ABSTRACT

Title of Document: AUTO-INDUCTION OF GREEN FLUORESCENT PROTEIN USING THE ARA OPERON IN HIGH CELL DENSITY FERMENTATION.

Paul Henry Shriner, Doctor of Philosophy, 2012

Directed By: Associate Professor Nam Sun Wang, Department of Chemical and Biomolecular Engineering

In the classical protocol for inducing protein expression, bacteria are grown to early or mid-log phase and protein expression is chemically induced. Post induction, cell formation is reduced or ceases altogether as the cellular energy is dedicated to the production of a target protein. For maximum protein expression, cells are induced prior to cell saturation, requiring careful monitoring of the culture's growth. Cells are then harvested before media is depleted to avoid cellular metabolism shifting towards degradation of protein. In the auto-induction protocol for protein expression, glucose is fed to obtain high cell mass while simultaneously minimizing recombinant protein expression due to catabolite repression. Once glucose is depleted, repression ceases and the operon is induced by constituents remaining in the media. The cells continue to grow

via an alternate carbon source, and a targeted protein of interest is expressed. Auto-induction protocols using lactose to induce T7 RNA polymerase and unblock the T7lac promoter are well developed. Auto-induction protocols using arabinose to induce T7 RNA polymerase and lactose to unblock the T7lac promoter have also been reported.

In this study media for auto-induction of the *ara* operon without *lac* is presented. *Ara* was selected for auto-induction due to several potential advantages over the *lac* system. First, in contrast to the *lac* operon gene regulation by *ara* repression is tight. Second, high level protein expression by arabinose induction does not preclude further cell growth. Further, induction can be cycled on and off during cell growth using glucose to stimulate catabolite repression. Lastly, the *ara* operon exhibits all or nothing expression, where mutant transport systems can be used to tightly control protein production.

The utility of this *ara* auto-induction protocol is demonstrated by producing green fluorescent protein (GFP) in fed-batch fermentation. Auto-induction media in fed-batch fermentation appears to be a novel application of the auto-induction protocol. Yields of GFP were found to exceed 250% of high cell density batch cultures in less than 12 hours. The combined fed-batch and auto-induction provides a high cell density and high protein expression system that avoids the need for human monitoring and intervention.

AUTO-INDUCTION OF GREEN FLUORESCENT PROTEIN USING THE ARA OPERON IN HIGH CELL DENSITY FERMENTATION.

By

Paul Henry Shriner.

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2012

Advisory Committee:
Professor Nam S. Wang, Chair
Professor William Bentley
Professor William Weigand
Professor Jeffrey Klauda
Professor Adam Hsieh, Dean's Representative

Acknowledgements

This dissertation would not be complete without first thanking my advisor, Dr. Dr. Nam Sun Wang, for seeing my potential and pushing me to complete both my Masters and my Doctoral degrees. Without his guidance, invaluable suggestions, and unrelenting support during years of competing priorities, it wouldn't have been possible to complete my work.

Similarly, the support of family and friends is crucial to success in this kind of effort: Mom and Dad, Daniel, Alex, Jan, Erik, Pat and Beth, Lenny and Kris, and Nick. I especially appreciate Elena for her patience and understanding while I was doing lab work, and Maria for giving me a venue free of distractions to do my initial writing.

Many others contributed to this project, but I would like to thank some individuals in particular: Nate Carlberg for his recommendations and advice in HPAE instrumentation; Barney "Ben" Woodard, Director of the Bioprocess Scale-Up Facility (BSF), for use of fermentation facilities and product analysis equipment; Amy Beaven, Director of the Imaging Core Facility for the Department of Cell Biology and Molecular Genetics, for fluorescent photography; Kenneth Class, Director of the Flow Cytometry and Cell Sorting Facility at the Maryland Pathogen Research Institute, for flow cytometry. I also wish to thank Dr. Rachael M. Kiss and Dr. John E. Cronan from the University of Illinois for providing strains encoding the mutant LacYA177C transporter.

I am also appreciative of my employer the U.S. Environmental Protection Agency for partially funding my education, including my managers Sheila Frace, Mary T. Smith, Tom Wall, Julie Hewitt, and Janet Goodwin, all of whom provided moral support throughout this project.

Finally, I would also like to thank long-time mentors, professors, and committee members Dr. Willian E. Bentley and Dr. William Weigand, as well as Dr. Sheryl Ehrman, who each encouraged me to persevere in my studies; Dr. Adam Hsieh and Dr. Jeffery Klauda for generously agreeing to review my dissertation and serve on my committee with relatively short notice; Barbara Ferguson of the Graduate School for her understanding and support; and all of the members of the Department of Chemical Engineering main office, past and present.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction and Overview

1.1 Overview

To achieve high levels of protein or metabolic products from a microbiological culture, a gene of interest is often expressed from a plasmid that replicates in a stable manner and maintains itself in the cell (Watson 2006; Glick 2010). The ensuing high levels of production can be used for many purposes: preparation of fermented foods such as cheese, pickles, soy sauce, sauerkraut, wine and yogurt (Johnson 2006; McNeil 1990); production of therapeutic proteins including insulin, growth factors, or antibodies (Graumann 2006; Bailey 1986); or characterization of a protein biochemically and structurally (Ahmed 2005).

Commonly used microorganisms include prokaryotes such as bacteria and eukaryotes such as filamentous fungi and yeast (Nelson 2008; White 1999; Lengeler 1999). In general, prokaryotes have a larger surface-area-to-volume ratio giving them higher metabolic and growth rates compared to eukaryotes (Campbell 2003). Bacterial systems in particular are attractive due to low cost, high productivity, and rapid use (Terpe 2006; Shuler 2002; McNeil 1990). Consequently bacteria are often the microorganism of choice used to produce high levels of recombinant proteins or their metabolic products (Ferrer 2009; Connell 2001). The most commonly used bacteria are *Bacillus, Caulobacter* and *Escherichia coli* (Shallmey 2004; Lee 1996; Bernard 2001). Despite the challenges of expression and recovery in *E. coli* such as protein insolubility,

cytotoxicity, post-translational modifications, or inefficient translation, *E. coli* remains the most commonly used host bacterium for the large-scale production of recombinant proteins (Lee 1996; Berrow 2006).

Protein production via gene expression consists of the processes transcription and translation, where the information from genes is used to make the protein (Watson 2006; Watson 2007; Mathews 1996). During the process of transcription, the information stored in a gene's DNA (deoxyribonucleic acid) is transferred to RNA (ribonucleic acid) in the cell nucleus (Watson 2006; Schleif 2003). In translation the messenger RNA interacts with a ribosome complex, which reads the sequence of RNA bases. Each sequence of three bases, called a codon, codes for one particular amino acid. Transfer RNA (tRNA) assembles the protein, one amino acid at a time. Protein assembly continues until the ribosome encounters a "stop" codon (Watson 2007).

In order for the transcription to take place, the enzyme that synthesizes RNA, known as RNA polymerase, must attach to the DNA (Schleif 1993). Promoters contain specific DNA sequences and response elements which provide a binding site for RNA polymerase. Regulated promoters are useful tools for many aspects related to recombinant gene expression in bacteria, including high-level expression of proteins. It is important to select an appropriate promoter for high-level protein expression (Schleif 1993; Lehninger 1999), and generally, an inducible promoter is more preferable than a constitutive promoter (Brautaset 2008).

Recent advances in protein production include the use of high density cell cultures (Lee 1996; Yee 1992; Gosset 1993; Lim 2000), and the use of auto-induction media

(Studier 2005; Wei 2009; Grabski 2005; Fox 2009; Blommel 2007). High cell density is obtained in batch culture through carefully designed rich media (Studier 2005; Fox 2009; Blommel 2007). High cell density is more readily achieved in bioreactors using fed-batch operation with either a linear or an exponential feed profile (Rozkov 2006; Ramirez 1995). There are numerous advantages to fed-batch fermentation over batch:

- Degradation of target protein by proteases often encountered in the stationary phase of batch growth can be circumvented by fed-batch (Yoon 1994).
- Small-scale bioreactors have been successfully used to provide automated feeding profiles (Paalme 1994; McNeil 1990).
- Sometimes feed control during fed-batch is important due to catabolic repression (Agrawal 1989).
- Since fed-batch extends the operating time, high cell concentrations can be more easily achieved and thereby improve overall fermentation productivity (Longobardi 1994).
- Fed-batch might be the only option for fermentations with toxic products or low solubility substrates (Longobardi 1994).

Auto-induction media has a selective composition and relies on the resulting culture conditions to induce the culture (Studier 2005). For example, as initial levels of glucose are depleted, cAMP levels increase and catabolite repression no longer occurs (White 1999; Schleif 2000; Schleif 2003). The method eliminates the need to monitor the culture for optimal timing of inducer addition, and allows production to proceed

unsupervised. This method has been adopted to produce somewhat higher cell densities and considerably higher levels of protein expression compared to traditional induction. Several commercially available protein expression systems take advantage of auto-induction; these are further discussed in Chapter 3.

The experimental work described here begins with formulation of growth media for auto-inducing the *ara* operon, followed by further enhancements to obtain high cell densities. This work contrasts with previously reported uses of the *ara* operon for auto-induction because in this instance the *ara* operon is used independently, and not in combination with a *lac* inducible T7 promoter. As a novel application of auto-induction media protocols, fed-batch strategies were combined with auto-induction media for the *ara* operon. Initial results demonstrate cell density and protein levels more than twice that obtained by auto-induction media alone in batch fermentation of high cell density cultures, and an order of magnitude higher than obtained via traditional induction of batch cultures. These findings support future work and ongoing studies for a novel feed profile for fed-batch fermentation.

1.2 Problem Description

Classical induction of bacterial cell culture is the process where a chemical is added to initiate or enhance expression of a protein by orders of magnitude. Classical induction requires monitoring of the culture to a desired stage of growth before the inducing chemical is added (Schleif 1993). Additional monitoring may also be necessary

to optimize harvest time of the culture (Shuler 2002). Such levels of oversight are not preferred by production facilities (Studier 2006; Terpe 2006; personal communications), and are often less economical than automated bioreactor operation. Auto-induction, as first thoroughly described with the *lac* operon (Studier 2005), utilizes growth media designed to induce the cells as the media conditions change during growth. The auto-induction strategy successfully exploits bacteria's preference for one carbon source over another including the phenomenon of catabolite repression by glucose (White 1999; Zhang 1998; Guzman 1995; Mosher 2002; Schleif 2006). The experimental work described here focuses on auto-induction media intended specifically for use with the *ara* operon to express the reporter protein green fluorescent protein (GFP) in high cell density cultures. Once the media has been designed, the auto-induction and fed-batch fermentation techniques were combined. The desired outcome is a fully automated bioreactor that can produce high cell density cultures and correspondingly high levels of recombinant protein.

1.3 Significance

Section 6 of Chapter 4 will be of special interest to those interested in methods for measuring specific carbohydrates in complex media. The literature does not frequently refer to the methods employed here, instead relying primarily on hydrolysis and gas chromatography (GC) for measurement of individual sugars (Wrolstad 2005, Hedge 1962). In some instances, use of simple media further allows deduction of observed HPLC peaks (Blommel 2007). Similarly, current environmental lab protocols use GC.

Through personal communications and industry contacts, it was found the food and agricultural industries avoid time consuming derivatization and additional sample preparations necessary for GC, and instead use a simple and direct advanced high performance liquid chromatography (HPLC) method known as high performance anion exchange with pulsed amperometric detection (HPAE-PAD).

Chapter 5 is of key significance because the experimental work discussed in this dissertation focuses on the *ara* operon for auto-induction, whereas the majority of the work previously reported uses the *lac* operon. In those few reported instances where the *ara* operon was used, it was used along with *lac* induction of the T7 promoter and was primarily used to negate leaky expression of the *lac* operon.

The experimental work in Chapter 5 is a further contribution because it is the first reported instance of auto-induction and fed-batch techniques combined into one bioprocess. This is where the problem description is addressed, and the desired outcome of a fully automated bioreactor that can produce high cell density cultures and correspondingly high levels of recombinant protein is achieved.

1.4 Organization of this dissertation

Chapter 1 provides a brief overview of this work, and briefly describes the literature searches conducted in support of this dissertation.

Chapter 2 provides more background on the key elements of the experimental research. The background is divided into five sections:

1. gene regulation in bacteria, with a focus on lac and ara and the role of cyclic AMP;

2. autoinduction, summarizing selective media composition as the means of inducing a promoter;

3. Green Fluorescent Protein, and its use as a reporter gene;

4. all or nothing expression, the autocatalytic gene expression observed with the lac and ara operons, and;

5. large scale production of recombinant proteins, especially fermentation methods employing high density cell cultures.

Chapter 3 presents a more complete literature review on auto-induction in bacterial systems. This key review provides the basis for the first part of the experimental work. In further support of this dissertation, this review serves as the foundation for a review paper to be submitted for publication.

Chapter 4 describes the materials and methods used in the experimental work. The chapter is divided into four sections:

1. the cell line, plasmids, and medium tested for cell culture;

2. methods for fluorescence measurement;

3. methods for carbohydrate measurement, and;

4. fermentor setup and instrumentation.

Chapter 5 covers the overall findings and highlights of this research, organized into five main sections:

1. screening of cell strains and plasmids

2. demonstration of auto-induction

3. evaluation of growth media for improved auto-induction

4. cell homogeneity of GFP expression

5. auto-induced fed-batch cultures.

Chapter 6 discusses future studies that can be performed to further optimize the methods described in Chapter 5, and poses questions that can be pursued to gain additional insight into this topic.

1.5 Summary of Literature Search Parameters

This section briefly highlights the literature reviews conducted in support of this work. Initial literature reviews focused on the lactose and arabinose operons. Table 1 provides examples of Google Scholar search terms and shows how more narrow searches eventually led to the focus of this dissertation. Additional articles were reviewed but were not directly cited in this work; therefore these references are not listed in the Bibliography to this thesis.

Table 1.1 “Google Scholar” search results

“Google Scholar” Search Term	Number of Hits (Note: some results provide reference to the same article)
arabinose	106,000
arabinose operon	21,500
lactose operon	39,900
“all or nothing” gene expression	8,010

autoinduction and bacteria	4,720
arabinose and “all or nothing”	128
arabinose and autoinduction	110
fed-batch and microbial	27,900
autoinduction and fed-batch	133
autoinduction and fed-batch and arabinose	11

Access to PubMed Central[1] yields 8,643 hits for the term arabinose and 2,749 for arabinose induction, compared to 20,356 for lactose and 4,546 for lactose induction. A search of the U.S patent database shows 1,897 patents including the terms arabinose and induction since the year1976, compared to 15,067 patents addressing lactose and induction. These patents were found to generally address the commercially available protein expression systems. No further trademarks for “arabinose” or its operon were found.

[1] PubMed Central is a free digital archive of biomedical and life sciences journal literature at the U.S. National Institutes of Health (NIH), developed and managed by NIH's National Center for Biotechnology Information (NCBI) in the National Library of Medicine (NLM).

Chapter 2: Background

2.1 - Gene Regulation in Bacteria

In the 1970's it was discovered that cells do not produce the proteins necessary to facilitate cell transport or metabolism of a specific energy source when that particular energy source is not available (Wiseman 1975). It was also shown that when multiple energy sources are present, cells prioritize and consume one energy source before another (Gendron 1974). For example glucose is a preferred energy source for *E. coli*. Glucose is a highly abundant carbohydrate, and therefore it is believed metabolic pathways have evolved for efficient oxidation of glucose (Lengeler 1999). Catabolite repression is the mechanism that allows cells to adapt to glucose as the preferred energy source when other carbon sources are present (Schleif 1993). Interestingly, it is thought even the human cell expresses no more than 5 percent of its genes at a given time (Gregory 2006). In *E. coli*., the carbon sources lactose, galactose, and arabinose are only used when glucose is absent (Blommel 2007) and the strain is capable of importing and digesting the particular sugar (Lukjancenko 2010).

A repressible gene regulatory system is on except in the presence of a molecule that suppresses gene expression (Savageau 1998). In contrast, an inducible system is off except in the presence of a molecule that allows for gene expression (Watson 2006). Thus induction of gene expression can be negative, where removing a constraint allows gene expression, or positive, where some aspect must be provided to stimulate expression. Some genes are always expressed, such as those genes encoding the basic

essential functions of DNA replication, RNA transcription, or protein translation (Schleif 1993; Watson 2007). However, each known genome of *Escherichia coli (E. coli)* DNA codes for 4,000 to 5,500 genes and over 4,000 proteins (Lukjancenko 2010). Bacteria have been shown to adapt to their environment by using regulatory proteins to control the expression of groups of genes (Nelson 2008; Schleif 2003; Glick 2010; Watson 2006). Many of the genes expressed by bacteria appear to be expressed according to the food (i.e. carbon and energy) sources available in the environment. The genes are further regulated by virtue of their organization into an operon, often oriented in the same direction and immediately adjacent to each other on the chromosome (Schleif 1993; Lengeler 1999). The entire operon and thus all relevent control genes are controlled simultaneously.

