$ value indicates that the model explains 99.5% of the variation that was observed in OD measurements. The standard error of estimate (SEE) is the standard deviation of the observations around the model surface.

The model is linear and assumes that the values of G, N, and Y are between –1 and +1. The model gives no indication of curvature because the midpoints for the treatments were not measured.

The linear response surfaces for both glucose and glycerol are shown in Figure 3.4. The scales have been kept the same in order to provide a convenient visual comparison.

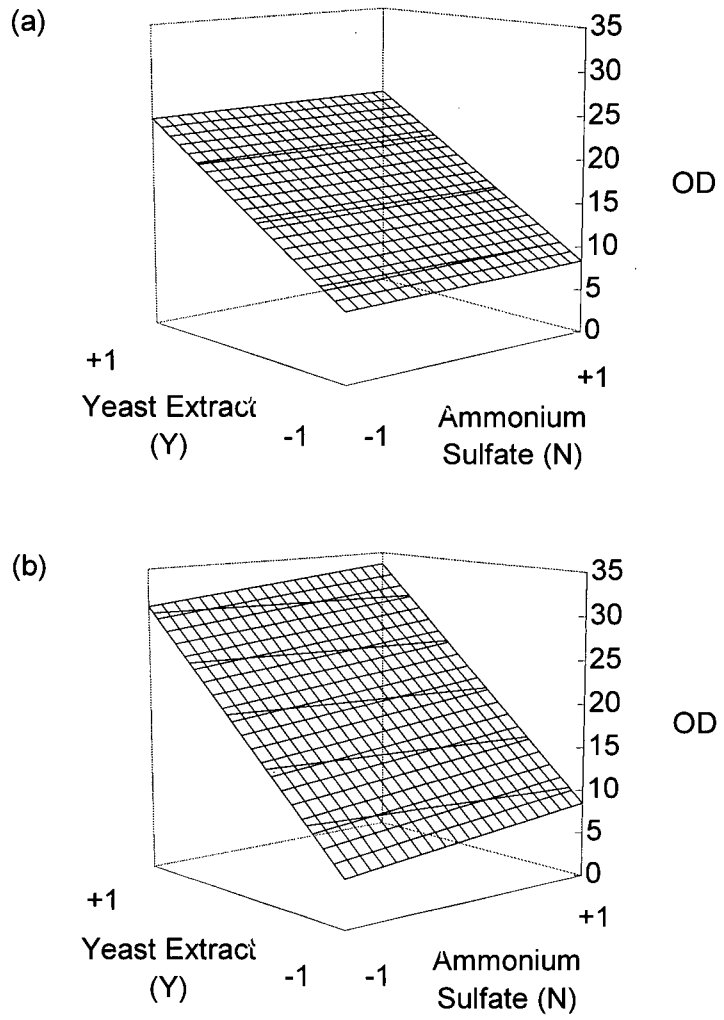


Figure 3.4. Response surfaces for **(a)** glucose and **(b)** glycerol. When glucose was used as the carbon source, the effect of yeast extract (Y) was significant but the effect of ammonium sulfate (N) was not. Both Y and N were significant when glucose was replaced by glycerol.

## 3.1.4 Discussion

### *Growth Curves*

The growth curves shown in Figure 3.1 are typical of *P. pastoris* cultures. Rapid growth was observed for nearly 24 hours and was followed by a stationary phase. A death phase was observed for all of the flasks except for G and GN (glycerol and glycerol plus ammonium sulfate).

The stationary phase was relatively brief for flasks that contained yeast extract. These flasks had a significantly higher cell density and would have consumed the substrate faster than flasks that had a lower cell density.

Flasks that included glycerol had longer stationary phases, especially in the absence of yeast extract. Since the mass concentration of glycerol and glucose were equivalent, glycerol metabolism may be more efficient.

There was a marked difference between cell densities in flasks that had yeast extract. The difference seemed to be much greater than would have been expected knowing that yeast grow nearly as well in an appropriate defined media as they do in complex media (Wegner 1983). It was later discovered during the bioreactor experiments that the cells used for these experiments were auxotrophic for histidine ($His^-$). The medium used for the bioreactor did not contain histidine (see Chapter 4 or Appendix II for the recipe of the Basal Salts medium and PTM1 Trace Salts solution). In the bioreactor, the cells did not grow until L-histidine was added to the bioreactor (data not shown).

The inability to synthesize histidine was likely due to transplacement of the histidinol dehydrogenase gene (*HIS4*) during electroporation. For histidine deficient strains, 40 $mg \cdot l^{-1}$ (Invitrogen 1998) to 1000 $mg \cdot l^{-1}$ of L-histidine has been recommended for high cell density cultures (Lesnicki 1999). The BMM medium contains only 20 $mg \cdot l^{-1}$ of L-histidine. Since yeast extract contains 1.8% (w/w) histidine, the media that had 10 $g \cdot l^{-1}$ of yeast extract would have had approximately 180 $mg \cdot l^{-1}$ of histidine (an adequate supply for $His^-$ transformants). Note also that histidine can tolerate the high temperatures of autoclaving (121°C) (Sherman et al. 1986). Therefore, the low cell yields in the flasks that did not have yeast extract can be attributed to a histidine deficiency of the medium.

A simple experiment was performed at a later date to investigate the effect of L-histidine concentration on cell density. The results are presented in Figure 3.5.

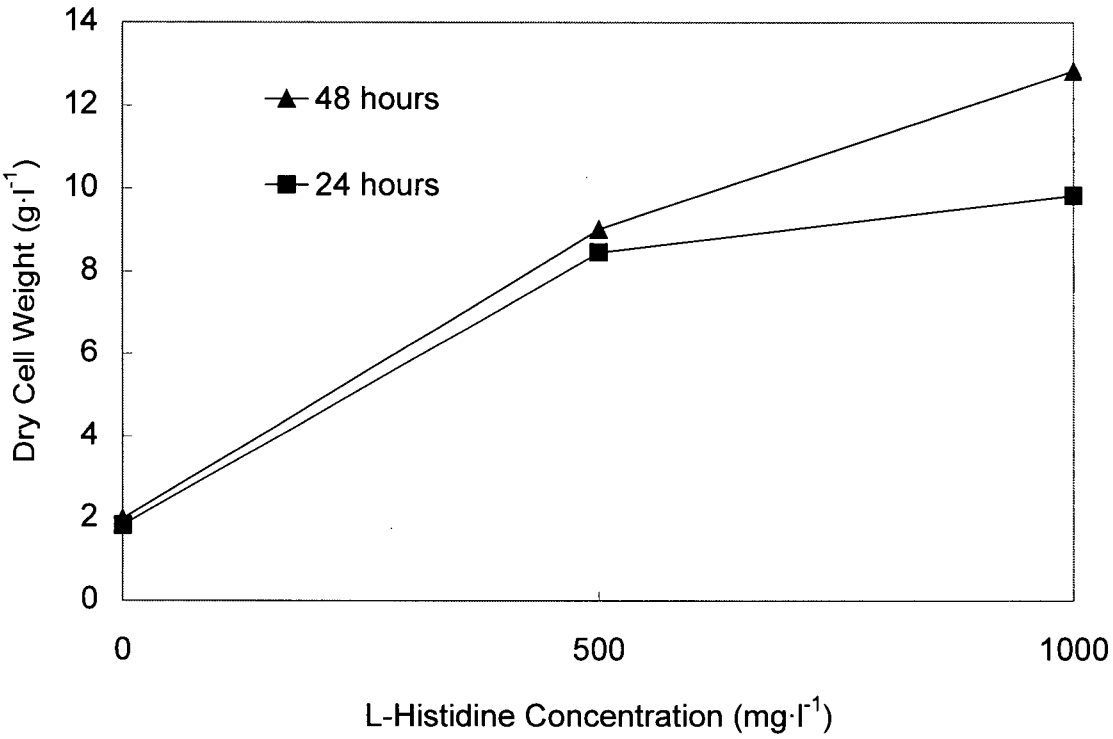


Figure 3.5. Dry cell weight as a function of L-histidine concentration in shake-flask cultures of a His$^-$ *P. pastoris* transformant. Each of the three shake-flasks contained 25 ml of Basal Salts medium + 0.11 ml of PTM1 Trace Salts solution (see recipe in Appendix II). This medium was used in the bioreactor and contains no histidine. An appropriate amount of L-histidine was added to each flask. Wet cell weights were measured and converted to dry cell weights using the procedure described in Chapter 4. The cell densities were measured at 24 and 48 hours after an equal volume of inoculation. The flasks were not replicated.

It is evident that the L-histidine concentration had a significant effect on the cell density that was achieved in the shake-flasks. After 24 hours, the flask that had no histidine had negligible growth. The flasks that had L-histidine concentrations of 500 and 1000 mg·l$^{-1}$ had cell densities of 8.5 and 9.0 g dcw·l$^{-1}$ respectively. During the first 24 hours, it appears that L-histidine concentrations greater than 500 mg·l$^{-1}$ do not have a significant effect on the cell growth. However, after 48 hours, a noticeably higher cell density was observed for the flask that had 1000 mg·l$^{-1}$ of L-histidine compared to the flasks that had 0 and 500 mg·l$^{-1}$.

The increased cell density after 24 hours was likely achieved by catabolism of excess histidine. Experiments in the bioreactor indicated that the cell yield from glycerol

was 0.2 g dcw·$g^{-1}$ glycerol. Since the Basal Salts medium contained a glycerol concentration of 40 g·$l^{-1}$, the expected cell concentration was 8.0 g dcw·$l^{-1}$. As seen in Figure 3.5, the cell concentration after 24 hours was approximately 8 - 9 g dcw·$l^{-1}$ for both the flasks that had L-histidine concentrations of 500 mg·$l^{-1}$ and 1000 mg·$l^{-1}$. Therefore, the cell densities observed indicated that a carbon source other than glycerol was being used for growth.

Based on these experiments, an L-histidine concentration of 500 mg·$l^{-1}$ could be recommended for shake-flask cultures of histidine auxotrophic strains of *P. pastoris*. Higher cell density cultures, such as those achieved in the bioreactor, will required more L-histidine.

### *Statistical Analysis*

Two assumptions for the analysis of variance are 1) the standard deviations are not dependent upon the treatment conditions and 2) the experimental errors are distributed normally. As shown in Figure 3.1, the standard deviations of the observations were similar for each treatment. In addition, Figure 3.2 indicates that the frequencies of experimental errors had a distribution that was approximately normal. This observation was verified with a 95% confidence Chi test having a degree of freedom of n – 3 = 15 (where n is the number of intervals used to evaluate the normal approximation) (data not shown). Therefore, the assumptions for the analysis of variance are justified.

As indicated in Table 3.4, the main effects for carbon source, nitrogen, and yeast extract were positive and statistically significant. However, since the interactions for GN and GY were also significant, it is evident that the effects of the individual factors are not independent. Prior to interpreting the main effects, the interactions have to be identified.

According to Figure 3.4, the effects of ammonium sulfate and yeast extract were dependent upon the level of carbon source. This observation explains the significance of the GN and GY interactions. In both cases, the effect was greater for glycerol (G = +1) than for glucose (G = -1). It is likely that glycerol metabolism is more efficient

than glucose metabolism. Also, glucose metabolism may have resulted in the production of toxic levels of ethanol.

The NY interaction was not significant. Therefore, the effect of ammonium sulfate was not dependent upon the level of yeast extract. It seems that ammonium sulfate and yeast extract act independently to increase the cell density of the culture.

### *Response Surfaces*

The calculated response surface can provide an estimate of the optical density within the range of the factors that were studied. However, the response surface assumes that there was a linear trend between the high and the low levels of each factor. Therefore, if the model were to be used for prediction, more information would be required about the curvature of the response surface.

The linear response surfaces shown in Figure 3.4 indicate that the positive effects of ammonium sulfate and yeast extract were more significant for glycerol than for glucose. As indicated by these surfaces, the highest optical density was achieved using glycerol, coupled with the highest levels of ammonium sulfate and yeast extract (GNY).

## 3.1.5 Conclusions and Recommendations

The factorial experiments were able to indicate the major trends and interactions for each of the factors. For the growth of *P. pastoris* in shake-flask cultures, the results indicate that 1) glycerol was superior to glucose, 2) ammonium sulfate was effective with glycerol but not with glucose, and 3) yeast extract increased cell growth for both glycerol and glucose. These conclusions only apply to the actual levels for each factor that was investigated. Extrapolation and interpolation may yield different results.

In addition, it was discovered that the BMM medium is histidine-limited for $His^-$ *P. pastoris* transformants. For such strains, 500 $mg \cdot l^{-1}$ of L-histidine should be added to shake-flask cultures.

## 3.2 Factorial Experiments to Optimize the Nitrogen Source

### 3.2.1 Introduction

Preliminary experiments with glucose, glycerol, ammonium sulfate, and yeast extract revealed that the nitrogen source is an important factor for attaining high cell densities in cultures of *P. pastoris*. Therefore, the relationship between cell density and the type of nitrogen source as well as its concentration was investigated in more detail.

Three simple nitrogen sources that can be assimilated by *P. pastoris* are ammonium sulfate, ammonium phosphate, and ammonium hydroxide. Two organic sources are yeast extract and yeast nitrogen base. Each of these nitrogen sources serves a particular function in *P. pastoris* cultures. For example, ammonium phosphate is most commonly used as a buffer (20 $g \cdot l^{-1}$ of $(NH_4)_2HPO_4$ plus 20 $g \cdot l^{-1}$ of $NH_4H_2PO_4$) (Tsujikawa et al. 1996). Ammonium phosphate was not selected for investigation because its concentration should remain fixed in order to perform its buffering task. Ammonium hydroxide serves the dual purpose of increase the pH of *P. pastoris* cultures and concurrently providing a source of nitrogen (Invitrogen 1998). Complications can arise if the cells require more nitrogen, but the pH must be held constant. Therefore, it is best to dedicate the function of ammonium hydroxide to pH adjustment and supplement the medium with an independent source of nitrogen. Ammonium sulfate is commonly used to maintain an adequate carbon to nitrogen ratio (Duff and Murray 1988). Ammonium sulfate was selected as a simple nitrogen source for further investigation.

Yeast extract is a complex organic nitrogen source that is a common ingredient for media that feed many types of microorganisms. As mentioned previously, its addition to culture media is most appropriate for intracellular expression of recombinant proteins. In the case of extracellular expression, the cystatin would have to be separated from the complex mixture of proteins and amino acids that are present in yeast extract. For that reason, yeast extract was not selected for further investigation. Yeast extract had served its purpose in the previous experiments (section 3.1) by identifying a deficiency in the BMM medium.

Yeast nitrogen base is a highly recommended nitrogen source for *P. pastoris* media (Invitrogen 1998). Its formulation was originally developed to test carbon assimilation by yeasts in the presence of various vitamins (Difco 1953). With the addition of a carbon source, yeast nitrogen base provides a complete medium for yeast growth. Yeast nitrogen base contains vitamins, minerals, and salts and is available with and without amino acids and with and without ammonium sulfate. Yeast nitrogen base without amino acids and without ammonium sulfate is considered to be a satisfactory nitrogen source for general growth media (Sigma, St. Louis). A list of its ingredients is provided in Appendix II. Yeast nitrogen base is more expensive than the other nitrogen sources that have been mentioned, but it does not complicate the medium like yeast extract. Yeast nitrogen base is commonly added to shake-flask cultures of *P. pastoris* at a concentration of 3.4 $g \cdot l^{-1}$ (Döring et al. 1997; Henry et al. 1997; Shen et al. 1998; Waterham et al. 1997). Higher concentrations of yeast nitrogen base (6.7 $g \cdot l^{-1}$) have also been effective (Kim et al. 1997; Tschopp et al. 1987a). Yeast nitrogen base was selected as the second potential nitrogen source.

To investigate the effects of ammonium sulfate and yeast nitrogen base, factorial experiments at three levels and response surface methods were used. Only two factors were considered because three-level factorial experiments require more experimental units than two-level factorial experiments. In fact, n factors at two levels require $2^n$ experimental units whereas n factors at three levels require $3^n$ experimental units.

## 3.2.2 Materials and Methods

### *Microorganism*

The microorganism used was *Pichia pastoris* X-33 ($Mut^+$, $His^+$) (Invitrogen, San Diego, CA). Nontransformed cells displaying the $Mut^+$, $His^+$ phenotype were selected to avoid histidine-limitations that were experienced for the previous experiments that were described in section 3.1.

### *Culture Media*

The culture media were the same as those described in section 3.1.2. YPD agar plates were used to maintain the cells, and the BMM medium was the base medium for

the factorial experiments. Glycerol (20 $g \cdot l^{-1}$) (Fisher, New Jersey) was used as the carbon source and ammonium sulfate (Fisher, New Jersey) and yeast nitrogen base (without amino acids and without ammonium sulfate) (Sigma, St. Louis) were added to the BMM medium in accordance with the experimental design.

### *Apparatus, Cell Cultivation, and Optical Density Measurements*

Refer to section 3.1.2.

### *Experimental Design*

The effects of ammonium sulfate and yeast nitrogen base on cell densities were studied at three levels using factorial experiments and a completely randomized experimental design. Three levels were selected to enable second-order surfaces to be fitted to the response data. These levels were denoted -1, 0, and +1 (low, mid, and high). Therefore, the total number of treatments for two factors was $3^2 = 9$. The levels for the two factors are shown in Table 3.6.

Table 3.6. Factors and their levels for the factorial experiments.

| Factor | Symbol | Coded Level | | |
|---|---|---|---|---|
| | | -1 | 0 | 1 |
| Ammonium Sulfate Concentration | $NH_4$ | 10 $g \cdot l^{-1}$ | 15 $g \cdot l^{-1}$ | 20 $g \cdot l^{-1}$ |
| Yeast Nitrogen Base Concentration | YNB | 3.4 $g \cdot l^{-1}$ | 5.1 $g \cdot l^{-1}$ | 6.8 $g \cdot l^{-1}$ |

The levels for ammonium sulfate were selected based on the two-level factorial experiments that indicated a positive effect on cell density with the addition of 4.3 $g \cdot l^{-1}$ of ammonium sulfate. The concentration range for ammonium sulfate was chosen to see if this trend was maintained at higher concentrations. The low level was based on a commonly used concentration of 10 $g \cdot l^{-1}$ (Invitrogen 1998) and the high level of 20 $g \cdot l^{-1}$ was such that excess ammonium sulfate was available for the cells (Loewen et al. 1997).

The commonly used concentration of 3.4 $g \cdot l^{-1}$ for yeast nitrogen base was selected as the low level for this factor (Invitrogen 1998). The high level was two fold greater.

For interpretation of the results, the level of significance, α, was selected to be 0.1. This is less strict than the typical value of 0.05. However, since fewer repetitions were performed compared to experiments in section 3.1, the degrees of freedom for the experiments were reduced. As a result, both critical F-values and p-values were increased. Since these values served as the benchmark for statistical significance, an effect had to be greater in magnitude to be considered statistically significant. In this case, the p-values are especially useful because they allow the reader some subjectivity when interpreting the significance of the results.

### *Factorial Arrangement*

The preliminary factorial experiments described in section 3.1 indicated that the experimental error was not significant. Therefore, replicates were performed only for the mid-point conditions. A total of eleven flasks were required for the nine experimental conditions plus the two additional mid-point replications. The complete factorial arrangement is shown in Table 3.7.

Table 3.7. Factorial arrangement of treatments and optical density data for the first three-level factorial experiments.

| Flask Number | Ammonium Sulfate ($NH_4$) | Yeast Nitrogen Base (YNB) | Optical Density[a] (600 nm) |
|---|---|---|---|
| 1 | -1 | -1 | 9.8 |
| 2 | 0 | -1 | 9.2 |
| 3 | 1 | -1 | 8.2 |
| 4 | -1 | 0 | 12.6 |
| 5 | 0 | 0 | 10.8 |
| 6 | 1 | 0 | 12.2 |
| 7 | -1 | 1 | 13.5 |
| 8 | 0 | 1 | 13.5 |
| 9 | 1 | 1 | 12.2 |
| 10 | 0 | 0 | 10.7 |
| 11 | 0 | 0 | 11.4 |

[a]The optical density was measured 26 hours after inoculation

A completely randomized design was selected for the experiments. Each of the flasks was placed randomly within the environmental shaker. Blocking was not necessary because the environmental conditions in the shaker and the buffered minimal medium were considered homogeneous for each experimental unit.

## 3.2.3 Results

### *Optical Density Measurements at 26 Hours*

The optical density measurements were used to compare the cell growth in each of the flasks. The OD values recorded at 26 hours are shown in Table 3.7.

### *Analysis of Variance*

Before a curve was fitted to the data, the effects of the factors and their significances were determined by an analysis of variance. The results were verified by SYSTAT 8.0 and are shown in Table 3.8.

Table 3.8. Analysis of variance based on the first set of experiments investigating the effect of ammonium sulfate and yeast nitrogen base concentrations at three levels.

| Source | Effect | df | SS | MS | F-value | p-value |
|---|---|---|---|---|---|---|
| $NH_4$ | -1.45 | 2 | 1.95 | 0.97 | 6.88 | 0.13 |
| YNB | 3.85* | 2 | 25.54 | 12.77 | 90.28 | 0.01 |
| $NH_4$xYNB | 0.15 | 4 | 2.60 | 0.65 | 4.59 | 0.19 |
| Error | | 2 | 0.28 | 0.14 | | |
| Total | | 10 | 30.37 | | | |

SS = sum of squares, df = degrees of freedom, MS = mean squares
F-value = $MS_{TREATMENT}/MS_{ERROR}$
p-value is the probability that the effect was due to chance alone
* indicates that the effect was significant with at least 90% confidence

Within the concentration ranges of investigation, yeast nitrogen base had the most significant effect. As indicated by its p-value, there was only a 1.1% chance that the effect was due to chance alone. Its main effect was positive indicating that increasing the yeast nitrogen base concentration increased the optical density. Its magnitude of 3.85 is interpreted as the average increase in optical density for flasks that had the high concentration (6.8 $g \cdot l^{-1}$) of yeast nitrogen base compared to the low concentration (3.4 $g \cdot l^{-1}$).

Ammonium sulfate had a negative effect indicating that the cell density was lower for higher concentrations of ammonium sulfate. The p-value of 0.13 indicates that the effect was on the verge of statistical significance (p-value < 0.1). It could be presupposed that the ammonium sulfate effect would have been statistically significant had more replications been performed.

The interaction effect was positive indicating that the detrimental effect of ammonium sulfate was more apparent at the lower level of yeast nitrogen base. However, the interaction effect was not considered to be statistically significant.

### *Response Surface*

A least squares regression of the data was performed using Microsoft Excel and verified with SYSTAT 8.0. Since the factors were studied at three-levels, the regression equation has quadratic terms. Within the concentration ranges that were investigated, the equation estimates the optical density after 26 hours of culture time. The arguments in the equation are the concentrations of the factors in units of $g \cdot l^{-1}$.

$$Y_i = 1.83 - 0.29(NH_4) + 0.01(NH_4)^2 + 3.91(YNB) - 0.27(YNB)^2$$

$$R^2 = 0.90; \quad SEE = 0.76$$

The $R^2$ value indicates that the model explains 90% of the variation that was observed in the OD measurements. The standard error of estimate (SEE) is the standard deviation of the observations around the regression surface.

The response surface for the full two-factor second-order polynomial is shown in Figure 3.6. The significant effect of yeast nitrogen base can be appreciated.

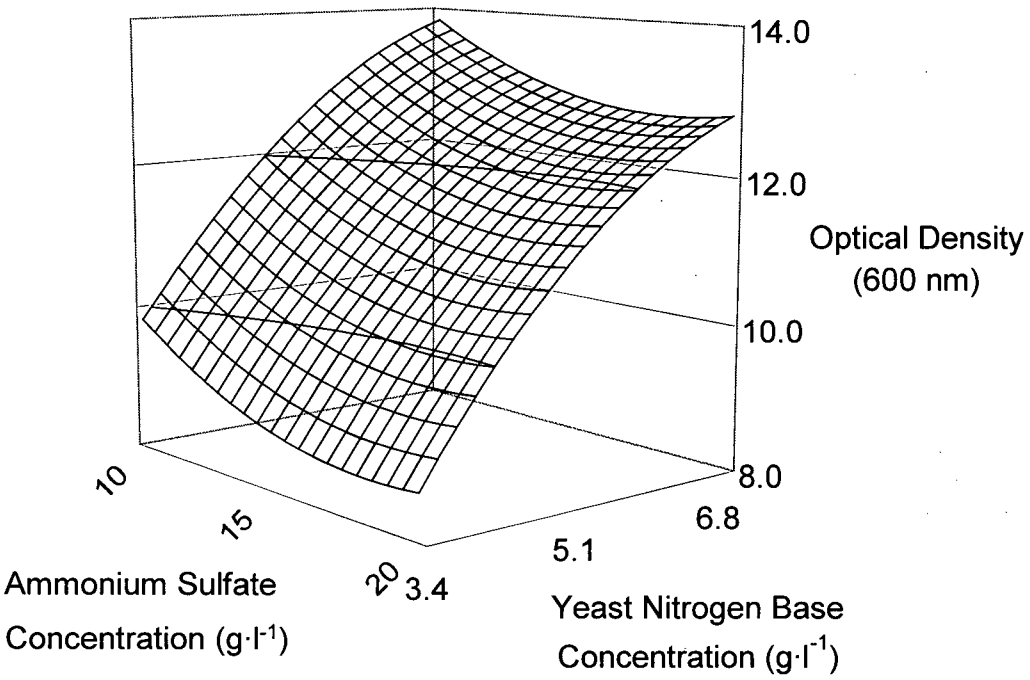


Figure 3.6. Second-order polynomial for the optical density data that were collected 26 hours after inoculation. The three levels that were used for the two factors are displayed on the figure. The two horizontal lines on the fitted surface are iso-optical density curves for optical densities of 10 and 12.

For each of the coefficients in the regression equation, the standard error, t-score, and p-value were calculated using SYSTAT 8.0. The statistical analysis of the coefficients reveals how the factors affected the optical density of the culture (i.e. linear versus quadratic). The coefficients and their statistics are presented in Table 3.9.

Table 3.9. Statistical analysis of the two-factor response surface.

| Effect | Coefficient | Standard Error | t-score | p-value (2-tail) |
|---|---|---|---|---|
| Constant | 1.83 | 4.20 | 0.44 | 0.68 |
| $NH_4$ | -0.29* | 0.14 | -2.01 | 0.10 |
| YNB | 3.91* | 1.82 | 2.14 | 0.09 |
| $(NH_4)^2$ | 0.01 | 0.01 | 0.70 | 0.52 |
| $(YNB)^2$ | -0.27 | 0.17 | -1.59 | 0.17 |
| $(NH_4)(YNB)$ | 0.00 | 0.04 | 0.06 | 0.96 |

the p-value is the probability that the effect was due to chance alone
* indicates that the coefficient was significant with at least 90% confidence

The linear affect of ammonium sulfate and yeast nitrogen base were both significant at the 90% confidence level. The quadratic affect of yeast nitrogen base was not as significant, but still worthy of consideration. By eliminating the terms that bear little statistical significance, the response surface equation is simplified:

$$Y_i = -0.29(NH_4) + 3.91(YNB) - 0.27(YNB)^2$$

The reduced equation predicts a ridge of maximum optical density for a yeast nitrogen base concentration of 7.2 $g \cdot l^{-1}$. The negative coefficient associated with the ammonium sulfate term indicates that its concentration should be reduced to achieve higher cell densities.

## *Response Surface Method*

A subsequent set of experiments was performed to encompass the suspected optimum concentrations for the factors (7.2 $g \cdot l^{-1}$ for YNB). The concentration range for ammonium sulfate was reduced to 5 – 15 $g \cdot l^{-1}$ and for yeast nitrogen base the range was increased to 6.8 $g \cdot l^{-1}$ – 8.8 $g \cdot l^{-1}$. This design allowed for some overlap between the two factorial experiments. The levels for the two factors are given in Table 3.10.

Table 3.10. Factors and their levels for the second factorial experiments.

| Factor | Symbol | Coded Level | | |
|---|---|---|---|---|
| | | -1 | 0 | 1 |
| Ammonium Sulfate Concentration | $NH_4$ | 5 $g \cdot l^{-1}$ | 10 $g \cdot l^{-1}$ | 15 $g \cdot l^{-1}$ |
| Yeast Nitrogen Base Concentration | YNB | 6.8 $g \cdot l^{-1}$ | 7.8 $g \cdot l^{-1}$ | 8.8 $g \cdot l^{-1}$ |

## *Factorial Arrangement*

A total of eleven flasks were required for the nine experimental conditions plus the two additional mid-point replications. The complete factorial arrangement is shown in Table 3.11. The levels for the coded factors are indicated in Table 3.10.

Table 3.11. Factorial arrangement of treatments and optical density data for the second three-level factorial experiments.

| Flask Number | Ammonium Sulfate ($NH_4$) | Yeast Nitrogen Base (YNB) | Optical Density[a] (600 nm) |
|---|---|---|---|
| 1 | -1 | -1 | 13.6 |
| 2 | 0 | -1 | 13.5 |
| 3 | 1 | -1 | 12.4 |
| 4 | -1 | 0 | 15.1 |
| 5 | 0 | 0 | 14.5 |
| 6 | 1 | 0 | 13.4 |
| 7 | -1 | 1 | 15.4 |
| 8 | 0 | 1 | 14.5 |
| 9 | 1 | 1 | 13.9 |
| 10 | 0 | 0 | 14.1 |
| 11 | 0 | 0 | 13.6 |

[a]The optical density was measured 26 hours after inoculation

A completely randomized design was selected for the experiments. Each of the flasks was placed randomly within the environmental shaker. Blocking was not necessary because the environmental conditions in the shaker and the buffered minimal medium were considered homogeneous for each experimental unit.

## *Analysis of Variance*

As with the first set of experiments, the effects and their significances were determined by an analysis of variance. The results are presented in Table 3.12.

Table 3.12. Analysis of variance based on the second set of experiments investigating the effect of ammonium sulfate and yeast nitrogen base concentrations at three levels.

| Source | Effect | df | SS | MS | F-value | p-value |
|---|---|---|---|---|---|---|
| $NH_4$ | -1.35 | 2 | 3.24 | 1.62 | 7.97 | 0.11 |
| YNB | 1.65 | 2 | 3.26 | 1.63 | 8.01 | 0.11 |
| $NH_4$xYNB | -0.15 | 4 | 0.26 | 0.07 | 0.33 | 0.85 |
| Error | | 2 | 0.41 | 0.20 | | |
| Total | | 10 | 7.17 | | | |

SS = sum of squares, df = degrees of freedom, MS = mean squares
F-value = $MS_{TREATMENT}/MS_{ERROR}$
p-value is the probability that the effect is due to experimental error
* indicates that the effect was significant with at least 90% confidence

Both ammonium sulfate and yeast nitrogen base maintained negative and positive effects respectively. The p-values indicate that both of their main effects were nearly significant at the 90% confidence level.

### *Response Surface*

A least squares regression on the collected data was performed using Microsoft Excel and SYSTAT 8.0. Within the concentration ranges that were investigated, the polynomial equation estimates the optical density after 26 hours of culture time. The arguments in the equation are the concentrations of the factors in units of $g \cdot l^{-1}$.

$$Y_i = -5.21 - 0.01(NH_4) - 0.001(NH_4)^2 + 4.44(YNB) - 0.23(YNB)^2 - 0.015(NH_4)(YNB)$$

$$R^2 = 0.90; \quad SEE = 0.37$$

The response surfaces for both the first and the second set of factorial experiments are shown in Figure 3.7. This figure provides a visualization of the response surface methodology.

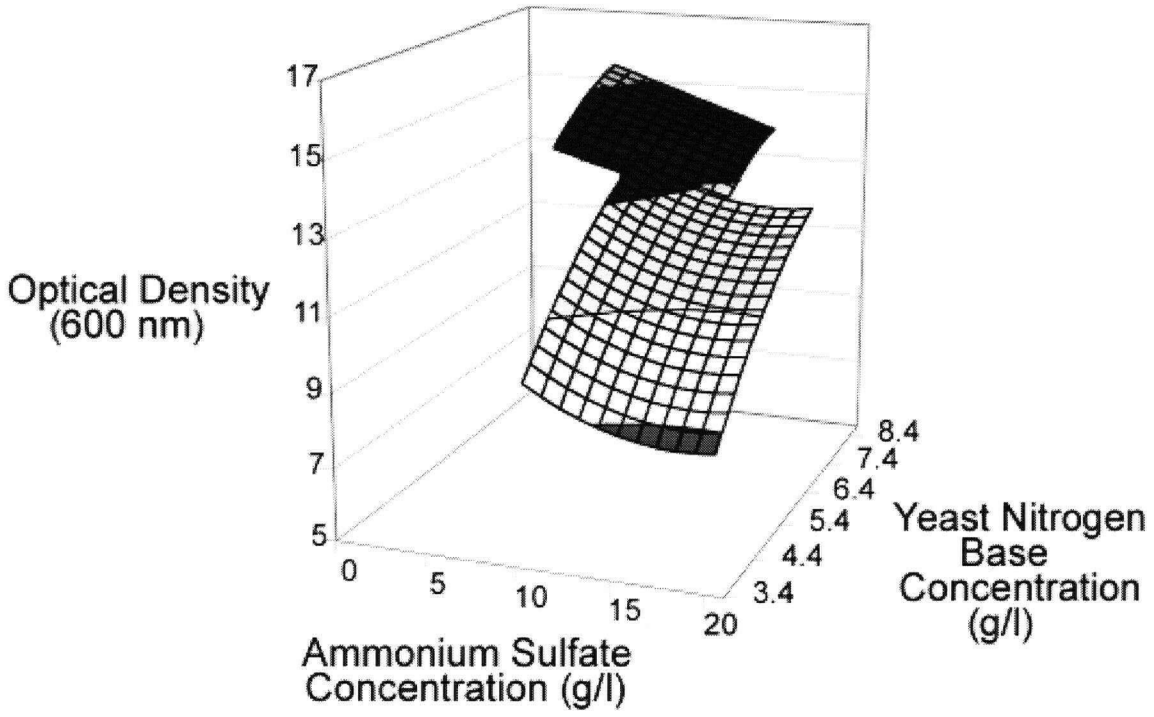


Figure 3.7. Response surfaces for optical density as a function of ammonium sulfate and yeast nitrogen base concentrations. The shaded regions are iso-optical density contours having a width of 2 OD units. The larger surface is a regression based on the first factorial experiments. The smaller surface is a regression based on the second factorial experiments. Neither of the factors was statistically significant in the second set of experiments.

According to the analysis of variance, the effects were not significant. This observation was investigated further by analyzing the regression coefficients. For each coefficient, the standard error, t-score, and p-value were calculated using SYSTAT 8.0. The coefficients and their statistics are presented in Table 3.13. None of the coefficients were statistically significant with 90% confidence.

Table 3.13. Statistical analysis of the two-factor response surface for the second set of experiments.

| Effect | Coefficient | Standard Error | t-score | p-value (2-tail) |
|---|---|---|---|---|
| Constant | -5.21 | 13.80 | -0.38 | 0.72 |
| $NH_4$ | 0.00 | 0.34 | -0.02 | 0.99 |
| YNB | 4.44 | 3.55 | 1.25 | 0.27 |
| $(NH_4)^2$ | 0.00 | 0.01 | -0.13 | 0.90 |
| $(YNB)^2$ | -0.23 | 0.23 | -1.01 | 0.36 |
| $(NH_4)(YNB)$ | -0.02 | 0.04 | -0.42 | 0.69 |

p-value is the probability that the effect was due to chance alone

Within the range of concentrations for ammonium sulfate and yeast nitrogen base, the coefficients of the regression polynomial are not statistically significant. However, it is evident in Figure 3.7 that decreasing the ammonium sulfate concentration from 15 to 5 $g \cdot l^{-1}$ and/or increasing the yeast nitrogen base concentration from 6.8 to 8.8 $g \cdot l^{-1}$ would likely increase the optical density. However, as a result of a significant experimental error, the statistics cannot support this inference.

### 3.2.4 Discussion

It is evident that a nitrogen source, as well as its concentration is important for attaining high cell density cultures of *P. pastoris*. Within the concentration ranges investigated, ammonium sulfate had a negative effect on cell density and yeast nitrogen base had a positive effect. The effects of these two factors were independent of each other.

The two- and three-level factorial experiments have approached the optimum ammonium sulfate concentration from both sides. The factorial experiments at two levels indicated that increasing the ammonium sulfate concentration from 0 to 4.3 $g \cdot l^{-1}$ has a positive effect on cell density. The experiments at three levels indicated that increasing the ammonium sulfate concentration above 10 $g \cdot l^{-1}$ had a negative effect on cell density. Invitrogen recommends an ammonium sulfate concentration of 10 $g \cdot l^{-1}$ (Invitrogen 1998). Based on these observations, and Invitrogen's recommendations, it seems prudent to maintain the ammonium sulfate concentration between 4.3 and 10 $g \cdot l^{-1}$. The three-level experiments that covered the range from 5 to 15 $g \cdot l^{-1}$ indicated

that decreasing the ammonium sulfate concentration to 5 $g \cdot l^{-1}$ may increase cell density, but the trend was not statistically significant.

Increasing the yeast nitrogen base concentration from 3.4 to 6.8 $g \cdot l^{-1}$ had a significant effect on increasing the cell density. However, the experiments have shown that increasing the yeast nitrogen base concentration above 6.8 $g \cdot l^{-1}$ did not result in a statistically significant increase in cell yield. When the yeast nitrogen base was doubled from 3.4 to 6.8 $g \cdot l^{-1}$, the average OD increased from 9 to 13. This represents a 44% increase in OD for a 100% increase in yeast nitrogen base compared to the recommended concentration (Invitrogen 1998). The retail price for yeast nitrogen base (without amino acids) is $\$518.20 \cdot kg^{-1}$ (Sigma, St. Louis, MO). Therefore, increasing the concentration from the recommended 3.4 $g \cdot l^{-1}$ to 6.8 $g \cdot l^{-1}$ incurs and increased medium expense of $\$1.76 \cdot l^{-1}$. Therefore, the value of one unit of OD would have to be greater than $\$0.44 \cdot l^{-1}$ to make this modification cost effective. It is likely that the increase in cell density would more than compensate for the increased cost of the medium.

### 3.2.5 Conclusions and Recommendations

Using factorial experiments and response surface methods, the trends and interactions for ammonium sulfate and yeast nitrogen base in the BMM medium were identified. For the growth of *P. pastoris* in shake-flask cultures, the results indicate that 1) increasing yeast nitrogen base concentrations up to 6.8 $g \cdot l^{-1}$ had an increasing effect on cell densities and 2) ammonium sulfate concentrations above 10 $g \cdot l^{-1}$ decreased cell densities.

It is likely that there is an optimal ammonium sulfate concentration between 4.3 and 10 $g \cdot l^{-1}$. A one-factor, three-level experiment with ammonium sulfate concentrations of 4.3, 7.1 and 10 $g \cdot l^{-1}$ would help to identify the optimal concentration. Since the interaction effects of ammonium sulfate and yeast nitrogen base were insignificant in the experiments that have been performed, the yeast nitrogen base concentration could be maintained at 6.8 $g \cdot l^{-1}$. Each of the three levels for ammonium sulfate should be replicated two or three times. This would allow for a greater amount of confidence when

interpreting the results compared to the experiments that have been presented here in which only the midpoint conditions were replicated.

However, in order to perform shake-flask experiments with results that are easily comparable with fellow users of the *P. pastoris* expression system, it would be advisable to replace the BMM medium with a medium that has been recommended by Invitrogen (Invitrogen 1998). An example is the MGY medium that is composed of 3.4 $g \cdot l^{-1}$ yeast nitrogen base, 10 $g \cdot l^{-1}$ ammonium sulfate, 10 $g \cdot l^{-1}$ glycerol, and 0.4 $mg \cdot l^{-1}$ biotin. Yeast nitrogen base supplies all of the vitamins, minerals, and salts that are present in the BMM medium. The BMM medium requires the preparation of separate stock solutions for vitamins, minerals, and salts. Therefore, compared to the BMM medium, the MGY medium is much more convenient to prepare and results based on experiments using this medium would be more appreciated by fellow researchers.

# Chapter 4

## Genetic Optimization of *Pichia pastoris* Transformants Expressing a Cystatin C Variant

This chapter describes the transformation of the cystatin gene into *P. pastoris* and the experiments that were performed to select an appropriate transformant for further experimentation. Section 4.1 describes the transformation procedure and the measurement technique that was used to quantify the recombinant cystatin. In addition, experiments are described that investigated the effect of gene copy number on cystatin expression. Section 4.2 investigates the stability of cystatin in the extracellular environment. In particular, the effect of storage time on cystatin activity is considered.

## 4.1 Electroporation and Selection of Multi-Copy Transformants According to Zeocin™ Resistance

### 4.1.1 Introduction

Human cystatin C is vulnerable to oxidation (Berti et al. 1997), proteolysis (Lenarcic et al. 1991), and dimerization (Abrahamson and Grubb 1994b; Ekiel and Abrahamson 1996), all of which reduce its stability. Site-specific N-linked glycosylation of mouse cystatin C was shown to increase protein stability and its resistance to heat and proteolysis while maintaining its biological activity (Nakamura et al. 1998a). For these reasons, Dr. Ogawa of Food, Nutrition, and Health at UBC, has used site-directed mutagenesis to introduce a consensus sequence for N-linked glycosylation (N-X-S/T) at two sites in the human cystatin C gene.

The methylotrophic yeast *P. pastoris* was selected as a host for the expression of this variant of human cystatin C. As previously discussed, *P. pastoris* is known for its high-level expression of heterologous proteins (Clare et al. 1991a; Tschopp et al. 1987a) and its tightly-regulated alcohol oxidase 1 (*AOX1*) gene promoter (Tschopp et al. 1987b) and its sub-cellular organelles such as the endoplasmic reticulum and Golgi

apparatus that allow for post-translational modifications such as folding, disulfide bridge formation and glycosylation (Cregg et al. 1993; Vedvick et al. 1991). In addition, the oligosaccharides of glycoproteins expressed by *P. pastoris* contain α-1,2-linked mannose residues and have a close resemblance to oligosaccharides observed in glycoproteins of higher eukaryotes (Tschopp et al. 1987a).

## *Selection of Multi-Copy Transformants*

In order to optimize heterologous protein expression in *Pichia pastoris* it is recommended to investigate the effect of gene copy number (Cregg et al. 1993; Invitrogen 1998; Romanos et al. 1992). Typically, a higher copy number results in a greater expression of recombinant protein. Since the *AOX1* gene promoter is under transcriptional control, more copies of the gene are transcribed for each pass of the multi-copy reading frame.

However, the experimental evidence has indicated that the optimal gene copy number is not necessarily the maximum possible gene copy number (Romanos 1995). In rare cases, a decrease in expression has been observed for higher gene copy numbers (Thill et al. 1990). It is hypothesized that the secretion pathways become blocked or ineffective at higher rates of expression in multi-copy transformants. Therefore, in the case of a secreted protein, a lower copy number may be prudent. This may be especially true if the protein requires extensive post-translational modifications. If the protein remains intracellular, then the secretion bottleneck is not a factor for protein production. In this case, selecting a transformant with the highest copy number would usually be the best choice, but should be verified empirically for each protein.

Experiments have been performed to examine the relationship between gene copy number and protein expression (Clare et al. 1991a). As shown in Figure 4.1, the level of expression was proportional to the number of gene copies.

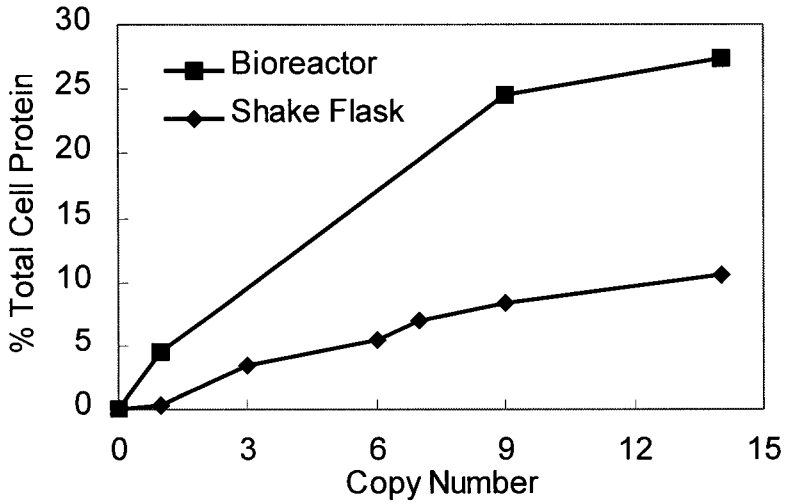


Figure 4.1. Percent of total soluble cell protein as a function of gene copy number in a Mut$^s$ strain of *P. pastoris* expressing intracellular tetanus toxin fragment C. Production of recombinant protein is typically higher in a bioreactor because of the increased aeration and mixing as well as better control of pH and methanol concentration. Source: (Clare et al. 1991a)

Since these results were published, the majority of the researchers have selected transformants that contain multiple copies of the expression cassette. As shown Table 4.1, all of the experiments with multiple copies were published after 1991.

Table 4.1. Production of notable recombinant proteins in *P. pastoris*.
Source: (Faber et al. 1995)

| Protein | Mut[a] | Copy #[b] | Protein Expression[c] | | Process[d] | Year |
|---|---|---|---|---|---|---|
| | | | % tcp | $g \cdot l^{-1}$ | $g \cdot l^{-1}$ | |
| Thrombomodulin | s | 1 | N/A | 0.3 | CC: 248 | 1996 |
| Thrombomodulin | s | 1 | N/A | 0.36 | FB: 115 | 1996 |
| Lysozyme | + | 1 | N/A | 0.35 | CC: 100 | 1989 |
| Lysozyme | + | 1 | 60 | 0.55 | FB: 120 | 1989 |
| D-ala Cpase-Bs | s | 10 | N/A | 0.8 | FB: N/A | 1993 |
| HIV gp120 | + | 12 | 2.5 | 1.25 | FB: N/A | 1993 |
| B-lactoglobulin | + | 1 | N/A | 1.5 | FB: 85 | 1997 |
| Invertase-Sc | s | 1 | 80 | 2.5 | FB: N/A | 1987 |
| Pertactin-Bp | + | 21 | N/A | 3 | FB:130 | 1991 |
| Human TNF | s | multi | 32 | 10 | FB:100 | 1993 |
| Tetanus toxin fr. C | + and s | 14 | 27 | 12 | FB:130 | 1991 |

[a]Methanol utilization phenotype (s = slow, + = plus)
[b]Integrated number of expression cassettes
[c]Protein expression expressed as percent of total cell protein (soluble) (% tcp) and grams per liter of culture medium. N/A indicates that the data are not available
[d]Continuous Culture (CC) or Fed-Batch (FB) and dry cell weight ($g \cdot l^{-1}$)

Before multi-copy transformants can be selected, they have to be generated. Depending upon the method of introducing recombinant DNA into *P. pastoris*, multiple insertion events will occur spontaneously at about 1 to 10% of the single insertion events (Invitrogen 1998). The methods of transformation for *P. pastoris* include the lithium chloride method (Ito et al. 1983), the spheroplast method (Cregg et al. 1985) and by electroporation (Becker and Guarente 1992). The experiments that are described here made use of electroporation as a means for introducing the mutated cystatin gene into *P. pastoris* cells. Electroporation allows for direct selection of transformants based on Zeocin™ resistance (Invitrogen 1998). The spheroplast method is not as convenient for transformant selection because damage to the cell wall leads to excessive sensitivity to Zeocin™. Therefore, when the goal is to isolate *P. pastoris* transformants with multiple copies of the recombinant gene, electroporation is strongly recommended.

It is also possible to generate multi-copy clones by introducing multiple copies of the expression cassette into a single vector prior to transformation (Cregg et al. 1993). The first step is to insert the foreign gene into the appropriate site of the vector. The

expression vector is then digested and the expression cassette is inserted into a vector that already has a single copy of the expression cassette. The reinsertion process is repeated until the desired number of copies has been inserted in a tandem arrangement within the expression vector.

Regardless of the method of transformation, it is often necessary to take specific action to isolate multi-copy transformants. Due to the low frequency of multi-copy transformants that are randomly generated by electroporation, screening thousands of transformants may be required to isolate a sufficient range of multi-copy transformants (Invitrogen 1998). The search can be simplified if transformation is performed using a vector that houses the gene for Zeocin™ resistance. Zeocin™ is an antibiotic that has been isolated from *Streptomyces* that shows strong toxicity against bacteria, fungi (including yeast), plants, and mammalian cells. Zeocin™ causes cell death by binding to and then cleaving the host's DNA (Berdy 1980). The Zeocin™ resistant protein, a product of the *Sh ble* gene (*Streptoalloteichus hindustanus* bleomycin gene) binds to Zeocin™ and renders it inactive. The protein product of the Zeocin™ resistance gene functions stoichiometrically, not enzymatically, and therefore provides a more efficient copy number correlation compared to enzymatic markers. In fact Zeocin™ resistance is approximately correlated to gene copy number (Brierly 1999). Therefore, transformants that exhibit Zeocin™ hyperresistance (resistance to Zeocin™ concentrations greater than 100 $\mu g \cdot ml^{-1}$) are likely to have multiple copies of the expression vector (Higgins and Cregg 1998). However, Zeocin™ can become inactive in the presence of salt concentrations above 5 $g \cdot l^{-1}$ (Invitrogen 1998). Therefore, low salt media are required for effective selection of hyperresistant transformants.

For Zeocin™ selection, Invitrogen (San Diego, CA) markets the pPICZ expression vectors. These vectors were designed for simple cloning, selection, high-level expression, and rapid purification. Incorporation of the vector into the *P. pastoris* genome allows for increased resistance to the Zeocin™ antibiotic.

Once the suspected multi-copy transformants have been isolated, it may be helpful to quantify the number of gene copies. This can be accomplished using quantitative DNA dot blots or by Southern hybridization (Clare et al. 1991a). These procedures require isolation of the genomic DNA, not only from the multi-copy

transformants, but also from the untransformed host strain, the host strain transformed with the parent vector (no copies of the recombinant gene) and the host strain transformed with one copy of the gene (Invitrogen 1998). The various DNA isolated from these additional cell types serve as a basis to quantify the recombinant gene multiplicity in the transformants under investigation.

Quantitative DNA dot blotting is the most accurate method for identifying the gene copy number. This method involves binding the chromosomal DNA to a radiolabeled probe and then comparing the extent of binding with a control sample that has been identified as a single-copy transformant. A greater binding relative to the single-copy control indicates the presence of multiple gene copies. The amount of binding is quantified by densitometry of the blot or by using a β-scanner if the chromosomal DNA is radiolabeled.

The objective of the experiments presented in this chapter was to determine the relationship between cystatin expression and Zeocin™ resistance. Primarily because of time constraints for this project, the experiments did not attempt to accurately quantify the cystatin gene copy number. Instead, the resistance to Zeocin™ was used as a basis for determining the relative multiplicity of gene copies in the various transformants.

## 4.1.2 Materials and Methods

### *Microorganisms*

*Pichia pastoris* X-33 incorporating the pPICZα-A vector was used as a host for extracelluar expression of a mutated variant of human cystatin C. *Escherichia coli* TOP10F' was used for plasmid amplification. Microorganisms and plasmids were purchased from Invitrogen (San Diego, CA).

The host cells for site-directed mutagenesis were *Epicurian coli* XL1-Blue supercompetent cells (Strategene, La Jolla, CA). T4 DNA Ligase, alkaline phosphatase and restriction enzymes were purchased from Promega (Madison, WI). All other chemicals were of analytical grade for biochemical use.

## *Synthesis of Human Cystatin C Gene*

A synthetic gene encoding human cystatin C was designed on the basis of its amino acid sequence (Grubb and Löfberg 1982) and *P. pastoris* preferred codons (Sreekrishna and Kropp 1996). Four oligonucleotides were synthesized, hybridized, and ligated. The synthetic gene included an *Xho* I restriction site at the 5'-end and an *Xba* I restriction site at the 3'-end.

The *Xho I/Xba* I segment, including the synthetic gene, was ligated with the pPICZα-A vector (Invitrogen, San Diego, CA). The gene was then sequenced on both strands with an α-factor sequencing primer (5'-TACTATTGCCAGCATTGCTGC-3') and a 3'*AOX1* sequencing primer (5'-GCAAATGGCATTCTGACATCC-3') by the dideoxynucleotide protocol (Sanger et al. 1977). The DNA sequence of the gene was verified to be homologous to the human cystatin C gene.

## *Mutation of Human Cystatin C*

To enhance the stability of recombinant human cystatin C, the gene was mutated to introduce a consensus sequence for N-linked glycosylation (N-X-S/T) at two sites on the gene. The *Hind III/Xba I* segment of the pPICZα-A carrying the recombinant human cystatin C gene was ligated with the PUC19 vector (Boehringer Mannheim, Laval, QU) and was subjected to site-directed mutagenesis (QuikChange Site-Directed Mutagenesis Kit, Stratagene, La Jolla, CA). The oligonucleotides used for mutagenesis were: a) 5'-GAG TAC AAC AAG TCC TCT AAC GAC ATG-3' and 5'-CAT GTC GTT AGA GGA CTT GTT GTA CTC-3' for Ala37 replaced with Ser37 and b) 5'-CTC AAC CAA ACT TGA CTA ACT GTC CAT TCC ACG AC-3' and 5'-GTC GTG GAA TGG ACA GTT AGT CAA GTT TGG TTG AG-3' for Asp81 replaced with Thr81. The DNA sequences of pUC19 carrying the mutated cystatin C gene were verified with the α-factor sequencing primer.

## *Transformation of Mutated Cystatin C into P. pastoris by Electroporation*

The competent *P. pastoris* host cells and the linearized DNA containing the mutated cystatin gene were prepared by Dr. Masahiro Ogawa in Food, Nutrition, and Health at UBC.

*Host Cell Preparation*

Nontransformed X-33 cells were added to 5 ml of YPD medium (10 $g \cdot l^{-1}$ yeast extract, 20 $g \cdot l^{-1}$ peptone, 20 $g \cdot l^{-1}$ dextrose) and grown overnight at 30°C. This culture was added to 500 ml of YPD medium in a 2-liter flask and grown overnight at 30°C. The culture was then centrifuged at 1500xg for 5 minutes at 4°C. The pellet was resuspended with 500 ml of ice-cold sterile water and centrifuged again. The pellet was then resuspended with 250 ml of ice-cold sterile water and centrifuged once again. Two more centrifuge cycles were performed with 20 ml and then 1 ml of ice-cold 1 M sorbitol. The final volume of the suspension was approximately 1.5 ml. The cells were kept on ice and used for electroporation that same day.

*Linearized DNA Preparation*

The *Xho* I/*Xba* I segment of the expression cassette, including the mutated DNA insert, was ligated with pPICZα-A to allow for secretion of the mutated cystatin C and selection of transformants by Zeocin™ resistance. Approximately 10 μg of the vector was linearized with the restriction enzyme *Dra* I. The solution was centrifuged to pellet the DNA and then suspended with 80% ethanol. The suspension was air-dried and then resuspended in 10 μl of sterile deionized water. The suspension was used for electroporation that same day.

*Electroporation Protocol*

The electroporation protocol was performed according to the EasySelect™ *Pichia* Expression Kit (Invitrogen 1998). 80 μl of the sorbitol-suspended cells were added to an ice-cold electroporation cuvette. 10 μl of the linearized DNA suspension was then added and mixed with the pipette. After 5 minutes incubation on ice, the cuvette was pulsed with a Biorad GenePulser according to the manufacturer's instructions. The settings for the Biorad GenePulser were 1500 V, 25 μF, and 200 Ω. Immediately after pulsing, 1 ml of ice-cold 1 M sorbitol was added to the cuvette and its contents were transferred to a sterile 15 ml tube. The tube was incubated at 30°C for 2 hours without shaking.

Since plating at low cell densities favours efficient Zeocin™ selection, aliquots of 20 μl were spread on YPDS plates (10 $g \cdot l^{-1}$ yeast extract, 20 $g \cdot l^{-1}$ peptone, 20 $g \cdot l^{-1}$ dextrose, 182 $g \cdot l^{-1}$ sorbitol, 20 $g \cdot l^{-1}$ agar) that contained either 100, 500, 1000, or 2000 $\mu g \cdot ml^{-1}$ Zeocin™. The plates were incubated at 30°C until colonies formed (approximately 3 days). A total of seven colonies were selected from the plates and were streaked onto fresh YPD plates, incubated once again, and stored at 4°C. In addition, one colony of nontransformed X-33 cells was plated and stored in the same manner.

Integration of the mutated cystatin C gene into *P. pastoris* was confirmed by genomic PCR (polymerase chain reaction) using a 5'AOX1 primer paired with a 3'AOX1 primer, 5'-GAC TGG TTC CAA TTG ACA AGC-3' and 5'-GCA AAT GGC ATT CTG ACA TCC-3'. The PCR products showed that all of the clones tested incorporated the mutated cystatin C gene.

## *Evaluating Zeocin™ Resistance*

In order to correlate Zeocin™ resistance to protein expression, the threshold of Zeocin™ resistance was determined for each of the seven colonies that were selected after electroporation as well as for the nontransformed control. The Zeocin™ threshold was defined as the highest concentration in which accumulation of cells on the bottom of the well was visible as an opaque film (wells in which no growth was observed remained translucent). YPD (+ Zeocin™) liquid media were prepared with Zeocin™ concentrations of 0, 100, 250, 500, 1000, 1500, and 2000 $\mu g \cdot ml^{-1}$. 200 μl of each YPD (+ Zeocin™) media was added to a 96-well ELISA plate. Enough wells were filled to allow for duplicates of each colony in each Zeocin™ concentration (a total of 112 wells). Since a low cell concentration is favourable for Zeocin™ selection, a 5-μl inoculum was used for each well that contained 200 μl of YPD (+ Zeocin™) medium. The ELISA plates were incubated at 30°C and observations of cell growth were made on a daily basis.

## *Evaluating Cystatin Expression*

The cystatin expression experiments were performed in baffled shake flasks containing 35 ml of buffered YPG media (10 $g \cdot l^{-1}$ yeast extract, 20 $g \cdot l^{-1}$ peptone, 10 $g \cdot l^{-1}$glycerol, 100 mM potassium phosphate buffer, pH 7.0). Methanol feeding was initiated at 48 hours and was performed once every 24 hours. Methanol was fed at a dosage of 0.5% (v/v).

## *Cystatin Assay*

Active cystatin expressed by the *P. pastoris* transformants was measured using a papain inhibition assay with Nα-benzoyl-DL-arginine p-nitroanilide hydrochloride as the substrate (Barrett 1981). The amount of nitroaniline liberated by residual papain activity was measured by spectrophotometry at 410 nm (Milton Roy, Spectronic 20D). Papain (EC 3.4.22.2) and BAPNA (lot 127H0760) were from Sigma (St. Louis, MO). The activity of cystatin was compared to blanks that were prepared by heating the sample to 100°C for 10 minutes to denature the cystatin (Nakamura et al. 1998a). Samples and blanks were measured in duplicate.

The assay was designed for cystatin concentrations above $10^{-6}$ Molar. This is equivalent to a concentration of approximately 1.3 $mg \cdot l^{-1}$ (13.4 kDa) of cystatin C in the culture medium. Higher concentrations were diluted with distilled water and lower concentrations were concentrated by filtration using a YM-3 membrane in a 50-ml ultrafiltration cell (Amicon Inc., MA) under 30 psi of nitrogen pressure.

### *Cystatin Assay Reagents*

I. The substrate stock solution was Nα-benzoyl-DL-arginine p-nitroanilide hydrochloride (BAPNA) (Sigma, St. Louis, MO) dissolved in dimethyl sulfoxide (DMSO) (Fisher, Fair Lawn, NJ). The concentration was 43.4 $mg \cdot ml^{-1}$. This solution was stored at 4°C.

II. The assay buffer was a 1.0 M sodium phosphate buffer at pH 6.8 containing 2 mM disodium EDTA (ethylene-diamine-tetra-acetic acid) (all from Fisher, St. Louis, MO). Cysteine base (Sigma, St. Louis, MO) was added freshly before use to a final concentration of 5 mM (0.61 $g \cdot l^{-1}$).

III. The papain solution was prepared by making a 50X dilution of papain stock in distilled water. The papain stock solution was 300 $\mu g \cdot ml^{-1}$ of papain in 0.5 mM $HgCl_2$ and 1.0 mM EDTA (Sigma, St. Louis, MO). The activity of the papain was 19 $units \cdot mg^{-1}$ protein. One unit will hydrolyze 1.0 μmole Nα-benzoyl-L-arginine ethyl ester per minute at pH 6.2 and 25°C.

IV. The stopping reagent was glacial acetic acid (Fisher, Nepean, ON).

*Cystatin Assay Procedure*

A schematic diagram of the assay procedure is shown in Figure 4.2 and is described as follows:

Into each 1.5 ml Eppendorf centrifuge tube was added 0.5 ml of assay buffer, 0.1 ml of papain stock, and 0.1 ml of the cystatin sample. The reaction was started by the addition of 50 μl of stock substrate solution. The tubes were incubated for 15 minutes at 37°C. After incubation, the reaction was stopped by the addition of 0.32 ml of the acetic acid such that its final concentration was 30% (v/v). The nitroaniline that was liberated by enzymatic activity was quantified by spectrophotometry at 410 nm. At this wavelength, the product of papain activity (the decomposition of BAPNA) can be detected. Therefore, lower concentrations of product indicated higher concentrations of cystatin. The percentage of inhibition of papain was calculated by comparing the absorbance of the sample assay solution to the absorbance of the blank assay solution.

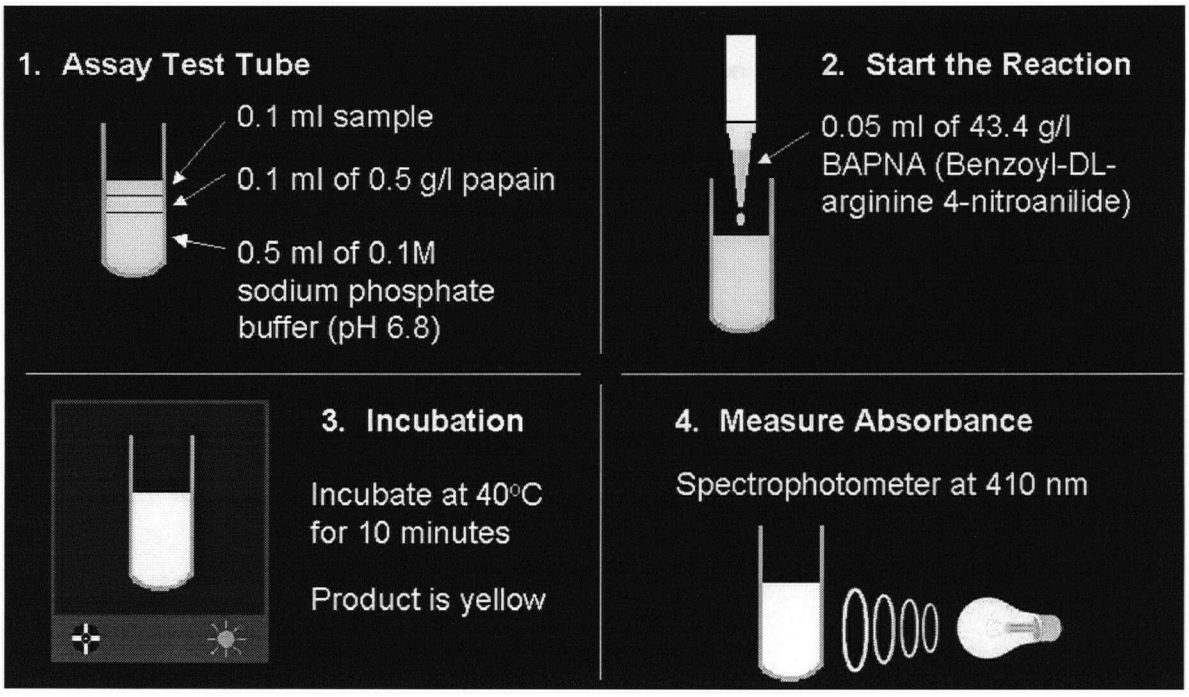


Figure 4.2. Schematic diagram for the cystatin assay by papain inhibition.