Transcription of genes starts with the binding of RNA polymerase, a DNA-binding protein, which binds to a specific DNA binding site (the promoter) (Watson 2006, Watson 2007, Nelson 2008). Transcription results in a single mRNA for all of the genes coded by that operon. A repressor is a protein that inhibits RNA polymerase binding. An inducer promotes RNA polymerase binding (Schleif 2003). For example, in the *lac* operon, the lacI gene product is a repressor (Glick 2010; Brautaset 2008; Simons 1987; Sung 2007). The *ara* operon uses the araC protein to function as a either a repressor or an inducer depending on whether arabinose is present in the cell (Guzman 1995, Schleif 2000; Schleif 1993). Both of these specific operons are discussed further in the following sections.

Subsection 2.1.1 - Lactose and the lac operon

Lactose is a disaccharide found in milk. In the presence of lactose, *Escherichia coli* produces the enzymes needed to digest the sugar; lactose is hydrolysed to glucose and galactose (Simons 1987). The molecular formula for lactose is $C_{12}H_{22}O_{11}$, and the molar mass is 342.30 g/mol (Weast 1981). The official chemical name of lactose is 4-O-β-D-galactopyranosyl, D-glucopyranose (IUPAC nomenclature).

The *lac* operon is required for the transport and metabolism of lactose in *Escherichia coli* (Watson 2007; Donovan 2000). It consists of a promoter, a terminator, and an operator, plus the three structural genes Z, Y, and A that function in metabolism of lactose (Schleif 1993)(see Figure 1.2.1). *lacZ* encodes β-galactosidase (LacZ), an intracellular enzyme that breaks down lactose into glucose and galactose. *lacY* encodes β-galactoside permease (LacY), a protein that transports lactose into the cell. *lacA* encodes β-galactoside transacetylase (LacA), an enzyme that transfers an acetyl group from acetyl-CoA to β-galactosides.

The *lac* operon is regulated by several factors including the availability of glucose and lactose (Matthews 1996; Nelson 2008). The proteins for catabolism of lactose are not produced when lactose is unavailable as a carbon source (Schleif 1993; Matthews 1996; Nelson 2008). Because gene regulation by the *lac* operon was the first genetic regulatory mechanism to be thoroughly studied (and consequently well understood) it is often used as a benchmark when discussing and comparing other systems of gene regulation.

Figure 1.2.1 Structure of the *lac* operon. Image adapted from /wiki/Image:Lac_operon1.png

The genetic regulatory response to lactose first requires an intracellular regulatory protein. The *lacI* gene is always (i.e. constitutively) expressed (Novotny 1996; Schleif 1993; Matthews 1996; Nelson 2008). The *lac* regulatory protein is a repressor because it keeps RNA polymerase from transcribing the operon. If lactose is missing from the growth medium, the repressor binds very tightly to a short DNA sequence just downstream of the promoter called the *lac* operator (Schleif 1993; Matthews 1996; Nelson 2008). Repressor bound to the operator interferes with binding of polymerase to the promoter. Consequently, mRNA encoding LacZ and LacY is made at very low levels. The repressor effect is also known as negative regulation (Schleif 1993).

When lactose is present, a small amount of a lactose isomer called allolactose binds to the repressor, causing a change in its shape (Kolodrubetz 1981; Schleif 1993; Matthews 1996; Nelson 2008). The repressor is unable to bind to the operator, allowing polymerase to transcribe the *lac* genes. Allolactose is an inducer because it turns on expression of *lac* genes. When the lacZ and lacY genes are produced, they break down both lactose and allolactose. Eventually repressor is released and additional mRNA

synthesis halts. The remaining mRNA breaks down in a very short time; bacterial mRNA has a half-life of 1 to 10 minutes (Gilbert 2006).

Another lactose analog used for inducing the *lac* operon is isopropyl B-D-thio-galactoside (IPTG); unlike allolactose, IPTG is not metabolized (Schleif 1993). This may be an advantage since the concentration of IPTG is constant and the rate of expression of *lac* controlled genes can be restricted. IPTG intake is dependent on the action of lactose permease (Hanson 1998). One disadvantage of IPTG is the toxicological properties are not well understood (Sigma 2010), and to date IPTG is still handled as a potential human toxin. Further, IPTG is two orders of magnitude expensive relative to lactose (Cole-Parmer online ordering, accessed 2011).

2.1.2- Regulation by cyclic AMP

Whenever glucose is present, it is preferentially metabolized by *E. coli*, i.e. before other sugars (Watson 2006). The enzymes for glucose metabolism are always present in the cell. However, the amount of glucose present is inverse to the concentration of cyclic AMP (Mosher 2002; Zhang 1998; Schleif 2006). The *cya* gene encodes adenylate cyclase, which synthesizes cyclic AMP from ATP obtained in the glucose metabolic pathway (Timmes 2004). A second gene, *crp*, encodes a protein called catabolite activator protein (CAP) or cAMP receptor protein (CRP) (Sung 2007; Zhang 1998; Schleif 2003; Schleif 2006). Cyclic AMP binds with the catabolite activator protein that binds to the *lac* promoter and facilitates transcription. When glucose levels are high, cAMP levels are low, the catabolite activator protein is not bound, RNA polymerase

binding is poor, and transcription is low. Thus when both glucose and lactose are present in the medium, metabolism of lactose does not occur because the β-galactosidase is not produced. Conversely when glucose is low, cAMP accumulates. The cAMP and activator protein bind to the *lac* promoter, RNA polymerase binding is enhanced, and transcription proceeds. Figure 1.2.2 depicts these conditions.

Note the enzymes for lactose metabolism are made in small quantities even in the presence of both glucose and lactose (Matthew 2006) due to lactose inhibiting LacI from binding to the operator. This phenomenon of basal expression is called "leaky" expression. This basal expression may be undesirable in recombinant protein expression systems. However, leaky expression is necessary in order for the metabolism of some small amount of lactose after the glucose is consumed, but before *lac* expression is fully activated (Schleif 2003). Nevertheless, there is a delay in the log growth phases reflecting the time needed to produce these metabolizing enzymes in large quantities (Megerle 2008). This delay results in the multiple log growth phases observed when cultures are grown on media containing multiple carbon sources. These multiple log phase characteristics of cultures with two or more substrates are also referred to as "diauxic" growth (Bailey 1986; Shuler 2002).

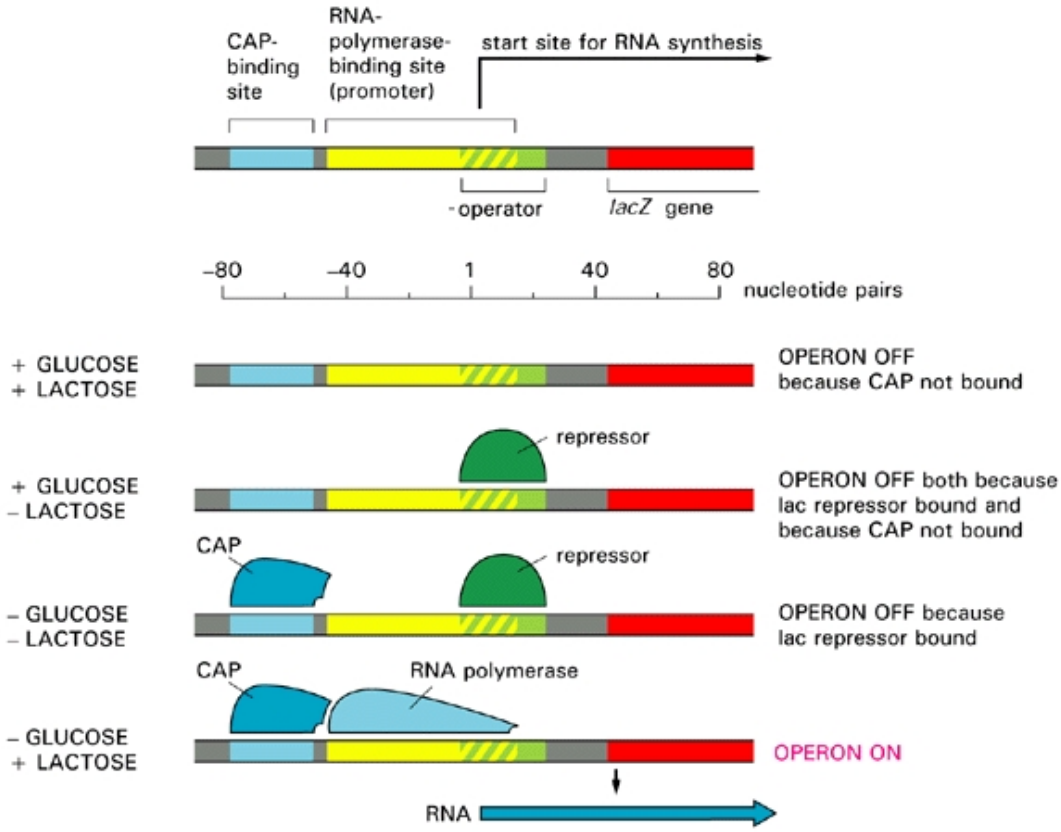


Figure 2.1.1 Regulatory control by the lactose operon. Image adapted from “The Truth About Genetics” published to the wiki by NIH.

2.1.3 - Arabinose and the ara operon

The sugar L-arabinose is a five carbon monosaccharide that can be used by bacteria as a source of carbon and energy. It has chemical formula $C_5H_{10}O_5$ and a molar mass of 150.13 g/mol. Sugars are almost always more abundant in nature as the "D" form, or structurally analogous to D-(+)-glyceraldehyde (Weast 1981). However, L-arabinose is the more common form of arabinose in nature.

The *ara* operon codes for three structural genes: ribulokinase, encoded by *araB,* phosphorylates ribulose; arabinose isomerase, encoded by *araA,* coverts arabinose to ribulose; and ribulose-5-phosphate epimerase, encoded by *araD,* converts ribulose-5-phosphate to xylulose-5-phosphate which can then be metabolized via the pentose phosphate pathway (White 1999; Schleif 2003; Ogden 1980; Schleif 2006).
In addition to *araBAD,* the *ara* operon contains two promoter sites, Pc and Pbad; two operator sites, *O1* and *O2;* and an inducer site *ara I* (Brautaset 2008; Schleif 2006). Figure 2.1.2 shows the three structural genes are arranged in an operon that is regulated by the *araC* gene protein. AraC bound to the operator site represses its own transcription from the promoter (Schleif 2000; Ogden 1980). Thus, it is an autoregulator of its own expression. In the presence of arabinose, AraC bound at this site helps to activate expression of the P_{BAD} promoter. If the concentration falls too low then transcription of *araC* resumes until the amount of AraC is sufficient to prevent more transcription again (Schleif 2000). AraC bound at the operator site simultaneously binds to the *initiator* site to repress transcription from the P_{BAD} promoter. When arabinose is present, it binds to AraC and allosterically induces it to bind to *araI* instead $araO_2$. If glucose is also absent, then the presence of CRP bound to its site between $araO_1$ and *araI* helps to break the DNA loop (Schleif 1988; Timmes 2004; Schleif 2002; Lobell 1991; Schleif 2006). Figure 2.1.3 shows the steps for expression of the arabinose operon.

Prior to metabolism, the sugar must first be transported to the cell's cytoplasm from the growth medium. Two transport systems are responsible for this uptake of arabinose. Promoter P_E controls the low affinity transport system area (Novotny 1966).

Promoter P_{FGH} controls the high affinity transport system araFGH (Kolodrubetz 1981). These transport operons are induced rapidly, within 15 to 30 seconds of arabinose addition (Johnson 1995; Guzman 1995; Schleif 2000). Induction occurs at arabinose concentrations as low as 0.53 mM.

The *ara* operon demonstrates both negative and positive control (Ogden 1986). The operon has different functions for the cAMP receptor protein (CRP). CRP does not cause RNA polymerase to bind to the promoter. In the presence of arabinose, the receptor protein promotes the rearrangement of AraC to activate transcription of the P_{BAD} promoter (Timmes et. al. 2004). The transformed AraC does this by binding simultaneously to an operator and the inducer site (Schleif 1988).

The araC protein is a dimer in solution, and it might be expected that a dimer binds to araI and another dimer binds to $araO_2$. Interestingly, one dimer binds to both sites and generates the DNA loop (Schleif 1992; Schleif 1988; Timmes 2004; Schleif 2002; Schleif 2006). It is believed DNA looping allows lower concentrations of protein to saturate their DNA-binding sites. At any rate, the intervening DNA loop prevents access to the P_{BAD} promoter (Greenblatt 1971; Lobell 1991). The P_{BAD} promoter is a very weak promoter compared to the *lac* promoter.

The arabinose operon also exhibits catabolite repression (Zhang 1998, Schleif 2003). In contrast to the lac operon, basal expression of the ara operon is quite low (Nielsen 2007). A cAMP-CRP complex must be present for expression of the arabinose operon. High levels of glucose in the environment will repress the arabinose operon due

to low levels of the cAMP molecule (Schleif 2003). This is similar to the previously discussed conditions necessary for lactose to be utilized as a carbon source. In the absence of arabinose, the uninduced expression is about 1/300 the induced level (Schleif 2000). Inductions occurs at arabinose concentrations as low as 0.53 mM (Johnson 1995).

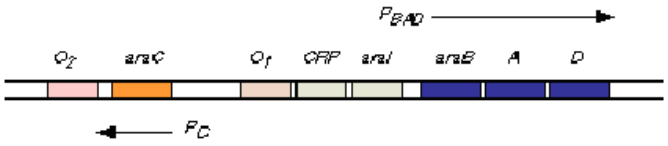


Figure 2.1.2 Structure of the Arabinose Operon. This image adapted from from http://info.bio.cmu.edu/Courses/03441/AraGal/AraGal.html.

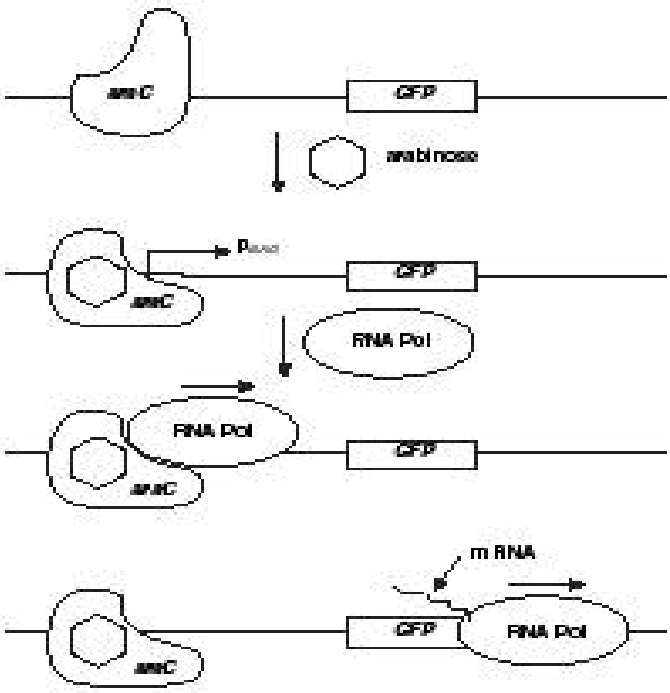


Figure 2.1.3 Expression of GFP using the Arabinose Operon. This image is adapted from "Biotechnology Explorer" by BioRad.

In summary, the key differences between the *lac* and *ara* operons are:

1. the positive control of the protein araC serves the dual function of repressor and activator;

2. araC represses its own transcription;

3. the *ara* operon regulates via DNA looping.

The pGLO plasmid (BioRad) has re-engineered the *ara* operon to delete the genes coding for arabinose catabolism (araB, A, and D). In place of araBAD, the genes for a mutant green fluorescent protein (GFP) are present. In the presence of arabinose, araC promotes RNA polymerase binding and GFP is expressed (Life Science Education, Biorad). While transport of arabinose alone is sufficient to achieve repression of its own operon, metabolism of arabinose must occur to achieve permanent catabolite repression of the *ara* operon (Gendron 1974). This effect is called "self-catabolite repression." The pGLO plasmid does not consume arabinose, therefore self-catabolite repression does not occur.

Section 2.2 – Autoinduction

As previously discussed in Section 2.1 Gene Regulation, an inducible system is turned off except in the presence of a molecule that allows for gene expression. Typically, after induction has occurred the cells will use most of their resources for the production of the target protein, cell growth rates decrease, and in some instances will not grow much further. Therefore conventional induction occurs by growing the cells to the desired level, and then adding inducer or otherwise changing culture conditions to induce the target protein.

One commonly used gene expression system consists of the lac operon in BL21 E. coli cells with IPTG addition to induce protein expression (Connell 2001; Keasling 1990; Sung 2007; Sorenson 2005; Nielson 2007). Bacteria are grown to mid-log phase, and expression is induced by the IPTG addition. Cultures are monitored frequently to

prevent saturation of the culture prior to induction. Cell multiplication is largely discontinued as the cells energy is dedicated to protein production under control of the T7 promoter. Another limitation of this system is the plasmids are unstable in BL21 cells grown on LB medium. In response to these limitations, Studier and colleagues developed tailored growth media that led to the auto-induction methods (Studier 2005).