## *Cell Density*

The wet cell weight (wcw, $g \cdot l^{-1}$) was used as a measure of cell density within the flasks as well as in the bioreactor. For high cell density cultures, wet cell weights have the advantage of convenience compared to optical density measurements. Measuring the wet cell weight does not require that the samples be diluted.

A 1.0 ml sample was centrifuged at 14,000 rpm for 3 minutes (Eppendorf Centrifuge 5415 C). The supernatant was stored at 4°C and the pellet was weighed. The wet cell weights were measured in duplicate. The dry cell weights (dcw, $g \cdot l^{-1}$) were calculated based on a calibration curve that is shown in Figure 4.3 ($R^2 = 0.99$):

$$dcw = 0.31 \times wcw - 3.47$$

The calibration curve was prepared using 30 1-ml samples that were dried for 24 hours at 105°C.

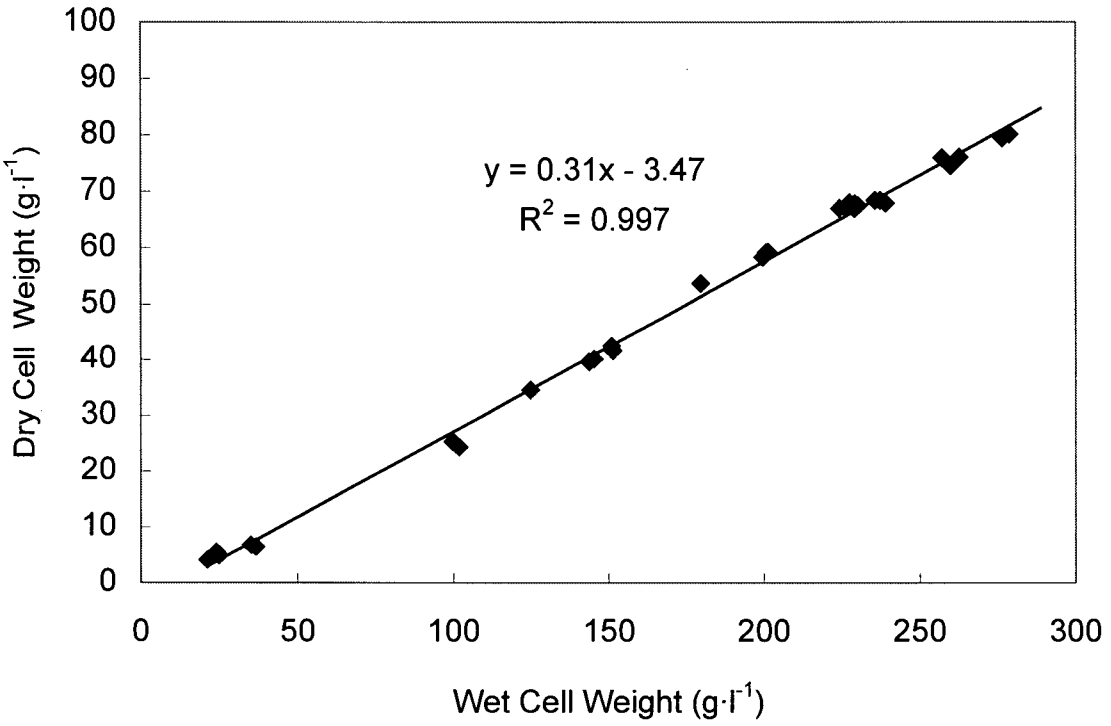


Figure 4.3. Calibration curve for dry cell weight as a function of wet cell weight. The calibration curve was based on 30 1-ml samples that were dried for 24 hours at 105°C.

## 4.1.3 Results

### *Evaluating Zeocin™ Resistance*

The electroporated cells were spread onto YPD agar plates that had Zeocin™ concentrations of 0, 100, 500, 1000, or 2000 $\mu g \cdot ml^{-1}$. 20 $\mu l$ of the electroporated suspension was spread onto duplicate plates for each concentration of Zeocin™. In addition, the untransformed X-33 wild-type *P. pastoris* was spread onto plates that had 0, 100, and 500 $\mu g \cdot ml^{-1}$ of Zeocin™. Colonies from the electroporated suspension completely covered the 0 and 100 $\mu g \cdot ml^{-1}$ plates, but only appeared in sparse numbers on the 500 $\mu g \cdot ml^{-1}$ plates. No colonies grew on the 1000 or 2000 $\mu g \cdot ml^{-1}$ plates. Wild-type colonies completely covered the plates without Zeocin™, and a few colonies were seen on the 100 $\mu g \cdot ml^{-1}$ plates.

A total of 8 colonies were selected to determine their threshold resistances to Zeocin™. One of these colonies was the nontransformed wild strain of *P. pastoris*, two

colonies were selected from the 100 $\mu g \cdot ml^{-1}$ plates, and 5 colonies were selected from the 500 $\mu g \cdot ml^{-1}$ plates. A 200 µl inoculum for each colony was grown in the wells of an ELISA plate in YPD medium (without Zeocin™). After 24 hours of growth, 5 µl aliquots were added, in duplicate, to wells that contained 200 µl of YPD (+ Zeocin™) media. The Zeocin™ concentrations were 0, 100, 250, 500, 1000, 1500, and 2000 $\mu g \cdot ml^{-1}$. The Zeocin™ threshold was defined as the highest concentration in which accumulation of cells on the bottom of the well was visible as an opaque film. Table 4.2 indicates the Zeocin™ threshold concentration that was observed in the wells of the ELISA plate. The table also indicates the Zeocin™ concentration of the original plate from which the colony was selected.

Table 4.2. Zeocin™ resistance as a function of colony number.

| | | Colony | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **Wild** | **1** | **2** | **3** | **4** | **5** | **6** | **7** |
| Zeocin™ in original plate ($\mu g \cdot ml^{-1}$) | 0 | 100 | 100 | 500 | 500 | 500 | 500 | 500 |
| Zeocin™ threshold ($\mu g \cdot ml^{-1}$) | 0 | 100 | 100 | 500 | 500 | 1000 | 500 | 1000 |

As indicated in Table 4.2, all of the colonies grew in the liquid media that had Zeocin™ concentrations that were greater than or equal to the concentration in the plate from which they were selected. However, some of the colonies showed a higher Zeocin™ threshold compared to the concentration of Zeocin™ in the original plate. Colonies 5 and 7 were taken from a plate containing 500 $\mu g \cdot ml^{-1}$ of Zeocin™ and each had Zeocin™ thresholds of 1000 $\mu g \cdot ml^{-1}$ when tested in the liquid media.

## *Evaluating Cystatin Expression*

The 8 colonies that were selected from the agar plates exhibited a range of Zeocin™ thresholds. The next step was to determine if a higher Zeocin™ threshold could be correlated with a higher expression of biologically active cystatin.

Each of the colonies was grown in duplicate in 250 ml shake flasks. In order to facilitate sampling, no more than 5 colonies were investigated at any one time. The first set of experiments tested the nontransformed wild-type as well as colonies 2, 4, 5, and 7. The second experiments repeated the wild-type as well as colony 2, and tested colonies 1, 3, and 6. As shown in Figure 4.4, the maximum cell densities were reached after approximately 48 hours. Methanol feeding was initiated at 48 hours. Note that the dry cell weights remained fairly constant during the methanol-feeding phase. However, there was a steady yet slight decline in cell density after 48 hours.

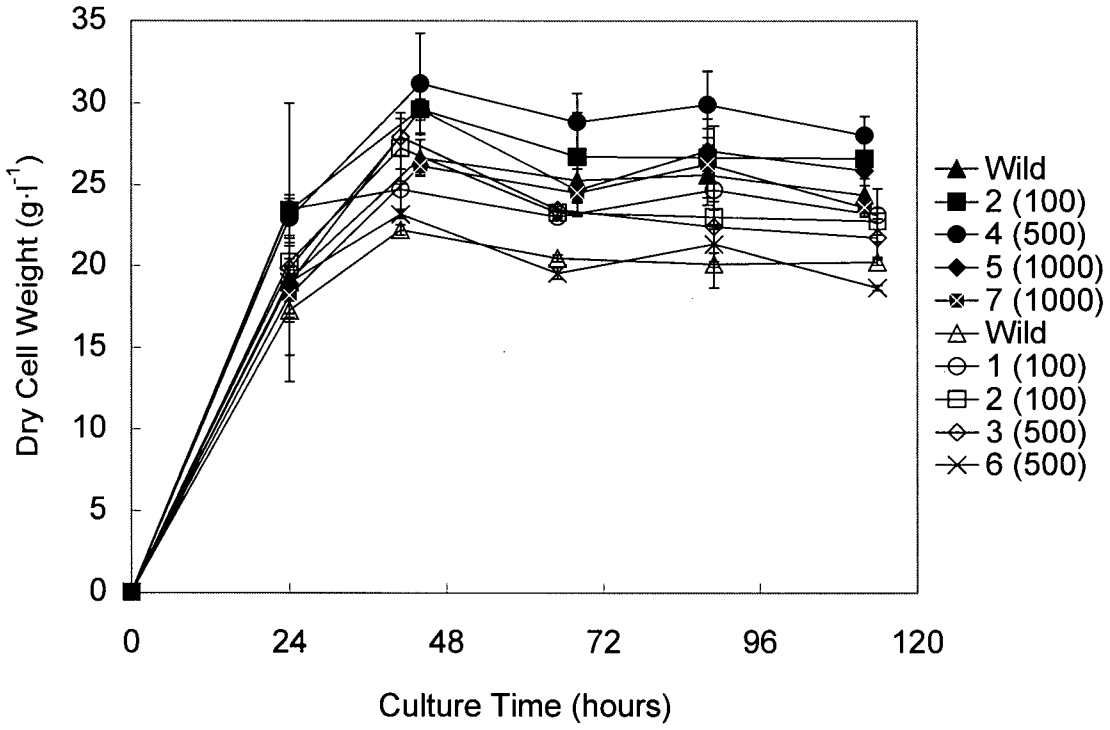


Figure 4.4. Dry cell weight as a function of culture time. Solid symbols indicate the first set of experiments and open symbols indicate the second set. Each experiment had two replications for each flask. A 1-ml sample was taken for wet cell weight measurements and the dry cell weights were calculated based on an experimentally determined calibration curve. The error bars are the standard deviations of the dry cell weights for the two flasks.

Biologically active cystatin in the culture supernatants was measured using the papain inhibition assay. The cystatin activities were measured as soon as possible after the samples were taken from the flasks. The yield of active cystatin was calculated by

dividing the molar concentration of cystatin by the cell concentration. The cystatin yields versus culture time for both sets of experiments are plotted in Figure 4.5.

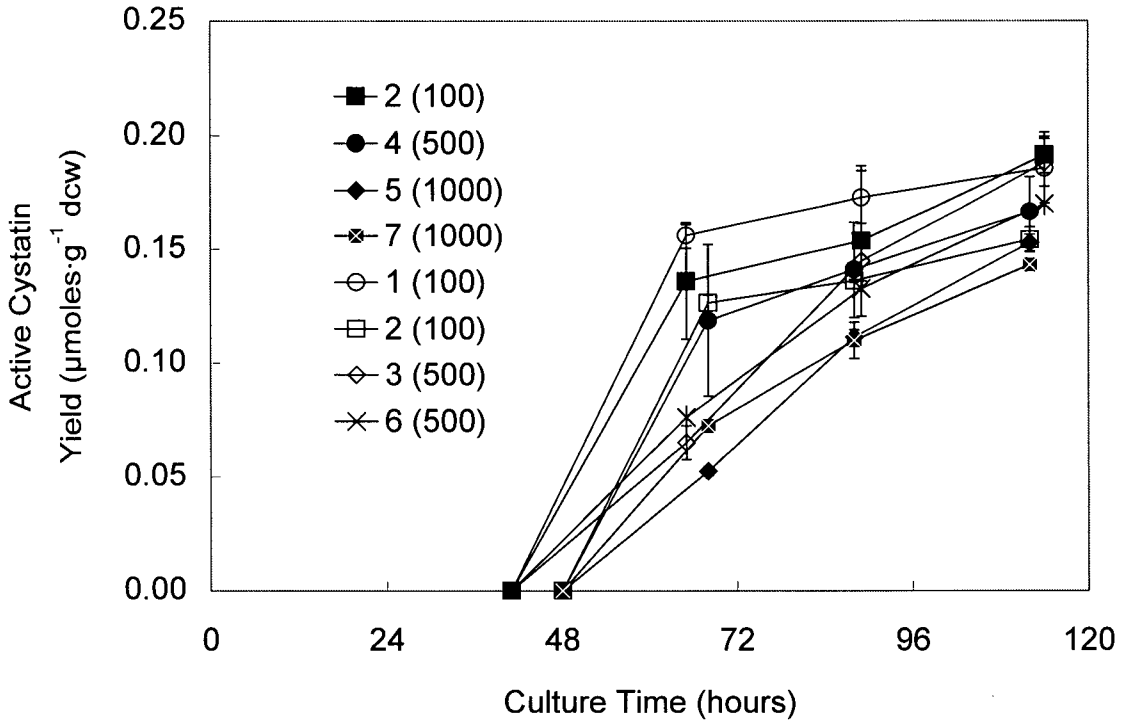


Figure 4.5. Yield of active cystatin as a function of culture time and Zeocin™ resistance. The numbers that appear in parenthesis in the legend correspond to the threshold of Zeocin™ resistance for the transformant (measured in $\mu g \cdot ml^{-1}$ of Zeocin™). Presented are the average values for the two replicates. The error bars are the standard deviations.

After 3 days of induction, colony 2 achieved the highest cystatin yield of 0.19 $\mu moles \cdot g^{-1}$. Colonies with lower resistances to Zeocin™ had a rapid increase in cystatin during the first day of induction and then a slowing increase during the subsequent days. Colonies with higher resistances to Zeocin™ exhibited a relatively constant increase in yield during the entire induction period. After the 3-day induction period there did not appear to be a significant difference between cystatin yield and Zeocin™ resistance. The cystatin yields for each of the colonies were all within the range of $0.16 \pm 0.03$ $\mu moles \cdot g^{-1}$.

## 4.1.4 Discussion

### *Evaluating Zeocin™ Resistance*

Transformation by electroporation resulted in colonies that exhibited various resistances to the Zeocin™ antibiotic. Direct plating onto YPDS (+ Zeocin™) agar resulted in the isolation of colonies exhibiting resistance to either 100 or 500 $\mu g \cdot ml^{-1}$ of Zeocin™. This level of Zeocin™ resistance is low compared to reported resistances of up to 2000 $\mu g \cdot ml^{-1}$ (Higgins and Cregg 1998). These researchers found that the number of colonies appearing on plates containing Zeocin™ at 2000 $\mu g \cdot ml^{-1}$ was only about 1% of that appearing on 100 $\mu g \cdot ml^{-1}$ plates. Based on this comparison, it is likely that the transformants isolated on the 500 $\mu g \cdot ml^{-1}$ plate contain only a small number of gene copies. But resistance to Zeocin™ concentrations above 100 $\mu g \cdot ml^{-1}$ does infer greater than one gene copy.

Evaluating the Zeocin™ resistance threshold in liquid media indicated a slightly higher resistance to Zeocin™. There are several possibilities for this observation. The cell density may have been too high for efficient selection, the Zeocin™ may have lost activity, or the transformant may have had a higher copy number than was required for growth on the 500 $\mu g \cdot ml^{-1}$ plate. However, in the case of the control colony, the Zeocin™ was effective. The wild-type colony, which proliferated in the absence of Zeocin™, exhibited only limited growth in the well that had 100 $\mu g \cdot ml^{-1}$ of Zeocin™. However, since some growth was in fact observed for the wild-type colony in the presence of Zeocin™, its antibiotic effect may have been reduced.

High cell densities have been shown to decrease the antibiotic effect of Zeocin™ (Higgins and Cregg 1998). When colonies are streaked onto solid media, the relatively thick layer of cells placed on the media surface results in the growth of false-positive colonies. The same is true in the case of liquid media if a large quantity of inoculum is used. The reason is that dead cells titrate out the available Zeocin™ from the media, and allow the remaining cells to grow and appear resistant and/or hyperresistant. Therefore, dead cells in the inoculum, or cells that died due to Zeocin™, may have absorbed enough of the Zeocin™ to allow for the remaining cells to grow. However,

since a mere 5 μl inoculum was added to 200 μl of medium, it is not likely that dead or dying cells effected the Zeocin™ concentration.

The Zeocin™ may have lost activity due to a salt concentration above 5 $g \cdot l^{-1}$, a pH less than 7.5, or exposure to too much light (Invitrogen 1998). The salt concentration of the YPD (+ Zeocin™) liquid media is less than 5 $g \cdot l^{-1}$ but the pH may have dropped below its initial value of 7.0. A decrease in pH from 7.0 to 3.5 has been observed for high cell density shake-flask cultures of *P. pastoris* that were grown in the same unbuffered YPD media (data not shown). However, the cell densities in the ELISA plate were significantly less because the plate was incubated without shaking and only a small inoculum was used. The final pH values in the wells were not measured, but it is likely that there was not a significant enough decrease in pH to inhibit the action of Zeocin™. Covering the ELISA plates in aluminum foil minimized deactivation of Zeocin™ by light exposure.

It is also possible that the transformants that exhibited higher Zeocin™ resistances in the liquid media had a higher gene copy number than was required for survival on the 500 $\mu g \cdot ml^{-1}$ plate. These colonies (5 and 7) were not subsequently streaked onto plates that had higher concentrations of Zeocin™ because of the likely result of observing false-positives due to a thick layer of cells that would titrate out the Zeocin™.

## *Evaluating Cystatin Expression*

The expression experiments indicated that there might have been an inverse relationship between Zeocin™ resistance and cystatin yield. This was especially evident during the first 24 hours of expression. This trend has been observed by other researchers and was attributed to secretion pathways that became blocked or rendered ineffective (Thill et al. 1990). Because cystatin is a relatively small molecule (13.4 kDa), it is unlikely that its size would have been responsible for blocking secretion pathways. However, the variant of cystatin had two consensus sequences for N-linked glycosylation that may have interfered with the secretion pathways. It is interesting to note that the colony that exhibited the greatest resistance to Zeocin™ (colony 7, that exhibited growth in 1000 $\mu g \cdot ml^{-1}$ of Zeocin™) expressed the majority of cystatin in its

nonglycosylated form when grown in the bioreactor (see Chapter 5). Conversely, experiments by Dr. Ogawa generally indicated a higher proportion of glycosylated cystatin. His experiments were performed in shake-flasks and with colonies taken from plates with only 300 or 500 $\mu g \cdot ml^{-1}$ of Zeocin™. Therefore, the transformants that expressed a higher proportion of glycosylated cystatin may have a lower gene copy number. It could be speculated that the secretion pathways associated with glycosylation of cystatin C become ineffective for transformants that have multiple copies of the expression cassette.

The most likely bottleneck for secretion pathway inefficiencies is the endoplasmic reticulum (ER). In the ER of eukaryotic cells, proteins fold and are assembled into oligomers, exiting the ER only upon attaining native structure. This pathway provides a direct means of producing correctly folded and assembled proteins, since the ER's proofreading apparatus ensures that misfolded proteins are not released into the medium. Accumulation of recombinant proteins in the ER has been reported for yeast cells (Biemans et al. 1991) and Chinese hamster ovary cells (Gennaro et al. 1991). It has also been observed that glycosylation can inhibit the proper folding of glycoproteins in the endoplasmic reticulum (Helenius et al. 1992). Therefore, it is possible that attempts to glycosylate cystatin interfered with its folding and/or disulfide bridge formation. As a result, glycosylated cystatin may have been retained in the ER. Nonglycosylated cystatin may have been able to achieve its authentic structure and was thus released from the endoplasmic reticulum and directed for secretion.

Bovine pancreatic trypsin inhibitor (BPTI) expressed by *S. cerevisiae* was observed to accumulate intracellularly in an inactive form when expressed by multi-copy transformants (Parekh et al. 1995). Since the actual level of secretion was similar for both single- and multi-copy transformants, it was hypothesized that high-level expression saturates available chaperones and foldases in the secretory pathway. As a result, extensive BPTI aggregation was observed. It was reported later that increases in BPTI transcription could actually decrease secretion due to titration of available ER foldases and chaperones (Parekh and Wittrup 1997). It was noted that secretion of BPTI was not effected by the presence or absence of N-linked glycosylation.

A secretory saturation was also noted for mammalian cells secreting immunoglobulins, whereby excessive secretion impaired cell replication (Reff 1993).

### 4.1.5 Conclusions and Recommendations

These experiments have indicated that 1) electroporation results in transformants that have a range of thresholds to Zeocin™ resistance, 2) only a small portion of the transformants exhibited hyperresistant phenotypes for Zeocin™, and 3) there appears to be a weak inverse relationship between Zeocin™ resistance and cystatin expression. It is hypothesized that the secretion pathways became blocked or inefficient for high levels of protein expression associated with multiple insertions of the expression cassette.

Colony 7 was suspected to have a higher gene copy number than colony 2 because it exhibited a greater resistance to Zeocin™ (1000 $\mu g \cdot ml^{-1}$ compared to 100 $\mu g \cdot ml^{-1}$). Despite colony 2 yielding the greatest amount of biologically active cystatin, colony 7 was selected for the bioreactor experiments that are presented in Chapter 5. The selection of colony 7 was based on several factors. Primarily, an overwhelming majority of the published reports have indicated that higher gene copy numbers resulted in higher levels of expression. Since the experiments presented in this chapter were the first expression experiments performed in Bio-Resource Engineering, we could not eliminate the possibility that unknown factors and inexperience resulted in colony 2 expressing a greater yield of active cystatin than colony 7. However, we were very confident that there was a significant difference in their Zeocin™ resistance. Therefore, without a strong argument to support the conclusion of an inverse relationship between Zeocin™ resistance and cystatin expression (see Figure 4.5), the colony that was suspected to have the highest copy number was selected. In addition, an increased expression for multi-copy transformants has been shown to be more pronounced in bioreactor cultures to those in shake-flasks (see Figure 4.1). Therefore, it was hypothesized that colony 7 would express higher levels of cystatin in the bioreactor compared to colony 2. However, due to time constraints, this hypothesis was not verified in the bioreactor.

A more thorough investigation of the effects of gene copy number would have been advantageous. An unusual inverse relationship between Zeocin™ resistance and recombinant protein expression was observed, but too few experiments were performed to impart confidence in these observations. Also, the colonies that were investigated likely represented a limited range of gene copy numbers. Electroporation should have been performed several times and many more Zeocin™ plates should have been used for isolating the transformants. This would have allowed for transformant selection based on a much greater range of gene copy numbers. Finally, but most importantly, the gene copy numbers for each of the transformants should have been quantified by DNA dot blots. Without knowing the actual copy number, the relationship between gene copy number and cystatin expression cannot be clearly defined.

## 4.2 Cystatin Stability in the Extracellular Environment

### 4.2.1 Introduction

As previously discussed, human cystatin C is vulnerable to oxidation (Berti et al. 1997), proteolysis (Lenarcic et al. 1991), and dimerization (Abrahamson and Grubb 1994b; Ekiel and Abrahamson 1996), all of which reduce its stability. During the bioreactor experiments that are presented in Chapter 5, it became evident that the papain inhibitory activity of the samples decreased during storage. The first data collection strategy was to collect samples from the bioreactor for the entire culture time (at least 4 days), store them at 4°C, and then measure all of the samples together. The intent was to avoid day-to-day measurement variations due to factors such as assay buffer composition (that was made fresh each day) and spectrophotometer calibration. However, some measurements for active cystatin were performed directly after sampling. When these same samples were measured several days later, it became evident that the activity decreased during storage. The storage affect was more thoroughly investigated by performing daily measurements of samples that were stored at 4°C. The objective of these experiments was to tabulate cystatin activity as a function of storage time. This section also suggests some measures that could be used to maintain cystatin activity during storage. These experiments are presented in this Chapter because of their applicability to the analytical procedures that were used throughout the experiments in the following chapter.

### 4.2.2 Materials and Methods

#### *Sampling from the Bioreactor*

A detailed description about the operation of the bioreactor is provided in Chapter 5. The bioreactor was equipped with an aseptic sampling port from which approximately 10 ml of the culture was collected during sampling. From this 10-ml sample, 1 ml was added to each of 3 Eppendorf centrifuge tubes. The 1-ml samples were centrifuged at 14,000 rpm for 3 minutes (Eppendorf Centrifuge 5415 C). The

supernatant from each sample was poured into separate centrifuge tubes and were stored at 4°C.

The samples were measured for cystatin activity on a daily basis. This procedure was performed for various sampling times from the bioreactor as well as for various bioreactor cultures. As a result, a pool of data was collected that tabulated cystatin activity as a function of storage time.

### *Analytical Procedures*

Cell densities and cystatin concentrations were measured as described in section 4.1.2.

## 4.2.3 Results

The pool of cystatin activity data was plotted as a function of storage time. The results are shown in Figure 4.6.

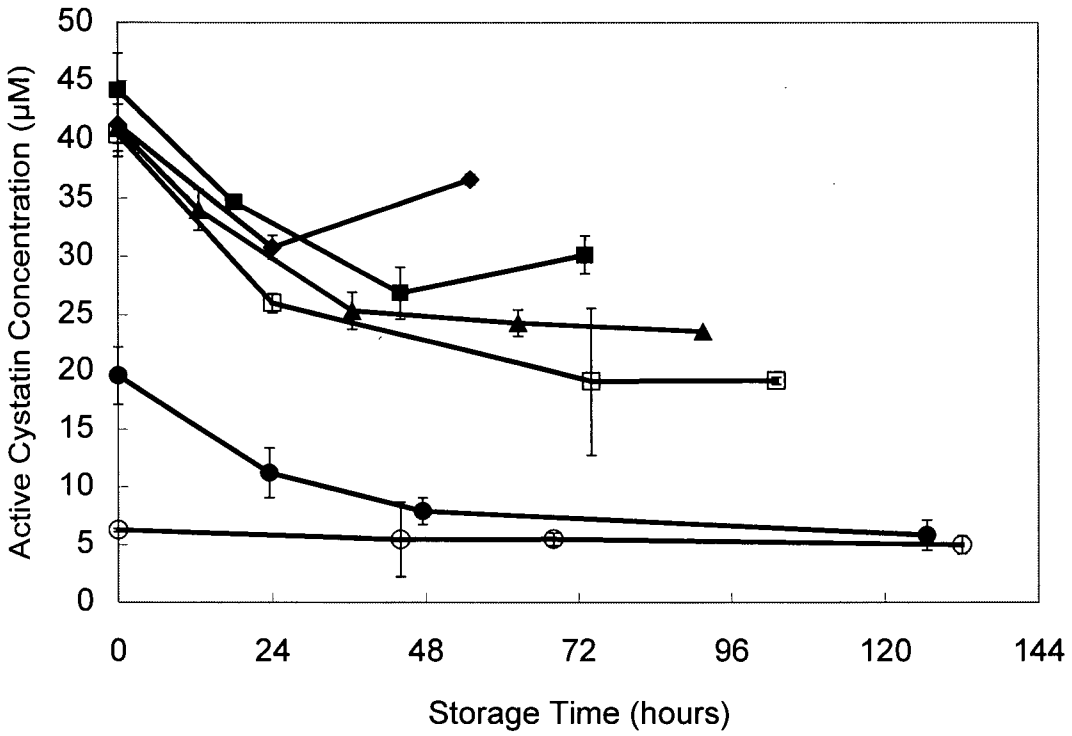


Figure 4.6. Measured activity for cystatin as a function of storage time. Samples were taken from the reactor, centrifuged, and the supernatant was stored in Eppendorf centrifuge tubes at 4°C. Each set of points represents a sample taken from the bioreactor at a particular time. Cystatin activities were measured in duplicate. The error bars are the standard deviations of the measurements.

The measured data were used as a basis for estimating a series of empirical deactivation curves. The deactivation curves are shown in Figure 4.7.

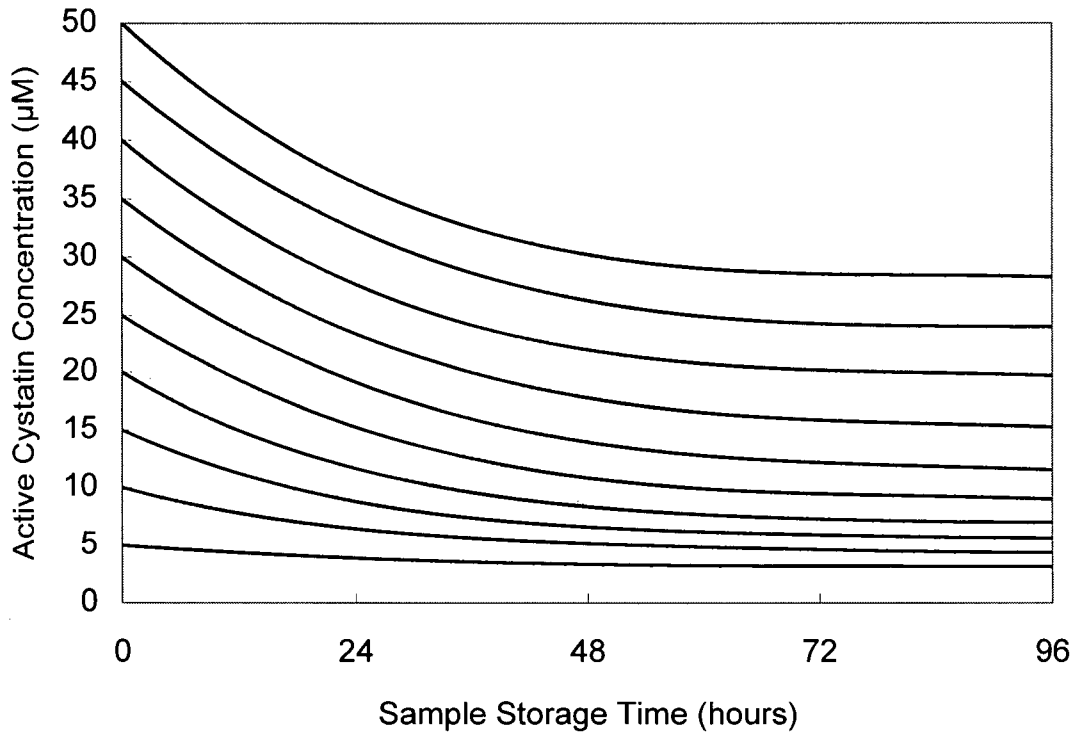


Figure 4.7. Cystatin deactivation curves for the storage of samples taken from the bioreactor and stored at 4°C. The data presented in Figure 4.6 were used as a basis for generating 10 hypothetical data sets that were each fitted with a third-order regression line using Microsoft Excel. This figure was used to estimate concentrations of active cystatin in the bioreactor based on the time lapse between sampling and cystatin measurement.

The curves presented in Figure 4.7 were estimated empirically and were not derived from a model for deactivation. The curves were used to adjust the measured cystatin activities if they were stored prior to measurement. For consistency with the first bioreactor experiments, a better method of storage was not used for the remainder of the experiments. However, for all subsequent experiments, an effort was made to measure the activities as soon as possible after sampling.

### 4.2.4 Discussion

It is evident that the method of sample storage did not provide optimal stability for the cystatin variant. The most significant decrease in measured activity was observed during the first 24 hours of storage. This was especially true for higher concentrations. After 24 hours, the activity remained fairly stable. Protease enzymes, aggregation, or

media composition are factors that may have contributed to the loss of activity during storage time.

Cystatin is most stable at moderate temperatures. In fact, with the addition of a sodium phosphate buffer (50 mM, pH 6.7) as well as sodium chloride (0.1 M), cystatin activity has been maintained for 3 months at room temperature (Ekiel and Abrahamson 1996). As well, a temperature of 4°C is commonly chosen to perform purification processes for recombinant cystatin 4°C (Dalboge et al. 1989; Ni et al. 1997; Sotiropoulou et al. 1997). Therefore, it seems evident that factors other than a storage temperature of 4°C were responsible for the loss of activity that was observed.

Cystatin expressed intracellularly in *E. coli* has been shown to be susceptible to degradation by protease enzymes (Hall et al. 1993). To overcome this problem, these researchers added a cocktail of protease inhibitors to assist in the preservation of cystatin. Protease enzymes from *P. pastoris* are also known to degrade recombinant proteins (Brierly et al. 1994). However, in the supernatant of *P. pastoris* cultures, protease enzymes are fewer in number as well as variety compared to the periplasm of *E. coli* extracts (Cregg et al. 1993). Therefore, protease inhibitors were not added to the bioreactor or to the stored samples. Instead, a storage temperature of 4°C was relied upon to reduce the activities of any protease enzymes that may have been in the sample. However, even at 4°C, the protease enzymes may have had enough activity to hydrolyze the cystatin. The effects of protease activity can be reduced by 1) the addition of peptone (20 $g \cdot l^{-1}$) or casamino acids (5 $g \cdot l^{-1}$) to the culture medium that appear to reduce degradation by acting as excess substrates for protease enzymes (Chen et al. 2000; Clare et al. 1991b) or 2) the use of a protease-deficient *P. pastoris* host strain (Invitrogen 1998). However, these protease deficient strains are not as robust as the wild-type strains, and require greater care in growth and storage (Higgins and Cregg 1998).

Cystatin C can be rendered inactive if it becomes aggregated (Ekiel and Abrahamson 1996). The process of aggregation is a function of temperature, pH, and concentration. However, cystatin C is stable within a large temperature range. In an appropriate buffer, temperature-associated aggregation occurs above 75°C (Nakamura et al. 1998a). Cystatin is stable at low temperatures and is commonly stored at -20°C or

-70°C (Hall et al. 1993; Lindahl et al. 1992). However, the freezing process has been shown to cause partial inactivation in the case of chicken cystatin (Barrett 1981). In order to prevent freezing-associated inactivation, a sufficient amount of sucrose or glycerol can be added to the sample to prevent freezing at -20°C

It has been shown that below pH 5, human cystatin C forms an inactive dimer when incubated for 11 hours in a sodium phosphate buffer (50 mM, pH 6.7) that includes sodium chloride (0.1 M) (Ekiel and Abrahamson 1996). Cystatin samples were taken from the bioreactor when it was operated at pH 6, but they were not stored in a pH buffer or in the presence of a stabilizing salt. It is possible that the pH of the supernatant could have dropped below pH 5 during storage, but this was not verified. In addition, in the absence of a stabilizing salt, it is possible that dimerization occurs at higher pH values than pH 5.

Apart from temperature and pH, high concentrations of cystatin encourage aggregation. Even in the presence of a sodium phosphate buffer, and a stabilizing salt, concentrations above 75 μM (1 $mg \cdot ml^{-1}$) lead to aggregation and subsequent precipitation of cystatin (Ekiel and Abrahamson 1996). This concentration is significantly higher than concentrations in blood serum (0.08 μM) and cerebrospinal fluid (0.4 μM) (Barrett et al. 1984), but is not much higher than the maximum expression of cystatin that was achieved in the bioreactor (55 μM) (see Chapter 5). However, in the absence of a buffer and a stabilizing salt, the concentration required for aggregation may be much less than 75 μM. The trends in Figure 4.6 and Figure 4.7 indicate that a greater loss of activity was observed for higher initial concentrations of active cystatin. This may be indicative of concentration-associated aggregation.

### 4.2.5 Conclusions and Recommendations

These experiments have revealed that 1) cystatin activities were not maintained in supernatant samples that were stored at 4°C and 2) the rate of loss of activity was greater for higher concentrations of cystatin.

Experiments performed in the future should aim to reduce proteolytic activity and prevent aggregation of cystatin C. To reduce proteolysis of cystatin, the addition of peptone (up to 20 $g \cdot l^{-1}$) (Clare et al. 1991b) to the bioreactor should be considered in

preference to using a protease-deficient strain of *P. pastoris* (SMD1163). As previously mentioned, protease deficient strains are not as robust as the wild-type strains, and require greater care in growth and storage. To prevent aggregation, a pH buffer and a stabilizing salt should be added to samples taken from the bioreactor. The final concentration of sodium phosphate should be 50 mM (pH 6.7) and sodium chloride should be 0.1 M (Ekiel and Abrahamson 1996). The addition of a buffer and a stabilizing salt would also serve to dilute the concentration of cystatin in the sample and help to prevent concentration-associated aggregation. The sample could be diluted by 50% with an addition of an equal part of 0.2 M sodium chloride buffered with 100 mM sodium phosphate (pH 6.7).

# Chapter 5

## Bioreactor Expression of a Cystatin C Variant by *Pichia pastoris*

The experiments presented in Chapter 2 investigated the effects of culture medium on the cell density in shake-flask cultures of *P. pastoris*. The experiments in Chapter 4 investigated the effect of gene multiplicity on cystatin expression as well as the activity of cystatin upon storage. This chapter describes cystatin expression in a bench-scale bioreactor. Section 5.1 describes the optimization efforts with a $Mut^s$ transformant. The effects of pH as well as a mixed feed of glycerol and methanol during induction are described. Section 5.2 compares the $Mut^s$ transformant from section 5.1 to a $Mut^+$ transformant for expression of cystatin in the bioreactor.

## 5.1 Bioreactor Expression of a Cystatin C Variant

### 5.1.1 Introduction

#### *Process Configurations*

The choice of batch, fed-batch, or continuous processes can make a significant difference to the yield of recombinant proteins. In the case of *P. pastoris*, the debate is whether a particular process scheme is more effective than the others in terms of protein productivity and the stability of the host.

As with most pioneering cultures, *P. pastoris* was first grown in batch reactors (Patel et al. 1983). With the use of methanol as the sole carbon source, the cells were grown to allow for the study of its endogenous proteins.

The first patent was issued to Wegner in 1983 for a fed-batch high cell density fermentation of *P. pastoris* (Wegner 1983). Using the $Mut^+$ phenotype, the cells were grown to 130 $g \cdot l^{-1}$ based on dry cell weight (dcw). Wegner's later efforts with a similar configuration attained cell densities of 105 $g \cdot l^{-1}$ and had growth rates of 11.6 g $dcw \cdot l^{-1} \cdot h^{-1}$

(Shay et al. 1987). Attempts by fellow researchers to scale-up the patented system from 1 to 14 liters were met with disappointing low cell yields and sporadic protein synthesis (Siegel and Brierley 1988). At the 1-liter scale, methanol was fed at 10 $g \cdot l^{-1} \cdot h^{-1}$ (dilution rate of 0.05 $h^{-1}$) and the recombinant lysozyme productivity was 15 $mg \cdot l^{-1} \cdot h^{-1}$.

According to Wegner's patent, the initial cell mass was grown in a batch phase with excess glycerol to allow for growth and repression of the recombinant protein. Upon exhaustion of the substrate (indicated by an increase in DO, known as a 'DO spike'), nutrients and methanol were continuously fed to the reactor. The $Mut^+$ cells were then harvested at the completion of the fed-batch cycle.

Early experiments with the $Mut^s$ strain did not show promising results (Tschopp et al. 1987b). Wegner's fed-batch approach was used, but a low cell concentration of 40 g $dcw \cdot l^{-1}$ did not compare favourably with the results of the $Mut^+$ strain. Subsequent experiments made use of the $Mut^+$ strain (Tschopp et al. 1987a).

In order to increase the cellular productivity, a continuous process using high cell densities was investigated (Digan et al. 1989). In order to have a fast growth rate to accommodate the continuous process, the $Mut^+$ strain was selected. Methanol was feed constantly at 12.5 $g \cdot l^{-1} \cdot h^{-1}$ to maintain a cell density of 100 g $dcw \cdot l^{-1}$ and achieve a lysozyme c2 productivity of $13 \pm 2$ $mg \cdot l^{-1} \cdot h^{-1}$. At the time it was the highest production rate ever reported for a mammalian protein secreted from a recombinant microorganism. Experiments with fed-batch cultures gave the same protein productivity with a cell concentration of 120 g $dcw \cdot l^{-1}$. Soon after these experiments, it was also suggested that the methanol utilization phenotype did not have a significant effect on productivity (Clare et al. 1991a).

The potential of the continuous culture led to a more detailed investigation of the effects of dilution rate on productivity and plasmid stability in recombinant yeasts (DaSilva and Bailey 1991). Although the studies were targeted at non-methylotrophic yeasts, some relevant issues were addressed. Dilution rates were tested ranging from 0.1 to 0.26 $h^{-1}$ with the result of increased productivity for decreasing dilution rate. However, lower dilution rates had higher cell concentrations that more than offset the lower volumetric flow rate. Another issue that was raised was that continuous

production might select for faster growing genetic variants that have a reduction in plasmid stability. Since plasmids tend to be lost from generation to generation, the variants may grow faster without the burden of recombinant protein production.

A reactor configuration that has not received attention is a multistage system with repressed cells in the first stage, and induced cells in the second (or later) stages (DaSilva and Bailey 1991). This may allow for higher productivities when using the $Mut^s$ strain of *P. pastoris*. The repressed cells could be grown rapidly on a repressing carbon source such a glucose or glycerol and this supply of cells could be continuously fed to a number of parallel vessels where induction with methanol would occur at lower dilution rates.

Recently, it has been shown that it is possible to utilize the convenience of fed-batch control without having to waste time on fermentation cycles (Sreekrishna et al. 1997). This method also takes advantage of the genetic stability of transformants with a copy of the expression cassette integrated into the host's genome. The study used a methanol + sorbitol feed to prevent repression that would otherwise result if glycerol or glucose were to be used. After harvesting 90% of the culture from the first cycle, the bioreactor was refilled with an appropriate volume of growth medium containing sorbitol as the carbon source and supplemented with yeast extract and peptone. These supplements may be in response to a suggestion that long periods of methanol-limited culture may cause an imbalance in the amino acids and result in decreasing protein production (Chiruvolu et al. 1997). After 16 to 20 hours of growth in the supplemented medium, the methanol + sorbitol feeding resumed for up to 72 hours. Using this cycling method and the $Mut^s$ strain, 25 liters of MMP-2 supernatant (matrix metalloproteinase-2) was collected from a 4-liter bioreactor in less than 4 weeks.

Fed-batch cultures have remained the most popular choice among the *P. pastoris* community of researchers. This has been especially evident since the mid 1990s, when the $Mut^s$ phenotype became popular. Fed-batch more easily accommodates the slower growing strain. Because of its simplicity and reliability, a fed-batch configuration was selected for the production of the cystatin C variant by *P. pastoris*.

## *Bioreactor Feeding Strategies*

A Mut$^s$ (methanol utilization slow) strain of *P. pastoris* was chosen for these experiments. The Mut$^s$ phenotype was selected over the Mut$^+$ phenotype because of the latter's increased sensitivity to excess methanol and higher oxygen requirement that can result in oxygen-deficient conditions within the bioreactor (Couderc and Barratti 1980). Mut$^+$ strains consume methanol much faster than Mut$^s$ and produce a proportionately greater amount of formaldehyde, the first product of methanol metabolism. In the presence of excess oxygen and methanol, formaldehyde rises to toxic levels and "pickles" the cells. More detail about the advantages and disadvantages of the two phenotypes is provided in section 5.2.

For Mut$^s$ strains, the methanol concentration should be kept between 0.4 and 3% (Murray et al. 1989). A residual methanol concentration of 1% has been shown to provide the best growth conditions. The culture strategy for Mut$^s$ is similar to Mut$^+$ except that a lower methanol feed rate is required to maintain the desired methanol concentration (Cregg and Madden 1988).

Protein expression with Mut$^s$ strains requires long induction times (100 hours) for maximal protein expression (Chen et al. 1997). However, a mixed feed of methanol and glycerol has been shown to reduce the induction time by increasing the cell density (Loewen et al. 1997) and in some cases, increasing the cellular productivity of heterologous protein (Katakura et al. 1998). In fact, an increased cell density may more than compensate for a reduced cellular productivity that may result from mixed feeding (Brierly et al. 1990).

The first mixed-feed experiments began with methanol to glycerol ratios of 1:4 (0.9 g methanol$\cdot l^{-1}\cdot h^{-1}$: 3.6 g glycerol$\cdot l^{-1}\cdot h^{-1}$) and 1:2 (0.9 g methanol$\cdot l^{-1}\cdot h^{-1}$: 1.8 g glycerol$\cdot l^{-1}\cdot h^{-1}$) for the expression of bovine lysozyme by a Mut$^s$ strain of *P. pastoris* (Brierly et al. 1990). The specific yields of lysozyme were less for mixed feeding compared to methanol feeding, but the lysozyme volumetric productivities were greater due to higher cell densities and reduced induction times. In addition, a higher lysozyme concentration was observed when the lesser amount of glycerol was fed (1:2). Experiments were then performed in a step-wise manner by beginning with the 1:2 ratio, and then gradually increasing the proportion of methanol in the feed until methanol

began to accumulate. This latter feeding protocol achieved the highest concentration of lysozyme and shortened the induction time from 175 to 45 hours. The volumetric methanol feed rate was 4 fold greater compared to feeding methanol as the sole carbon source. For all of the mixed-feed protocols, and especially for higher feed rates of glycerol, the production of inhibitory levels of ethanol (>100 $mg \cdot l^{-1}$) was observed.

The mixed feeding strategy has been gaining popularity. Induction with a methanol to glycerol ratio of 1:4 (1 g methanol:4 g glycerol) increased the production of fish antifreeze protein by 15 fold compared to methanol induction (Loewen et al. 1997). The increased expression was attributed to an increase in cell density achieved by mixed feeding.

Despite glycerol's ability to maintain the productivity of biomass during induction, it has been shown that glycerol concentrations above 2% (w/v) repress the *AOX1* gene promoter (Gellissen et al. 1995). Therefore, alternative carbon sources have been investigated for mixed-feed strategies. Researchers at the Phillips Petroleum Company have recommended sorbitol and alanine as alternative carbon sources that support cell growth, but do not repress the *AOX1* gene promoter (Sreekrishna et al. 1997).

To confirm these claims, mixed-feed experiments that compared glycerol and sorbitol have been performed (Thorpe et al. 1999). The methanol feed rate was adjusted to maintain methanol concentrations between 0.1 and 0.2% (w/v). It was found that the cell yields from sorbitol were less than from glycerol, but the specific rate of recombinant protein production was higher, resulting in volumetric expression levels that were slightly higher for sorbitol feeding. In addition, the sorbitol was allowed to accumulate to 5 $g \cdot l^{-1}$ without an observable effect on protein yield. Therefore, the complications associated with having tight control of glycerol concentrations could be avoided by using sorbitol as a supplemental substrate.

The effects of glycerol and sorbitol on the *AOX1* gene promoter have been studied (Sreekrishna and Kropp 1996). It was found that the highest inductions of the *AOX1* gene promoter occurred using methanol for $Mut^+$ cells or methanol + sorbitol or alanine for $Mut^S$ cells. Lower levels of induction were seen with methanol + glycerol feed combinations.

This section describes the effects of pH and mixed feeding of methanol + glycerol on cell growth rate and cellular productivity of a variant of human cystatin C by *Pichia pastoris*. This work complements other optimization efforts that have investigated the effects of temperature, glycerol fed rate, pH, and mixed feeding of methanol and glycerol for the expression of recombinant proteins in *P. pastoris* (Chiruvolu et al. 1997).

## 5.1.2 Materials and Methods

### *Microorganism*

*Pichia pastoris* X-33 incorporating the pPICZα-A vector (as described in Chapter 4, section 4.1.2) was used as a host for extracelluar expression of a mutated variant of human cystatin C. A transformant exhibiting a $Mut^s$ phenotype was selected.

### *Pichia pastoris Expression of Cystatin*

The expression experiments were performed in an Inceltech LH bioreactor with a 2-liter working volume and control modules for pH, temperature, and dissolved oxygen. A photo of the bioreactor apparatus is shown in Figure 5.1

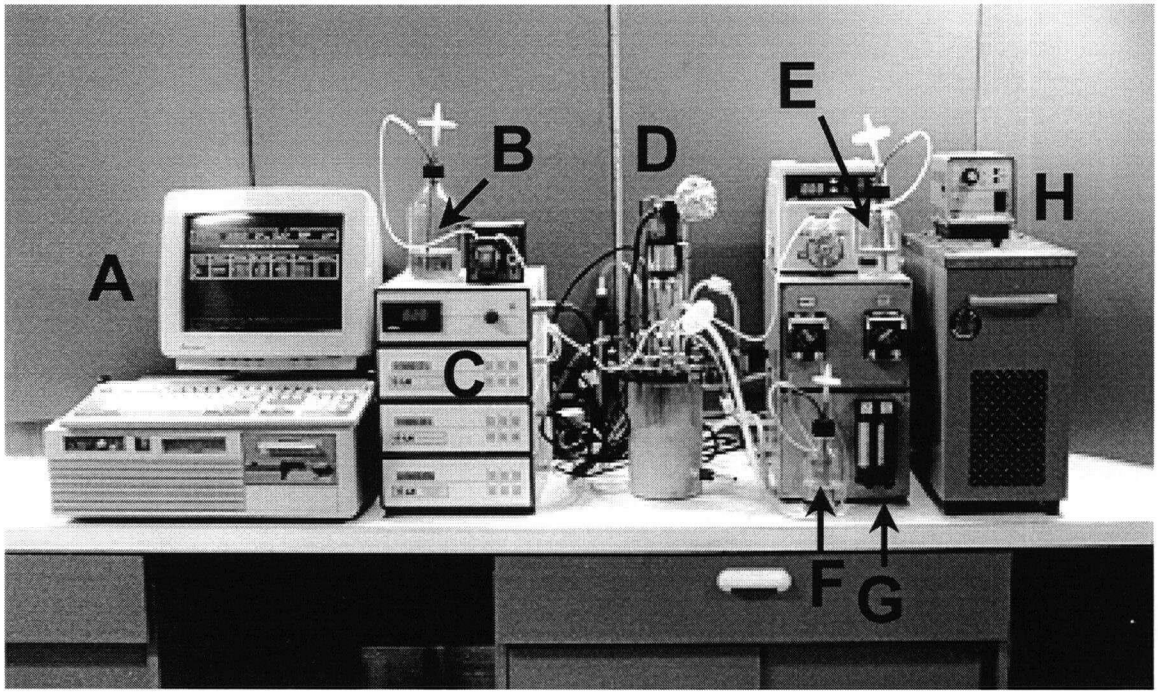


Figure 5.1. Bioreactor apparatus. From left to right and top to bottom are **(A)** the data logging computer, **(B)** the methanol feeding reservoir and pump, **(C)** the control modules, **(D)** the Inceltech 2-liter bioreactor, **(E)** the glycerol feeding reservoir and pump, **(F)** the ammonium hydroxide reservoir and pump, **(G)** the air/oxygen flow meters, and **(H)** the cooling water tank.

The inoculum was grown for 24 hours at 30°C in baffled shake flasks (250 rpm) in a complex medium containing 10 $g \cdot l^{-1}$ yeast extract, 20 $g \cdot l^{-1}$ peptone, and 20 $g \cdot l^{-1}$ glycerol. 100 ml of inoculum was added to 1.0 liter of Basal Salts medium (26.7 $ml \cdot l^{-1}$ 85% o-phosphoric acid, 0.93 $g \cdot l^{-1}$ calcium sulfate$\cdot 2H_2O$, 18.2 $g \cdot l^{-1}$ potassium sulfate, 14.9 $g \cdot l^{-1}$ magnesium sulfate$\cdot 7H_2O$, 4.13 $g \cdot l^{-1}$ potassium hydroxide, 40 $g \cdot l^{-1}$ glycerol) and 4.35 $ml \cdot l^{-1}$ of PTM1 Trace Salts (6.0 $g \cdot l^{-1}$ cupric sulfate$\cdot 5H_2O$, 0.08 $g \cdot l^{-1}$ sodium iodide, 3.0 $g \cdot l^{-1}$ manganese sulfate$\cdot H_2O$, 0.2 $g \cdot l^{-1}$ sodium molybdate$\cdot 2H_2O$, 0.02 $g \cdot l^{-1}$ boric acid, 0.5 $g \cdot l^{-1}$ cobalt chloride, 20.0 $g \cdot l^{-1}$ zinc chloride, 65.0 $g \cdot l^{-1}$ ferrous sulfate$\cdot 7H_2O$, 5.0 ml sulfuric acid, and 0.2 $g \cdot l^{-1}$ d-biotin). Prior to inoculation, 0.1 $g \cdot l^{-1}$ of Antifoam 289 (Sigma, St. Louis, MO) was added and the pH was adjusted to 5.0 by the addition of 30% (w/w) ammonium hydroxide. The pH was measured with a Mettler Toledo pH electrode (405-DPAS-SC). The temperature was maintained at 30°C and the impeller speed was set to 1000 rpm.

Medical grade oxygen was fed to the bioreactor at approximately 0.3 $l \cdot min^{-1}$ to maintain the dissolved oxygen concentration above 30% of air saturation. The flow rate was controlled with a solenoid valve (Valley Instrument Corp., PA) and the oxygen concentration was measured with an Ingold electrode (322 756702/74091).

The bioreactor was operated in batch mode until all of the glycerol was consumed (approximately 22 hours). A glycerol feed (50% (v/v) glycerol + 12 ml PTM1$\cdot l^{-1}$ glycerol) was then maintained for 4 hours at a growth-limiting rate of 15 $ml \cdot l^{-1} \cdot h^{-1}$ (9.5 g $glycerol \cdot l^{-1} \cdot h^{-1}$) (the feed rates were based on the volume in the bioreactor during the batch phase). The pH was adjusted to the desired set point during the glycerol fed-batch phase by automatic feeding of 30% (w/w) ammonium hydroxide. A methanol feed (100% methanol + 12 ml PTM1$\cdot l^{-1}$ methanol) was then maintained for as long as 96 hours at a constant rate of 1.8 $g \cdot l^{-1} \cdot h^{-1}$ (~2.25 $ml \cdot l^{-1} \cdot h^{-1}$). During mixed-feed experiments, the glycerol feed rate was adjusted according to the desired ratio of methanol to glycerol as indicated in Table 5.1.

### *Analytical Procedures*

Cell densities and cystatin concentrations were measured as described in section 4.1.2 of Chapter 4.

### *SDS-PAGE*

The samples from the culture supernatant were concentrated ten fold using a YM-3 membrane in a 50-ml ultrafiltration cell (Amicon Inc., MA) under 30 psi of nitrogen pressure. SDS-polyacrylamide gel electrophoresis was then carried out using a 15% acrylamide separating gel and a 3% stacking gel containing 1% SDS (Laemmli 1970). The gel sheets were stained with 0.025% Coomassie Brilliant Blue R-250.

## 5.1.3 Results

### *Optimization of pH*

Bioreactor cultures of *P. pastoris* were maintained at pH 5 for the initial batch phase of growth. This phase lasted for approximately 22 hours. During the 4-hour glycerol fed-batch phase, the pH was adjusted to the desired value and then maintained

there for the remainder of the experiment. After 24 hours of methanol fed-batch, the yields of cystatin were measured by inhibition of papain activity. As shown in Figure 5.2, the highest yield of biologically active cystatin was 0.18 μmoles·g$^{-1}$ dcw and was observed at pH 6. All subsequent experiments were performed at pH 6.

The dry cell weights 24 hours after induction are also shown in Figure 5.2. It is evident that higher pH values resulted in higher cell yields. The pH was adjusted with 30% (w/w) ammonium hydroxide that provided a nitrogen source for *P. pastoris* cells.

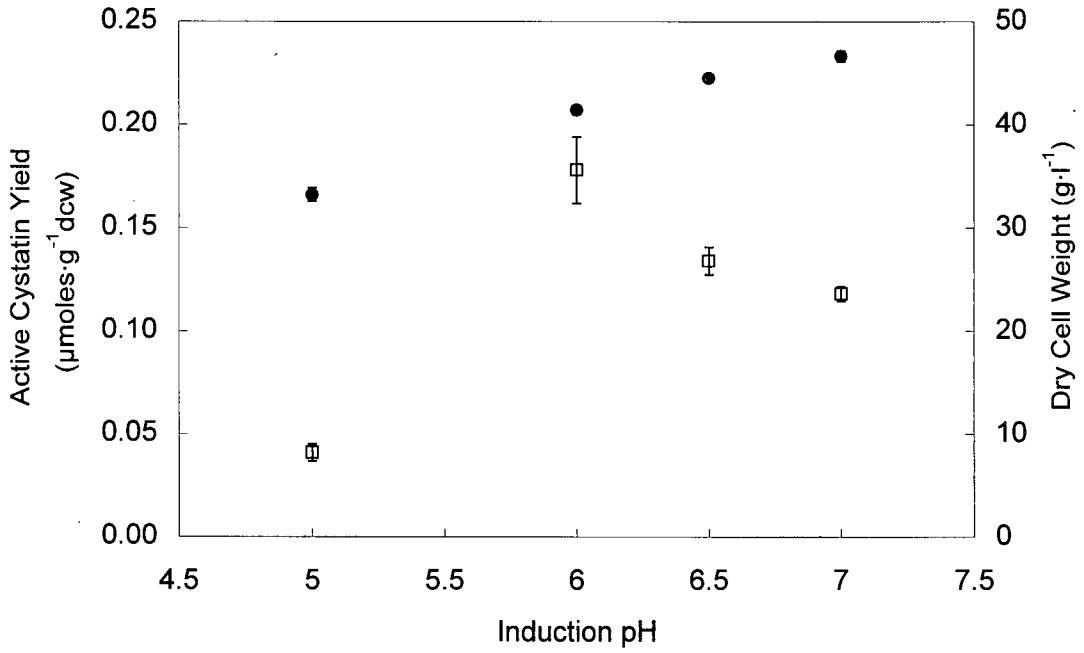


Figure 5.2. Active cystatin yield (□) and dry cell weight (●) as a function of bioreactor pH. Measurements were taken 24 hours after methanol induction. Cell weights were measured in triplicate and cystatin activities were measured in duplicate. The error bars are the standard deviations of the measurements. An analysis of variance indicated a p-value < 0.05 for the effects of pH on active cystatin yield and dry cell weight. In addition, Tukey's test indicated that observations between each pH level were significantly different (data not shown).

## *Methanol Induction*

The control experiment was performed at pH 6 without glycerol feeding during induction. Methanol was fed at 1.8 g·l$^{-1}$·h$^{-1}$ during the induction phase. The dissolved oxygen concentration and the dry cell weight as a function of culture time are shown in Figure 5.3.

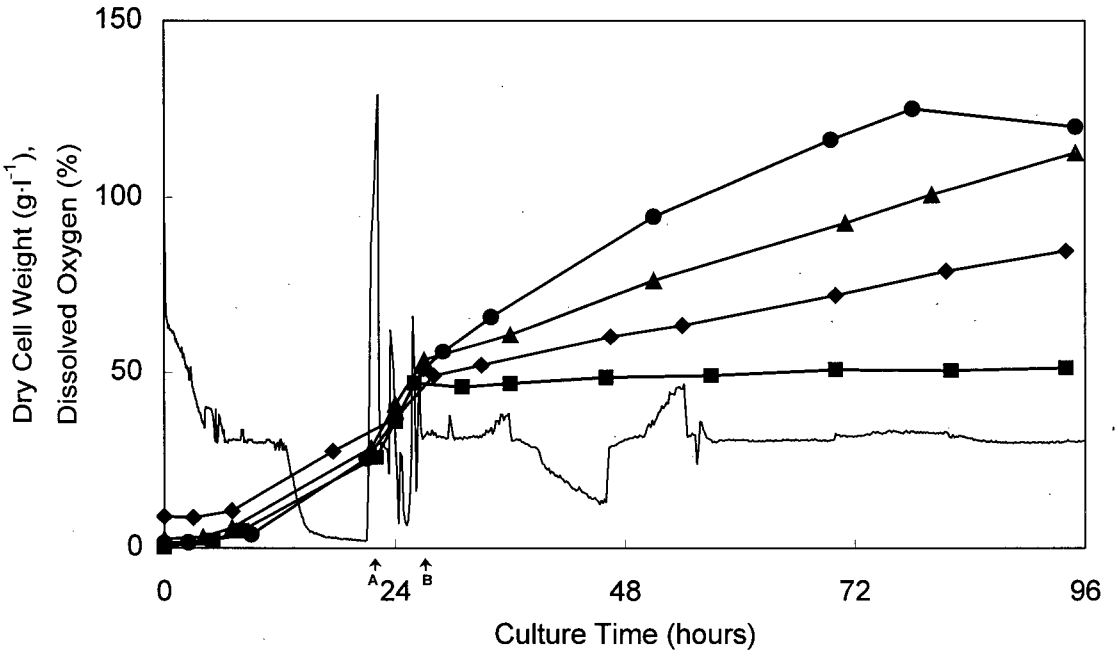


Figure 5.3. Dry cell weight and dissolved oxygen concentration as a function of culture time and glycerol feed rate. Time **A** indicates the initiation of the glycerol fed-batch phase at 9.5 $g \cdot l^{-1} \cdot h^{-1}$. The glycerol feed was adjusted at time **B** to 0 (■), 2.1 (◆), 3.5 (▲), 6.4 (●) $g \cdot l^{-1} \cdot h^{-1}$ for the duration of the experiment. Methanol feeding was initiated at time **B** and was maintained at 1.8 $g \cdot l^{-1} \cdot h^{-1}$. The dissolved oxygen profile is based on data collected for the control experiment that received no glycerol feeding after time **B**.

The total culture time was approximately 5 days. The batch phase was 22 hours, the glycerol fed-batch phase was 4 hours, and the methanol fed-batch phase was maintained up to 96 hours. The dissolved oxygen concentration (DO) served as an indicator to identify the end of the batch phase. As shown in Figure 5.3, there was a sharp increase in dissolved oxygen concentration 22 hours after inoculation. The "DO spike" was used as an indicator of glycerol depletion at the end of the batch phase. The glycerol fed-batch phase was then initiated to further increase the cell density prior to induction with methanol. The cell growth rate dwindled significantly when glycerol feeding was replaced by methanol feeding. As indicated in Table 5.1, the average growth rate during the induction phase was only 0.08 $g \cdot l^{-1} \cdot h^{-1}$ for the culture with no glycerol feed. However, the oxygen consumption rate was considerable and necessitated a feed of 100% oxygen. Two streams of pure oxygen were fed to the bioreactor. A constant feed was maintained at 0.2 $l \cdot min^{-1}$ and a solenoid valve

controlled a supplement feed of 0.1 $l \cdot min^{-1}$. The rates of oxygen feeding were increased manually if the dissolved oxygen concentration could not be maintained at 30% of air saturation.

Table 5.1. Cell growth rates and cystatin expression for various feeding conditions.

| MeOH:Gly[a] (w:w)[b] | Glycerol Feed ($g \cdot l^{-1} \cdot h^{-1}$) | Cell Growth Rate[c] ($g\ dcw \cdot l^{-1} \cdot h^{-1}$) | Cystatin Concentration[d] ($\mu moles \cdot l^{-1}$) | Cystatin Yield[d] ($\mu moles \cdot g^{-1}$ dcw) |
|---|---|---|---|---|
| 1:0 | 0 | 0.08 | 54 | 0.99 |
| 1:1.2 | 2.1 | 0.48 | 45 | 0.52 |
| 1:1.9 | 3.5 | 0.81 | 34 | 0.33 |
| 1:3.6 | 6.4 | 1.38 | 15 | 0.16 |

[a]MeOH = methanol, Gly = glycerol
[b](w:w) = mass ratio of methanol to glycerol during mixed feeding
[c]average growth rate during the first 48 hours of induction
[d]maximum cystatin concentration and yield taken from Figure 5.4

The cystatin concentrations are plotted as a function of culture time in Figure 5.4(a). A linear increase in cystatin concentration was observed for the control. The concentration of cystatin after 4 days of methanol feeding was 54 $\mu moles \cdot l^{-1}$. Presumably, the yield would have been higher if the experiment were to have continued beyond 4 days.