The auto-induction occurs by using a selective media composition including glucose and lactose. More specifically, normally the lactose present in the media would have induced the cells. The presence of glucose prevents the induction by lactose. (As previously discussed, glucose causes low cAMP levels and triggers catabolite repression.) Glucose is transported into the cells and consumed first as it is the preferred carbon source. Once the glucose is consumed, lactose is transported into the cells whereby the lactose is able to induce the cells. In the case of *lac*, expression of T7 polymerase is induced only in the late log-phase growth, after glucose is depleted and leaving lactose to be imported as the carbon source. The method avoids addition of IPTG which is relatively expensive and potentially toxic; it is no longer necessary to monitor the culture; and recombinant protein production can occur unsupervised, such as overnight. This method has been reported to allow higher cell densities and higher levels of protein expression (See Chapter 3 for examples). The method was adopted in protocols using several commercially available protein expression systems. For example, Novagen's pET expression vectors use the T7 expression system is often employed for auto-induction. Auto-induction media and additives are also commercially available, such as

Formedium's AIM for use with Luria broth/ Luria-Bertani medium/ lysogeny broth (LB), Yeast Extract and Tryptone Broth (YT), Terrific Broth (TB), and Super Broth (SB) base.

Section 2.3 - Green Fluorescent Protein

Green fluorescent protein (GFP) was selected as the reporter gene in this study. GFP and its corresponding gene were originally isolated from the jellyfish, *Aequorea Victoria* (see Figure 2.3.1). GFP is fairly small at 27 kDa, and requires no cofactors or substrates (Chalfie 1994). GFP absorbs at 395 nm and 470 nm, and emits green fluorescence at 509 nm (Biotek). Therefore active GFP expression can be quickly confirmed by visual observation with a handheld ultraviolet (UV) lamp, and non-invasively monitored with UV spectrometry.

Figure 2.3.1 Fluorescence photograph of the jellyfish *Aequorea Victoria,* courtesy of AP BioTexan.

Once produced, GFP is very stable, with fluorescence maintained well beyond 24 hours after the stationary phase of cell growth (unpublished data). Maxtgen Inc. (Sanata Clara, California) introduced specific mutations to enhance fluorescence. This modified GFP gene is used by Bio-Rad in the pGLO plasmid for educational purposes (see Figure 2.3.2) to teach a variety of recombinant techniques and phenomena such as catabolite repression (Mosher 2000), plasmid transformation (Life Science Education pGLO Bacterial Transformation Kit), and phenotype identificaiton (Biotechnology Explorer by Biorad). Other sources of plasmids encoding GFP are numerous with commercially available sources including Addgene, Altogen, Invivogen, and Clontech. Besides GFP, the available fluorescent proteins include BFP (Blue Fluorescent Protein), CFP (Cyan Fluorescent Protein), YFP (Yellow Fluorescent Protein), Orange Fluorescent Protein and Photoswitchable Fluorescent Protein. A series of eight customized red fluorescent proteins and a dual expressing GFP and Red fusion protein has been commercially available since 2008 (Clontech).

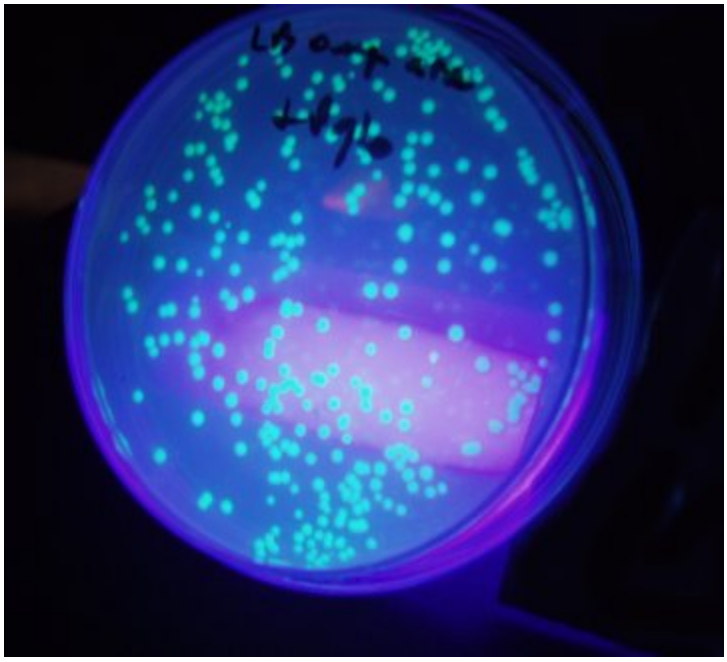


Figure 2.3.2 *E. coli* expression of pGLO.
Photo by Professor Auclair Jean Jacques

According to Wikipedia (accessed in 2010), the GFP gene has been introduced and expressed in bacteria, yeast and other fungi, zebrafish, plants, flies, and mammalian cells, including human cell lines. The reporter protein GFP has been used as a quantitative, real time and on-line indicator for protein production in multiple hosts, most relevant of which includes quantitative monitoring of fluorescence of the GFP-fusion protein in *E. coli* (Albano 1996). More recently, the time delay between inducer addition and GFP fluorescence was studied at the individual cell level (Megerle 2008). The optimal inducer concentration for maximum GFP protein production was found to correlate with protein activity (i.e. GFP fluorescence) rather than expression levels

(Geertsma 2005). In all cases, intermediate levels of the inducer L-arabinose provided the highest protein expression.

GFP has been used extensively in the Bioscaleup Facility (BSF) at the University of Maryland. Prior work has shown GFP can be produced in one day under batch conditions, and pGLO is highly suitable for high cell density cultures [unpublished data]. It was also found that reasonably high levels of fluorescence could be measured in refrigerated cell cultures for several months after harvest, demonstrating long term stability [unpublished data]. Previous studies have shown the amount of GFP protein is linearly proportional to the fluorescent intensity (Richards 2003; Remans 2000). This factor might allow near instantaneous measurement of GFP expression as an indicator of protein production under the variable growth conditions explored in this work. However, the time course studies previously mentioned show approximately 1.5 hour lag between protein expression and fluorescence emissions (Albano 1996; Lu 2002). The lengthy time lag would suggest native GFP expression as indicated by fluorescence is a poor indicator of gene expression in bioprocesses.

A further result of extensive prior studies involving GFP is that filter sets specific to GFP are readily available from commercial sources including Lightools, Nikon, Biotek,Chroma, Horiba, Motic, and Semrock. This ease of availability facilitates monitoring tools such as the fluorescence microscopy and flow cytometry discussed in Chapter 4.

Section 2.4 – All or Nothing Expression

For expression systems in which high-level gene expression is desired, expression can be induced to a maximal level in all cells of the population. For other applications, intermediate expression levels are necessary to reduce metabolic burden or to achieve specific intracellular conditions. In these cases, it may be preferable to have little or no expression in the absence of inducer, and controlled expression that varies directly with the level of inducer (Keasling 1999). The ability to control protein expression over several orders of magnitude by varying inducer concentration is common to both the *lac* and *ara* systems (Guzman 1995).

However, low level expression of wild type *lac* and *ara* does not represent a homogenous population of low expressing cells; low expression reflects a heterologous population of cells. More than 50 years ago studies explored expression of the *lac* operon in the presence of inducer concentrations less than that needed for maximal induction (subsaturating concentrations) (Novick 1957). It was demonstrated that a fraction of cells in the population was fully induced while the remainder was uninduced. Moreover, the number of fully induced cells varied directly with the concentration of inducer. It should be further noted the response of *ara* to inducer is not linear (Khlebnikov 2000). Expression systems demonstrating all-or-nothing expression contain the genes encoding the transporter under the control of the inducer. This mechanism is referred to as "all-or-none" or autocatalytic gene expression. Some expression systems based in *E. coli* exhibit this all-or-nothing (autocatalytic) induction of expression (Ozbudak 2004).

Gene expression from plasmids containing the *araBAD* promoter at low inducer concentrations exhibits all-or-nothing expression (Siegele 1997; Geertsma 2005). More specifically, under conditions of subsaturating levels of arabinose, some of the cells of the population are fully induced, whereas other cells remain uninduced (Megerle 2008). Plasmid vectors have been engineered to carry only the *araC and* P_{BAD} fragment from the *ara* operon; the genes for transport proteins have been deleted. These plasmids have been used successfully in *Escherichia coli* (Guzman 2005). In this case the plasmid demonstrates tight repression in the absence of inducer, and induction over a 1,000-fold range in the presence of inducer. Expression can also be modulated with the *ara* promoter, allowing rapid turning off and on of synthesized proteins (Guzman 1995; Siegele 1997).

For the most part, all-or-nothing expression is not important as inducer levels can be identified where high level expression is induced in all cells. In those cases where controlled low level expression is desirable, a transport gene controlled independently of the inducer can be designed. This includes use of a transport deficient strain, or a plasmid carrying an independently regulated transporter. For example, use of a mutant lactose transporter allows facilitated diffusion of arabinose, resulting in homogenous expression (Morgan-Kiss 2002), or use of IPTG inducible Ptac or Ptaclac promoter to control the araE transport protein (Khlebnikov 2000).

Section 2.5 - Large Scale Production of Recombinant Proteins

This section briefly describes batch, fed-batch, and chemostat fermentation methods of growing bacterial cultures.

2.5.1 Batch Fermentation

In the typical batch fermentation, an inoculum of the desired culture is introduced into sterile growth media (Shuler 2001). The fermentation proceeds to the desired harvest time. The microorganisms proceed through a lag phase, followed by log (or exponential) growing, then a stationary phase, and finally a decline or death phase (Figure 2.5.1) (Bailey 1986; Shuler 2001). Cells used as the starting seed may have entered a stationary phase wherein the cells have stopped growing due to substrate limitations or product inhibition. Therefore, after the initial inoculation, there is little observed growth while cells adapt to their environmental conditions. Further some new metabolic pathways may be induced upon inoculation, in which case the cells need time to readjust to the introduction of fresh or new medium (Glick 2010; Shuler 2001). Under optimal conditions, an actively growing culture used as an inoculum in fresh but identical medium may exhibit very little lag (Glick 2010).

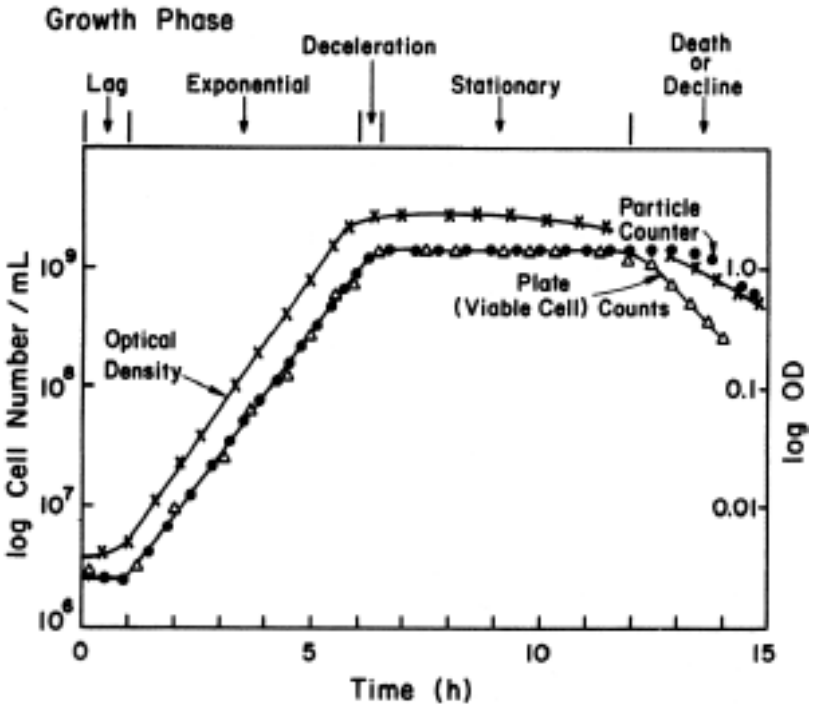


Figure 2.5.1 Typical batch fermentation cell growth curve.
Figure adopted from Shuler 2002.

With excess nutrients and no product inhibition, the specific growth rate of the cell culture is a function of the limiting substrate concentration (Bailey 1986). The rate of cell mass "X" increase over time "t" may be described as $dX/dt = uX$ where the specific growth rate $u = u_{max} S / (K_s + S)$ with a substrate concentration of "S" and a substrate specific constant Ks. For *E. coli*, u_{max} approaches 2.1, representing a doubling time of approximately 20 minutes (Nelson 2008); for comparison purposes the doubling time in LB is closer to 45 minutes (Blaber 1998), and a few hours when grown in minimal media. Therefore, after just a few hours of growth *E.coli* cells are healthy and in an ideal state to produce proteins in the early log phase. The K_S for glucose is about 1 mg/L, whereas starting glucose levels in growth media may be a thousand times higher. Thus substrate

levels equivalent to Ks will not usually be encountered until late into the fermentation (Nelson 2008). This condition is likely to be encountered much sooner in the case of auto-induction, as glucose levels initially present in the media will be much as an order of magnitude lower than in batch fermentation where glucose is the primary carbon source. Is is noteworthy that the growth rate is not linear, substrate concentrations may be low, and therefore nonlinear responses to various auto-induction media is expected.

E.coli cells grown in batch mode are often harvested at the late log phases for protein production (Bernard 2001). In the stationary phase, metabolic products may accumulate to such a high level that they are inhibitory to the cell growth or outright toxic. As the stationary phase proceeds, microorganisms can produce proteases that degrade synthesized proteins (Svensater 2001). Fewer cell resources are available to produce most proteins of interest, and the fraction of energy dedicated to protein production is lesser still, therefore this part of the batch cycle is generally avoided. With some exceptions, harvesting cells later than the log phase usually lowers protein yield and shows significant protein degradation (Zhou 1998). It is unusual to wait until the death phase for cell harvest, where energy reserves are depleted and there has been a break-down of the desired product.

2.5.2 Fed-batch

Yeast producers observed that in the presence of high concentrations of malt, ethanol was produced, and the yeast growth was restricted. The problem was then solved by a controlled feeding regime, so that yeast growth remained substrate limited (McNeil

1990). The concept was extended to the production of other products, such as some enzymes, antibiotics, growth hormones, microbial cells, vitamins, amino acids and other organic acids (Meszaros 1992; McNeil 1990). Cells are grown under a batch strategy for some time, often until close to the end of the exponential growth phase. At this point, the reactor is fed with a solution of substrates, without the removal of any culture fluid. This balanced feed addition keeps the growth of the microorganisms at a desired specific growth rate and reduces by-products that otherwise might inhibit growth, reduce target protein, or otherwise make the production system less effective (Ramirez 1995). For example, these by-products may affect the culture environment in such a way that leads to early cell death even though sufficient nutrients are available.

Fed-batch feeding strategies consist of periodic addition of substrate or a continuous feeding (Yee 1982). The feeding strategy may initially require monitoring of cell growth so that an increasing amount of substrate may be added as the cell density increases. Such metabolically based induction in fed-batch has been shown to reduce cellular stress responses (Ramirez 1995). Phased fed-batch feeding has also been shown to suppress generation of inhibitory by-products (Reisenberg 1991). Thus fed-batch is useful in achieving high concentration of products as a result of high concentration of cells for a relative large span of time. Two basic approaches can be considered: the production of a growth associated product and the production of a non-growth associated product (Fierrera 2010). In the first case, the growth phase is extended as much as possible while minimizing the changes in the fermentor that could affect specific growth rate, recombinant protein formation, and by-products. For non-growth associated

products, the fed-batch has a growth phase in which the cells are grown to the required concentration, followed by the production phase in which carbon source and other requirements for production are fed.

The case of non-growth associated products is also of particular interest for classical induction of recombinant protein, where cells are initially grown to high concentrations and then induced to express the recombinant protein (Yee 1992). For example, classical induction of protein expression in the fed batch occurs in the early log phase, at an OD_{600} of about 0.5 to 0.7. In comparison, both high cell density cultures and fed-batch cultures are induced much later in the log phase than in batch fermentation, corresponding to an OD_{600} as high as 90 (Jeung 1999).

Fed-batch fermentation processes are now commonly used in many industrial applications (Meszaros 1992). Notable examples of products derived from automated addition of nutrients to fed batch culture include:

- insulin B peptide (Gosset 1993);
- IFN-gamma (Lim 1998); and
- the monoclonal antibody tetanus toxoid (Donovan 2000).

One reason for the wide use of this technique is the highest possible cell density is most readily obtained in fed-batch. The incremental addition of nutrients during the fermentation prolongs both the log (and stationary) phases (Hewitt 2007). Thus the time window to express protein is significantly extended. Generally, fed-batch can increase recombinant protein yields by 25% to 1000% (Ramirez 1995). Improved process

control, especially computer-coupled systems, likely contributed to adoption of fed-batch fermentation.

2.5.3 Chemostat

The chemostat is a device used to maintain continuous culture. In the chemostat, fresh media is added continuously with an equivalent volume of spent media removed. By changing the rate with which media is added to the bioreactor the growth rate of the cells can be controlled. Therefore bacterial cultures can be held in the log-phase of growth for extended periods of time. The chemostat is often used to collect steady state data in order to elucidate metabolic processes (Shuler).