The yield of cystatin as a function of time is shown in Figure 5.4(b). The yield was calculated by dividing the molar concentration of active cystatin by the dry cell weight corresponding to the sampling time. The control experiment achieved a yield of cystatin equal to 1.0 $\mu mole \cdot g^{-1}$ dcw.

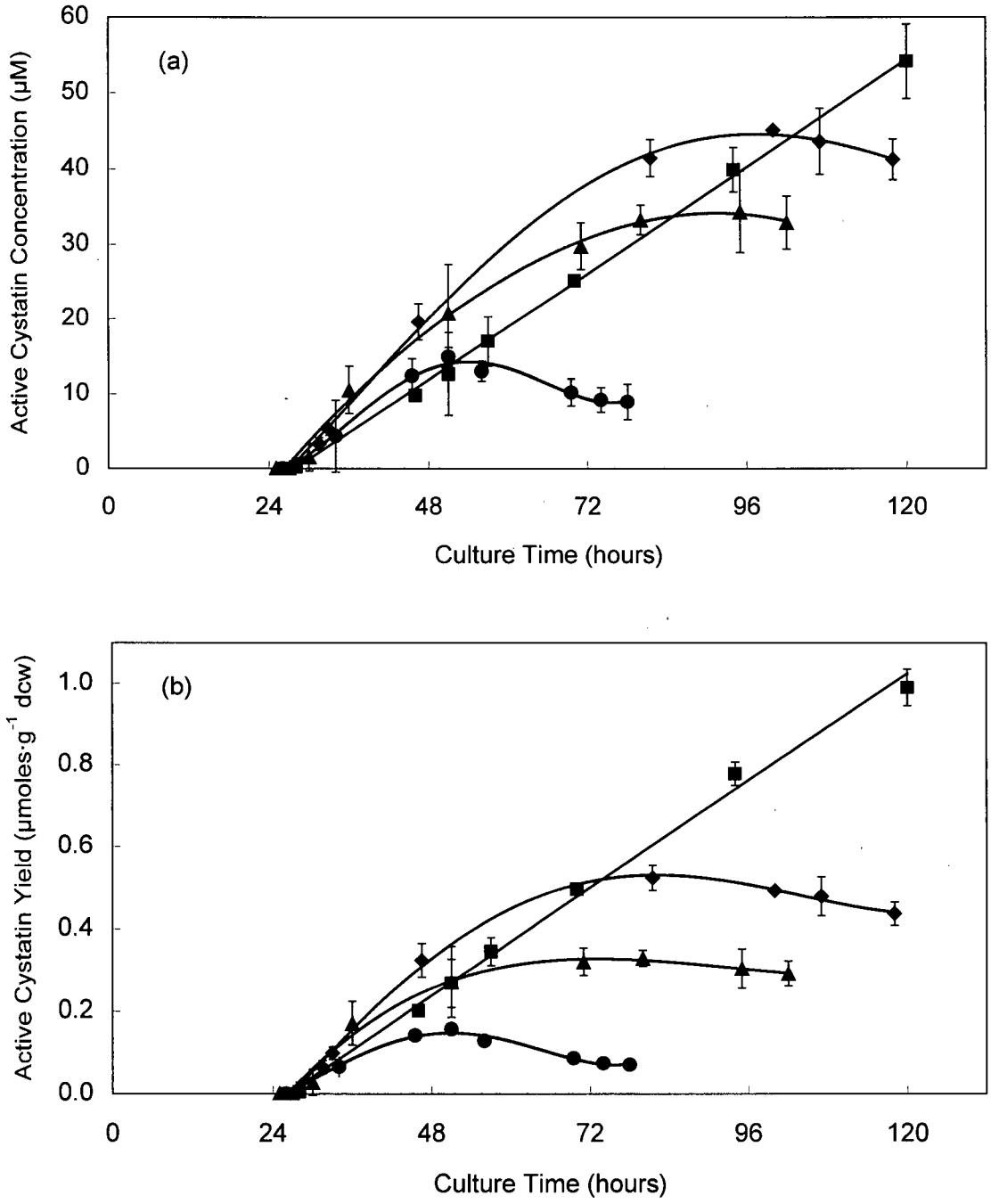


Figure 5.4. Active cystatin concentration **(a)** and active cystatin yield **(b)** as a function of culture time and glycerol feed rate. After 24 hours, glycerol feeding was adjusted to 0 (■), 2.1 (◆), 3.5 (▲), 6.4 (●) $g \cdot l^{-1} \cdot h^{-1}$ and methanol feeding was initiated and then maintained at 1.8 $g \cdot l^{-1} \cdot h^{-1}$. The assay measurements for cystatin were performed in duplicate. The error bars are the standard deviations of the measurements.

## *Mixed Feeding Induction*

In an attempt to reduce the induction time by increasing the cell growth rate and cellular productivity, glycerol was fed to the bioreactor during the methanol fed-batch phase. The methanol feed rate was kept constant at 1.8 $ml \cdot l^{-1} \cdot h^{-1}$ and various feed rates of glycerol were investigated. Glycerol was fed at rates of 2.1, 3.5, and 6.4 $g \cdot l^{-1} \cdot h^{-1}$, corresponding to mass ratios of methanol to glycerol of 1:1.2, 1:1.9, and 1:3.6 respectively. As seen in Figure 5.3, the cell growth rates increased with increasing glycerol feed rates. During the mixed-feed phase, the cell growth rates were steady and were linearly proportional to the feed rate of glycerol. The average growth rates during induction are listed in Table 5.1. The exception to the linear trend was observed for the highest glycerol feed rate. In this case, the cell growth rate decreased over time and the cell density reached a maximum of 125 g $dcw \cdot l^{-1}$ after 2 days of mixed feeding. During induction, the average cell yield for all glycerol feed rates was 0.2 g $dcw \cdot g^{-1}$ glycerol. The cell yield from methanol was only 0.045 g $dcw \cdot g^{-1}$ methanol.

As shown in Figure 5.4(a), glycerol feeding had the effect of greater cystatin expression during the first 24 hours of induction. However, after 24 hours it was observed that higher rates of glycerol feeding reduced the rate of accumulation of cystatin. The maximum cystatin concentrations for each of the mixed-feed experiments are listed in Table 5.1.

The rate of cystatin yield began to decline after 24 hours of mixed feeding (see Figure 5.4(b)). For higher rates of glycerol feeding, the decline was more prominent, and gave rise to a lower maximum yield of cystatin (see Table 5.1). Compared to the linear increase in yield for the control experiment, glycerol feeding gave rise to a decreasing yield of cystatin over time.

## *SDS-PAGE*

SDS-PAGE was used to examine a sample from the bioreactor that was taken 48 hours after induction with methanol as the sole carbon source. Three cystatin bands labeled A, B, and C are indicated in Figure 5.5. The presence of cystatin variants within these bands was verified by western blot analysis using rabbit antiserum that was raised against human cystatin C (DAKO Corp., CA) (data not shown). The most intense

protein band was observed at an apparent molecular weight of 16 kDa (labeled C in Figure 5.5). The estimated molecular weight of this band is consistent with that reported for nonglycosylated cystatin. Despite human cystatin C's calculated molecular weight of 13.4 kDa, electrophoresis systems show a band between 15 and 16 kDa (Abrahamson 1994a). Two less intense and diffuse bands (labeled A and B in Figure 5.5) likely indicate the expression of glycosylated cystatin. The glycosylated cystatin had molecular weights that were variable, but corresponded to apparent molecular weights of approximately 20 and 30 kDa. Compared to nonglycosylated cystatin, the increases in apparent molecular weights were 4 and 14 kDa respectively.

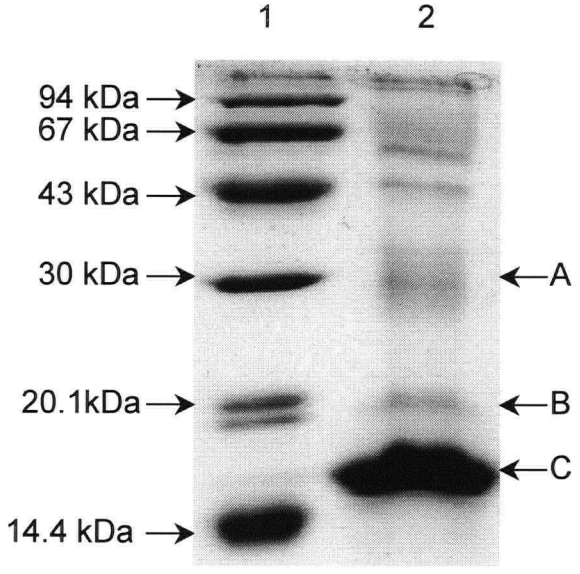


Figure 5.5. SDS-polyacrylamide gel electrophoresis of cystatin C. SDS-PAGE was carried out as indicated in the section 5.1.2. The sample was taken 48 hours after induction with methanol as the sole carbon source. The molecular weight markers are shown in lane 1. The culture supernatant is shown in lane 2. Points **A** and **B** likely indicate bands of glycosylated cystatin with molecular weights of approximately 30 kDa and 20 kDa respectively. Point **C** indicates nonglycosylated cystatin at 16 kDa.

## 5.1.4 Discussion

A biologically active variant of human cystatin C has been expressed at an extracellular concentration of 54 $\mu$moles$\cdot l^{-1}$ by a genetically modified strain of *Pichia pastoris*. This expression is equivalent to 0.72 g$\cdot l^{-1}$ of human cystatin C (13.4 kDa). According to the SDS-PAGE analysis, it is likely that both glycosylated and nonglycosylated cystatin were expressed. However, the intensities of the protein bands

indicate that the majority of the expressed cystatin was nonglycosylated. Future studies in Food, Nutrition, and Health at UBC will investigate the expression of glycosylated cystatin and characterize its oligosaccharides. Assuming that points A and B in Figure 5.5 represent glycosylated cystatin, the average increases in apparent molecular weights due to glycosylation were either 4 or 14 kDa. Since two consensus sequences for glycosylation were introduced into the human cystatin C gene, it is likely that the increases in molecular weights corresponded to glycosylation at a single site and at both sites.

Optimizing the culture pH was the first step to achieving high-level expression of biologically active cystatin. Inadequate pH control was a limiting factor for cystatin expression in shake flask cultures of *P. pastoris*. As a consequence of high cell density cultures, acidic byproducts often exceeded the buffering capacity of the 100 mM potassium phosphate (data not shown). A bench-scale bioreactor allowed for monitoring and maintaining the pH at a desired set point.

The expression of cystatin was investigated within the recommended range of pH 3 – 7 for *P. pastoris* cultures (Wegner 1983). The highest expression of biologically active cystatin was observed at pH 6. Expression at pH 5 resulted in low cystatin activity. It has been shown that below pH 5, human cystatin C forms an inactive dimer that is not broken in the presence of papain (Ekiel and Abrahamson 1996). Therefore, it is likely that the inhibitory activity of cystatin expressed at pH 5 was limited due to dimerization. The reduced yield of active cystatin above pH 6 may have been due to degradation of cystatin by protease enzymes. It has been shown that the activity of protease enzymes is significantly greater for pH values above 6 (Brierly et al. 1994). In the case of mouse cystatin C, glycosylation conferred additional stability to proteolytic degradation and thermal inactivation compared to the nonglycosylated protein (Nakamura et al. 1998a). However, in this work, the potential for the variant of cystatin C to confer additional stability was not realized because the majority of the expressed cystatin was nonglycosylated.

The addition of ammonium hydroxide served a dual function of increasing the pH and supplying a nitrogen source for cell growth and protein expression. It was observed that cell yields were increased when larger volumes of ammonium hydroxide were

required to adjust the pH. Since the pH range used in this study has been shown to have little effect on the growth rate (Inan et al. 1999; Wegner 1983), it is likely that the increased cell yields for higher pH values can be attributed to a greater supply of nitrogen from ammonium hydroxide. This finding is evidence that a nitrogen-limitation may exist in *P. pastoris* cultures that depend on ammonium hydroxide as the nitrogen source. Such cultures, especially those operated at low pH values, may benefit from a supplemental nitrogen source such as ammonium sulfate or ammonium phosphate. In fact, the addition of ammonium salts has been shown to increase cell density as well as inhibit protease enzymes (Tsujikawa et al. 1996).

After determining a stable pH value for cystatin expression, an attempt was made to increase the cystatin production. The production of recombinant proteins is dependent upon the rate of expression by the cells and the overall cell density in the bioreactor (Romanos et al. 1992). Relative to growth on methanol as the sole carbon source, we have observed significantly higher growth rates and cell densities when glycerol was used as a carbon source. Therefore, the effect of glycerol feeding during induction was investigated. Mass ratios of methanol to glycerol of 1:0, 1:1.2, 1:1.9, and 1:3.6 were compared for cystatin expression.

The highest expression was achieved without glycerol feeding during the induction phase. However, during the early stages of induction, glycerol feeding did show potential to significantly increase cystatin production. This was especially true for the lowest feed rates of glycerol. The highest productivity during the first 48 hours of induction was achieved for a 1:1.2 mass ratio (1.8 g methanol$\cdot l^{-1}\cdot h^{-1}$ and 2.1 g glycerol$\cdot l^{-1}\cdot h^{-1}$). If the initial rate of cystatin yield could be maintained, the increased productivity of cystatin could significantly reduce the induction time for cystatin expression by *P. pastoris*.

Brierly et al. (1990) observed a similar trend when feeding methanol and glycerol during induction of bovine lysozyme. Mass ratios of methanol to glycerol of 1:2 and 1:4 were compared for cystatin expression. While both feed ratios gave higher volumetric productivities of lysozyme, only the lower feed rate of glycerol resulted in a higher final concentration compared methanol feeding as the sole carbon source for induction. The highest productivity and concentration of lysozyme was achieved for non-limiting

methanol feeding that was increased step-wise from 2 to 5 $g \cdot l^{-1} \cdot h^{-1}$ and a constant glycerol feed of 1.8 $g \cdot l^{-1} \cdot h^{-1}$. These optimum feed rates were very close to those of this work that gave maximum cystatin expression during the first 48 hours of induction (1.8 g methanol$\cdot l^{-1} \cdot h^{-1}$ and 2.1 g glycerol$\cdot l^{-1} \cdot h^{-1}$). However, the methanol feed rate used by Brierly et al. (1990) was increased step-wise to a rate that was 2.8 times greater than that which was used for cystatin expression. Therefore, a methanol limitation may have occurred during the expression of cystatin. A constant methanol feed of 1.8 $g \cdot l^{-1} \cdot h^{-1}$ may not have been adequate to maintain induction of the *AOX1* promoter, especially under high cell density conditions.

To investigate the possibility of a methanol limitation, an experiment was performed in which the methanol feed rate was doubled during the entire induction period from 1.8 $g \cdot l^{-1} \cdot h^{-1}$ to 3.6 $g \cdot l^{-1} \cdot h^{-1}$. As shown in Figure 5.6, the result was a higher initial productivity of cystatin. However, after just 24 hours of induction, both the cell growth rate and productivity decreased. This observation was attributed to methanol toxicity. As recommended by Brierly et al. (1990), a step-wise increase in methanol feed rate would have been more appropriate. A step-wise procedure balances the accumulation of cells with the optimal expression from the *AOX1* promoter. They too observed a rapid decrease in lysozyme productivity when the methanol feed rate was increased too rapidly.

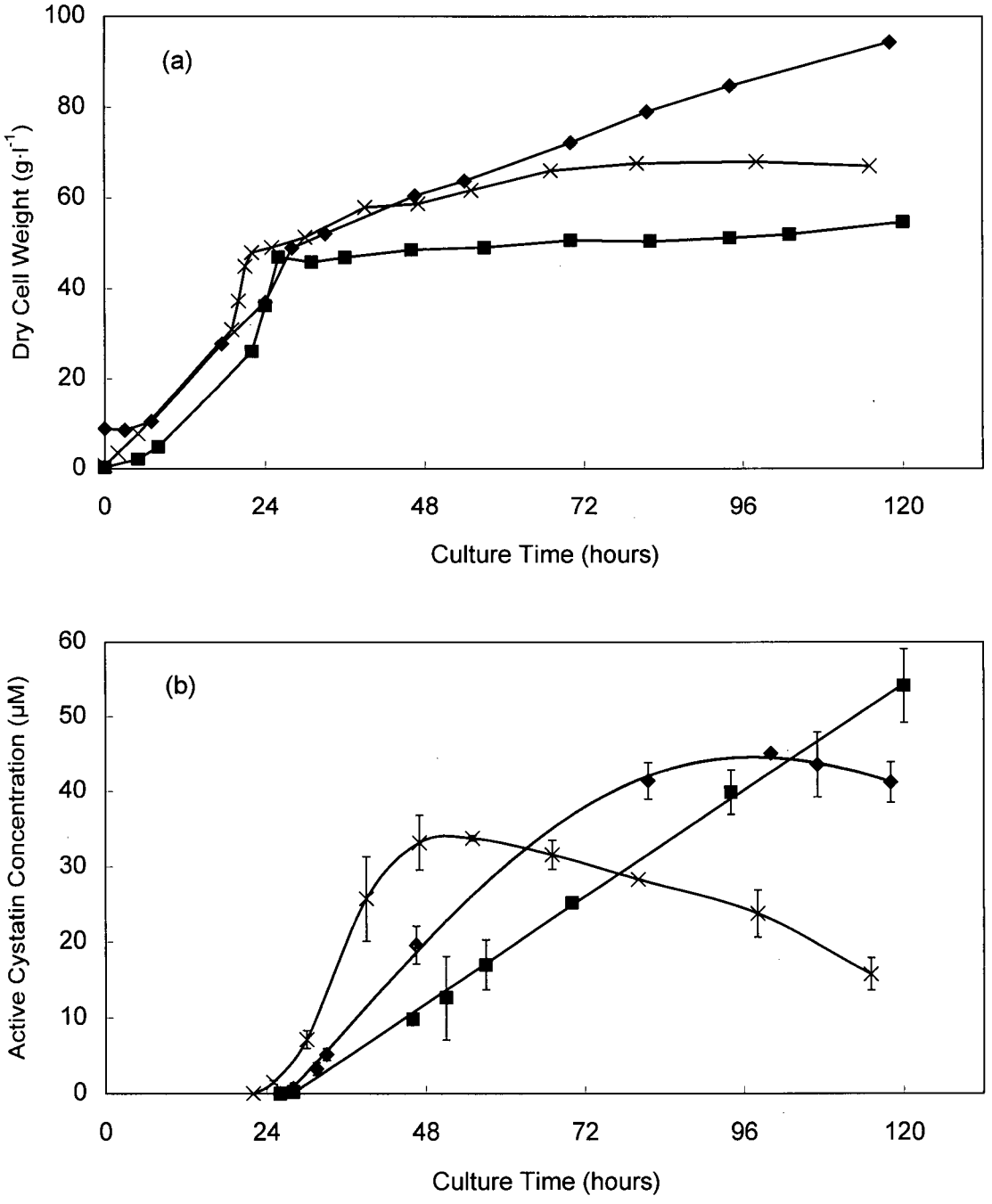


Figure 5.6. Dry cell weight **(a)** and active cystatin concentration **(b)** as a function of culture time, glycerol feed rate, and methanol feed rate. The glycerol feed rates during induction were 0 (■), 2.1 (◆), and 2.1 (×) $g \cdot l^{-1} \cdot h^{-1}$. The methanol feed rates during induction were 1.8 $g \cdot l^{-1} \cdot h^{-1}$ (■), 1.8 $g \cdot l^{-1} \cdot h^{-1}$ (◆), and 3.6 $g \cdot l^{-1} \cdot h^{-1}$ (×). The assay measurements for cystatin were performed in duplicate. The error bars are the standard deviations of the measurements.

As indicated in Figure 5.6, an excessive feed rate of methanol decreased cystatin expression. Therefore, the mixed-feed experiments that received a methanol feed of 1.8 $g \cdot l^{-1} \cdot h^{-1}$ may have experienced methanol toxicity that resulted in low final concentrations of cystatin. Preliminary experiments with the $Mut^s$ strain of *P. pastoris* revealed that in the absence of glycerol feeding, a methanol feed above the recommended rate of 2.4 $g \cdot l^{-1} \cdot h^{-1}$ (~3 $ml \cdot l^{-1} \cdot h^{-1}$) (Invitrogen 1999), resulted in a decline in cell density, cystatin expression, and oxygen consumption rate (data not shown). During mixed feeding however, no decline in oxygen consumption rates were observed, and with the exception of the highest feed rate of glycerol, there were no decreases in cell density. In addition, the steady growth rates and their linear proportionality to glycerol feed rates indicated that methanol did not likely accumulate to toxic concentrations in the bioreactor.

In addition to a methanol limitation, the repression of the *AOX1* promoter may explain the reduced cystatin yield at higher glycerol feed rates. As previously discussed, the *AOX1* promoter is inducible by methanol and is repressed by glucose or glycerol. In the case of glycerol, concentrations above 2% (w/v) repress the *AOX1* promoter (Gellissen et al. 1995). Glycerol accumulation likely occurred for high glycerol feed rates and towards the end of the culture time for low glycerol feed rates. The highest glycerol feed rate of 6.4 $g \cdot l^{-1} \cdot h^{-1}$ achieved a very high cell density, but the growth rate began to decrease after 24 hours of mixed feeding. Since the glycerol feeding was kept constant, it can be assumed that glycerol accumulation ensued when the growth rate diminished. It is not too surprising that the decline in cystatin yield began at the same time as the glycerol accumulation ensued. By this same argument, it is likely that the delayed decline in cystatin yield for lower feed rates of glycerol was attributable to a longer time required for glycerol accumulation and subsequent repression of the *AOX1* promoter.

It is also possible that in addition to a glycerol accumulation, ethanol in the bioreactor may have contributed to the repression of the *AOX1* promoter. Ethanol is a metabolic byproduct and can repress the *AOX1* promoter at concentrations greater than 100 $mg \cdot l^{-1}$. Glycerol feed rates greater than 1.8 $g \cdot l^{-1} \cdot h^{-1}$ have been observed to result in the accumulation of more than 100 $mg \cdot l^{-1}$ of ethanol (Brierly et al. 1990; Inan et al.

1999). All of the glycerol feed rates during cystatin expression were above 1.8 $g \cdot l^{-1} \cdot h^{-1}$. However, ethanol repression can only be speculated in this work because the ethanol concentration in the bioreactor was not monitored.

### 5.1.5 Conclusions and Recommendations

The results of these experiments have indicated that 1) glycerol feeding showed potential to increase both cell density and yield of cystatin, but 2) repression of the *AOX1* gene promoter due to glycerol accumulation and/or a methanol limitation in the bioreactor may have impeded cystatin production. The maximum concentration of cystatin in the culture supernatant was 54 $\mu moles \cdot l^{-1}$, equivalent to 0.72 $g \cdot l^{-1}$ of human cystatin C. The maximum yield of cystatin was 1.0 $\mu mole \cdot g^{-1}$ dry cell weight.

The optimal pH for cystatin activity was investigated. These experiments were the first that were performed in the bioreactor. As a result of using these experiments as a means of learning how to operate the bioreactor, the operational parameters such as methanol feed rate and dissolved oxygen concentration were not optimal. In addition, it was during these first experiments that the method of sample storage was found to be unsatisfactory. All of these factors may have interfered with the interpretation of the effect of pH on cystatin expression.

As with the growth medium experiments, the investigation into the substrate feeding strategy began without adequate knowledge of similar efforts made by other researchers. The experiments could have been significantly streamlined by following the recommendations of Brierly et al. (1990). For example, higher methanol feed rates and lower glycerol feed rates should have been used. Instead of using a constant methanol feed rate of 1.8 $g \cdot l^{-1} \cdot h^{-1}$, Brierly et al. (1990) increased the feed rate in a step-wise manner from 2 $g \cdot l^{-1} \cdot h^{-1}$ to a maximum of 5 $g \cdot l^{-1} \cdot h^{-1}$. In future experiments, these feed rates should be investigated, with the rate of increase in methanol feeding from 2 to 5 $g \cdot l^{-1} \cdot h^{-1}$ being proportional to the observed increase in cell density. However, if an on-line method for measuring methanol is available, such as that reported by Guarna et al. (1997), the methanol concentration should be maintained between 1 and 2 $g \cdot l^{-1}$ as recommended by Thorpe et al. (1999). In terms of glycerol feeding, the experiments reported by Brierly et al. (1990) and the experiments presented here observed the

greatest expression of recombinant protein for a glycerol feed rate of approximately 2 $g \cdot l^{-1} \cdot h^{-1}$. Higher glycerol feed rates may result in increases in ethanol production and proteolysis as well as accumulation of glycerol that can repress the *AOX1* gene promoter. At a feed rate of 2 $g \cdot l^{-1} \cdot h^{-1}$, the accumulation of glycerol is unlikely (Brierly et al. 1990; Loewen et al. 1997), but it should be monitored to ensure that its concentration is as low as possible.

## 5.2 Comparison of Methanol Utilization Phenotype

### 5.2.1 Introduction

An ongoing debate within the *P. pastoris* research community involves the desired methanol utilization capabilities of the strain selected for recombinant protein expression. As a methylotrophic yeast, *P. pastoris* has the ability to grow on methanol as its sole carbon source. *P. pastoris* expresses the alcohol oxidase 1 enzyme (AOX1) that allows for the first step of the oxidation of methanol (Veenhuis et al. 1983). This native strain of *P. pastoris* is referred to as having a $Mut^+$ phenotype (methanol utilization plus). The expression of AOX1 is strongly regulated by the *AOX1* gene promoter, which can produce AOX1 in amounts up to 30% of the total intracellular soluble protein (Cregg et al. 1993). The excessive production of AOX1 compensates for its poor affinity for oxygen (Invitrogen 1998). Transformation by gene insertion leaves the *AOX1* gene intact and the $Mut^+$ phenotype is retained.

If the vector containing the recombinant gene is integrated into the *AOX1* gene locus, the transformant cannot synthesize AOX1. This method of recombinant gene integration is known as transplacement. However, methanol utilization is not completely eliminated because *P. pastoris* produces a similar enzyme to AOX1 called alcohol oxidase 2 (AOX2). AOX2 is 97% identical to AOX1 and has the same specific activity (Cregg et al. 1989). Since growth on methanol is much slower when only AOX2 is available for metabolism, the resulting phenotype is $Mut^s$ (methanol utilization slow).

Successful expression of recombinant proteins has been achieved with both $Mut^+$ (Digan et al. 1989; Eldin et al. 1997; Guarna et al. 1997; Henry et al. 1997; Katakura et al. 1998; Kim et al. 1997) and $Mut^s$ (Chen et al. 1996; Clare et al. 1991a; Cregg et al. 1987; Döring et al. 1997; Payne et al. 1995; Waterham et al. 1997) strains of *P. pastoris*.

Some direct comparisons between $Mut^+$ and $Mut^s$ phenotypes have proven useful to clarify the advantages and disadvantages of each phenotype. For the secretion of invertase, the $Mut^s$, KM71 strain of *P. pastoris* expressed higher concentrations of protein than the $Mut^+$, GS115 strain (Tschopp et al. 1987a). A similar

observation was made for the expression of the *lacZ* gene (Tschopp et al. 1987b). However, despite its higher protein concentration, expression with $Mut^s$ required long induction times of greater than 100 hours to reach maximum cell density and protein concentration. In fact, the cell concentration had only reached its maximum (40 g dcw·$l^{-1}$) after 256 hours of induction. Subsequently, a $Mut^+$ strain was shown to be more productive due to its shorter induction time compared to $Mut^s$. Only 25 hours of induction were required to achieve a productivity of 13 ± 2 mg·$l^{-1}$·$h^{-1}$ of lysozyme c2 (Digan et al. 1989). This productivity was six times higher than that which was achieved using a $Mut^s$ strain (Brierly et al. 1990). The high productivity was attributed to the increased growth rate and cell concentration of the $Mut^+$ over the $Mut^s$ strain.

A broader understanding of the different strains of *P. pastoris* and their expression potential was achieved through a series of experiments that compared the $Mut^+$ and $Mut^s$ stains, the location of gene integration (*AOX1* vs *HIS4*), the gene copy number, and the method of integration of the recombinant gene (insertion vs transplacement) for the production of tetanus toxin fragment C (Clare et al. 1991a). The expression of the heterologous protein was relatively independent of the integration site, the type of integrant, and the methanol utilization phenotype of the cells. The most significant factor for protein production was the gene copy number. By repeatedly performing the single crossover recombination, they were able to obtain multiple integrated copies of the expression cassette. The result was a $Mut^s$ strain with 14 gene copies that expressed fragment C at 12 g·$l^{-1}$ (27% of the total protein in the medium). At these high copy numbers, they found that the $Mut^s$ strain was 20% more productive than the $Mut^+$ strain.

Contrary to the findings that $Mut^+$ have faster cycle times for lysozyme c2 (Digan et al. 1989), β-lactoglobulin was detected one day after induction for a $Mut^s$ strain whereas it required two days for expression in a $Mut^+$ strain (Kim et al. 1997).

More recently, better control of the methanol concentration has been achieved to allow for better expression for both $Mut^+$ and $Mut^s$ strains. The use of an on-line methanol sensor has been shown to increase cell densities compared to daily feedings of methanol (Guarna et al. 1997). A $Mut^+$ strain was selected and the methanol concentration was maintained at 0.3% (v/v). The final dry cell density was 130 g dcw·$l^{-1}$,

and despite the low methanol concentration, the cell density had doubled during the 3-day induction period.

Because the Mut$^{S}$ strain consumes methanol more slowly than the Mut$^{+}$ strain, the methanol feed rate is lower and bears less risk of reaching toxic levels. However, due to its slow growth, dissolved oxygen (DO) spikes are slow and do not allow for a convenient means of indicating methanol depletion. Mut$^{+}$ phenotypes on the other hand will display a DO spike when methanol is depleted. In the absence of on-line measurement of methanol, DO spikes can be used to ensure that the methanol does not reach toxic levels (Cregg et al. 1993).

The mode of integration might influence the level of foreign gene expression. Mut$^{S}$ strains, where the *AOX1* gene is disrupted by transplacement, do not simultaneously produce high levels of alcohol oxidase and the recombinant protein. For pertactin production in *P. pastoris*, a Mut$^{S}$ strain expressed the recombinant protein at 10% of total soluble cell protein while the Mut$^{+}$ strain reached only a 5% level (Gellissen et al. 1995). Therefore, a higher percentage of total soluble cell protein in Mut$^{S}$ may aid in the purification process. Other advantages of the Mut$^{S}$ and some disadvantages are presented in Table 5.2.

Table 5.2. A summary of some of the factors to consider when selecting the methanol utilization phenotype for the *P. pastoris* expression host. Despite the many differences indicated, the actual expression of the foreign protein might or might not be effected by the phenotype.

| **Criterion** | **Mut$^+$** | **Mut$^s$** |
|---|---|---|
| Methanol utilization | Yes | Slow |
| Methanol concentration | Less than 1% (but as low as possible) | Less than 1% |
| Induction time | 25+ hours | 70+ hours |
| Oxygen consumption | Need pure oxygen | Air with oxygen supplement |
| Protein purification | More complicated if intracellular due to high concentration of AOX1 | Purity is high for intracellular and extracellular protein |
| Gene copy number | Up to 10, but low frequency | Up to 30, and higher frequency |
| Scale-up | More difficult due to MeOH control and oxygen requirements | Easier to scale-up but longer cycle times |

Despite that fact that the preference has shifted in favour of the Mut$^s$ phenotype, the final selection should be based on experimental data for the protein of interest. The experiments described in this section compared a Mut$^+$ and a Mut$^s$ transformant for expression of cystatin in the 2-liter Inceltech bioreactor. The objective was to determine which phenotype expressed the higher concentration of active cystatin. All of the data for the Mut$^s$ transformant are taken from experiments that were presented in section 5.1.

## 5.2.2 Materials and Methods

### *Microorganisms*

*Pichia pastoris* X-33 incorporating the pPICZα-A vector was used as a host for extracelluar expression of mutated cystatin C. The Mut$^s$ transformant that was used for experiments presented in section 5.1 was compared to a Mut$^+$ transformant that was isolated from a subsequent electroporation experiment.

Both yeast phenotypes were transformed by electroporation according to the procedure described in section 4.1.2 in Chapter 4. Both transformants harboured the mutated cystatin gene. However, it was not known if the gene copy number was the same for each of the transformants.

### *Bioreactor Operation*

The bioreactor was operated in the same manner as described in section 5.1.2. During the batch phase, the pH was maintained at 5 by feeding 30% (w/w) ammonium hydroxide. The pH was increased to 6 during the glycerol fed-batch phase. However, during the first experiment with the $Mut^+$ strain, the pH was not increased to 6 until a culture time of 72 hours.

After the batch and fed-batch phases on glycerol, the methanol fed-batch phase was initiated for both the $Mut^S$ and $Mut^+$ phenotypes. As the control, the $Mut^S$ transformant received a constant feed rate of 1.8 $g \cdot l^{-1} \cdot h^{-1}$ (~2.25 $ml \cdot l^{-1} \cdot h^{-1}$) of 100% methanol + 12 ml PTM1$\cdot l^{-1}$ methanol. The $Mut^+$ transformant received methanol feeding at growth-limiting rates. The growth-limiting rate was determined by monitoring the dissolved oxygen spikes. The methanol feed was temporarily stopped and the time required for the DO spike was noted. The methanol feed rate was adjusted so that the DO spikes were observed within 1 minute.

### *Analytical Procedures*

Cell densities and cystatin concentrations were measured as described in section 4.1.2 of Chapter 4.

## 5.2.3 Results

The dry cell weights as functions of culture time are shown in Figure 5.7. The $Mut^+$ transformant achieved much higher cell densities than the $Mut^S$ transformant. The results for the $Mut^S$ transformant have been taken from the control experiment in section 5.1.3.

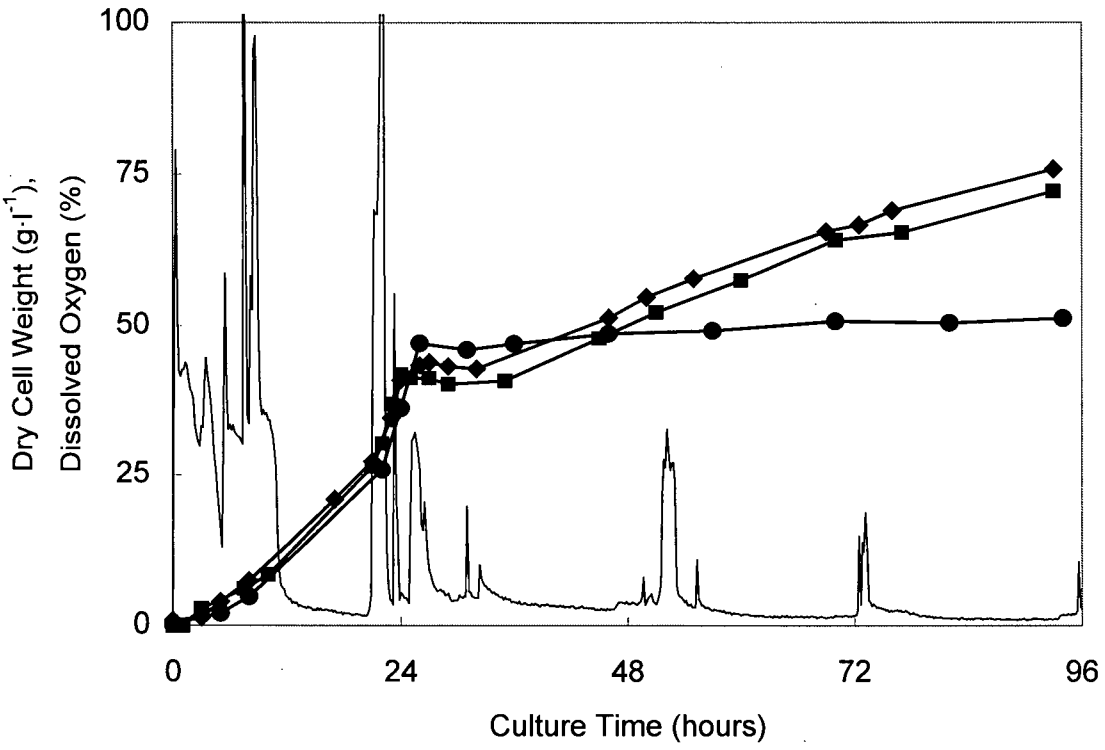


Figure 5.7. Dry cell weights and Mut$^+$ dissolved oxygen profile as functions of culture time. The dry cell weight data sets are for Mut$^s$ at pH 6 (●), Mut$^+$ at pH 6 (◆), and Mut$^+$ at pH 5 that was changed to pH 6 at 72 hours (■). The dissolved oxygen profile (—) is for Mut$^+$ at pH 6. After the glycerol batch and fed-batch phases, methanol feeding was initiated at approximately 22 hours. Methanol was fed at growth-limiting rates for Mut$^+$ (maximum rate of 6.2 g·l$^{-1}$·h$^{-1}$) and at a constant rate of 1.8 g·l$^{-1}$·h$^{-1}$ for Mut$^s$. The feed rate for Mut$^+$ was adjusted depending on the required time for a dissolved oxygen spike. Cell weights were measured in triplicate and the standard deviations are not visible due to their small magnitude.

The two growth curves for the Mut$^+$ phenotype are very similar. However, when the pH was maintained at 5 for the first 72 hours, the cell densities were lower. This supports the observations in section 5.1.3 that indicated higher cell densities for higher pH values. This observation was attributed to an increase in the available nitrogen from ammonium hydroxide. The amount of ammonium hydroxide that was fed during each experiments was not recorded.

The average growth rate for the Mut$^+$ transformant was 0.54 g·l$^{-1}$·h$^{-1}$ compared to only 0.08 g·l$^{-1}$·h$^{-1}$ for the Mut$^s$ transformant. In addition, the yield of cells from methanol was nearly two times greater for Mut$^+$ compared to Mut$^s$ (0.088 g dcw·g$^{-1}$ methanol

compared to 0.045 g dcw·$g^{-1}$ methanol). This suggests that methanol metabolism may be more efficient for the $Mut^+$ transformant.

The oxygen consumption rate was considerably higher for the $Mut^+$ transformant. As shown in Figure 5.7, the oxygen concentration could not be maintained at the desired 30% of air saturation. In fact, an oxygen mass transfer limitation in the bioreactor was observed. Even at very high oxygen feed rates, only an increase in impeller speed was able to increase the dissolved oxygen concentration (the maximum impeller speed was 1000 rpm). For the $Mut^s$ strain, oxygen feeding did not have to be increased beyond 0.3 l·$min^{-1}$. In the case of the $Mut^+$ transformant, oxygen flow rates as high as 1.2 l·$min^{-1}$ were attempted in order to increase the dissolved oxygen concentration (this was the full-scale oxygen flow rate for the apparatus).

The higher oxygen consumption rate of the $Mut^+$ transformant corresponded to a higher methanol consumption rate. The average methanol feed rate for $Mut^+$ was 6.2 g·$l^{-1}$·$h^{-1}$ compared to 1.8 g·$l^{-1}$·$h^{-1}$ for $Mut^s$. However, even 6.2 g·$l^{-1}$·$h^{-1}$ is less than the feed rate of 8.6 g·$l^{-1}$·$h^{-1}$ that is typically recommended for $Mut^+$ transformants (Invitrogen 1999). It is evident that the growth conditions for the $Mut^+$ transformant were not optimal.

The cystatin concentrations are plotted as a function of culture time in Figure 5.8(a). A linear increase in cystatin concentration was observed for the $Mut^s$ transformant. The concentration of cystatin after 4 days of methanol feeding was 54 μmoles·$l^{-1}$. Presumably, the yield would have been higher if the experiment were to have continued beyond 4 days. The maximum concentration of cystatin for the $Mut^+$ transformant was 21 μmoles·$l^{-1}$.

The yield of cystatin as a function of time is shown in Figure 5.8(b). The yield was calculated by dividing the molar concentration of active cystatin by the dry cell weight corresponding to the sampling time. The $Mut^s$ transformant achieved a yield of cystatin equal to 1.0 μmole·$g^{-1}$ dcw. The maximum yield for the $Mut^+$ transformant was 0.3 μmole·$g^{-1}$ dcw.

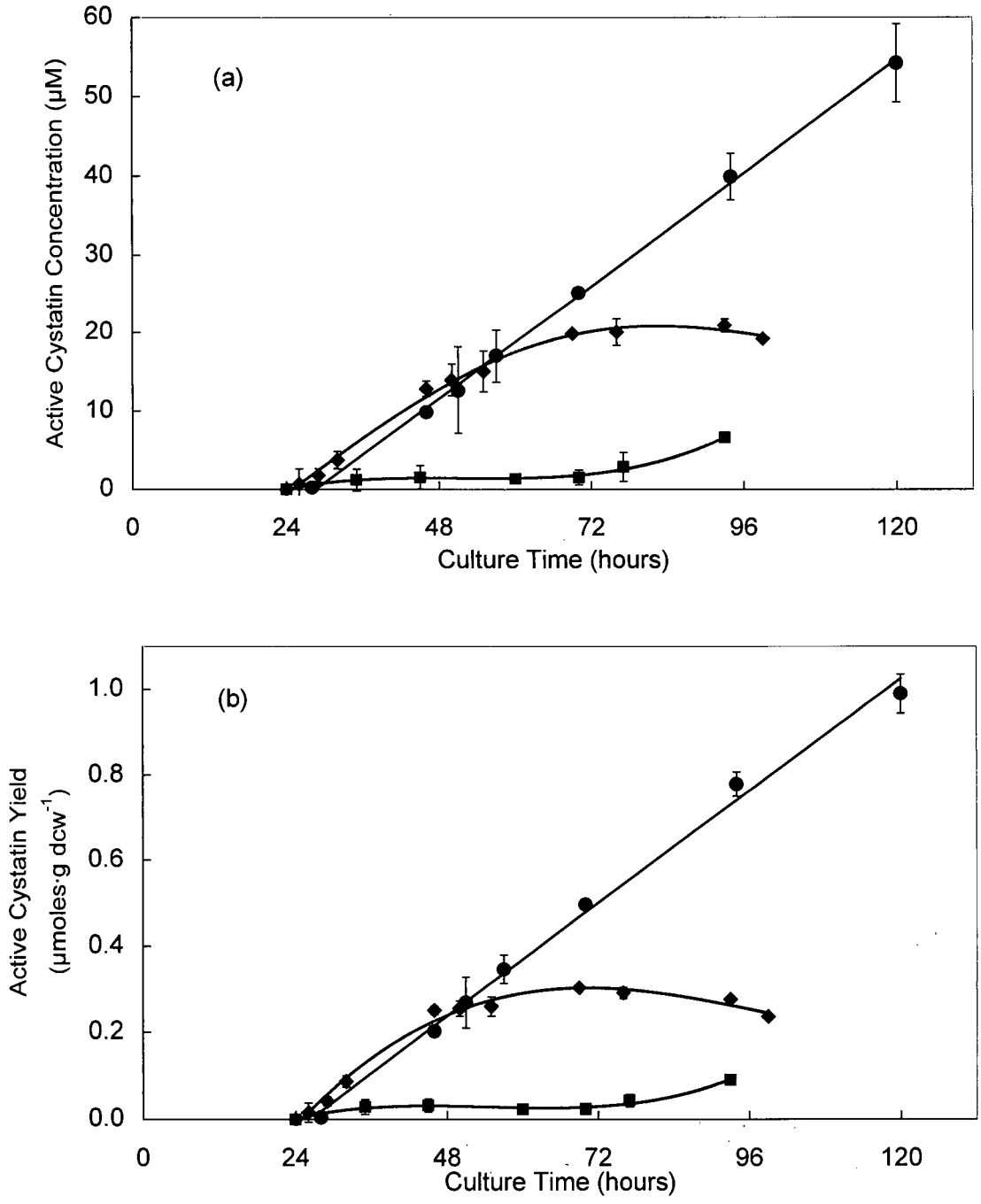


Figure 5.8. Active cystatin concentration **(a)** and active cystatin yield **(b)** as functions of culture time and methanol utilization phenotype. The three data sets are for $Mut^s$ at pH 6 (●), $Mut^+$ at pH 6 (◆), and $Mut^+$ at pH 5 (■). The assay measurements for cystatin were performed in duplicate. The error bars are the standard deviations of the measurement error.

As shown in Figure 5.8, the $Mut^s$ transformant achieved much higher cystatin concentrations and cystatin yields. The $Mut^+$ that was maintained at pH 6 performed equally well for the first 24 hours of induction, but reached a maximum expression level after 48 hours of induction (corresponding to a 72 hour culture time). The $Mut^+$ transformant that was grown at pH 5 until 72 hours and then increased to pH 6 had very low expression levels for active cystatin. However, it is evident that the change from pH 5 to pH 6 did in fact result in an increase in measured cystatin activity.

### 5.2.4 Discussion

The different characteristics of the methanol utilization phenotype were very evident in these experiments. The $Mut^+$ transformant had a higher methanol consumption rate, a faster growth rate, and a higher oxygen consumption rate. The oxygen consumption rate was such that mass transfer limitations were observed in the Inceltech 2-liter bioreactor. As a result, the $Mut^+$ transformant was operated under oxygen-limited conditions for much of the culture time. The oxygen limitation did not seem to affect its steady growth rate, but it may have had a detrimental effect on cystatin expression.

During the first 24 hours, the performance of both phenotypes was comparable. Cystatin concentrations and yields increased fairly linearly for both $Mut^+$ and $Mut^s$ transformants. These results confirm experiments that have indicated that the methanol utilization phenotype did not have a significant effect on expression levels (Digan et al. 1989). However, the trend in the experiments presented here was only maintained for the first 24 hours. After 24 hours of induction, the rate of accumulation of active cystatin began to decrease for the $Mut^+$ transformant. By 48 hours, cystatin had reached its maximum concentration. $Mut^+$ transformants are known for reaching maximum expression levels as soon as 25 hours after induction (Clare et al. 1991a), but these short induction times corresponded to a greater initial expression rate for $Mut^+$ compared to $Mut^s$ strains. In the experiments presented here, the initial expression rates were not significantly different.

The maximum active cystatin concentrations and yields were significantly higher for $Mut^s$ compared to $Mut^+$. Similar results were observed for the maximum

concentrations and yields of invertase (Tschopp et al. 1987a) and the *lacZ* gene (Tschopp et al. 1987b). The difference in expression levels was likely due to oxygen-limited conditions experienced by the $Mut^+$ transformants. As indicated by its lower than expected maximum methanol feed rate (6.2 vs 8.6 $g \cdot l^{-1} \cdot h^{-1}$), the $Mut^+$ transformant did not have the benefit of optimal growth conditions in the Inceltech 2-liter bioreactor.

### 5.2.5 Conclusions and Recommendations

These bioreactor experiments have indicated that 1) the $Mut^s$ phenotype expressed cystatin at higher concentrations and higher yields than the $Mut^+$ phenotype and 2) oxygen mass transfer limitations in the Inceltech 2-liter bioreactor were observed for the $Mut^+$ phenotype.

In order to provide a reliable comparison between the $Mut^s$ and $Mut^+$ phenotype, the bioreactor will have to be modified to increase the mass transfer coefficient for oxygen ($k_La$). The oxygen mass transfer coefficient can be increased with a higher Reynolds number and/or a smaller oxygen bubble size. Using an impeller with a larger diameter will increase the Reynolds number, and replacing the ring-shaped gas sparger with a porous diffuser will reduce the bubble size.

In addition, the gene copy number for the $Mut^s$ and the $Mut^+$ transformants should be verified to be the same. This can be performed using DNA dot blots.

# Chapter 6

# Conclusions and Recommendations

## 6.1 Conclusions

The overall objective of this project was to maximize the expression of a recombinant variant of human cystatin C using the methylotrophic yeast *Pichia pastoris.* This goal was achieved by experiments to investigate 1) the growth medium, 2) the cystatin gene copy number, 3) the pH for expression, 4) the substrate feeding strategy used in the bioreactor, and 5) the *P. pastoris* phenotype.

The growth medium was investigated using shake-flask factorial experiments. These experiments provided an excellent means of becoming familiar with the growth of *P. pastoris.* The most relevant observation based on these experiments was that the BMM medium is histidine-limited for His$^-$ *P. pastoris* transformants. For such auxotrophic strains, 500 mg·l$^{-1}$ of L-histidine should be added to shake-flask cultures. The importance of a nitrogen source was also realized in the shake-flask cultures. These small-scale cultures do not benefit from the addition of ammonium hydroxide as a means for pH control and nitrogen addition. However, the actual optimization efforts of the shake-flask experiments were not appreciated because a different medium formulation was selected for use in the bioreactor.

The effect of gene copy number in the Mut$^S$ transformant was investigated by comparing Zeocin™ resistance to expression of biologically active cystatin in shake-flask cultures. During the first 24 hours of induction, there was evidence of an inverse relationship between the yield of secreted cystatin and the resistance to Zeocin™ exhibited by the transformant. However, after 72 hours of induction, the yields were similar for all of the transformants. Because the experiments did not elucidate a clear relationship between Zeocin™ resistance and cystatin expression, a transformant exhibiting the highest resistance to Zeocin™ was selected for the bioreactor experiments. This decision was based on the majority of the reports indicating that a higher gene copy number increased protein expression in the bioreactor environment.

The optimum pH for cystatin activity was investigated in a 2-liter Inceltech bioreactor. The bioreactor allowed for precise control and monitoring of the pH. A pH of 6 was found to be optimal, with a change in pH of ±1 resulting in significantly lower yields of biologically active cystatin. This was especially true for low pH values in which low activities of cystatin were attributed to self-aggregation.

A substrate feeding strategy was also investigated in the bioreactor. During the induction phase, both methanol and glycerol were fed to the culture. The ratios of methanol to glycerol were similar to those that have been reported to increase the production of recombinant proteins in *P. pastoris* (Brierly et al. 1990). Although the maximum concentrations and yields of cystatin were less for the mixed-feed experiments compared to methanol-fed experiments, the mixed-feed strategy increased cystatin productivity during the first 48 hours of induction. The sensitivity of the *AOX1* gene promoter to methanol and glycerol concentrations became very evident in these experiments. Too much methanol was seen to be toxic to the cells and too little was suspected to have decreased the expression of cystatin. It is possible that glycerol accumulation in the bioreactor acted to repress the *AOX1* gene promoter.

A comparison between the methanol utilization phenotype of the wild-type strain of *P. pastoris* was performed in the bioreactor. Cystatin productivities were very similar during the first 24 hours of induction, but the $Mut^s$ phenotype expressed the highest levels of biologically active cystatin. An extracellular concentration of 54 $\mu moles \cdot l^{-1}$ and a yield of 1.0 $\mu mole \cdot g^{-1}$ dcw were achieved in a culture time of 120 hours. The $Mut^+$ phenotype achieved a maximum cystatin concentration and yield that were only one-half and one-third respectively, those of the $Mut^s$ phenotype. The maximum expression for the $Mut^+$ phenotype was observed after a culture time of 72 hours. These experiments revealed that the growth medium and the bioreactor apparatus must be tailored to the phenotype that has been selected.

Generally, the experimental designs used in this work proved to be adequate, but the initial conditions were often far from optimal. Good designs allowed for the experiments to move efficiently toward the optimal conditions. However, convergence of these experiments to optimal conditions was compromised by the poor choice of initial conditions coupled with the obvious time restraints of this project. A more

thorough review of the relevant literature would have allowed for the selection of better initial conditions that would have resulted in a rewarding convergence to optimal expression of cystatin. However, these experiments have identified and quantified some of the most important factors for cystatin expression by *P. pastoris.* Some excellent results have been achieved and many new experiments have been justified.

## 6.2 Recommendations

The *Pichia pastoris* expression system has shown great potential for the production of cystatin C and its genetic variants. The preliminary experiments have been performed and the limitations and advantages of this system have been realized. Experiments can now be designed to converge to optimal expression conditions.

### *Bioreactor Experiments*

Experiments to investigate the substrate feeding strategy in the bioreactor should be continued. However, sorbitol should be investigated as a replacement for glycerol. Sorbitol is less repressing than glycerol and has been shown to give rise to higher protein yields (Thorpe et al. 1999). However, sorbitol is more expensive than glycerol (3.6¢ versus 2.0¢ per gram of carbon (Sigma, St. Louis, MO)), has a lower cell yield coefficient, and inhibits cell growth at high concentrations (50 $g \cdot l^{-1}$). Therefore, glycerol should be used for the batch phase, and upon exhaustion, methanol and sorbitol should be fed to the bioreactor during the induction phase. The methanol concentration should be maintained between 1 and 2 $g \cdot l^{-1}$ or increased step-wise from 2 to 5 $g \cdot l^{-1} \cdot h^{-1}$ according to the recommendations of Brierly et al. (1990). Thorpe et al. (1999) have suggested that the sorbitol feed rate should be maintained below 0.9 $g \cdot l^{-1} \cdot h^{-1}$.

However, proteolysis may be increased if the cells are grown sequentially on several different substrates. In the case of *S. cerevisiae*, diauxic growth has been reported to increase proteolysis (Chen et al. 2000). There are no reported studies of increased proteolysis in *P. pastoris* cultures due to diauxic growth, but an investigation is certainly warranted. Diauxic growth refers to the sequential metabolism of two different substrates. Increased proteolysis resulting from diauxic growth is attributed to the nutritional stress associated with a reorganization of the cell metabolism. In the case of *P. pastoris*, diauxic growth is experienced when glycerol feeding is replaced by methanol feeding. Increased proteolysis may result from a feeding scheme that uses glycerol to generate cell mass followed by a mixed feed of sorbitol and methanol. As a means of reducing proteolysis, the addition of peptone (up to 20 $g \cdot l^{-1}$) or casamino acids (up to 5 $g \cdot l^{-1}$) should be considered (Clare et al. 1991b).

Both the shake-flask and bioreactor experiments have indicated that *P. pastoris* cultures may be prone to a nitrogen limitation. Not only does a nitrogen-limitation impede cell growth and recombinant protein production, but for the case of *S. cerevisiae*, proteolysis increases several fold during nitrogen starvation (Chen et al. 2000). The two recommended media for bioreactor cultures of *P. pastoris* differ in their means of supplying nitrogen (Invitrogen 1998). The Basal Salts medium relies upon the addition of ammonium hydroxide to control pH as well as supply a source of nitrogen. The FM22 medium includes 5 $g \cdot l^{-1}$ of ammonium sulfate to supply nitrogen and uses potassium hydroxide to adjust the pH. It seems that neither of these media are ideal in terms of nitrogen supply. In the case of the Basal Salts medium, the nitrogen supply is coupled with the pH set point. In the FM22 medium, high cell densities may exhaust the fixed supply of nitrogen. Therefore, an experiment should be performed in which 5 $g \cdot l^{-1}$ of ammonium sulfate is added to the Basal Salts medium and ammonium hydroxide is used to adjust the pH. This experiment should be well received by the *P. pastoris* research community because it would be based on two commonly used media recipes.

In order to interpret the results from these proposed experiments, the concentrations of substrates and byproducts will have to be monitored. Knowing the concentrations of methanol, ethanol, glycerol, and sorbitol would not only be very advantageous from an operational point of view, but essential in order to make a valid comparison between glycerol and sorbitol. Methanol and ethanol can be measured by gas chromatography. Glycerol and sorbitol can be measured by HPLC. In addition, an on-line methanol sensor like that reported by Guarna et al. (1997) would be an asset to any subsequent experiments involving the methylotrophic yeast *P. pastoris.*

## *Protein Characterization and Genetic Experiments*

Some observations have been made indicating that the secretion pathway for cystatin expression may have been a bottleneck. It was noted that a higher resistance to Zeocin™ (inferring a higher gene copy number) might have resulted in less glycosylation of the mutated variant of cystatin. It may have also been the case that a higher rate of expression in the bioreactor resulted in a reduced ratio of glycosylated to nonglycosylated expression. An efficient means of characterizing the oligosaccharides

of glycosylated cystatin will be necessary to elucidate these observations. These experiments will be performed in Food, Nutrition, and Health at UBC. It will also be important to measure both intracellular and extracellular expression of cystatin. In this case, a *P. pastoris* host expressing intracellular cystatin will be beneficial for investigating an efficient means of cell lysis.

The *P. pastoris* host that expresses intracellular cystatin could be transformed with the pAO815 vector. The pAO815 is specifically designed for intracellular expression and allows for multiple copies of the recombinant gene to be inserted into a single vector. With this vector, the number of copies of the cystatin gene can be predetermined. As a result, a more thorough analysis of the effect of gene copy number on cystatin expression could be afforded.

# Abbreviations

| | |
|---|---|
| AOX | alcohol oxidase |
| dcw | dry cell weight ($g \cdot l^{-1}$) |
| DNA | deoxyribonucleic acid |
| GAP | glyceraldehyde-3-phospate |
| Gly | glycerol |
| HPLC | high performance liquid chromatography |
| $k_La$ | oxygen mass transfer coefficient ($s^{-1}$) |
| MeOH | methanol |
| $Mut^+$ | methanol utilization positive |
| $Mut^s$ | methanol utilization slow |
| N | asparagine (Asn) |
| S | serine (Ser) |
| T | threonine (Thr) |
| wcw | wet cell weight ($g \cdot l^{-1}$) |

# References


1. Abrahamson, M. (1994a) Cystatins. *Methods Enzymol.* **244,** 685-700.
2. Abrahamson, M., Barrett, A. J., Salvesen, G. and Grubb, A. (1986) Isolation of six cysteine proteinase inhibitors from human urine. Their physicochemical and enzyme kinetic properties and concentrations in biological fluids. *J. Biol. Chem.* **261,** 11282-11289.
3. Abrahamson, M., Dalboge, H., Olafsson, I., Carlsen, S. and Grubb, A. (1988) Efficient production of native, biologically active human cystatin C by *Escherichia coli*. *FEBS Lett.* **236,** 14-18.
4. Abrahamson, M. and Grubb, A. (1994b) Increased body temperature accelerates aggregation of the Leu-68 → Gln mutated cystatin C, the amyloid-forming protein in hereditary cystatin C amyloid angiopathy. *Proc. Natl. Acad. Sci. U.S.A.* **91,** 1416-1420.
5. Abrahamson, M., Grubb, A., Olafsson, I. and Lundwall, Å. (1987) Molecular cloning and sequence analysis of cDNA coding for the precursor of the human cysteine proteinase inhibitor cystatin C. *FEBS Lett.* **216,** 229-233.
6. Abrahamson, M., Mason, R. W., Hansson, H., Buttle, D. J., Grubb, A. and Ohlsson, K. (1991) Human cystatin C: Role of the N-terminal segment in the inhibition of human cysteine proteinases and its inactivation by leucocyte elastase. *J. Biochem.* **273,** 621-626.
7. Arakawa, T., Chong, D. K. X. and Langridge, W. H. R. (1998) Efficacy of a food plant-based oral cholera toxin B subunit vaccine. *Nature Biotechnol.* **16,** 292-298.
8. Barford, J. P. (1990) A general model for aerobic yeast growth: batch growth. *Biotech. Bioeng.* **35,** 907-920.
9. Barr, K. A., Hopkins, S. A. and Sreekrishna, K. (1992) Protocol for efficient secretion of HSA developed from *Pichia pastoris*. *Pharm. Eng.* **12,** 48-51.
10. Barrett, A. J. (1981) Cystatin, the egg white inhibitor of cysteine proteinases. *Methods Enzymol.* **80,** 771-778.

11. Barrett, A. J., Davies, M. E. and Grubb, A. (1984) The place of human γ-trace (cystatin C) amongst the cysteine proteinase inhibitors. *Biochem. Biophys. Res. Commun.* **120,** 931-636.
12. Becker, D. M. and Guarente, L. (1992) Protocol for high efficiency yeast transformation in Chang, D. C., Chassy, B. M., Saunders, J. A. and Sowers, A. E. (Eds), *Guide to Electroporation and Electrofusion*, Academic Press, pp. 501-505.
13. Berdy, J. (1980) Bleomycin-type antibiotics in Berdy, J. (Ed), *Amino Acid and Peptide Antibiotics*, CRC Press, pp. 459-497.
14. Berti, P. J., Ekiel, I., Lindahl, P., Abrahamson, M. and Storer, A. C. (1997) Affinity purification and elimination of methionine oxidiation in recombinant human cystatin C. *Protein Expr. Purif.* **11,** 111-118.
15. Beste, H., Fackeldey, M., Willems, M., Stockmann, L. and Weuster-Bolz, D. (1997) Optimization of fermentation medium composition in substrate-controlled continuous stirred tank reactors. *Chem. Eng. Technol.* **20,** 403-413.
16. Biemans, R., Thines, D., Rutgers, T., DeWilde, M. and Cabezon, T. (1991) The large surface protein of hepatitis B virus is retained in the yeast endoplasmic reticulum and provokes its unique enlargement. *DNA Cell Biol.* **10,** 191-200.
17. Björck, L., Åkesson, P., Bohus, M., Trojnar, J., Abrahamson, M., Olafsson, I. and Grubb, A. (1989) Bacterial growth blocked by a synthetic peptide based on the structure of a human proteinase inhibitor. *Nature* **337,** 385-386.
18. Björck, L., Grubb, A. and Kjellén, L. (1990) Cystatin C, a human proteinase inhibitor, blocks replication of herpes simplex virus. *J. Virol.* **62,** 941-943.
19. Blankenvoorde, M. F., Hof, W. v. t., Walgreen-Waterings, E., Steenbergen, T. J. v., Brand, H. S., Veerman, E. C. and Amerongen, A. V. N. (1998) Cystatin and cystatin-derived peptides have antibacterial activity against the pathogen *Porphyromonas gingivalis*. *J. Biol. Chem.* **379,** 1371-1375.
20. Bode, W., Engh, R., Musil, D., Thiele, U., Huber, R., Karshikov, A., Brzin, J., Kos, J. and Turk, V. (1988) The 2.0 Å X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *EMBO J.* **7,** 2593-2599.

21. Bokenkamp, A., Domanetzki, M., Zinck, R., Schumann, G. and Brodehl, J. (1998a) Reference values for cystatin C serum concentrations in children. *Pediatr. Nephrol.* **12,** 125-9.
22. Brierly, R. A. (1999) Optimization and scale-up of a recombinant protein production process for manufacturing of recombinant human insulin-like growth factor-I (IGF-I) using the methylotrophic yeast, *Pichia pastoris. Current Topics in Gene Expression Systems. San Diego, CA.*
23. Brierly, R. A., Bussinau, C., Kosson, R., Melton, A. and Siegel, R. S. (1990) Fermentation development of recombinant *Pichia pastoris* expressing the heterologous gene bovine lysozyme. *Anns. N.Y. Acad. Sci.* **589,** 350-362.
24. Brierly, R. A., Davis, G. R. and Holtz, G. C. (1994) Production of insulin-like growth factor-1 in methylotrophic yeast cells. US Patent 5,324,639.
25. Buckholz, R. G. and Gleeson, M. A. G. (1991) Yeast systems for the commercial production of heterologous proteins. *Bio/Technol.* **9,** 1067-1071.
26. Buttle, D. J., Abrahamson, M. and Barrett, A. J. (1986) The biochemistry of the action of chymopapain in relief of sciatica. *Spine* **11,** 688-694.
27. Cereghino, G. P. L. and Cregg, J. M. (1999) Applications of yeast in biotechnology: protein production and genetic analysis. *Curr. Opin. Biotechnol.* **10,** 422-427.
28. Chen, D., Wang, B., Chou, P. and Kuo, T. (2000) Asparagine as a nitrogen source for improving the secretion of mouse alpha-amylase in *Saccharomyces cervisiae* protease A-dificient strains. *Yeast* **16,** 207-217.
29. Chen, Y., Cino, J., Hart, G., Freedman, D., White, C. and Komives, E. A. (1997) High protein expression in fermentation of recombinant *Pichia pastoris* by a fed-batch process. *Process Biochem.* **32,** 107-111.
30. Chen, Y., Krol, J., Cino, J. and Freedman, D. (1996) Continous production of thrombomodulin from a *Pichia pastoris* fermentation. *J. Chem. Technol. Biotechnol.* **1996,** 143-148.
31. Chiruvolu, V., Gregg, J. M. and Meagher, M. M. (1997) Recombinant protein production in an alcohol oxidase-defective strain of *Pichia pastoris* in fedbatch fermentations. *Enzyme Microb. Technol.* **21,** 277-283.