As an example of continuous protein production, recombinant interferon-a in *E. coli* was stably produced using a trp promoter and unrepressed by casamino acid addition (Jung 2005). The feed of glucose was varied at different dilution rates in an attempt to increase cell mass without generating excessive acetic acid. While cell concentrations were relatively low, induction strategies were circumvented and continuous production was achieved.

2.5.4 High density cell culture media formulation

Frequently the primary goal of fermentation is maximum volumetric productivity – in other words obtaining the largest amount of product in a given volume in as short a time as possible. High cell density is necessary for the highest productivity, often requiring optimized growth media to achieve the high density (Lee 1996; Reisenber

1999). However, simple fixed media often pose cell density limitations; for example, glucose is inhibitory above 50 mg/L, ammonia above 3 mg/L, and phosphorus limits growth above 10 mg/L.

Complex media such as yeast extract and tryptone can vary in exact composition, resulting in fermentations that may not be reproducible. However high density growth media are needed to support high density cell growth. In a shaker flask under normal aeration conditions, LB media can support *E.coli* growth up to a cell density of OD_{600} up to 3. TB media can grow E.coli to OD_{600} up to 8. Proprietary high density bacterial growth media can grow *E.coli* to a cell density of OD_{600} as high as 50 (Expression Technologies 2003). This is over ten times higher than LB alone. In fed-batch fermentation, cell densities exceeding OD_{600} 300 at harvest have been reported (Lee 1993; Lee 1996).

On the other hand, complex media especially those containing yeast extract provide trace metals, vitamins, and amino acids that are limited or unavailable in simple media (Lim et. al. 2000). Yeast extract may also suppress recombinant protein degradation (Yoon et. al. 1994). Glycerol is frequently preferred as a carbon source in high-cell density culture media as it does not contribute to catabolite repression of P_{araB} (Lim et. al. 2000).

The pH of the growth medium is important for *E.coli* growth rate and even more so for high cell density cultures (Don 2008). *E.coli* cells can use sugars such as glucose or glycerol as carbon and energy sources. When cells use sugars as carbon sources, they will produce acetic acid which lowers the pH (Matthews 1996). Low aeration conditions also

cause acid production. Glucose in particular is known to cause acid conditions, quickly overwhelming any buffer present at the onset of the fermentation. The optimal growth pH for *E.coli* is near neutral (Venter 2008). The minimum and maximum growth pH for *E.coli* is pH 4.4 and 9.0 respectively. The most common buffered media use phosphate to maintain the desired pH. Phosphate buffer capacity will be quickly exhausted at a high cell density (OD600 > 10) (Sorensen 2005). Fed-batch media may therefore add buffer or limit glucose additions to better control pH.

Chapter 3 – A Literature Review of Bacterial Auto-induction

An inducible expression system is off except in the presence of a molecule that allows for gene expression. Conventional induction of the expression system occurs by growing the cells to early- or mid-log phase, and then adding inducer or otherwise changing culture conditions to induce expression of the target protein. In *E. coli* for example, induction of the *lac* operon is accomplished with inducer addition to the growth media at optimal phases in the growth cycle and in the absence of glucose (Schleif 1993). Cultures are monitored frequently to prevent saturation of the culture prior to induction, and harvest generally occurs prior to or early in the stationary phase (Sorenson 2005). Typically, after induction cells will use most of their resources for the production of the target protein and will not grow much further.

The recognition that lactose based induction can be controlled by compounds that are depleted during growth led to development of auto-induction media (Studier 2005). The first reported application of auto-induction media utilizes the *lac* operon and mixtures of glucose, glycerol, and lactose. In *E. coli*, glucose is preferentially used as a carbon source (Blommel 2007). The *lac* operon is positively controlled by a receptor protein complex (cyclic AMP and a cyclic AMP receptor protein), which is strongly influenced by the presence of glucose. This phenomenon is called catabolite repression (Grossman 1998); at high levels of glucose, the protein complex levels are low, and transcription from the lac promoter is low (Schleif 1993). As glucose is depleted, the cell shifts to consumption of lactose and glycerol. Lactose import into the cell results in

allolactose, which induces the *lac* operon. Auto-induction media were specifically formulated with these three carbon sources to produce target protein automatically, with no monitoring and no chemical inducer additions. Other operons have since been used with auto-induction media, such as *ara*, *trp*, *pro*, and *rha*.

This review tracks the recent evolution of auto-induction of gene expression in bacterial systems. While auto-induction phenomena were used for several years before its formal exploration, this review begins with the first exploration of optimized media for such purposes (Monteith 2004; Studier 2005). Thus, while modulation of gene expression from the arabinose inducible *ara* promoter and numerous studies of glucose mediated catabolite repression gives insight to auto-induction techniques (Guzman 1995; Albano 1996; Khlebnikov 2002), this review only includes literature published since optimized auto-induction media for the *lac* operon was first reported in 2004 and subsequently published in 2005. The works under review here are categorized by the expression system used, with an explanation of how the experimental work compares to the others. Division of works in this manner aids in the ensuing discussion of problems identified with each system, and perhaps most importantly reveals gaps in the literature. For purposes of this review, only chemical auto-induction of bacterial systems was explored; temperature induction, for example, was not reviewed. Table 3.1 identifies the pertinent literature, noting that patents identified as part of this review were based on work already listed in the table and therefore omitted from the table.

Table 3.1 Perspective of the Contributions

Author	Date	Expression System
Monteith	2004	lac
Studier	2005	lac
Y. Wei	2005	lac
Giomarelli	2005	lac
Grabski	2005	lac
Tyler	2005	lac
Sreenath	2005	lac
Berrow	2006	lac
Studier	2006	T7 lac / ara
Blommel	2007	lac
Lee	2008	pro
Xia	2008	lac
Deacon	2008	lac
Attrill	2009	lac
T. Wei	2009	lac
Dzafic	2009	lac
Nair	2010	lac
X. Wang	2010	lac
Neerathilinga	2010	lac
Rowe	2010	ara
Steinmetz	2011	rha
Z. Wang	2011	ara

The lac operon, in comparison to other operons, is a well understood gene expression system. As shown in Table 3.1, the majority of the auto-induction studies – almost three-fourths of the works reviewed - focus exclusively on the lac operon. Not surprisingly, bacterial expression systems specifically containing components of the lac operon comprise nearly 90 percent of the expression systems used to generate information for the Protein Data Bank (PDB) (Sorenson 2005). The auto-induction method was adopted in several commercially available protein expression systems. The

T7 controlled lac systems are dominated by the pET vectors sold by Novagen and available from Promega, Stratagene, Invitrogen, and others with more than 40 commercially available plasmids available as indicated by an online patent search.

One of the reviewed papers (Berrow 2006) conducts a comparative study of several techniques for screening large numbers of cultures on a small scale, including auto-induction protocols to generate recombinant proteins. While not a new study in its own right, the comparative review demonstrates the relatively fast adoption of the auto-induction method by the research community. Several of the works report protein labeling methods for characterization, such as selenomethionine, radio labeled nitrogen or carbon, and histidine tags. Each label has its own advantages and applications, but perhaps more relevant is that the protein yields were affected by numerous variables including induction strategy, fermentation temperature, and harvest time (Giomarelli 2006). This introduces the possibility that one universal auto-induction media is insufficient for all protein production. More than one works (Deacon 2008, Wang 2010) shows the harvest time may be particularly important to auto-induced cultures as cultures are scaled up in volume. A wide range of target proteins have been expressed with the auto-induction method, including the aforementioned radio labeled and His-tagged proteins, as well as toxic proteins (Wei 2009).

In the Studier method for auto-induction, the ZYP-5052 media formulation is a rich medium based on ZY (containing NZ amine and yeast extract) media, Luria-Bertani (containing tryptone and yeast extract) or Terrific Broth (tryptone, yeast extract, and glycerol) were also reported as viable. Supplemental sugar composition consists of 0.05%

glucose, 0.5% glycerol, and 0.2% lactose. Trace metals and a phosphate buffer complete the media composition. Using the T5-*lacI* expression plasmid with a terrific broth medium supplemented with an auto-induction mixture of 0.015% glucose, 0.8% glycerol, 0.5% lactose contributed to an ~5-fold increase in protein expression compared to previous reports (Blommel 2007). Note the increased glycerol and decreased glucose in the latter medium. Defined minimal media and rich media were also described, but these media promote cell growth to saturation with little or no induction. Since these media do not serve a role in auto-induction, these media will not be discussed further.

Interestingly, the factors affecting auto-induction success and receiving the most research focus are media and culture conditions, and yet there remain different views on the ideal culture and media formulation. Early work identified minimal media as offering higher plasmid stability and thus higher protein expression than observed in Luria Bertani or similar tryptone based media (Monteith 2004). Yeast extract and tryptone based media made it difficult to maintain tight repression (Studier 2005). The auto-induction media made IPTG additions unnecessary, avoided any culture monitoring requirements, and negated the need to re-transform bacterial cultures. To that end, brand and concentration of yeast extract was systematically reviewed (Nair 2010). It is not surprising that tryptone based media are problematic when controlling the *lac* operon since tryptone is a digest of casein, or milk protein. Laboratory protocols often caution workers that lactose, a disaccharide sugar found in milk, may also be found in tryptone digests. However, since lactose is not a part of other gene expression systems, one might expect tryptone based media are better suited for auto-induction of *ara, trp,* or *rha*.

Four instances of auto-induction with a non *lac* operon were identified (Lee 2008; Rowe 2010; Steinmetz 2011; Wang 2011), including two instances of auto-induction of the *ara* operon . In the case of *ara*, auto-induction media consists of a previously defined media supplemented with carbons sources. The first instance reported a M9 minimal salts medium supplemented with 0.5 % glycerol, 0.048% glucose, and 0.5% arabinose. Early in the fermentation non-natural amino acids were added to the culture, and growth was continued overnight. In the second instance, LB medium was supplemented with 0.4% glucose. There is no indication optimized auto-induction media specific to the *ara* system has been explored. In contrast, auto-induction media specific to the propionate-inducible pPro system was developed (Lee 2008).

Experimental work shows choice of promoter and repressor, low oxygen levels, and high lacI levels also reduce the effectiveness of autoinduction (Blommel 2007). While initial work identified oxygen levels as a factor in auto-induction effectiveness (Studier 2005), only later was it documented that oxygen limitation allows higher lactose import rates. These shortcomings can be partially addressed by changing the ratios of carbon sources in the media formulation to obtain higher cell density, earlier cell uptake and use of lactose through lower initial glucose levels, and fuller utilization of all carbon sources.

Higher cell densities were further obtained using different levels of carbon sources (Grabski 2005), especially higher glycerol levels (Giomarelli 2005), or glycine (glyphosate) as a means for high-level introduction of nonstandard aromatic amino acids into a protein (Neerathilingam 2010). Throughout this review, it was discovered varied

approaches to pH control and adjustment as well as media salt concentrations were presented. While several papers addressing pH control and one evaluation of salt concentrations were described, upon review the results were variable and generally inconclusive with respect to optimized growth conditions. Indeed, the first reported use of auto-induction surmised fine tuning a rich media with so many competing factors is “not straight forwards” (Studier 2005). This could well be due to competing effects of promoter and inducer selection, lacI level controls, acetate to glycerol ratios, or other factors.

In some cases literature focuses on modifications to the widely used *lac* expression system, where the T7lac system places the gene of interest under control of the T7 phage promoter, and the T7 RNA polymerase is under the control of *lac* promoter. The lacUV5 promoter is a mutation of the wild-type lac promoter that is less sensitive to cAMP, and therefore less sensitive to catabolite repression (Larry 1971). Alternatively, the promoter and repressor were modified such that arabinose induces T7 RNA polymerase and lactose unblocks the T7 *lac* promoter (Studier 2006). While glucose in the auto-induction media suppresses induction via catabolite expression, some basal expression still occurs. Thus the primary advantage of this *ara* and *lac* combination is a further reduction in basal expression found in *lac* only systems.

In one study the expression of the protein luciferase serves as a tool for measuring basal expression. Lactose in the medium did not significantly increase basal expression of luciferase, indicating that the effects of catabolite repression and inducer exclusion prevent induction of the *lac* operon. As mentioned earlier, other promoter system

modifications include use of lacUV5 (Nair 2010; Novy) to reduce basal activity. The lac-UV5 has an increased promoter strength, decreased dependency on the CAP/cAMP complex, and overall a reduced sensitivity to catabolite repression. Finally, an *ara* controlled T7 was used to demonstrate fully induced and homogenous expression at low level inducer concentrations (Wang 2011). Similar to an approach taken to control expression of the araBAD promoter by reducing specificity of the lactose transporter (Morgan-Kiss 2002), the araFGH transport gene for high arabinose affinity was deleted, the araE low affinity transporter was enhanced, and the glucose transport gene was deleted (Wang 2011).

Two reviewed papers explore more fundament changes insomuch as the *lac* or *lac/ara* is replaced by a different gene expression system. The propionate inducible system is one option for induction to varying degrees where toxicity to the cell or controlled metabolic burden of the protein of interest is critical (Lee 2008). The rhamnose promoter in lieu of the T7 promoter exhibits stable expression but at considerably lower levels of expression; this may be useful in toxic protein expression or when it is necessary to increase correct folding of certain proteins (Steinmetz 2011).

Highest cell densities were obtained using different levels of carbon sources (Grabski 2005), especially higher glycerol levels (Giomarelli 2005), or glycine (glyphosate) as a means for high-level introduction of nonstandard aromatic amino acids into a protein (Neerathilingam 2010). Throughout this review, varied approaches to pH control and adjustment as well as media salt concentrations were presented. While all papers addressing pH control and salt levels through media formulation were described, this

review found the results to be variable and generally inconclusive. Indeed, the first reported use of auto-induction surmised fine tuning a rich media with so many competing factors is “not straight forwards” (Studier 2005). This could well be due to competing effects of promoter and inducer selection, lacI level controls, acetate to glycerol ratios, or some other factor.

It is curious that a focused study on auto-induction of arabinose could not be identified. It is possible that the role of catabolite repression and cAMP are sufficiently similar between lac and ara that no further studies are warranted. It is also possible that since auto-induction with arabinose using Studier’s media for *lac* is effective, further study of the auto-induction of *ara* is not necessary. As previously discussed, several studies used *ara* to cycle gene expression off and on. Further, the two uses of *ara* auto-induction cited in this review – one for expressing spectrally shifted bioluminescent proteins, and one rendering T7 expression glucose insensitive - are not applications that necessarily require maximum induction. Also as previously discussed, some researchers may decide that the fine tuning of rich media for each desired application is too specific to justify reporting in the literature. The increasing use of auto-induction and the development of commercially available auto-induction resources suggest this is not the case. In summary, additional research on auto-induction may be useful, particularly alternatives to the *lac* system and further systematic evaluation of auto-induction media for those alternatives.

Chapter 4: Materials and Methods

4.1 Cultures and Plasmids

E. coli K-12 is a strain that does not consume arabinose (confirmation by plating an actively growing culture on minimal media supplemented with arabinose). In addition, the following strains of *E. coli* are commonly used in the university lab and were explored in the initial phase of the experimental work: EMB1, HB101, MM294, ATCC25404. Two additional strains were briefly considered in this work: 1. MCP which expresses GFP in the presence of D-arabinose rather than L-arabinose; and 2. CW2587 (from the parent strain WM2949) which lacks arabinose transport genes and was a generous gift from Dr. Rachel Morgan-Kiss and Dr. John Cronan.

Plasmids

Two commercially available plasmids were evaluated for this study. The pGREEN (Carolina Biological Supply Company) is a 3927 basepair plasmid with ampicillin resistance after the f1 origin, and a mutant GFP protein with the lacZ fragment following the colE1 origin (Hellens).

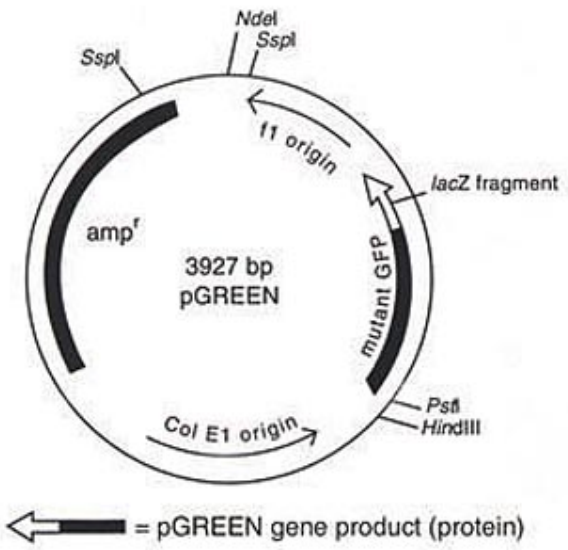


Figure 4.1.1 Map of pGREEN plasmid.
Image adapted from Carolina Biological Supply Company.

The second plasmid considered is pGLO (BioRad), a 5400 bp plasmid with pEF maintained for sugar transport, and the araC regulator for gene regulationIn pGLO the pBAD genes after the promoter (the genes for arabinose catabolism) have been deleted and replaced by the GFP gene.