32. Clare, J. J., Rayment, F. B., Ballantine, S. P., Sreekrishna, K. and Romanos, M. A. (1991a) High-level expression of tetanus toxin fragment C in *Pichia pastoris* strains containing multiple tandem integrations of the gene. *Bio/Technol.* **9,** 455-460.

33. Clare, J. J., Romanos, M. A., Rayment, F. B., Rowedder, J. E., Smith, M. A., Payne, M. M., Sreekrishna, K. and Henwood, C. A. (1991b) Production of mouse epidermal growth factor in yeast: high-level secretion using *Pichia pastoris* strain containing multiple gene copies. *Gene* **105,** 205-212.

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

35. Cox, J. L., Sexton, P. S., Green, T. J. and Darmani, N. A. (1999) Inhibition of B16 melanoma metastasis by overexpression of the cysteine proteinase inhibitor cystatin C. *Melanoma Res.* **9,** 369-374.

36. Cregg, J. M., Barringer, K. J., Hessler, A. Y. and Madden, K. R. (1985) *Pichia pastoris* as a host for the production of foreign proteins. *Dev. Ind. Microbiol.* **29,** 33-41.

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

38. Cregg, J. M., Madden, K. R., Barringer, K. J., Thill, G. P. and Stillman, C. A. (1989) Functional characteristics of the two alcohol oxidase and two other methanol regulatable genes from the yeast *Pichia pastoris*. *Mol. Cell. Biol.* **9,** 1316-1323.

39. Cregg, J. M., Tschopp, J. F., Stillman, C., Siegel, R., Akong, M., Craig, W. S., Buckholz, R. G., Madden, K. R., Kellaris, P. A., Davis, G. R., Smiley, B. L., Cruze, J., Torregrossa, T., Veliçelebi, G. and Thill, G. P. (1987) High-level expression and efficient assembly of hepatitis B surface antigen in the methylotrophic yeast, *Pichia pastoris*. *Bio/Technol.* **5,** 479-485.

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

41. Dalboge, H., Jensen, E. B., Hottrup, H., Grubb, A., Abrahamson, M., Olafsson, I. and Carlsen, S. (1989) High-level expression of active human cystatin C in *Escherichia coli*. *Gene* **79,** 325-332.
42. DaSilva, N. A. and Bailey, J. E. (1991) Influence of dilution rate and induction of cloned gene expression in continuous fermentations of recombinant yeast. *Biotechnol. Bioeng.* **37,** 309-317.
43. Difco (1953) *Difco Manual of Dehydrated Culture Media and Reagents for Microbiological and Clinical Laboratory Procedures*, 9th edn. Difco Laboratories.
44. Digan, M. E., Lair, S. V., Brierley, R. A., Siegel, R. S., Williams, M. E., Ellis, S. B., Kellaris, P. A., Provow, S. A., Craig, W. S., Veliçelebi, G., Harpold, M. M. and Thill, G. P. (1989) Continuous production of a novel lysozyme via secretion from the yeast, *Pichia pastoris*. *Bio/Technol.* **7,** 160-164.
45. Döring, F., Theis, S. and Daniel, H. (1997) Expression and functional characterization of the mammalian intestinal peptide transporter PepT1 in the methylotrophic yeast *Pichia pastoris*. *Biochem. Biophys. Res. Commun.* **232,** 656-662.
46. Duff, S. J. B. and Murray, W. D. (1988) Production and application of methylotrophic yeast *Pichia pastoris*. *Biotechnol. Bioeng.* **31,** 44-49.
47. Duman, J. G., Miele, R. G., Liang, H., Grella, D. K., Sim, K. L., Castellino, F. J. and Bretthauer, R. K. (1998) O-Mannosylation of *Pichia pastoris* cellular and recombinant proteins. *Biotechnol. Appl. Biochem.* **28,** 39-45.
48. Ekiel, I. and Abrahamson, M. (1996) Folding-related dimerization of human cystatin C. *J. Biol. Chem.* **271,** 1314-1321.
49. Eldin, P., Pauza, M. E., Hieda, Y., Lin, G., Murtaugh, M. P., Pentel, P. R. and Pennell, C. A. (1997) High-level secretion of two antibody single chain Fv fragments by *Pichia pastoris*. *J. Immunol. Methods* **201,** 67-75.
50. Ellis, S. B., Brust, P. F., Koutz, P. J., Waters, A. F., Harpold, M. M. and Gineras, T. R. (1985) Isolation of alcohol oxidase and two other methanol regulatable genes from the yeast, *Pichia pastoris*. *Mol. Cell. Biol.* **5,** 1111-1121.
51. Esnard, A., Esnard, F., Faucher, D. and Gauthier, F. (1988) Two rat homologies of human cystatin C. *FEBS Lett.* **236,** 475-478.

52. Faber, K. N., Harder, W., AB, G. and Veenhuis, M. (1995) Review: methylotrophic yeasts as factories for the production of foreign proteins. *Yeast* **11,** 1331-1344.

53. Faber, K. N., Westra, S., Waterham, H. R., Keizer-Gunnicnk, I., Harder, W. and Veenhuis, G. A. (1996) Foreign gene expression in *Hansenula polymorpha*. A system for the synthesis of small functional peptides. *Appl. Microbiol. Biotechnol.* **45,** 72-79.

54. Gellissen, G., Hollenberg, C. P. and Janowicz, Z. A. (1995) Gene expression in methylotrophic yeasts in Smith, A. (Ed), *Gene Expression in Recombinant Microorganisms*, Marcel Dekker, pp. 195-239.

55. Gennaro, D. E., Hoffstein, S. T., Marks, G., Ramos, L., Oka, M. S., Reff, M. E., Hart, T. K. and Bugelsi, P. J. (1991) Quantitative immunocytochemical staining for recombinant tissue type plasminogen activator in transfected Chinese hamster ovary cells. *Proc. Soc. Exp. Biol. Med.* **198,** 591-598.

56. Grinna, L. S. and Tschopp, J. F. (1989) Size distribution and general structural features of N-linked oligosaccharides from the methylotrophic yeast, *Pichia pastoris*. *Yeast* **5,** 107-115.

57. Grubb, A. and Löfberg, H. (1982) Human gamma-trace, a basic microprotein: amino acid sequence presence in the adenohypophysis. *Proc. Natl. Acad. Sci. U.S.A.* **79,** 3024-3027.

58. Grubb, A. and Löfberg, H. (1988) Safer proteinase treatment of sciatica: a biochemical preview of chymopapain inhibitors. *Acta Orthop. Scand.* **59,** 63-65.

59. Guarna, M. M., Lesnicki, G. J., Tam, B. M., Robinson, J., Radziminski, C. Z., Hasenwinkle, D., Boraston, A., Jervis, E., MacGillivray, R. T. A., Turner, R. F. B. and Kilburn, D. G. (1997) On-line monitoring and control of methanol concentration in shake-flask cultures of *Pichia pastoris*. *Biotechnol. Bioeng.* **56,** 279-286.

60. Guebel, D. V., Cordenons, A., Nudel, B. C. and Giulietti, A. M. (1991) Influence of oxygen transfer rate and media composition on fermentation of D-xylose by *Pichia stipitis* NRRL Y-7124. *J. Ind. Microbiol.* **7,** 287-292.

61. Hall, A., Dalboge, H., Grubb, A. and Abrahamson, M. (1993) Importance of the evolutionarily conserved glycine residue in the N-terminal region of human cystatin C (Gly-11) for cysteine endopeptidase inhibition. *J. Biochem.* **291,** 123-129.

62. Hall, A., Håkansson, K., Mason, R. W., Grubb, A. and Abrahamson, M. (1995) Structural basis for the biological specificity of cystatin C. *J. Biol. Chem.* **270,** 5115-5121.
63. Helenius, A., Marquardt, T. and Braakman, I. (1992) The endoplasmic reticulum as a protein folding compartment. *Trends Cell Biol.* **2,** 227-231.
64. Henry, A., Masters, C. L., Beyreuther, K. and Cappai, R. (1997) Expression of human amyloid precursor protein ectodomains in *Pichia pastoris*: analysis of culture conditions, purification, and characterization. *Protein Expr. Pur.* **10,** 283-291.
65. Higgins, D. R. and Cregg, J. M. (1998) *Pichia protocols*. Humana Press, New Jersey.
66. Inan, M., Chiruvolu, V., K.Eskridge, Vlasuk, G., Dickerson, K., Brown, S. and Meagher, M. (1999) Optimization of temperature-glycerol-pH conditions for a fed-batch fermentation process for recombinant hookworm (*Ancylostoma caninum*) anticoagulant peptide (AcAP-5) production by *Pichia pastoris*. *Enzyme Microb. Technol.* **24,** 438-445.
67. Invitrogen (1998) EasySelect *Pichia* Expression Kit.
68. Invitrogen (1999) *Pichia* Fermentation Process Guidelines.
69. Ito, H., Fukuda, Y., Murata, K. and Kimura, A. (1983) Transformation of intact yeast cells treated with alkali cations. *J. Bacteriol.* **153,** 161-168.
70. Katakura, Y., Zhang, W., Zhuang, G., Omasa, T., Kishimoto, M., Goto, Y. and Suga, K. (1998) Effect of methanol concentration on the production of human B2-glycoprotein *I* domain *V* by a recombinant *Pichia pastoris*: a simple system for the control of methanol concentration using a semiconductor gas sensor. *J. Ferment. Bioeng.* **86,** 482-487.
71. Kim, T.-R., Goto, Y., Hirota, N., Kuwata, K., Denton, H., Wu, S.-Y., Sawyer, L. and Batt, C. A. (1997) High-level expression of bovine ß-lactoglobulin in *Pichia pastoris* and characterization of its physical properties. *Protein Eng.* **10,** 1339-1345.
72. Kornfeld, R. and Kornfeld, S. (1985) Assembly of Asparagine-Linked Oligosaccharides. *Annu. Rev. Biochem.* **54,** 631-664.

73. Kost, T. A. (1997) Expression systems: gene expression systems in the genomics era. *Curr. Opin. Biotechnol.* **8,** 539-541.

74. Koutz, P. J., Davis, G. R., Stillman, C., Barringer, K., Gregg, J. M. and Thill, G. (1989) Structural comparison of the *Pichia pastoris* alcohol oxidase genes. *Yeast* **5,** 167-177.

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

76. LeBlanc, F. E. (1991) Sciatica management: chemonucleolysis vs. surgical discectomy. *Todays O.R. Nurse* **13,** 18-23.

77. Lenarcic, B., Krasovec, M., Rionjia, A., Olafsson, I. and Turk, V. (1991) Inactiviation of human cystatin C and kininogen by human cathespsin D. *FEBS Lett.* **280,** 211-215.

78. Lesnicki, G. (1999) University of British Columbia Pilot Plant Manager.

79. Lindahl, P., Abrahamson, M. and Bjork, I. (1992) Interaction of recombinant human cystatin C with the cysteine proteinases papain and actinidin. *J. Biochem.* **281,** 49-55.

80. Loewen, M. C., Liu, X., Davies, P. L. and Daugulis, A. J. (1997) Biosynthetic production of type II fish antifreeze protein: fermentation by *Pichia pastoris. Appl. Microbiol. Biotechnol.* **48,** 480-486.

81. Mancini, G., Carbonara, A. O. and Heremans, J. F. (1965) Immunochemical quantitation of antigens by single radial immunodiffusion. *Immunochemistry* **2,** 235-254.

82. Montesino, R., Garcia, R., Quintero, O. and Cremata, J. A. (1998) Variation in N-linked oligosaccharide structures on heterologous proteins secreted by the methylotrophic yeast *Pichia pastoris. Protein Expr. Purif.* **14,** 197-207.

83. Murray, D. W., Duff, S. J. B. and Lanthier, P. H. (1989) Induction and stability of alcohol oxidase in the methylotrophic yeast *Pichia pastoris. Appl. Microbiol. Biotechnol.* **32,** 95-100.

84. Nakamura, S., Ogawa, M. and Nakai, S. (1998a) Effects of polymannosylation of recombinant cystatin C in yeast on its stability and activity. *J. Agric. Food Chem.* **46,** 2882-2887.

85. Nakamura, S., Takasaki, H., Kobayashi, K. and Kato, A. (1993) Hyperglycosylation of hen egg white lysozyme in yeast. *J. Biol. Chem.* **268,** 12706-12712.

86. Ni, J., Abrahamson, M., Zhang, M., Fernandez, M. A., Grubb, A., Su, J., Yu, G.-L., Li, Y., Parmelee, D., Xing, L., Coleman, T. A., Gentz, S., Thotakura, R., Nguyen, N., Hesselberg, M. and Gentz, R. (1997) Cystatin E is a novel human cysteine proteinase inhibitor with structural resemblance to family 2 cystatins. *J. Biol. Chem.* **272,** 10853-10858.

87. Ni, J., Fernandez, M. A., Danielsson, L., Chillakuru, R. A., Zhang, J., Grubb, A., Su, J., Gentz, R. and Abrahamson, M. (1998) Cystatin F is a glycosylated human low molecular weight cysteine proteinase inhibitor. *J. Biol. Chem.* **273,** 24797-24804.

88. Ogata, K., Nishikawa, H. and Ohsugi, M. (1969) A yeast capable of utilizing methanol. *Agric. Biol. Chem.* **33,** 1519-1520.

89. Parekh, R., Forrester, K. and Wittrup, D. (1995) Multicopy overexpression of bovine pancreatic trypsin inhibitor saturates the protein folding and secretory capacity of *Saccharomyces cerevisiae*. *Protein Expr. Purif.* **6,** 537-545.

90. Parekh, R. N. and Wittrup, K. D. (1997) Expression Level Tuning for Optimal Heterologous Protein Secretion in *Saccharomyces cerevisiae*. *Biotechnol. Prog.* **13,** 117-122.

91. Parnas, M., Sanchez-Monge, R., Gomez, L. and Salcedo, G. (1998) A chestnut seed cystatin differentially effective against cysteine proteinases from closely related pests. *Plant Mol. Biol.* **38,** 1235-1242.

92. Patel, R. N., Hou, C. T. and Derelanko, P. (1983) Microbial oxidation of methanol: purification and properties of formaldehyde dehydrogenase from a *Pichia* sp. NRRL-Y-11328. *Arch. Biochem. Biophys.* **221,** 135-142.

93. Payne, M. S., Petrillo, K. L., Gavagan, J. E., Wagner, L. W., DiCosimo, R. and Anton, D. L. (1995) High-level production of spinach glycosylate oxidase in the methylotrophic yeast *Pichia pastoris*: engineering a biocatalyst. *Gene* **167,** 215-219.

94. Pernas, M., Sanchez-Monge, R. and Salcedo, G. (2000) Biotic and abiotic stress can induce cystatin expression in chestnut. *FEBS Lett.* **467,** 206-210.

95. Prasanna, G. L. and Panda, T. (1996) Electroporation: basic principles, practical considerations and applications in molecular biology. *Bioprocess Eng.* **16,** 261-264.

96. Randers, E., Kristensen, J. H., Erlandsen, E. J. and Danielsen, H. (1998) Serum cystatin C as a marker of the renal function. *Scand. J. Clin. Lab. Invest.* **58,** 585-592.

97. Reff, M. E. (1993) High-level production of recombinant immunoglobulins in mammalian cells. *Curr. Opin. Biotechnol.* **4,** 573-576.

98. Romanos, M. (1995) Advances in the use of *Pichia pastoris* for high-level gene expression. *Curr. Opin. Biotechnol.* **6,** 527-533.

99. Romanos, M. A., Scorer, C. A. and Clare, J. J. (1992) Foreign gene expression in yeast: a review. *Yeast* **8,** 423-488.

100. Sanger, F., Nicklen, S. and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **74,** 5463-5467.

101. Sawin, P. D., Traynelis, V. C., Rich, G., Smith, B. A., Maves, T. J., Follett, K. A. and Moore, S. A. (1997) Chymopapain-induced reduction of proinflammatory phospholipase A2 activity and amelioration of neuropathic behavioral changes in an in vivo model of acute sciatica. *J. Neurosurg.* **86,** 998-1006.

102. Shay, L. K., Hunt, H. R. and Wegner, G. H. (1987) *J. Ind. Microbiol.* **2,** 79.

103. Shen, S., Sulter, G., Jefferies, T. W. and Gregg, J. M. (1998) A strong nitrogen source-regulated promoter for controlled expression of foreign genes in the yeast *Pichia pastoris*. *Gene* **216,** 93-102.

104. Sherman, F., Fink, G. R. and Hicks, J. B. (1986) *Methods in yeast genetics*. Cold Spring Harbor Laboratory Press, New York.

105. Shi, G. P., Sukhova, G. K., Grubb, A., Ducharme, A., Rhode, L. H., Lee, R. T., Ridker, P. M., Libby, P. and Chapman, H. A. (1999) Cystatin C deficiency in human atherosclerosis and aortic aneurysms. *J. Clin. Invest.* **104,** 1191-1197.

106. Shimode, K., Kobayashi, S., Imaoka, K., Umegae, N. and Nagai, A. (1996) Leukoencephalopahty-related cerebral amyloid angiopathy with cystatin C deposition. *Stroke* **27,** 1417-1419.

107. Siegel, R. S. and Brierley, R. A. (1988) Methylotrophic yeast *Pichia pastoris* produced in high-cell-density fermentations with high cell yields as vehicle for recombinant protein production. *Biotechnol. Bioeng.* **34,** 403-404.

108. Smith, G. E., Summers, M. D. and Fraser, M. J. (1983) Production of human ß-interferon in insect cells infected with a baculovirus expression vector. *Mol. Cell Biol.* **3,** 2156-2165.

109. Sotiropoulou, G., Anisowicz, A. and Sager, R. (1997) Indentification, cloning, and characterization of cystatin M, a novel cysteine proteinase inhibitor, down-regulated in breast cancer. *J. Biol. Chem.* **272,** 903-910.

110. Sreekrishna, K., Brankamp, R. G., Kropp, K. E., Blankenship, D. T., Tsay, J.-T., Smith, P. L., Wierschke, J. D., Subramaniam, A. and Birkenberger, L. A. (1997) Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast *Pichia pastoris*. *Gene* **190,** 55-62.

111. Sreekrishna, K. and Kropp, K. (1996) *Pichia pastoris* in Wolf, K. (Ed), *Non-Conventional Yeasts in Biotechnology*, Hoeshst Marion Roussel Inc., pp. 203-253.

112. Steinberg, F. M. and Raso, J. (1998) Biotech pharmaceuticals and biotherapy: an overview. *J. Pharm. Pharmaceut. Sci.* **1,** 48-59.

113. Sudbery, P. E. (1996) The expression of recombinant proteins in yeasts. *Curr. Opin. Biotechnol.* **7,** 517-524.

114. Thill, G. P., Davis, G. R., Stillman, G. R., Holtz, C., Brierly, R., Engel, M., Buckholz, R., Kinney, J., Provo, S., Vedvick, T. and Siegel, R. S. (1990) Positive and negative effects of multi-copy integrated expression vectors on protein expression in *Pichia pastoris. Proc. 6th Int. Symp. on Genetic of Industrial Microorganisms. Strasbourg*, pp. 477-490.

115. Thorpe, E. D., d'Anjou, M. C. and Daugulis, A. J. (1999) Sorbitol as a non-repressing carbon source for fed-batch fermentation of recombinant *Pichia pastoris. Biotechnol. Lett.* **21,** 669-672.

116. Tschopp, J. F., Brust, P. F., Cregg, J. M., Stillman, C. and Grinna, L. (1987b) Expression of the lacZ gene from two methanol regulated promoters in *Pichia pastoris. Biochim. Biophys. Acta.* **1252,** 28-34.

117. Tschopp, J. F., Sverlow, G., Kosson, R., Craig, W. and Grinna, L. (1987a) High-level secretion of glycosylated invertase in the methylotrophic yeast, *Pichia pastoris*. *Bio/Technol.* **5,** 1305-1308.

118. Tsujikawa, M., Okabayashi, K., Morita, M. and Tanabe, T. (1996) Secretion of a variant of human single-chain urokinase-type plasminogen activator without an N-glycosylation site in the methylotrophic yeast, *Pichia pastoris* and characterization of the secreted product. *Yeast* **12,** 541-553.

119. Vedvick, T., Buckholz, R. G., Engel, M., Urcan, M., Kinney, J., Provow, S., Siegel, R. S. and Thill, G. P. (1991) High-level secretion of biologically active aprotinin from the yeast *Pichia pastoris*. *J. Ind. Microbiol.* **7,** 197-202.

120. Veenhuis, M., Dijken, J. P. V. and Harder, W. (1983) The significance of peroxisomes in the metabolism of one-carbon compounds in yeasts. *Adv. Microb. Physiol.* **24,** 1-82.

121. Waterham, H. R., Digan, M. E., Koutz, P. J., Lair, S. V. and Cregg, J. M. (1997) Isolation of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter. *Gene* **186,** 37-44.

122. Wegner, E. H. (1983) Biochemical conversions by yeast fermentation at high cell densities. US Patent 4,414,329.

123. Wegner, G. (1990) Emerging applications of methylotrophic yeasts. *FEMS Microbiol. Rev.* **87,** 279-284.

124. Wei, L., Berman, Y., Castano, E. M., Cadene, M., Beavis, R. C., Devi, L. and Levy, E. (1998) Instability of the amyloidogenic cystatin C variant of hereditary cerebral hemorrhage with amyloidosis, Icelandic type. *J. Biol. Chem.* **273,** 11806-11814.

# Appendix I

## Recombinant Proteins

Table A.1. Some FDA approved pharmaceuticals derived from the biotechnology industry. Source: Steinberg et al. 1998.

| Product | Year of First U.S. Approval | Approved for: |
|---|---|---|
| recombinant human insulin | 1982 | diabetes mellitus |
| recombinant somatrem (human growth hormone) for injection | 1985 | human growth hormone (hGH) deficiency in children |
| recombinant interferon alfa-2b | 1986<br>1988<br>1988<br>1991<br>1992 | hairy cell leukemia<br>genital warts<br>Kaposi's sarcoma<br>hepatitis C<br>hepatitis B |
| recombinant interferon alfa-2a | 1986<br>1988 | hairy cell leukemia<br>Kaposi's sarcoma |
| Muromonab-CD3 | 1986<br>1993 | reversal of kidney transplant rejection<br>reversal of heart and liver transplant rejection |
| recombinant hepatitis B vaccine | 1986 | hepatitis B prevention |
| recombinant somatropin for injection | 1987 | human growth hormone (hGH) deficiency in children |
| Epoetin alfa (rEPO, Epogen) | 1989 | anemia of chronic renal failure |
| recombinant hepatitis B vaccine | 1989 | hepatitis B |
| interferon alfa-n3 | 1989 | genital warts |

| Product | Year of First U.S. Approval | Approved for: |
|---|---|---|
| adenosine deaminase | 1990 | severe immunodeficiency in infants |
| sargramostim (yeast-derived GM-CSF) | 1991 | bone marrow transplantation |
| Aldesleukin (interleukin-2) | 1992 | renal cell carcinoma |
| Staumonab pendetide (OncoScint) | 1992 | colorectal and ovarian cancers |
| recombinant antihemophiliac factor (rAHF) | 1992 | hemophilia A |
| dornase alpha (Pulmozyme) | 1993 | cystic fibrosis |
| Pegaspargase | 1994 | lymphoblastic leukemia |
| imiglucerase for injection (Cerezyme, recombinant lucocerebrosidase) | 1994 | Gaucher's disease |
| abciximab (ReoPro) | 1994 | prevention of blood clotting |
| Humatrope | 1996 | adult- or childhood-onset growth hormone deficiency |
| Serostim | 1996 | AIDS wasting associated with catabolism, weight loss, or cachexia |
| Saizen | 1996 | human growth hormone deficiency in children |
| Nutropin | 1996 | Turner's syndrome |
| Becaplermin (Regranex Gel) | 1997 | diabetic foot ulcers |
| daclizumab (Zenapax) | 1997 | acute renal allograft rejection |

Table A.2. Recombinant proteins that have been expressed in *Pichia pastoris*. Source: http://www.invitrogen.com (17/03/2000).

| Protein Expressed | Expression Level ($g \cdot l^{-1}$) | Reference |
|---|---|---|
| **Enzymes** | | |
| Invertase | 2.3 | Tschopp, J.F. (1987) *Bio/Technology* **5**: 1305-1308 |
| D-alanine carboxypeptidase | 0.8 | Despreaux, C. W and Manning, R.F. (1993) *Gene* **131**: 35-41 |
| Alpha amylase | 2.5 | Paifer, E. *et al.* (1994) *Yeast* **10**: 1415-1419 |
| Pectate Lyase | 0.004 | Guo, W.J. *et al* (1995) *Archives of Biochem. and Biophys.* **323**:352-360 |
| Recombinant enterokinase | 0.021 | Vozza, L.A. *et al.* (1996) *Bio/Technology* **14**:77-81 |
| **Proteases and protease Inhibitors** | | |
| Kunitz protease inhibitor (AbPP) | 1.0 | Wagner, S.L. *et al.* (1992) *Biochem. and Biophys. Res. Comm.* **186**: 1138-1145 |
| Kunitz protease inhibitor (APLP-2) | 1.0 | Van Nostrand, W.E. *et al.* (1994) *Biochimica et Biophysica Acta* **1209**: 165-170 |
| Tick Anticoagulant Protein | 1.7 | Laroche, Y. *et al.* (1994) *Bio/Technology* **12**:1119-1124 |
| Ghilanten | 0.01 | Brankamp, R.G. *et al.* (1995) *Protein Expression and Purification* **6**: 813-820 |
| Human proteinase inhibitor 6 | 0.05 | Sun, J.R. *et al.* (1995) *Biochimica et Biophysica Acta* **1252**: 28-34 |
| **Membrane Proteins** | | |
| Human CD38 (soluble portion) | 0.05 | Fryxell, K.B. *et al* (1995) *Protein Expression and Purification* **6**: 329-336 |
| Mouse serotonin receptor | 0.001 | Weiss, H.M. *et al.* (1995) *FEBS Letters* **377**: 451-456 |
| **Antigens** | | |
| Tetanus toxin fragment C | 12.0 | Clare, J.J. *et al.* (1991) *Bio/Technology* **9**: 455-460 |

| Protein Expressed | Expression Level ($g \cdot l^{-1}$) | Reference |
|---|---|---|
| Bordetella pertussis antigen P69 | 3.0 | Romanos, M.A. *et al.* (1991) *Vaccine* **9**: 901-906 |
| HIV-1 gp120 (intracellular) | 1.25 | Scorer, C.A. *et al.* (1993) *Gene* **136**: 111-119 |
| HIV-1 gp120 (secreted) | 0.02 | Scorer, C.A. *et al.* (1993) *Gene* **136**: 111-119 |
| Bm86 tick gut glycoprotein | 1.5 | Rodriguez, M. *et al.* (1994) *Journal of Biotechnology* **33**: 135-146 |
| Regulatory Proteins | 10.0 | Sreekrishna, K. *et al.* (1989) *Biochemistry* **28**: 4117-4125 |
| Tumor necrosis factor (TNF) | 0.08 | Hagenson, M.J. *et al.* (1989) *Enzyme Microbial Technology* **11**: 650-656 |
| Streptokinase (active) | 0.45 | Clare, J.J. *et al.* (1991) *Gene* **105**: 205-212 |
| Mouse epidermal growth factor (EGF) | 0.4 | Garcia, J.N. *et al.* (1995) *Yeast* **11**: S589 |
| Human IFN-a2b | 0.05 | Steinlein, L.M. *et al.* (1995) *Protein Expresion and Purification* **6**: 619-624 |
| N-terminus of human transferrin | 4.0 | Cregg, J.M. *et al.* (1993) *Bio/Technology* **11**: 905-910 |

# Appendix II

## Experimental Data

### Growth Medium for Batch Cultures of *Pichia pastoris*

*YPD*

10 $g \cdot l^{-1}$ yeast extract, 20 $g \cdot l^{-1}$ peptone, 10 $g \cdot l^{-1}$ D-glucose.
20 $g \cdot l^{-1}$ of agar is added when making solid medium.

*BMM*

The BMM medium is used for *S. cerevisiae* and *P. pastoris.* The complete medium contains (per liter): 20 g of glucose or 20 g of glycerol, 0.1 g $CaCl_2 \cdot 2H_2O$, 0.5 g $MgSO_4 \cdot H_2O$, 1.5 g $KH_2PO_4$, 0.1 mg KI, 0.06 mg $H_3BO_3$, 0.06 mg $MnSO_4 \cdot 5H_2O$, 0.3 mg $ZnSO_4 \cdot 7H_2O$, 0.04 mg $CuSO_4 \cdot 5H_2O$, 0.25 mg $FeCl_3 \cdot 6H_2O$, 0.025 mg Mo, 0.025 mg NaCl, 20 mg histidine, 0.4 mg biotin, 0.2 mg thiamine, 0.2 mg pridoxine, 0.2 mg nicotinic acid, 10 mg inositol, and 0.2 mg pantotenic acid.

*Yeast Nitrogen Base*

A 3.4 $g \cdot l^{-1}$ solution of yeast nitrogen base without amino acids and without ammonium sulfate contains (per liter): 0.002 mg biotin, 0.4 mg calcium pantothenate, 0.002 mg folic acid, 2.0 mg inositol, 0.4 mg niacin, 0.2 mg p-aminobenzoic acid, 0.4 mg pyridoxine hydrochloride, 0.2 mg riboflavin, 0.4 mg thiamine hydrochloride, 0.5 mg boric acid, 0.04 mg copper sulfate, 0.1 mg potassium iodide, 0.2 mg ferric chloride, 0.4 mg manganese sulfate, 0.2 mg sodium molybdate, 0.4 mg zinc sulfate, 1.0 g potassium phosphate monobasic, 0.5 g magnesium sulfate, 0.1 g sodium chloride, and 0.1 g calcium chloride (Difco 1953).

### Growth Medium for Bioreactor Cultures of *Pichia pastoris*

Recipes are taken from Higgins and Cregg (1998)

*Basic Media Stock*

*Basal Salts Medium*

1 liter contains: 26.7 ml 85% $H_3PO_4$, 0.93 g $CaSO_4 \cdot 2H_2O$, 18.2 g $K_2SO_4$, 14.9 g $MgSO_4 \cdot 7H_2O$, 4.13 g KOH, 40 g glycerol. The medium will have a pH of 1.0 to 1.5 after autoclaving. After the medium reaches a temperature of 30°C in the bioreactor, the pH is adjusted to pH 5.0 with 30% (w/w) $NH_4OH$. $NH_4OH$ also serves as a source of nitrogen.

*FM22*

1 liter contains: 42.9 g $KH_2PO_4$, 5 g $(NH_4)_2SO_4$, 1.0 g $CaSO_4 \cdot 2H_2O$, 14.3 g $K_2SO_4$, 11.7 g $MgSO_4 \cdot 7H_2O$, 40 g glycerol and adjust to pH 5.0 with KOH. The pH should be 4.5 after autoclaving.

## *Trace Element Solutions*

*PTM1 (for use with the Basal Salts medium)*

1 liter contains: 6.0 g $CuSO_4 \cdot 5H_2O$, 0.8 g KI, 3.0 g $MnSO_4 \cdot H_2O$, 0.2 g $Na_2MoO_4 \cdot 2H_2O$, 0.2 g $H_3BO_3$, 0.5 g $CaSO_4 \cdot 2H_2O$, 20 g $ZnCl_2$, 65 g $FeSO_4 \cdot 7H_2O$, 0.2 g biotin, 5 ml conc. $H_2SO_4$. Filter sterilize and add 2 $ml \cdot l^{-1}$ of Basal salts medium. The trace salt solution is also added to the methanol feed supply at 2 $ml \cdot l^{-1}$.