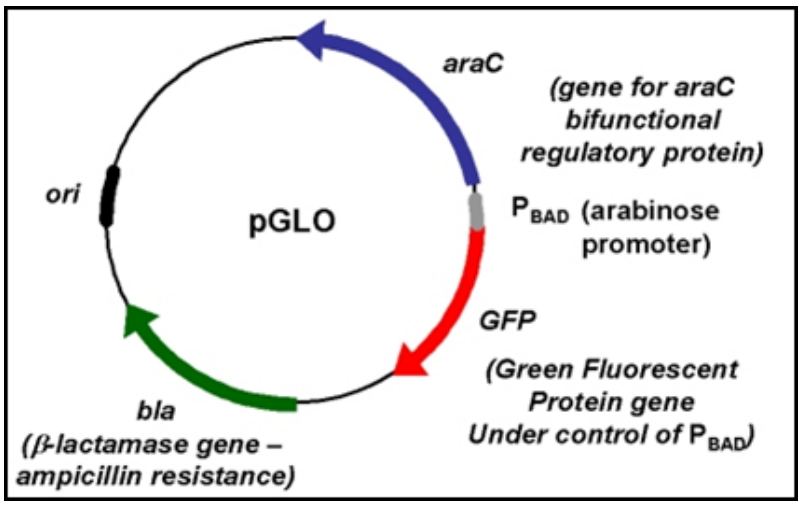


Figure 4.1.2 Map of pGLO plasmid. Image adapted from BioRad.

4.2 Media

The defined media listed here were used in one or more fermentations, and serve as the basis for media formulations evaluated for *ara* auto-induction. All media compositions are described on a per liter basis.

Lysogeny Broth (LB)

10 g tryptone
5 g yeast extract
10 g NaCl

The Lennox formulation of LB has a lower salt content (5 g/L NaCl), and is also referred to in this work as low sodium LB.

Terrific Broth (TB)

12g Tryptone
24g Yeast extract
4 ml Glycerol
100ml of a filter sterilized solution of 0.17M KH_2PO_4 and 0.72M K_2HPO_4

"x" 5052

where "x" is LB, TB, SB, ZY, or other growth media of choice
0.05% glucose
0.5% glycerol
0.2% lactose

Modified APS

15 g Martone
24 g yeast extract
9.4 g K2HPO4
2.2 g KH2PO4
5% glucose, unless otherwise specified

15-24-10

15 g Martone
24 g yeast extract
10 g tryptone
2 g NaCl
9.4 g K2HPO4
2.2 g KH2PO4

Minimal Media (MM)

1% casamino acids
0.4% glucose
2 mL of 1 M MgSO4 stock solution
0.1 mL of 1 M CaCl2
1x M9 salts

where M9 salts are diluted from 5x salts comprised as follows:

5x M9 salts
64 g Na2HPO4-7H2O
15 g KH2PO4
2.5 g NaCl
5 NH4Cl

4.3 Other chemical additions

Arabinose

L-Arabinose is the inducer for the *ara* operon. Arabinose stock solutions were filter sterilized and stored at 35 degrees C.

Glucose

Glucose (dextreose) is the preferred carbon source for *E. coli*. Glucose serves to inhibit *ara*. Glucose stock solutions were filter sterilized and stored at 35 degrees C.

IPTG

It was reported to our lab 0.5 mM IPTG added to cultures grown in LB (as well as to the modified APS media) would increase GFP production by the pGLO system (personal communication). However, at low arabinose levels IPTG is reported to be an inhibitor of the *ara* operon (Lee et. al. 2007). Initial experimental work (described in Chapter 5) yielded no discernable influence of IPTG on GFP expression, and IPTG was not further used in the growth media.

D-fucose

D-fucose is a nonmetabolizable analogue of L-arabinose. D-focuse inhibits induction of the L-arabinose operon (Wilcox 1974). Addition of L-arabinose to cells induced with d-fucose represses the operon, followed by catabolite repression (Beverin

et. al. 1971). In the initial experimental work, D-fucose did not result in GFP fluorescence, and was assumed to prevent *ara* induction.

4.4 Cell density

OD_{600} is frequently used to measure the *E. coli* cell density. The OD_{600} value is also used to determine if the *E. coli* cells are ready for making competent cells, cell stocks, induction, or for being harvested. Different cell strains may have different cell numbers at a given OD_{600} value, but $OD_{600} = 1$ usually means there are about 1×10^9 cells per ml culture. OD_{600} may also be used to calculate *E. coli* cell weight. The *E. coli* cell wet weight is measured by the cell pellet weight after the centrifugation of a cell culture (Expression Technologies 2003). Cell density at 1×10^9 per ml usually gives about 1 mg/ml cells or 1 gram/liter wet cell weight. The *E. coli* dry cell weight is usually a fourth of its wet cell weight. The following equation has been used for a general estimation of cell density, number and weight:

$OD_{600} = 1 \approx 1 \times 10^9$ cells/ml = 0.25 g/liter dry cell weight

4.5 Measuring Expression of the Green Fluorescent Protein

4.5.1 Fluorescence Photometry

UV fluorescence was measured photometrically in a Biotech Synergy HT microplate reader. The Synergy HT reader was controlled with a computer using KCL4 Data Analysis software.

Excitation and emission wavelength settings: As discussed in Chapter 2, the GFP absorbs light at 395 nm and 470 nm, and emits green fluorescence at 509. These wavelengths do not match the fixed filter settings of the BioTech Synergy HT. A culture expressing pGLO as confirmed by visual examination under a handheld UV lamp was used to systematically test each of the four emission and excitation settings on the unit. Table 4.5.1 shows peak fluorescence of the GFP was observed at the 485 / 528 setting. Results in Table 4.5.1 reflect the average of triplicate readings.

Table 4.5.1 Fluorescence of GFP at various excitation and wavelengths

	Excitation wavelength			
Emission wavelength	360 / 40	485 / 20	530 / 25	590 / 20
460 / 40	1	bandpass overlap	0	0
528 / 20	13	73	bandpass overlap	0
590 / 35	0	5	0	0
645 / 40	0	0	0	1

Sensitivity Setting: A sensitivity of 0 resulted in no measurable readings from visually fluorescing cultures. A sensitivity of 50 provided scatter of the chemically fluorescing standard with a higher variation between readings that could potentially reduce reproducible results. The default sensitivity setting of 25 was used in all subsequent measurements.

Internal Standard:

Internal Standard: 50 mM fluorescein (Sigma) in 0.1 M NaOH buffered with bicarbonate to pH 9.6 was used as a chemical fluorescence standard. Fluorescein has an excitation wavelength of 520 nm and an emission wavelength of 490 nm. Fluorescein decomposes slowly in solution, so standards were made fresh from a stock solution. All solutions were filtered to minimize from absorbing impurities. All solutions were kept refrigerated and tightly capped between uses.

The fluorescence reading was generally found to remain constant (less than 1% drift) over 24 hours of continuous Synergy HT instrument operation. However, the fluorescence measurements of the chemical standard from week to week were observed to fluctuate as much as 25%. To facilitate a comparison between fermentations, 1 mL of fluorescein was used as an internal standard.

4.5.2- Fluorescence Imaging

Initial images were taken using Zeiss Axiophot with a GFP filter, a digital camera, and 40x and 100x objective lenses. See Appendix A for the full settings. Figure 4.5.2.1 is an example of the GFP expressing E. coli cells (green) and non-expressing cells (gray).

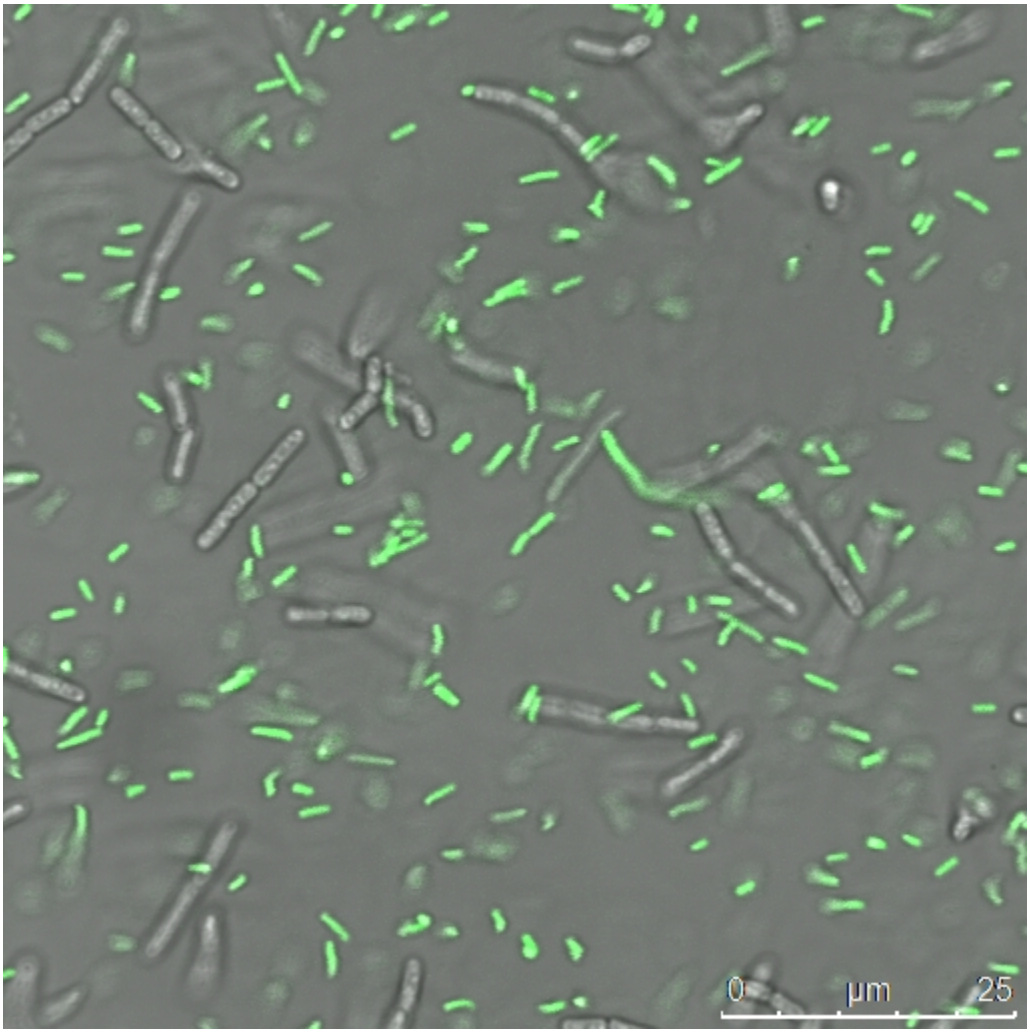


Figure 4.5.2.1 Fluorescent imaging of *E. coli* expressing GFP

4.5.3 Fluorescent Flow Cytometry

Flow Cytometry was performed using a Beckton Dickinson FACSAria (BD Biosciences). The FACSAria is a benchtop high-speed sorter equipped with three lasers providing excitation sources of 375-nm, 488-nm, and 633-nm wavelengths. Digital acquisition rates of up to 70,000 events/second are available. The laser and filter set for GFP has an excitation at 488 nm and emission at 508 nm. All samples were centrifuged and washed in phosphate buffered saline (PBS)(Fisher Scientific), and then stored cold in BD Falcon tubes (cat# 352235) in fresh PBS. Cells with no arabinose and no GFP expression were used as a control.

4.6 - Carbohydrate Measurements

4.6.1 Determination of Total Carbohydrate by Anthrone Method

The total carbohydrate content was determined by hydrolysing polysaccharides into simple sugars by acid hydrolysis and estimating the resultant concentration of monosaccharides. In the presence of anthrone, a green colored product with an absorption maximum at 630 nm is readily measured using a spectrophotometer (Hedge 1962). This method determines both reducing and non-reducing sugars because of the presence of the strongly oxidizing sulfuric acid. In this method a 1 ml sample of the solution with carbohydrate is added to 2 ml of anthrone reagent (2 g anthrone per L of 95% sulphuric acid) in a 25 mm diameter test tube. The solution is mixed by swirling, heated in a hot water bath for 10 minutes, and allowed to cool for 10 minutes. Parallel blanks containing

de-ionized and distilled water and standards of 1 to 6 mg of glucose per 100 ml of solution were run. There is a linear relationship between the absorbance and the amount of sugar that was present in the original sample. The amount of carbohydrate in the test sample is estimated from a standard curve (Figure 4.6.1).

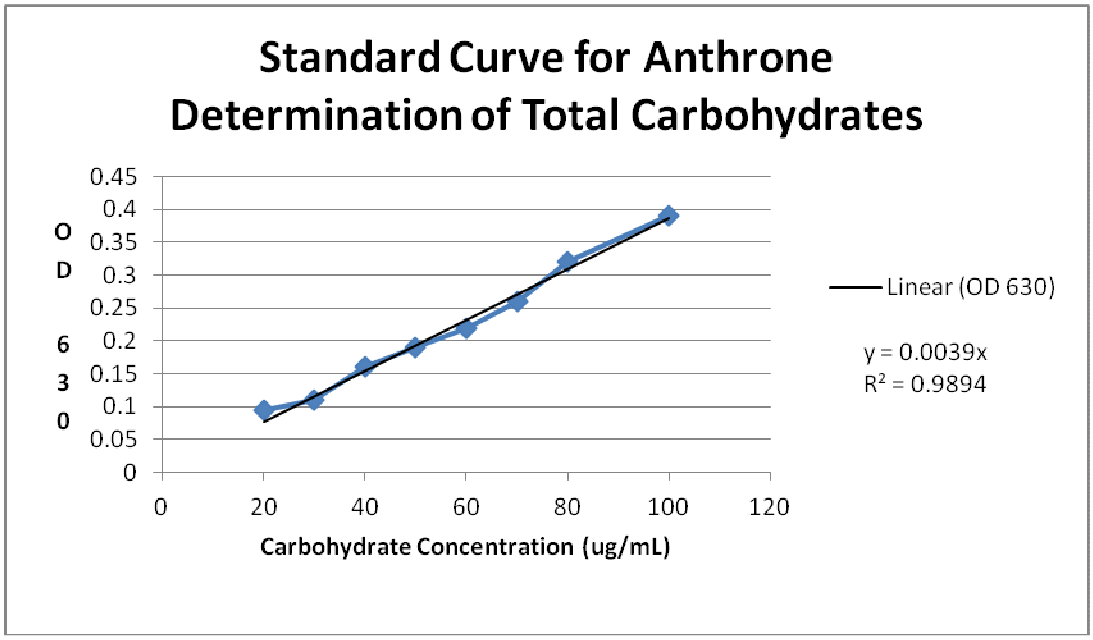


Figure 4.6.1 Standard Curve for Total Carbohydrate

4.6.2 Rapid Determination of Glucose

During early fermentations, it was desirable to rapidly identify decreases in glucose concentration. During fermentation, 2 mL aliquots were withdrawn. Glucose was measured using the YSI 2700 SELECT configured with YSI 2357 buffer and the YSI

2365 glucose oxidase enzyme membrane. Glucose concentrations that exceed 0.90% w/v (9.00 g/L) were diluted with de-ionized water (YSI Life Sciences). Results were reported as mg per liter of glucose.