*PTM4 (for use with the FM22 medium)*

1 liter contains: 2.0 g $CuSO_4 \cdot 5H_2O$, 0.08 g NaI, 3.0 g $MnSO_4 \cdot H_2O$, 0.2 g $Na_2MoO_4 \cdot 2H_2O$, 0.02 g $H_3BO_3$, 0.5 g $CaSO_4 \cdot 2H_2O$, 0.5 g $CoCl_2$, 7 g $ZnCl_2$, 22 g $FeSO_4 \cdot 7H_2O$, 0.2 g biotin, 1 ml conc. $H_2SO_4$. Filter sterilize and add 1 $ml \cdot l^{-1}$ of FM22 medium. The trace salt solution is also added to the methanol feed supply at 4 $ml \cdot l^{-1}$.

# Data for Chapter 3

Table A.3. Optical density data as a function of culture time and treatment condition for two-level factorial experiments. Each of the treatment conditions has two column corresponding to 2 replications for each flask. The optical densities measured at 26 hours were used as a basis for comparing the treatments.

| Time | Optical Density | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| (hours) | 1 | G | N | GN | 1 | G | N | GN |
| 0 | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 |
| 1.5 | 1.4 | 1.3 | 1.5 | 1.5 | 1.3 | 1.4 | 1.4 | 1.6 |
| 9.5 | 5.8 | 5.2 | 6.0 | 5.9 | 6.9 | 5.2 | 6.6 | 5.5 |
| 18 | 6.5 | 5.8 | 7.6 | 7.6 | 8.7 | 5.8 | 9.8 | 7.6 |
| 26 | 6.1 | 5.7 | 8.2 | 8.0 | 9.1 | 6.0 | 9.5 | 8.5 |
| 42 | 4.7 | 6.9 | 2.0 | 9.4 | 4.9 | 6.2 | 4.8 | 9.1 |
| 48 | 2.3 | 7.1 | 1.3 | 9.2 | 3.0 | 6.4 | 2.5 | 8.6 |
| 55 | 1.2 | 7.7 | 1.1 | 8.8 | 2.0 | 6.5 | 1.5 | 8.4 |
| 66 | 1.0 | 8.2 | 1.0 | 9.5 | 1.8 | 7.1 | 1.1 | 7.7 |

Table A.3 continued

| Time | Optical Density | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| (hours) | Y | GY | NY | GNY | Y | GY | NY | GNY |
| 0 | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 |
| 1.5 | 1.7 | 1.8 | 1.9 | 1.8 | 1.5 | 1.8 | 1.7 | 1.7 |
| 9.5 | 13.9 | 14.5 | 13.7 | 15.4 | 14.5 | 14.9 | 14.5 | 16.2 |
| 18 | 25.1 | 24.3 | 23.7 | 30.0 | 27.2 | 25.0 | 24.2 | 30.2 |
| 26 | 25.2 | 29.8 | 24.7 | 33.7 | 24.1 | 31.0 | 23.4 | 34.1 |
| 42 | 19.0 | 27.0 | 18.5 | 24.9 | 18.9 | 25.6 | 17.9 | 24.1 |
| 48 | 15.5 | 21.2 | 16.3 | 19.1 | 15.0 | 20.5 | 15.4 | 17.8 |
| 55 | 15.1 | 19.3 | 16.1 | 18.9 | 14.5 | 19.2 | 15.9 | 18.5 |
| 66 | 14.0 | 17.1 | 13.2 | 17.4 | 14.5 | 16.9 | 14.3 | 18.2 |

Table A.4. Dry cell weight as a function of shake-flask culture time and L-histidine concentration in the Basal Salts + PTM1 Trace Salts medium.

| Histidine ($mg \cdot l^{-1}$) | Dry Cell Weight ($g \cdot l^{-1}$) | | | |
|---|---|---|---|---|
| | Time (hours) | | | |
| | 0 hours | 24 hours | 48 hours | 72 hours |
| 0 | 1.8 | 1.8 | 2.0 | 2.0 |
| 500 | 1.8 | 8.5 | 9.0 | 7.6 |
| 1000 | 1.8 | 9.8 | 12.8 | 10.4 |

# Data for Chapter 4

## *Dry Cell Weight Calibration Data*

Table A.5. Calibration data for dry cell weight as a function of wet cell weight. 1-ml samples were centrifuged, the supernatants were poured off, the pellet was weighed, then dried overnight at 105°C, and then weighed again.

| **Wet Wt. ($g \cdot l^{-1}$)** | **Dry Wt. ($g \cdot l^{-1}$)** | **Wet Wt. ($g \cdot l^{-1}$)** | **Dry Wt. ($g \cdot l^{-1}$)** |
|---|---|---|---|
| 21.2 | 4.1 | 201.4 | 59.3 |
| 24 | 5.5 | 227.6 | 68 |
| 24.8 | 4.9 | 226.1 | 67.1 |
| 35.1 | 6.8 | 224.3 | 67 |
| 36.6 | 6.5 | 229.1 | 67 |
| 99.9 | 25.3 | 230.2 | 67.5 |
| 101.8 | 24.3 | 229.2 | 67.8 |
| 124.9 | 34.5 | 239.1 | 67.9 |
| 143.7 | 39.6 | 235.7 | 68.4 |
| 145.3 | 40.1 | 237.3 | 68.4 |
| 151.3 | 41.6 | 257 | 76 |
| 151 | 42.5 | 262.6 | 76.1 |
| 179.6 | 53.7 | 259.8 | 74.6 |
| 199.6 | 58.4 | 278.7 | 80.2 |
| 200.8 | 59.2 | 276.3 | 79.5 |
| | | 276.4 | 79.5 |

## *Cystatin Activity and Storage Data*

Table A.6. Cystatin activity measurements as a function of storage time. Sample time is the culture time corresponding to the sample that was measured. Measurement delay is the storage time for the sample prior to measuring the cystatin activity. A and B are the replicated measurements for cystatin activity.

| Sample | Measurement | Measured Inhibition of Papain (%) | | | |
|---|---|---|---|---|---|
| Time (hours) | Delay (hours) | A | B | Average | Std. Dev. |
| 118 | 0 | 39.6 | 36.1 | 37.9 | 2.5 |
| | 24 | 28.9 | 27.6 | 28.3 | 0.9 |
| | 55 | 33.3 | 33.9 | 33.6 | 0.4 |
| 100 | 0 | 30.1 | 34.2 | 32.2 | 2.9 |
| | 18 | 32 | 31.6 | 31.8 | 0.3 |
| | 44 | 26.1 | 23.2 | 24.7 | 2.1 |
| | 73 | 26.6 | 28.7 | 27.7 | 1.5 |
| 81.5 | 0 | 36.5 | 38.7 | 37.6 | 1.6 |
| | 12.5 | 32.5 | 29.9 | 31.2 | 1.8 |
| | 36.5 | 24.4 | 22.1 | 23.3 | 1.6 |
| | 62.5 | 23.3 | 21.2 | 22.3 | 1.5 |
| | 91.5 | 20.8 | 22.3 | 21.6 | 1.1 |
| 70 | 0 | 36.6 | 37.6 | 37.1 | 0.7 |
| | 24 | 24.4 | 23.3 | 23.9 | 0.8 |
| | 74 | 13.4 | 21.7 | 17.6 | 5.9 |
| | 103 | 17.8 | 17.4 | 17.6 | 0.3 |
| 46.5 | 0 | 19.6 | 16.4 | 18.0 | 2.3 |
| | 23.5 | 11.7 | 8.9 | 10.3 | 2.0 |
| | 47.5 | 8 | 6.5 | 7.3 | 1.1 |
| | 126.5 | 6.2 | 4.5 | 5.4 | 1.2 |
| | 13.5 | 26.4 | 21.3 | 23.9 | 3.6 |
| 33 | 140 | 12.3 | 9.7 | 11.0 | 1.8 |
| | 23 | 6.3 | 10.3 | 8.3 | 2.8 |
| 76 | 47 | 40.6 | 38.2 | 39.4 | 1.7 |
| | 228 | 38.3 | 38.9 | 38.6 | 0.4 |
| 55 | 0 | 29.2 | 28.2 | 28.7 | 0.7 |
| | 44 | 7.1 | 2.9 | 5.0 | 3.0 |
| | 68 | 26.9 | 23.4 | 25.2 | 2.5 |
| | 132 | 20.4 | 25.4 | 22.9 | 3.5 |
| | 312 | 12.3 | 11.1 | 11.7 | 0.8 |

# Data for Chapter 5

## *pH Optimization Data*

Table A.7. Wet cell weights as a function of bioreactor pH and culture time for the Mut$^s$ transformant that was selected for bioreactor experiments. The columns correspond to the induction pH and the date of inoculation of the culture.

| pH 5 (20/07/99) | | pH 6 (18/08/99) | | pH 6.5 (31/08/99) | | pH 7 (13/07/99) | |
|---|---|---|---|---|---|---|---|
| **Time (hours)** | **Wet Wt. ($g \cdot l^{-1}$)** | **Time (hours)** | **Wet Wt. ($g \cdot l^{-1}$)** | **Time (hours)** | **Wet Wt. ($g \cdot l^{-1}$)** | **Time (hours)** | **Wet Wt. ($g \cdot l^{-1}$)** |
| 0 | 15 | 0 | 13.9 | 0 | 13.8 | 0 | 19.8 |
| 0 | 12.6 | 0 | 11.8 | 0 | 13.7 | 0 | 18.4 |
| 3 | 14.7 | 2.5 | 14.8 | 2 | 15.5 | 6 | 23.7 |
| 3 | 14.6 | 2.5 | 14.6 | 2 | 16.6 | 6 | 21.7 |
| 6 | 24.3 | 5 | 18.3 | 6 | 25.8 | 13 | 24.2 |
| 6 | 20.9 | 5 | 19.4 | 6 | 26.2 | 13 | 26.6 |
| 16 | 70.8 | 10 | 35.9 | 9 | 34.6 | 24 | 85.8 |
| 16 | 71.3 | 10 | 35.6 | 9 | 37.4 | 24 | 85 |
| 22 | 89.2 | 20.5 | 67.4 | 23 | 104.9 | 27 | 121.4 |
| 22 | 86.7 | 20.5 | 64.9 | 23 | 112.8 | 27 | 120.3 |
| 24 | 93.8 | 24.0 | 71.7 | 24.5 | 139.9 | 27 | 121.9 |
| 24 | 98.7 | 24.0 | 72.7 | 24.5 | 139.5 | 30 | 123 |
| 26.5 | 113.3 | 25.0 | 77.9 | 25.5 | 149.6 | 30 | 124.2 |
| 26.5 | 115.5 | 25.0 | 75.8 | 25.5 | 150.2 | 30 | 124.5 |
| 30 | 126.9 | 28.5 | 90 | 26.5 | 158.4 | 32 | 150.4 |
| 30 | 123.7 | 28.5 | 92.3 | 26.5 | 159.1 | 32 | 150.3 |
| 40 | 123.7 | 30.5 | 125 | 29.5 | 165.9 | 32 | 151.4 |
| 40 | 121.5 | 30.5 | 123.4 | 29.5 | 164.4 | 34 | 172.4 |
| 45 | 118.7 | 31.5 | 129 | 32 | 159.7 | 34 | 172.9 |
| 45 | 120.7 | 31.5 | 130.7 | 32 | 162.7 | 34 | 172.5 |
| 50 | 118.9 | 33 | 137.7 | 32 | 164.6 | 49 | 164 |
| 50 | 118.7 | 33 | 135.4 | 46 | 156.9 | 49 | 165.3 |
| 64 | 120.7 | 44 | 144.5 | 46 | 155.7 | 49 | 163.3 |
| 64 | 117.4 | 44 | 142.8 | 46 | 158.6 | 56 | 157.6 |
| 70 | 116.8 | 49 | 148.2 | 49 | 157.7 | 56 | 157 |
| 70 | 115 | 49 | 146 | 49 | 155.9 | 56 | 159.4 |
| 78 | 103.8 | 69 | 142.2 | 49 | 157.4 | 74 | 157 |
| 78 | 103.6 | 69 | 142 | 55 | 163.8 | 74 | 157.5 |
| 89 | 104.2 | 73 | 143.6 | 55 | 163.1 | 74 | 156.5 |
| 89 | 105.3 | 73 | 145.4 | 55 | 161.2 | 80 | 154.1 |
| 92 | 95.5 | 73 | 143 | 70.5 | 165.5 | 80 | 153.9 |
| 92 | 94.8 | 76 | 138.1 | 70.5 | 158.9 | 80 | 152.8 |
| | | 76 | 139.2 | 70.5 | 161.2 | 98 | 142.7 |
| | | 76 | 138.8 | 74.5 | 158.9 | 98 | 140.2 |

| pH 5 (20/07/99) | | pH 6 (18/08/99) | | pH 6.5 (31/08/99) | | pH 7 (13/07/99) | |
|---|---|---|---|---|---|---|---|
| Time (hours) | Wet Wt. ($g \cdot l^{-1}$) | Time (hours) | Wet Wt. ($g \cdot l^{-1}$) | Time (hours) | Wet Wt. ($g \cdot l^{-1}$) | Time (hours) | Wet Wt. ($g \cdot l^{-1}$) |
| | | 94 | 139.4 | 74.5 | 159.5 | 98 | 142.5 |
| | | 94 | 133.1 | 74.5 | 157.9 | | |
| | | 94 | 133.8 | 78.5 | 160.5 | | |
| | | 104 | 119.5 | 78.5 | 159.4 | | |
| | | 104 | 125.1 | 78.5 | 159.9 | | |
| | | 104 | 120.8 | 95.5 | 151.5 | | |
| | | | | 95.5 | 147.7 | | |
| | | | | 95.5 | 150.2 | | |

Table A.8. Papain inhibition assay data for cystatin expression by Mut$^s$ *P. pastoris* and various induction pH values. The dates are the inoculation times and the MeOH:Gly ratios are the volumetric ratios of methanol to glycerol in the mixed feed. The inhibition of papain was measured in duplicate and the inhibition of papain was calculated relative to the blank samples that were treated by heating the sample to 100°C for 10 minutes to denature the cystatin and render it inactive against papain. The samples were diluted according to the dilution factor and the amount of diluted sample added to the assay mixture is indicated by the sample volume.

| pH 5, Inoculated on 20/07/99 | | | | | |
|---|---|---|---|---|---|
| Culture Time (hours) | Wet Wt. ($g \cdot l^{-1}$) | Sample Vol. (ml) | Dilution Factor | Inhibition of Papain (%) | |
| 28 | 126.0 | 0.2 | 1.0 | 0 | 0 |
| 40 | 122.6 | 0.2 | 1.0 | 2.3 | 1 |
| 45 | 119.7 | 0.2 | 1.0 | 16.4 | 15.4 |
| 50 | 118.8 | 0.2 | 1.0 | 11.4 | 10.7 |
| 64 | 119.1 | 0.2 | 1.0 | 18.2 | 20.1 |
| 70 | 115.9 | 0.2 | 1.0 | 7.2 | 14.1 |
| 78 | 103.7 | 0.2 | 1.0 | 0 | 6.6 |
| 89 | 104.8 | 0.2 | 1.0 | 0 | 0 |

Table A.8 con't

**pH 6, Inoculated on 18/08/99**

| Culture Time (hours) | Wet Wt. ($g \cdot l^{-1}$) | Sample Vol. (ml) | Dilution Factor | Inhibition of Papain (%) | |
|---|---|---|---|---|---|
| 33 | 136.6 | 0.2 | 1.0 | 0 | 0 |
| 44 | 143.7 | 0.2 | 1.0 | 64.2 | 57.9 |
| 49 | 147.1 | 0.2 | 1.0 | 70.9 | 71.8 |
| 69 | 142.1 | 0.2 | 1.0 | 58.9 | 59.5 |
| 73 | 144.0 | 0.2 | 1.0 | 45.1 | 51.6 |
| 76 | 138.7 | 0.2 | 1.0 | 32 | 37.6 |
| 94 | 135.4 | 0.2 | 1.0 | 3 | 0 |
| 104 | 121.8 | 0.2 | 1.0 | 6.4 | 10.4 |
| 118 | 117.9 | 0.2 | 1.0 | 10.4 | 8.5 |

Table A.8 con't

**pH 6.5, Inoculated on 31/08/99**

| Culture Time (hours) | Wet Wt. ($g \cdot l^{-1}$) | Sample Vol. (ml) | Dilution Factor | Inhibition of Papain (%) | |
|---|---|---|---|---|---|
| 27.5 | 158.8 | 0.2 | 1.0 | 0 | 0 |
| 29.5 | 165.2 | 0.2 | 1.0 | 9.3 | 6.8 |
| 32 | 162.3 | 0.2 | 1.0 | 9.5 | 7.2 |
| 46 | 157.1 | 0.2 | 1.0 | 60.2 | 60.4 |
| 49 | 157.0 | 0.2 | 1.0 | 58.7 | 58.5 |
| 55 | 162.7 | 0.2 | 1.0 | 50.7 | 50.4 |
| 70.5 | 161.9 | 0.2 | 1.0 | 39.6 | 39.9 |
| 74.5 | 158.8 | 0.2 | 1.0 | 38.1 | 37.6 |
| 78.5 | 159.9 | 0.2 | 1.0 | 33.1 | 34 |

Table A.8 con't

**pH 7, Inoculated on 13/07/99**

| Culture Time (hours) | Wet Wt. ($g \cdot l^{-1}$) | Sample Vol. (ml) | Dilution Factor | Inhibition of Papain (%) | |
|---|---|---|---|---|---|
| 49 | 172.6 | 0.2 | 1.0 | 0 | 0 |
| 56 | 158.0 | 0.2 | 1.0 | 47 | 49 |
| 74 | 157.0 | 0.2 | 1.0 | 50.9 | 51.5 |
| 80 | 153.6 | 0.2 | 1.0 | 55.2 | 53.4 |
| 98 | 141.8 | 0.2 | 1.0 | 19.6 | 16.4 |

## *Mixed Feeding Data*

Table A.9. Papain inhibition assay data for cystatin expression by Mut$^S$ *P. pastoris* and various mixed feeding ratios. The dates are the inoculation times and the MeOH:Gly ratios are the volumetric ratios of methanol to glycerol in the mixed feed. The inhibition of papain was measured in duplicate and the inhibition of papain was calculated relative to the blank samples that were treated by heating the sample to 100°C for 10 minutes to denature the cystatin and render it inactive against papain. The samples were diluted according to the dilution factor and the amount of diluted sample added to the assay mixture is indicated by the sample volume.

| **MeOH:Gly = 1:0.75 (vol:vol), pH 6, Inoculated on 08/09/99** | | | | | |
|---|---|---|---|---|---|
| **Culture Time** | **Wet Wt.** | **Sample Vol.** | **Dilution** | **Inhibition of Papain** | |
| **(hours)** | **($g \cdot l^{-1}$)** | **(ml)** | **Factor** | **(%)** | |
| 26 | 162.6 | 0.1 | 1.0 | 0 | 0 |
| 28 | 172.0 | 0.1 | 1.0 | 5.5 | 1 |
| 31.5 | 176.5 | 0.1 | 1.0 | 17.8 | 12.4 |
| 33 | 182.2 | 0.1 | 1.0 | 26.4 | 21.4 |
| 46.5 | 209.4 | 0.1 | 5.0 | 19.6 | 16.4 |
| 81.5 | 270.2 | 0.1 | 5.0 | 32.5 | 29.9 |
| 100 | 311.2 | 0.1 | 5.0 | 32 | 31.6 |
| 107 | 309.0 | 0.1 | 5.0 | 35.8 | 31.1 |
| 118 | 320.5 | 0.1 | 5.0 | 39.6 | 36.1 |
| 127 | 317.1 | 0.1 | 5.0 | 35.1 | 33.3 |

Table A.9 con't

| **MeOH:Gly = 1:1.25 (vol:vol), pH 6, Inoculated on 15/09/99** | | | | | |
|---|---|---|---|---|---|
| **Culture Time** | **Wet Wt.** | **Sample Vol.** | **Dilution** | **Inhibition of Papain** | |
| **(hours)** | **($g \cdot l^{-1}$)** | **(ml)** | **Factor** | **(%)** | |
| 25 | 161.8 | 0.1 | 1.0 | 0 | 0 |
| 27 | 187.2 | 0.1 | 1.0 | 5.5 | 1 |
| 30 | 194.3 | 0.1 | 1.0 | 17.8 | 12.4 |
| 36 | 211.6 | 0.1 | 5.0 | 26.4 | 21.4 |
| 51 | 261.7 | 0.1 | 5.0 | 19.6 | 16.4 |
| 71 | 314.9 | 0.1 | 5.0 | 32.5 | 29.9 |
| 80 | 341.3 | 0.1 | 5.0 | 32 | 31.6 |
| 95 | 379.9 | 0.1 | 5.0 | 35.8 | 31.1 |
| 102 | 379.8 | 0.1 | 5.0 | 39.6 | 36.1 |

Table A.9 con't

**MeOH:Gly = 1:2.25 (vol:vol), pH 6, Inoculated on 21/09/99**

| Culture Time (hours) | Wet Wt. ($g \cdot l^{-1}$) | Sample Vol. (ml) | Dilution Factor | Inhibition of Papain (%) | |
|---|---|---|---|---|---|
| 26 | 150.4 | 0.1 | 1.0 | 0 | 0 |
| 34 | 227.9 | 0.1 | 1.0 | 5.5 | 1 |
| 45.5 | 296.4 | 0.1 | 1.0 | 17.8 | 12.4 |
| 51 | 321.0 | 0.1 | 1.0 | 26.4 | 21.4 |
| 56 | 340.0 | 0.1 | 1.0 | 19.6 | 16.4 |
| 69.5 | 392.3 | 0.1 | 1.0 | 32.5 | 29.9 |
| 74 | 412.9 | 0.1 | 1.0 | 32 | 31.6 |
| 78 | 421.0 | 0.1 | 1.0 | 35.8 | 31.1 |

Table A.9 con't

**MeOH:Gly = 1:0 (vol:vol), pH 6.5, Inoculated on 31/08/99**

| Culture Time (hours) | Wet Wt. ($g \cdot l^{-1}$) | Sample Vol. (ml) | Dilution Factor | Inhibition of Papain (%) | |
|---|---|---|---|---|---|
| 33 | 136.6 | 0.2 | 1.0 | 0 | 0 |
| 35.5 | 138.0 | 0.2 | 1.0 | 5.5 | 1 |
| 37.5 | 138.0 | 0.2 | 1.0 | 17.8 | 12.4 |
| 44 | 143.7 | 0.2 | 1.0 | 26.4 | 21.4 |
| 49 | 147.1 | 0.2 | 1.0 | 19.6 | 16.4 |
| 69 | 142.1 | 0.2 | 1.0 | 32.5 | 29.9 |
| 73 | 144.0 | 0.2 | 1.0 | 32 | 31.6 |
| 104 | 121.8 | 0.2 | 1.0 | 35.8 | 31.1 |
| 118 | 117.9 | 0.2 | 1.0 | 39.6 | 36.1 |

Table A.9 con't

**MeOH:Gly = 1:0 (vol:vol), pH 6, Inoculated on 25/10/99**

| Culture Time (hours) | Wet Wt. ($g \cdot l^{-1}$) | Sample Vol. (ml) | Dilution Factor | Inhibition of Papain (%) | |
|---|---|---|---|---|---|
| 26 | 165.3 | 0.1 | 1.0 | 0 | 0 |
| 28 | 169.8 | 0.1 | 1.0 | 5.5 | 1 |
| 46 | 170.6 | 0.1 | 2.0 | 17.8 | 12.4 |
| 51 | 165.5 | 0.1 | 1.0 | 26.4 | 21.4 |
| 57 | 172.3 | 0.1 | 1.0 | 19.6 | 16.4 |
| 70 | 177.6 | 0.1 | 5.0 | 32.5 | 29.9 |
| 94 | 179.2 | 0.1 | 5.0 | 32 | 31.6 |
| 120 | 190.5 | 0.1 | 5.0 | 35.8 | 31.1 |

Table A.9 con't

| MeOH:Gly = 2:0.75 (vol:vol), pH 6, Inoculated on 31/10/99 | | | | | |
|---|---|---|---|---|---|
| **Culture Time (hours)** | **Wet Wt. ($g \cdot l^{-1}$)** | **Sample Vol. (ml)** | **Dilution Factor** | **Inhibition of Papain (%)** | |
| 22 | 168.5 | 0.1 | 1.0 | 0 | 0 |
| 25 | 172.3 | 0.1 | 1.0 | 5.5 | 1 |
| 30 | 179.6 | 0.1 | 2.0 | 17.8 | 12.4 |
| 39 | 201.2 | 0.1 | 5.0 | 26.4 | 21.4 |
| 47 | 203.7 | 0.1 | 5.0 | 19.6 | 16.4 |
| 55 | 229.7 | 0.1 | 5.0 | 32.5 | 29.9 |
| 67 | 227.6 | 0.1 | 5.0 | 32 | 31.6 |
| 80 | 232.9 | 0.1 | 5.0 | 35.8 | 31.1 |
| 98 | 250.0 | 0.1 | 5.0 | 39.6 | 36.1 |
| 103 | 261.1 | 0.1 | 5.0 | 35.1 | 33.3 |
| 115 | 230.9 | 0.1 | 5.0 | 35.1 | 33.3 |
| 91 | 203.7 | 0.1 | 5.0 | 35.1 | 33.3 |

## *Wet Cell Weight Data for Mut$^+$*

Table A.10. Wet cell weights for the Mut$^+$ phenotype as a function of bioreactor pH and culture time. The date refers to the date of inoculation. Measurements were taken in duplicate or triplicate as indicated by the replicated sampling times.

| **pH 5/6 (26/09/99)** | | **pH 6 (08/10/99)** | |
|---|---|---|---|
| **Time (hours)** | **Wet Wt. ($g \cdot l^{-1}$)** | **Time (hours)** | **Wet Wt. ($g \cdot l^{-1}$)** |
| 0 | 9 | 0 | 13.8 |
| 0 | 7.4 | 0 | 14.3 |
| 1 | 11.1 | 3 | 21.2 |
| 1 | 10.3 | 3 | 10.6 |
| 3 | 21.5 | 5 | 24 |
| 3 | 19.7 | 5 | 24.8 |
| 7.5 | 32.1 | 8 | 35.1 |
| 7.5 | 30.9 | 8 | 36.6 |
| 10 | 40.2 | 17 | 80 |
| 10 | 38.1 | 17 | 80 |
| 21 | 97.7 | 21 | 99.9 |
| 21 | 98 | 21 | 101.8 |
| 22 | 111.1 | 23 | 124.9 |

| pH 5/6 (26/09/99) | | pH 6 (08/10/99) | |
|---|---|---|---|
| Time (hours) | Wet Wt. ($g \cdot l^{-1}$) | Time (hours) | Wet Wt. ($g \cdot l^{-1}$) |
| 22 | 110 | 23 | 124.2 |
| 23 | 131 | 24 | 143.7 |
| 23 | 133 | 24 | 145.3 |
| 24 | 150 | 24 | 145.9 |
| 24 | 146.5 | 26 | 152.1 |
| 24 | 147.6 | 26 | 152.7 |
| 25 | 147.7 | 26 | 154.6 |
| 25 | 145.6 | 27 | 153.7 |
| 25 | 145.3 | 27 | 155.2 |
| 27 | 145.8 | 27 | 155.3 |
| 27 | 144.8 | 29 | 152.2 |
| 27 | 147.4 | 29 | 152.5 |
| 29 | 142.2 | 29 | 152.9 |
| 29 | 142.5 | 32 | 151.3 |
| 29 | 143.4 | 32 | 151.1 |
| 35 | 145 | 32 | 151 |
| 35 | 144.9 | 46 | 179 |
| 35 | 143.8 | 46 | 179.3 |
| 45 | 167.8 | 46 | 179.8 |
| 45 | 168.4 | 50 | 189.9 |
| 45 | 168 | 50 | 193.5 |
| 51 | 184.1 | 50 | 188.4 |
| 51 | 182.5 | 55 | 199.6 |
| 51 | 180.6 | 55 | 200.8 |
| 60 | 199.6 | 55 | 201.4 |
| 60 | 198.7 | 69 | 227.6 |
| 60 | 201 | 69 | 226.1 |
| 70 | 219.8 | 69 | 224.3 |
| 70 | 221 | 72.5 | 229.1 |
| 70 | 223.3 | 72.5 | 230.2 |
| 77 | 227.9 | 72.5 | 229.2 |
| 77 | 222.1 | 76 | 239.1 |
| 77 | 226.6 | 76 | 235.7 |
| 93 | 250.3 | 76 | 237.3 |
| 93 | 247 | 93 | 257 |
| 93 | 246.2 | 93 | 262.6 |
| | | 93 | 259.8 |
| | | 99 | 278.7 |

| pH 5/6 (26/09/99) | | pH 6 (08/10/99) | |
|---|---|---|---|
| Time (hours) | Wet Wt. ($g \cdot l^{-1}$) | Time (hours) | Wet Wt. ($g \cdot l^{-1}$) |
| | | 99 | 276.3 |
| | | 99 | 276.4 |

## *Cystatin Expression Data for Mut$^+$*

Table A.11. Papain inhibition assay data for cystatin expression by Mut$^+$ *P. pastoris*. The dates are the inoculation times and the MeOH:Gly ratios are the volumetric ratios of methanol to glycerol in the mixed feed. The inhibition of papain was measured in duplicate and the inhibition of papain was calculated relative to the blank samples that were treated by heating the sample to 100°C for 10 minutes to denature the cystatin and render it inactive against papain. The samples were diluted according to the dilution factor and the amount of diluted sample added to the assay mixture is indicated by the sample volume.

| MeOH:Gly = 1:0 (vol:vol), pH 5/6, Inoculated on 26/09/99 | | | | | |
|---|---|---|---|---|---|
| Culture Time (hours) | Wet Wt. ($g \cdot l^{-1}$) | Sample Vol. (ml) | Dilution Factor | Inhibition of Papain (%) | |
| 24 | 148.0 | 0.1 | 1.0 | 0 | 0 |
| 35 | 144.6 | 0.1 | 1.0 | 5.5 | 1 |
| 45 | 168.1 | 0.1 | 1.0 | 17.8 | 12.4 |
| 60 | 199.8 | 0.1 | 1.0 | 26.4 | 21.4 |
| 70 | 221.4 | 0.1 | 1.0 | 19.6 | 16.4 |
| 77 | 225.5 | 0.1 | 1.0 | 32.5 | 29.9 |
| 93 | 247.8 | 0.1 | 1.0 | 32 | 31.6 |

Table A.11 con't

**MeOH:Gly = 1:0 (vol:vol), pH 6, Inoculated on 08/10/99**

| Culture Time (hours) | Wet Wt. ($g \cdot l^{-1}$) | Sample Vol. (ml) | Dilution Factor | Inhibition of Papain (%) | |
|---|---|---|---|---|---|
| 24 | 145.0 | 0.1 | 1.0 | 0 | 0 |
| 26 | 153.1 | 0.1 | 1.0 | 5.5 | 1 |
| 29 | 152.5 | 0.1 | 1.0 | 17.8 | 12.4 |
| 32 | 151.1 | 0.1 | 1.0 | 26.4 | 21.4 |
| 46 | 179.4 | 0.1 | 1.0 | 19.6 | 16.4 |
| 50 | 190.6 | 0.1 | 1.0 | 32.5 | 29.9 |
| 55 | 201.0 | 0.1 | 1.0 | 32 | 31.6 |
| 69 | 226.0 | 0.1 | 1.0 | 35.8 | 31.1 |
| 76 | 237.4 | 0.1 | 1.0 | 39.6 | 36.1 |
| 93 | 259.8 | 0.1 | 1.0 | 35.1 | 33.3 |
| 99 | 277.1 | 0.1 | 1.0 | 35.1 | 33.3 |