4.6.3 Arabinose Measurement by Pulsed Amperometric Detection

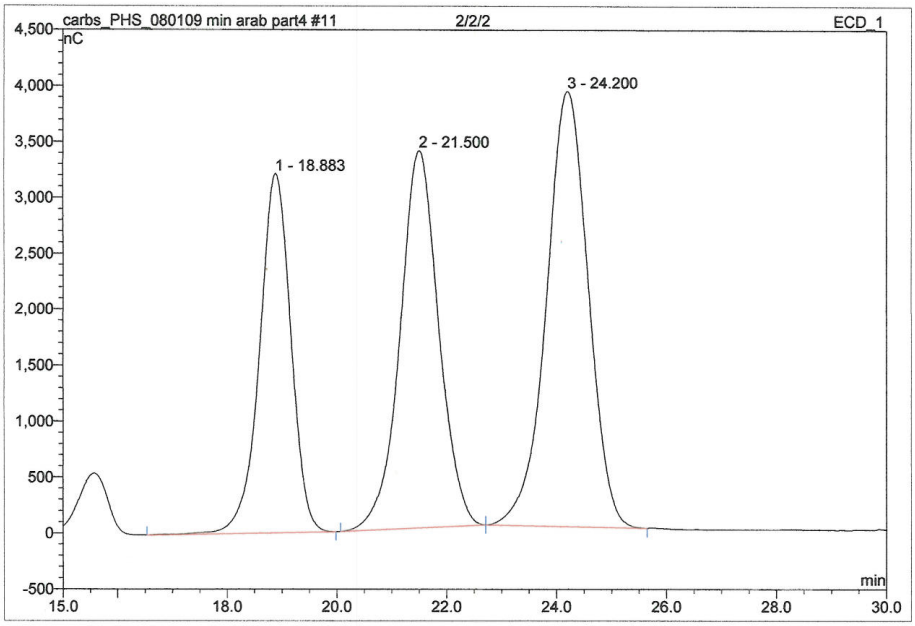
Fermentation broths are complex mixtures. Carbohydrates can be oxidized and subsequently detected by amperometry. High performance anion-exchange (HPAE) chromatography can separate individual carbohydrates in complex mixtures. Pulsed Amperometric Detection (PAD) in combination with HPAE chromatography (Dionex Technical Note 21) was used to measure mixed carbohydrates including glucose and arabinose. In this work an isocratic method was developed based on the methods for molasses characterization and beer fermentation (Dionex carboPac) and the method for carbohydrate, alcohol, and glycol in fermentation broths (Dionex Application Note 121). The HPAE instrument consisted of an autosampler AS50, the photodiode array detector PDA100, the electrochemical detector ED50, pump GS50, and the thermal compartment AS50. Chromeleon 6.60 software build 1428 was used to operate the instrument and to integrate peaks. To avoid contamination between samples, fresh vials (Dionex 055428) were used for all samples. The entire unit was cleaned and the gold electrodes (Dionex 060082) were replaced biweekly or whenever peak resolution was lost. The CarboPac MA1 column contains a resin specifically designed for sugar, sugar alcoholos, and glycol separations. The MA1 column was rinsed, capped, and then stored refrigerated between uses. 50% (w/w) sodium hydroxide was diluted with high resistivity water. Sugar standards were prepared from 10 g/L stock solutions. See Technical Note 20 for the

HPAE-PAD method of carbohydrate analysis, and Technical Note 21 for details concerning waveform A parameters used in this work. Each run consists of 0.4 ml/min of eluent A (480 nM NaOH, kept under a nitrogen blanket at 5-8 psi) for 45 minutes, followed by a step increase to 0.50 M NaOH for 5 minutes to reduce the retention time of residual carbohydrates, and concluding with a washing, regeneration, and re-equilibration cycle of 20 minutes. Chromeleon settings for carbohydrate measurements include an injection volume (sample volume) of 10 uL and a flush volume of100 uL.

Carbohydrate standards consist of glucose, arabinose, and galactose at varying concentrations (see Figure 4.6.3 for a sample chromatogram where each of the standards was mixed at 2 g/L). Pure samples of arabinose showed a column retention time of approximately 18 minutes, while pure glucose showed a retention time of 21 to 22 minutes. The salts elute in less than 10 minutes. Galactose elutes after 24 minutes and was used as an internal standard. An increase in retention time of up to 2 minutes was observed when the carbohydrate standards were mixed with a complex growth media such as LB (see Figure 4.6.3). Figure 4.6.4 shows a sample chromatogram where the internal standard was kept at 2 g/L and the remaining standards were diluted to 0.05 g/L. This generally represents the lower limit of detection for this work, as defined by visual observation of discrete peaks and where the peak height to peak width ratio was 3 or higher.

11 2/2/2

Sample Name:	**2/2/2**	*Injection Volume:*	**20.0**
Vial Number:	**1**	*Channel:*	**ECD_1**
Sample Type:	**standard**	*Wavelength:*	**n.a.**
Control Program:	**Carbs061309**	*Bandwidth:*	**n.a.**
Quantif. Method:	**gluc-arab**	*Dilution Factor:*	**1.0000**
Recording Time:	**8/2/2009 2:41**	*Sample Weight:*	**1.0000**
Run Time (min):	**15.00**	*Sample Amount:*	**1.0000**

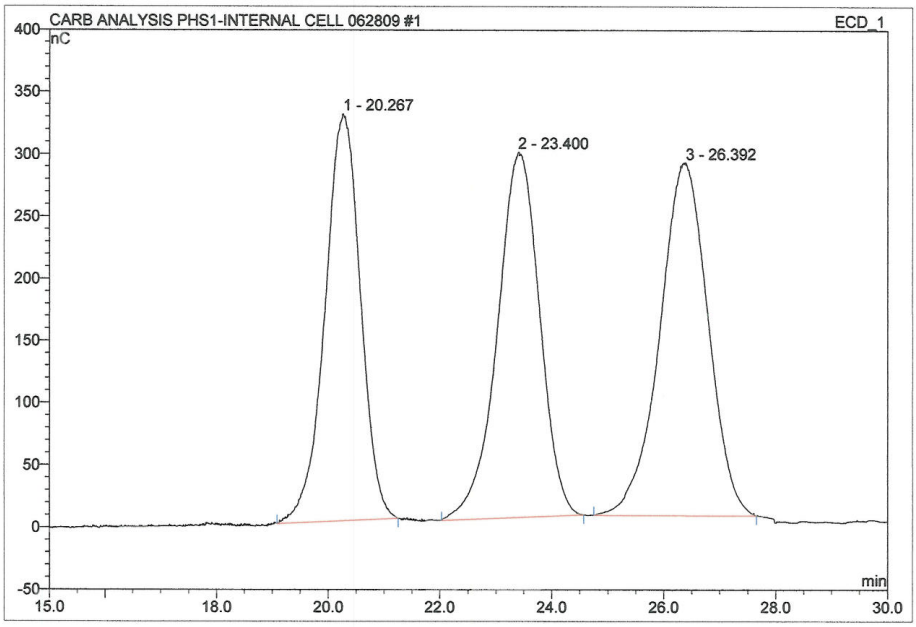


No.	Ret.Time min	Peak Name	Height nC	Area nC*min	Rel.Area %	Amount	Type
1	18.88	n.a.	3208.977	2126.560	25.58	n.a.	BMB
2	21.50	n.a.	3374.861	2752.512	33.12	n.a.	BMb
3	24.20	n.a.	3895.861	3432.842	41.30	n.a.	bMB
Total:			#########	8311.914	100.00	0.000	

Figure 4.6.2 Chromatogram for pure carbohydrate standards

1 2.0 gl, 2.0 ar, 2.0 ga in LB

Sample Name:	**2.0 gl, 2.0 ar, 2.0 ga in LB**	*Injection Volume:*	**20.0**
Vial Number:	**1**	*Channel:*	**ECD_1**
Sample Type:	**standard**	*Wavelength:*	**n.a.**
Control Program:	**Carbs061309**	*Bandwidth:*	**n.a.**
Quantif. Method:	**gluc-arab**	*Dilution Factor:*	**1.0000**
Recording Time:	**6/28/2009 13:00**	*Sample Weight:*	**1.0000**
Run Time (min):	**15.00**	*Sample Amount:*	**1.0000**



No.	Ret.Time min	Peak Name	Height nC	Area nC*min	Rel.Area %	Amount n.a.	Type
1	20.27	n.a.	327.132	240.337	30.77	n.a.	BMB
2	23.40	n.a.	294.020	257.114	32.91	n.a.	BMB
3	26.39	n.a.	284.108	283.721	36.32	n.a.	BMB
Total:			905.260	781.173	100.00	0.000	

Figure 4.6.3 Chromatogram for carbohydrate standards mixed with LB media

4.7 Fermentation Conditions and Instrumentation

4.7.1 Shaker Flask Fermentations

Baffled shaker flasks of 250, 500, and 1000 mL volumes were used for 25 mL, 50 mL, and 200 mL fermentations respectively. Flasks were capped with Bug Stopper (Whatman) and sterilized by autoclave. All shaker flask cultures were grown in a New Brunswick Innova 44 Floor Shaker incubator set to 37 degrees Celsius and 200 rpm agitation. See Figure 4.7.1.

Figure 4.7.1 New Brunswick Innova 44
(image courtesy of New Brunswick)

4.7.2 Benchtop Fermentations

The BioFlo 3000 (New Brunswick) was used for 5 L or larger cultures. The BioFlo 310 (New Brunswick) was used for 1 L and 2.5 L fermentations. See Figure 4.7.2.

Figure 4.7.2 New Brunswick Bioflo
(image courtesy of New Brunswick)

Time of Induction for Classical Protocol

In fermentations specifying manual or “classical” induction, cultures were monitored until the target OD_{600} = 0.5 was reached. A stock solution of 200 g per L of L-arabinose was used to achieve the specified media arabinose concentration.

4.7.3 Small Benchtop Chemostat Fermentations

A 1 L Erlenmeyer flask was used to design a chemostat with a 150 to 350 mL working volume. A tube for spent media withdrawal via a siphon was affixed at a specific height to provide the desired working volume. The exit pump was set to twice the inlet feed pump rate to ensure no overflow occurred. A perforated steel tube was inserted to provide aeration at all depths of the working volume. Oxygen was controlled with a pressure transducer to provide either 1 L per minute or 3 L per minute of filtered air. An exit tube with a filter allows constant pressure release. A vigorously turning two

inch magnetic stir bar was placed on the bottom of the flask to ensure rapid mixing of feed into the fermentor and to provide further aeration. pH probes were calibrated with buffer solutions of pH 7.0 and 10.0, and immersed in the spent feed withdrawal line. The flask and all feed lines were immersed in a constant temperature water bath equilibrated to 37 C. Feed pumps and aeration pumps were independently calibrated between each fermentation using a graduated cylinder and a stopwatch to measure flow rates. Figure 4.7.3 provides a schematic of this setup.

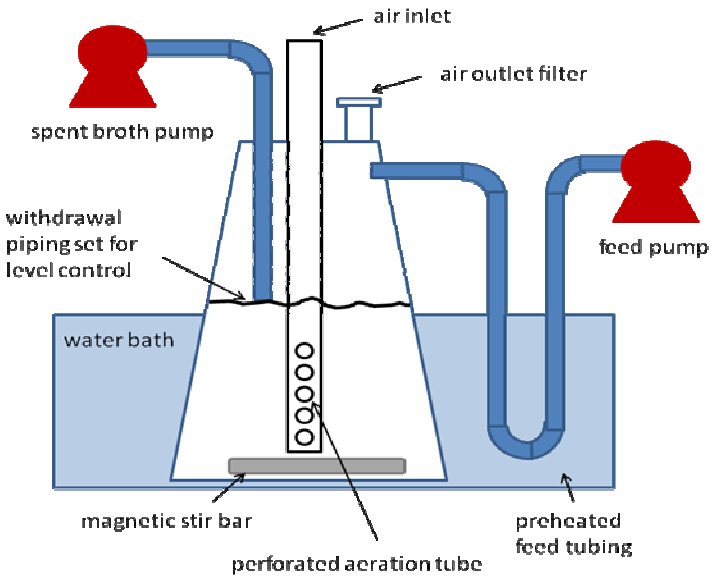


Figure 4.7.2 Chemostat Setup

4.7.4 24-well microplates

24-well microplates were aliquoted with 1.5 mL total volume of growth media and seed culture. Seed cultures were grown in 25 mL total volume in shaker flasks at 37 degrees C. When the seed culture reached the exponential phase, typically at about 3 hours, 150 to 200 uL of the seed culture was used to inoculate the microplates. Plates were kept at 37 degrees C and shaken continuously between fluorescence, glucose, and cell density (OD600) readings. All timed samples were maintained on ice until carbohydrate measurements were conducted post fermentation.

4.7.5 Fed-batch

Fed-batch fermentations were initially conducted by adding fresh growth media stepwise to a 1 mL volume microplate well. Fed-batch fermentation in shaker flasks was conducted by continuously pumping fresh media at a fixed rate into the shaker flask, or by stepwise timed additions to the shaker flask. All other growth medium and fermentation conditions are as previously described for batch growth.

4.8 Total Protein

Culture samples were diluted with four volumes of deionized water. After mixing 200 uL of diluted sample with 2 mL Bradford Reagent (BioRad), the sample was put in a 1 cm cuvette and the optical density at 595 nm was recorded. Protein was determined by comparing to a standard. Figure 4.8.1 shows the standard curve generated with 1.4 ug/mL stock bovine serum albumin (BSA).

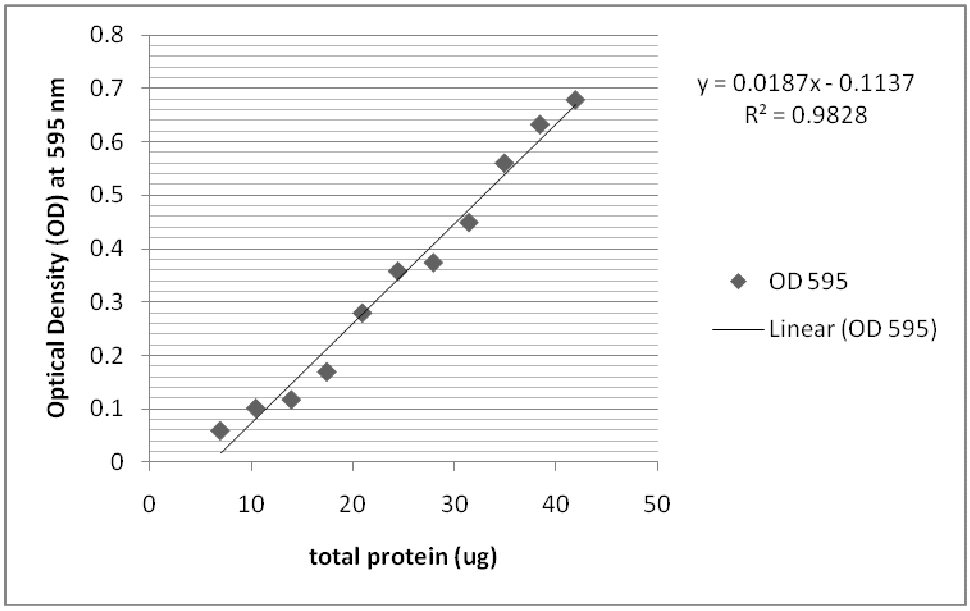


Figure 4.8.1 Standard curve for total protein using Bradford Assay.

Chapter 5: Results and Discussion

5.1 Initial Screening and Selection of Expression Systems

Each of the expression systems discussed in Chapter 4 were grown on LB agar with the appropriate antibiotics. Viable colonies were replicated on LB supplemented with 0.4% arabinose to identify GFP expression. To confirm the selected strains system would grow on glucose, each culture was grown on minimal media supplemented with 10% glucose. It was also confirmed whether each strain potentially consumed arabinose by growing each strain on minimal media supplemented with 10% arabinose. In Chapter 2, glucose was identified as preventing expression of the GFP protein even in the presence of arabinose. This was confirmed with replicates on LB agar supplemented with 0.4% arabinose and 10% glucose (data not shown).

All plates were visually observed for good colony formation and (where appropriate) GFP expression with a handheld UV lamp. Table 5.1.1 shows the observed results. Based on high levels of GFP expression in the presence of L-arabinose, no observable GFP expression in the presence of glucose, positive growth on glucose as the sole carbon source, and no growth on L-arabinose as the sole carbon source, the strain HB101 with the plasmid pGLO (BioRad) was selected for further studies.

Table 5.1.1 Visual observation of GFP for selecting Expression Systems

Strain / Plasmid	Fluorescence Observed?
MC-P	Weak*
CW2587	None
HB-pGLO	Strong
pLACy	None
pGREEN	Weak

*But expresses strongly with D-arabinose.

5.2 Demonstration of auto-induction using ara operon

pGLO was plated on LB with 0.4% arabinose, followed by replicates on LB and arabinose with a range of glucose concentrations. Growth and fluorescence were observed at 12, 24, and 36 hours. Using a handheld UV lamp, it was observed that full GFP expression occurred between 12 and 24 hours in the absence of glucose, whereas GFP expression was delayed by several hours with low to moderate levels of glucose, and no GFP expression was observed at all with the highest levels of glucose. Colony formation at each given time point was approximately the same across the range of glucose concentrations, with slightly larger colonies forming in the glucose supplemented agar. Table 5.2.1 shows the qualitative results of the rapid screening performed to confirm auto-induction.

Table 5.2.1 Visual fluorescence: quick screening to confirm autoinduction

Glucose Concentration	12 hours	24 hours	36 hours
0%	No GFP	GFP expressed	GFP expressed
0.1%	No GFP	GFP expressed	GFP expressed
0.5%	No GFP	No GFP	GFP expressed
1%	No GFP	No GFP	GFP expressed
2%	No GFP	No GFP	No GFP
5%	No GFP	No GFP	No GFP

In the next step, cell growth is measured under a variety of induction conditions. Figure 5.2.1 shows cell growth of a fermentation with HB101 cells grown on LB and using traditional induction, i.e. induction of the culture with arabinose when the OD_{600} reaches 0.5. Figure 5.2.2 shows cell growth using LB media supplemented with 0.4% glucose for auto-induction. Glucose is consumed first, and arabinose levels do not change. No fluorescence is observed until the remaining glucose levels approach 0.1%, in contrast to an almost immediate presence of fluorescence in when glucose is excluded from the growth media (see Table 5.2.2).

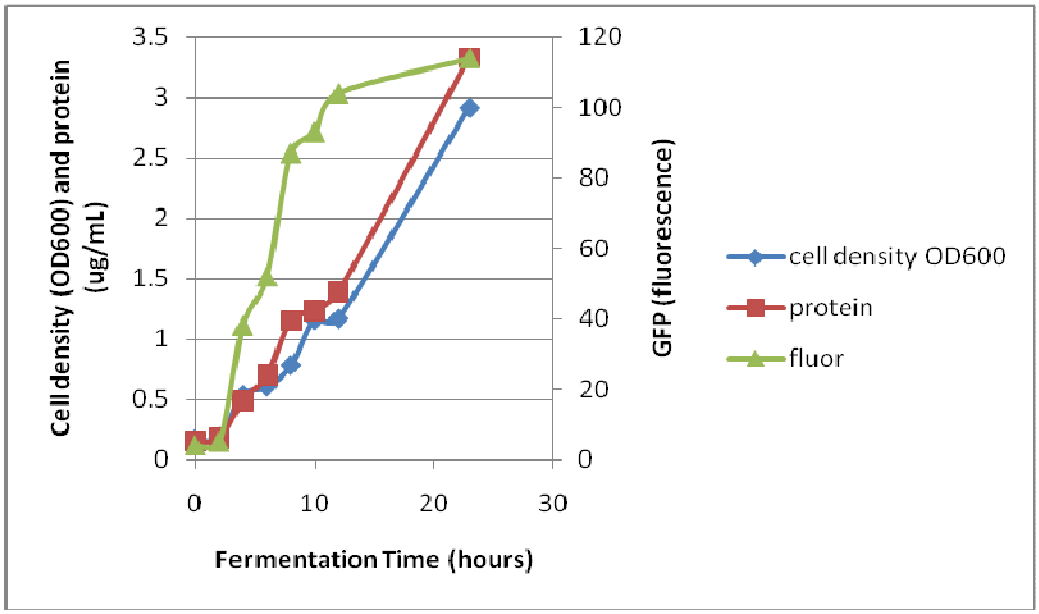


Figure 5.2.1 Classical induction of HB101 with 0.4% arabinose at 3 hours. Fermentation was performed in shaker flasks.

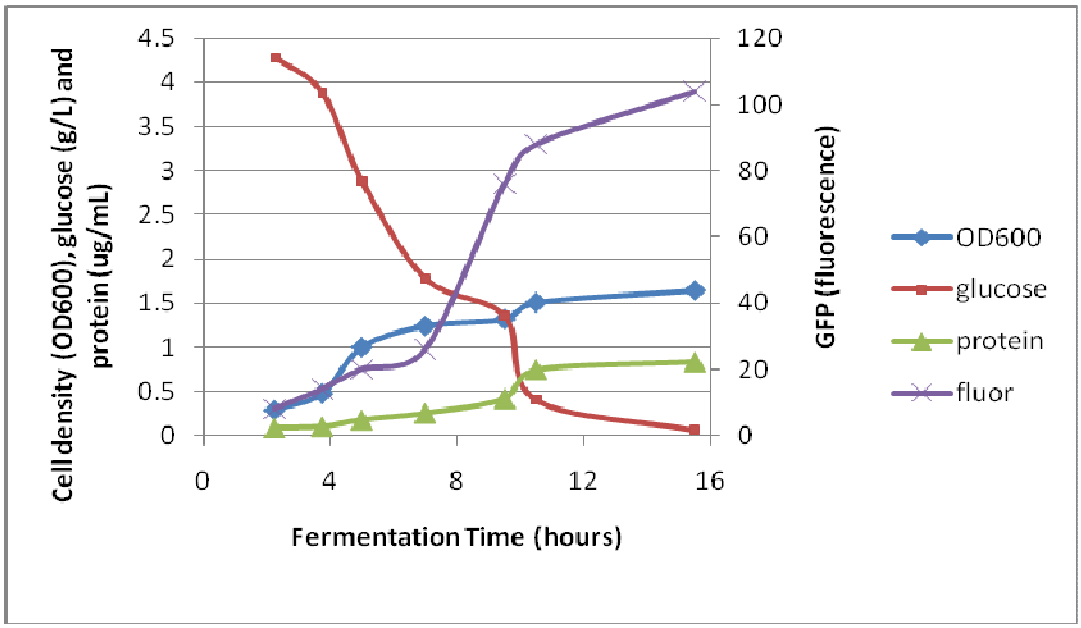


Figure 5.2.2 Auto-induction of HB101 using LB media.
Fermentation was performed in a 2L BioFlo, pH controlled to 7.0,
air flow at 4 L per minute, mixing at 275 rpm. One standard deviation is: OD600 +/- 0.59; glucose +/- 0.1; protein +/- 0.048; fluorescence +/- 16.

Table 5.2.2 compares the GFP production at harvest (t=12 hours) as a function of induction method. HB101 cells were grown in microwell plates on LB supplemented with glucose or arabinose as indicated. Note the GFP production where inducer is present immediately in the growth medium (column 3) was almost identical to the GFP production of classical induction (columns 4 and 5). This was an unexpected result. Upon further examination, it was found the starting OD_{600} after inoculation was .227, which is higher than the starting cell density of the previous BioFlo or shaker flask cultures. To rule out possible variations in inoculums as an explanation, three shaker flasks were inoculated with 5%, 10%, and 15% volume from a single seed culture. In all

three cases the measured cell density, protein, and fluorescence generally remained within 10% of each other, and never exceeded 14% variation. A likely explanation for immediate induction producing GFP similar to the classically induced log-phase induction is that the starting OD_{600} was so close to the target classical induction value of $OD_{600} = 0.5$.

Table 5.2.2 Comparison of classical induction, auto-induction, and immediate induction of HB101.

Induction Method	**None**	**Auto**	**Immediate (no lag in induction)**	**Auto**	**Induction at OD_{600} = .6**	**Induction at OD_{600} = .9**
Glucose	2%	0.1%	-	0.15%	-	-
Arabinose	-	0.1%	1%	1%	1%	1%
GFP (fluor)	21	109	247	189	246	248

Glucose and arabinose concentrations +/- 0.1 g/L; one standard deviation for fluorescence is +/- 16.

Figure 5.2.3 shows the GFP production as measured at 11 hours over a range of % arabinose levels. HB101 cells grown on 1.5 mL LB supplemented with arabinose to the final % w/v indicated. Cells grown on LB supplemented with 1% or 2% glucose did not express any active GFP (data not shown). In section 5.1 it was shown *E. coli* does not metabolize arabinose.

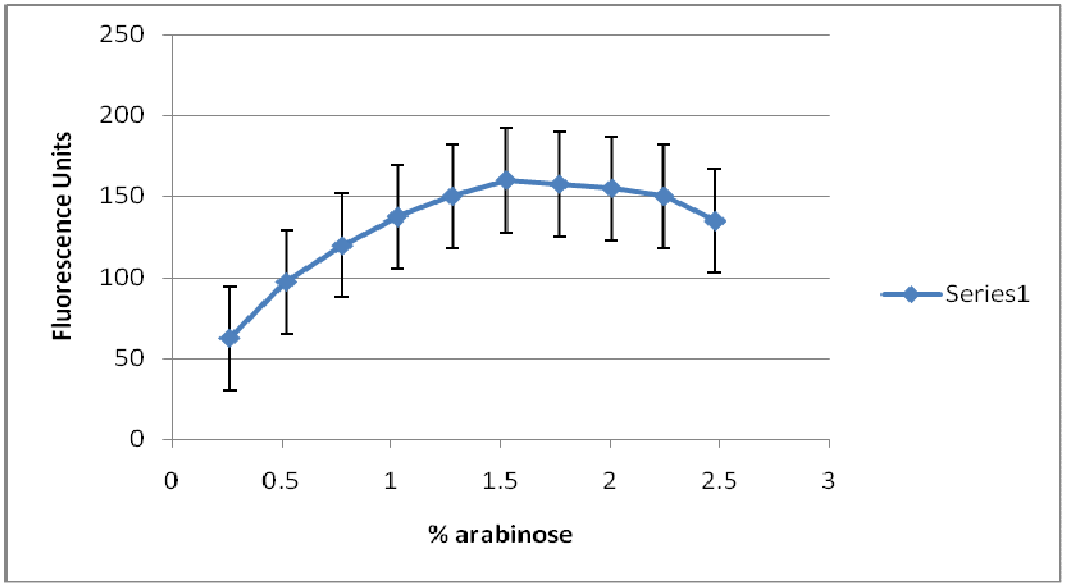


Figure 5.2.3 Effect of arabinose concentration on GFP production. Data points represent cell harvest at t=11 hours. GFP expression indicated by fluorescence. One standard deviation is shown.

Time to maximum expression

In this experiment series, LB media was supplemented with 0.1% arabinose. Duplicate shaker flasks were inoculated by 10%, 15%, and 20% volume of an actively growing culture. No additional induction was performed. As shown in Figure 5.2.4, the size of inoculum had minimal effect on the time it took to reach maximum fluorescence, and had no effect on the amount of GFP ultimately expressed. This is surprising, as it was expected larger inoculums would support shorter overall fermentation times. It may be that after approximately 4 hours, fermentor conditions had already become nutrient limited, and the size of the inoculum was therefore irrelevant.

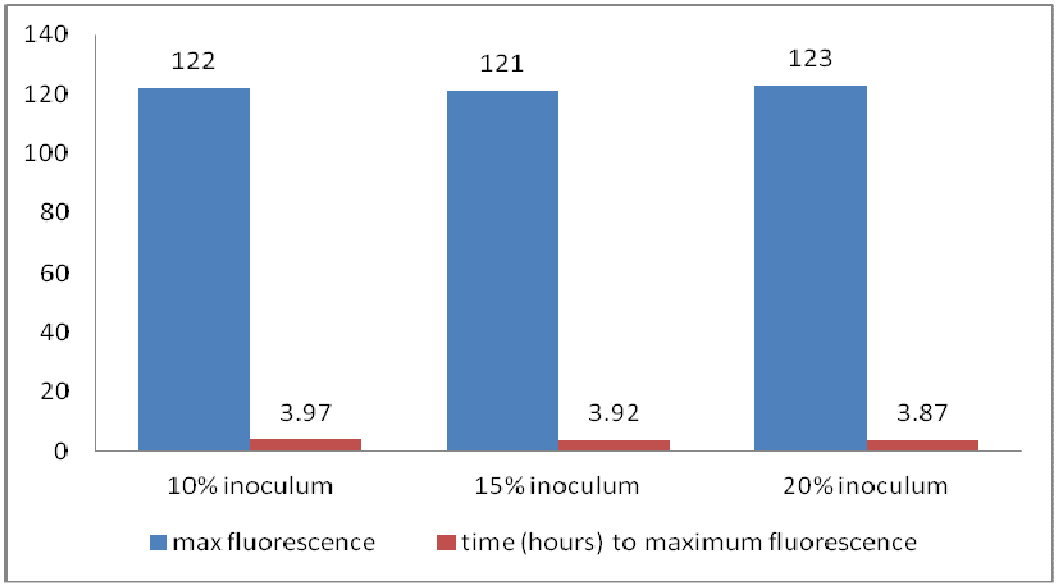


Figure 5.2.4 Effect of inoculum size on time to maximum GFP expression of immediately induced HB101 cells. One standard deviation of fluorescence is 2; time to maximum GFP expression is +/- 0.05 hrs.

Figure 5.2.5 shows the results of the same experiment when 0.2% glucose is added to the medium. In this case approximately six hours are needed to obtain the maximum fluorescence, and the fluorescence is considerably lower than the glucose free media. At first glance this is surprising. While it is expected that after cells deplete glucose, an additional lag phase will occur while the cells shift to a new carbon source, the final cell density is slightly higher and a higher level of fluorescence is expected. High density cultures are often induced later in the log phase than typical batch cultures, and in fed-batch mode a longer period of catabolite repression should be result in improved yields at harvest. Since cell growth (doubling time) on glucose is optimal, an

improved (shorter) time to harvest is also expected. These data also suggest LB alone will be unable to support a high cell density fed-batch culture.

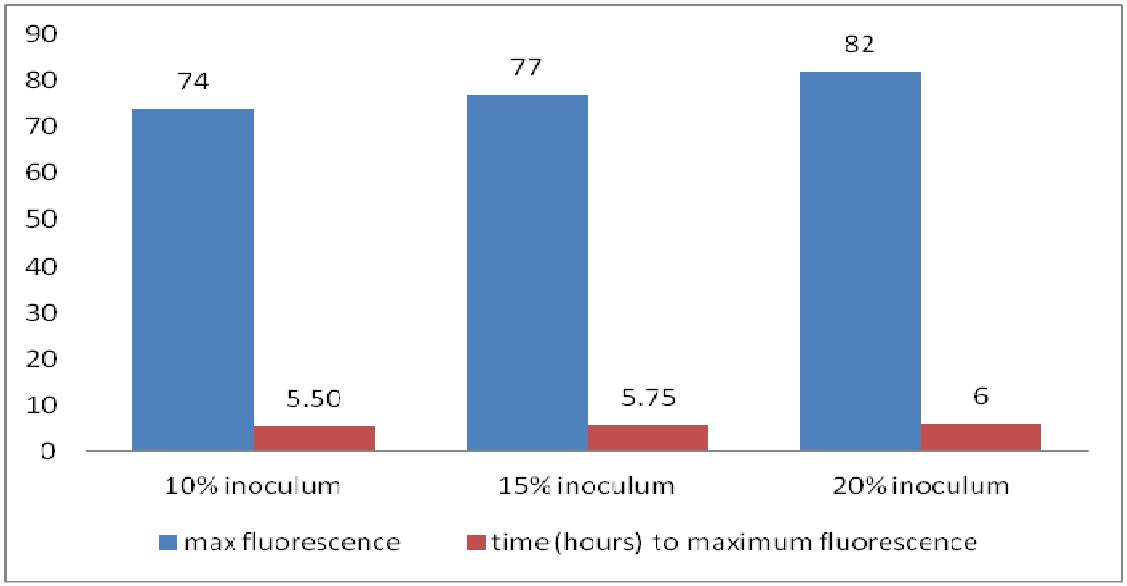


Figure 5.2.5 Effect of inoculum size on time to maximum GFP expression in auto-induced HB101 cells. One standard deviation of fluorescence is 18; time to maximum GFP expression is +/- 0.3 hrs.

5.3 Growth Media Formulation

A series of chemostat fermentations was conducted to estimate minimum arabinose concentrations necessary for full GFP expression. Dilution rates representing a residence time of 45 minutes to 2 hours were tested. No GFP expression could be obtained at any of these rates. Cells collected from the chemostat were placed in fresh LB media in shaker flasks shows GFP expression within several hours. As discussed in Chapter 2, there is a lag between GFP expression and fluorescence. It is likely the residence time was too short to allow full GFP expression. Arabinose and glucose

concentrations in the chemostat were measured using HPAE as discussed in Chapter 4. In all cases, the arabinose feed and the spent broth had identical arabinose concentrations. There was no apparent consumption or significant accumulation of arabinose in the biomass. These findings are mostly unremarkable and are therefore not presented here; the chromatography and a summary of the fermentation run conditions may be found in Appendix B. As a further result of conducting the chemostat fermentations, good agreement was found between glucose measurements obtained with the HPAE and those with the YSI glucose analyzer. In subsequent fermentations, the rapid glucose determination using the YSI was used to measure glucose decreases during fermentation.

After initial screening (see Section 5.1) and a review of the auto-induction literature for the *lac* operon (see Chapter 3), the strategy for formulating the auto-induction growth media was based on the following assumptions:

- low level glucose addition will be used to keep cAMP levels low and control expression of the *ara* operon
- glucose metabolic products increase acidity and lower overall cell densities; phosphate buffer may be useful but is unlikely to be adequate in high cell density cultures
- high salt concentrations may be detrimental to plasmid stability and may preclude high GFP expression
- moderate oxygenation is important

A matrix may be constructed to experimentally evaluate each of these assumptions. A series of batch fermentations with varying glucose levels tests the first

assumption. A comparison of growth media with and without glycerol and/or phosphate buffer was conducted to evaluate effects of pH on GFP production. Similarly, growth media with and without sodium chloride, or with varying levels of sodium chloride, were tested for effects on GFP expression. Finally, larger scale fermentations with and without oxygen cascade were compared. Results for each of these trials will be described below.

In section 5.1 and 5.2, it is demonstrated arabinose is not consumed by the chosen strain/plasmid. Therefore all cultures grown in this phase include a constant concentration of 0.4% arabinose. Next the time profile of fluorescence expression was reviewed for each of the media using 24 well microplates. The microplate reader was programmed to operate at 37 degrees with continuous agitation, with GFP fluorescence measured every 20 minutes. While the lag phase for GFP production varied in each case, the most notable observation from the protein expression curves is the maximum level of fluorescence. Figure 5.3.1 is typical of the GFP expression over time. Figure 5.3.2 displays multiple expression curves with a range of initial glucose concentrations in the medium.

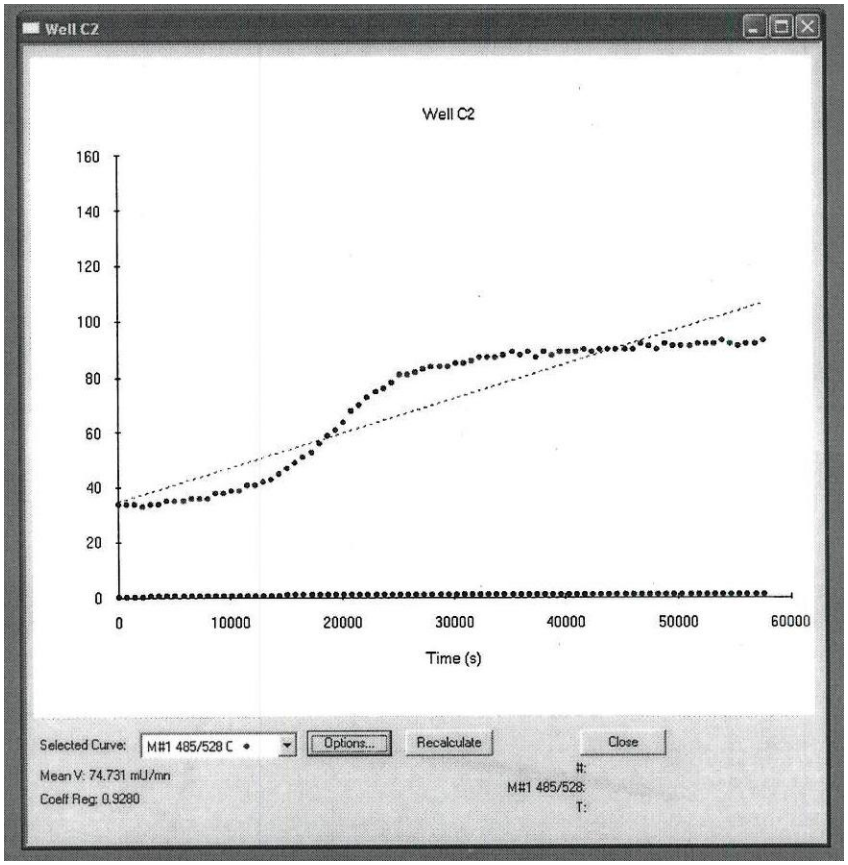


Figure 5.3.1 Typical plot of fluorescence as a function of time for tested auto-induction media.

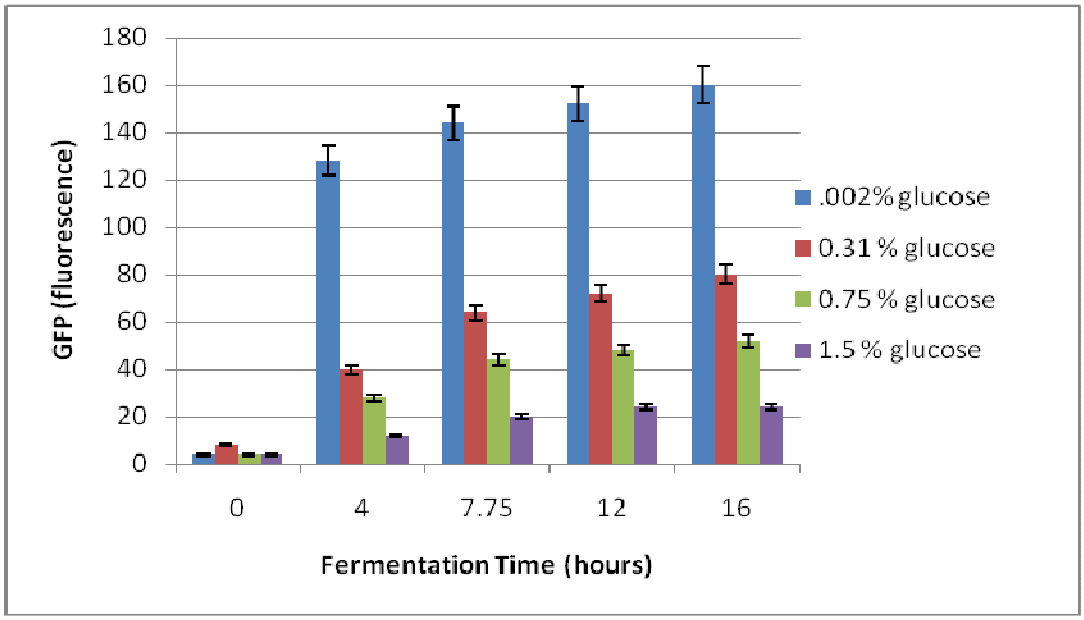


Figure 5.3.2 Fluorescence over time for several initial glucose concentrations. Error bars reflect one standard deviation.

It was already shown HB101 cell mass increases after induction. Figure 5.3.3 shows cell growth (biomass) increases in the presence of the inducer arabinose, specifically post auto-induction. A higher initial glucose concentration of 7 g/L was selected to better illustrate cell growth before and after GFP expression. Fluoresence is first measured at around 3 hours. This further illustrates successful auto-induction of the *ara* operon.

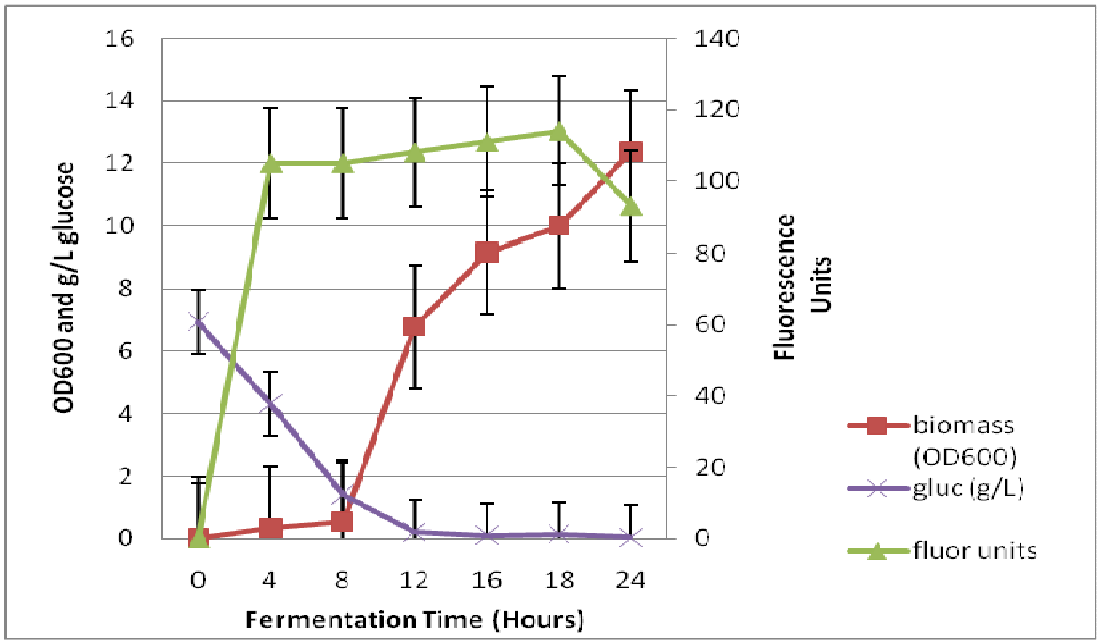


Figure 5.3.3 Cell growth in presence of arabinose. Standard error is shown with black lines with caps.

Approximate ranges of glucose and arabinose necessary for auto-induction are now identified. These two carbohydrates may be combined to supplement growth media and evaluate success and performance of these media as auto-inducing media. Therefore, cultures were grown in shaker flasks with 50 mL of various growth media. GFP expression was classically induced by the addition of 0.4% arabinose when the measured OD_{600} reached 0.5 as a benchmark for performance of auto-induction.

Statistical evaluation of effects on cell growth and GFP by yeast extract

In statistics, the *F*-distribution is a continuous probability distribution. In analysis of variance, the *F*-distribution is the distribution of the test statistic, and is used frequently to evaluate the null distribution of a test statistic (NIST, 2006). In comparing

independent samples, the F-test assumes the samples are of the same distribution with a single variance. The F-test may then be used to determine if the means of two or more populations are equal. In this section, changes in the growth media were evaluated for effects on biomass and GFP production. As an example, Table 5.3.1 provides the cell mass at harvest as measured by OD_{600}. Cell mass was observed for four treatments (i.e. four different concentrations of yeast extract added to modified APS media supplemented with 0.8 g/L glucose and 1.5 g/L arabinose) were each evaluated three times, thus there are 4 groups of treatments with 3 observations per group. All fermentations consist of 25 mL of media in baffled shaker flasks, 10% inoculum, 37 degrees C, and otherwise run in parallel under identical growth conditions. The overall average (over all 12 observations) is 7.52.

Table 5.3.1 Observations of cell density (OD600)

Yeast extract in g/L added to modified APS media	**Treatment A (2 g/L YE)**	**Treatment B (12 g/L YE)**	**Treatment C (24 g/L YE)**	**Treatment D (36 g/L YE)**
sample 1	7.16	7.03	7.24	8.27
sample 2	7.05	8.23	7.92	8.13
sample 3	7.01	6.85	8.37	7.02
Group Average (mean)	7.07	7.37	7.84	7.81
Standard Deviation	0.078	0.75	0.57	0.68
Relative Standard Deviation	0.011	0.10	0.073	0.088

Table 5.3.2 provides the individual squares of the deviation for each group. The group sum of squares (addition of all squares in Table 5.x.2) is 2.72. The between groups sum of squares value of 1.23 is calculated using the overall average of 7.52 as the mean, where each sum of squares is weighted by the number of observations 3.

Table 5.3.2 Square of deviation of each sample within the group

	Individual Squared Deviations-within treatment			
	Treatment A	Treatment B	Treatment C	Treatment D
sample 1	0.00751111	0.1156	0.36401111	0.2146778
sample 2	0.00054444	0.7396	0.00587778	0.1045444
sample 3	0.00401111	0.2704	0.27737778	0.6188444

In applying the F-test, the null hypothesis is defined such that the means of several normally distributed populations, all having the same standard deviation, are equal. Thus if the null hypothesis holds, the variability among the treatment averages is reasonably attributed to experimental error. Where variability is expressed in terms of the sum of squares, the ratio of between group and within group variance is compared to the F-distribution. The sums of squares divided by the degrees of freedom for both the between groups and within the groups provides the F-ratio as follows:

Between group SS / between group degrees of freedom = 1.23 / (3) = .409

Within group SS / within group degrees of freedom = 2.72 / (8) = .340

The F-ratio is thus .409 / .340 = 1.20.

The F-critical (F_α) value is obtained from the NIST Handbook for 5% significance using 3 degrees of freedom between groups and 8 degrees of freedom within group ($F_{0.05}$= 4.07). Since F-ratio is less than F-crit, we fail to reject the null hypothesis. A reasonable conclusion is all four treatment groups are likely the same, and thus varying yeast extract in modified APS media likely has no significant affect on cell density. (Additional experiments with different treatments reinforces the conclusion that biomass is not significantly affected by the concentration of growth media additive.)

The probability value of this F-test, given the F-value 1.20, numerator degrees of freedom 3, and denominator degrees of freedom 8, is P-value=0.37. This suggests a 37% chance that there could be differences. Therefore each treatment is compared in pairs, with the F-test and P-value provided in Table 5.3.3. The F-critical (F_α) value is 7.7. Again the F-ratio is less than F-crit, we fail to reject the null hypothesis, and all four groups of treatments are reasonably the same at the 95% confidence level.

Table 5.3.3 Comparison of cell mass for each of four treatments evaluated in pairs

Treatments Compared	A, B	A, C	A, D	B, C	B, C	C, D
overall average	7.22	7.46	7.44	7.61	7.59	7.83
F-ratio	0.464	5.395	3.396	0.758	0.554	0.005
P-value	0.57	0.14	0.21	0.48	0.54	0.95

The same statistical procedure is applied to GFP production as measured by fluorescence. The fluorescence is adjusted for the APS media without cells, and is further adjusted to 1 mL fluorescein as an internal standard. Table 5.3.4 shows the corrected observations of fluorescence for each of the four treatments. The overall average fluorescence is 194.

Table 5.3.4 Observations of GFP (fluorescence) by treatment

Yeast extract in g/L added to modified APS media	Treatment A (2 g/L YE)	Treatment B (12 g/L YE)	Treatment C (24 g/L YE)	Treatment D (36 g/L YE)
sample 1	166	166	215	249
sample 2	160	184	209	229
sample 3	140	167	228	215
Group Averages	155	172	217	231
Standard Deviation	11.0	8.4	8.2	14.0
Relative Standard Deviation	0.058	0.041	0.034	0.056

There are 3 observations per group, and four groups. The between group sum of squares is 11743. The within group sum of squares is 1321. Thus the F-ratio is 23.7, which is greater than the $F_{0.05}$ value of 4.07. The P-value for this test is 0.0002, and such a low P-value indicates we reject the null hypothesis and can reasonably conclude there are differences in GFP production by treatment. Next the F-Test is applied to each pairings of treatments to assess which of the pairs show a significant difference. Table

5.3.5 shows the F-test and P-value for each pairing. Where the P-value is higher than 0.1, the treatments are reasonably the same. When the P-value is below 0.05 the treatments are different. One interpretation is thus the two lowest concentrations of yeast extract each have roughly the same effect on GFP production, the two highest concentrations of yeast extract each have roughly the same effect on GFP production, but there is a difference between the lower and higher concentrations of yeast extract on GFP production.

Table 5.3.5 Comparison of GFP average, F-ratio, and P-value for each of four treatments evaluated in pairs

Treatments Compared	A, B	A, C	A, D	B, C	B, C	C, D
overall average	164	186	193	195	202	224
F-ratio	3.08	42.75	37.48	30.99	26.91	1.46
P-value	0.16	0.003	0.004	0.005	0.007	29.3

When additional experiments (for example, concentration of tryptone replace yeast extract as the variable) are conducted in duplicate with four possible treatments, the table value of $F_{0.05}$ is 6.6. Group means of 161, 162, 164, and 166 in GFP (fluorescence) result in the calculated F-ratio of 7. The null hypothesis is rejected, and there is a difference in GFP production by treatment. In this case, the P-value is 0.049. In subsequent trials, the variances are again treated as the same (same equipment, growth conditions, methods, et. al.) and the assumption that similar magnitude changes in GFP is

a likely indicator of treatment significance. Thus some treatment series were conducted without repeated trials. In these cases, preference is therefore given to those treatments conditions showing a difference of 4 or more points in fluorescence. Treatment conditions and results for single treatments may be found in Appendix F.

GFP expression in auto-induced cultures of common growth media

Table 5.3.6 provides the components of some common growth media. Saturation cell density as measured at harvest and the GFP expression as indicated by fluorescence at harvest are reported for each media.

Table 5.3.6. Saturation Density in Different Growth Media

	Media Component							Culture Performance	
Common Media Tested	tryptone (g/L)	yeast extract (g/L)	Martone (g/L)	NZ amine (g/L)	glycerol (w/v)	phosphate (mM)	NaCl (g/L)	cell density at harvest (OD600)	GFP (fluor)
LB (classical induction**)	10	5	-	-	-	-	5	3.23	137
LB	10	5	-	-	-	-	5	3.16	105
2xYT	16	10	-	-	-	-	5	7.15	173
TRB	12	24	-	-	.5%	89	-	5.39	149
TB	12	24	-	-	.5%	-	-	4.86	140
NZ	-	5	-	10	-	-	-	3.72	160
APS	-	24	15	--	-	100	-	5.93	135
modified APS	10	24	15	-	.5%	100	2	6.91	154

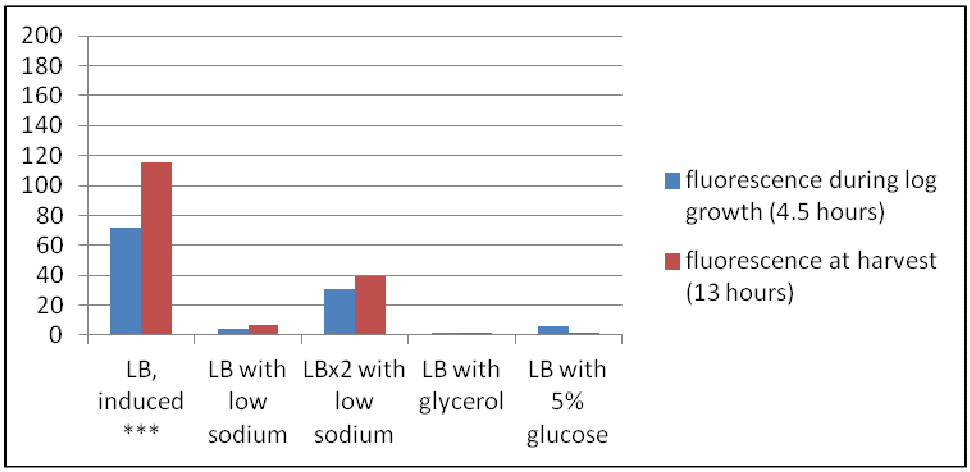
Harvest at 14 hours. Reported values of OD600 are +/- 1.6, fluorescence values are +/- 32.

GFP expression in auto-induced cultures of supplemented media

Figures 5.3.4, 5.3.5, and 5.3.6 show GFP expression over time (as measured by fluorescence) for each media. Classically induced GFP using LB is repeated on the far left of each of the second and third figures to allow direct comparison amongst all trials.

In addition, the x-axis of each figure has been scaled to allow comparison of fluorescence expression between each set of media.

Figure 5.3.4 Evaluation of salt (sodium chloride) on fluorescence expression.

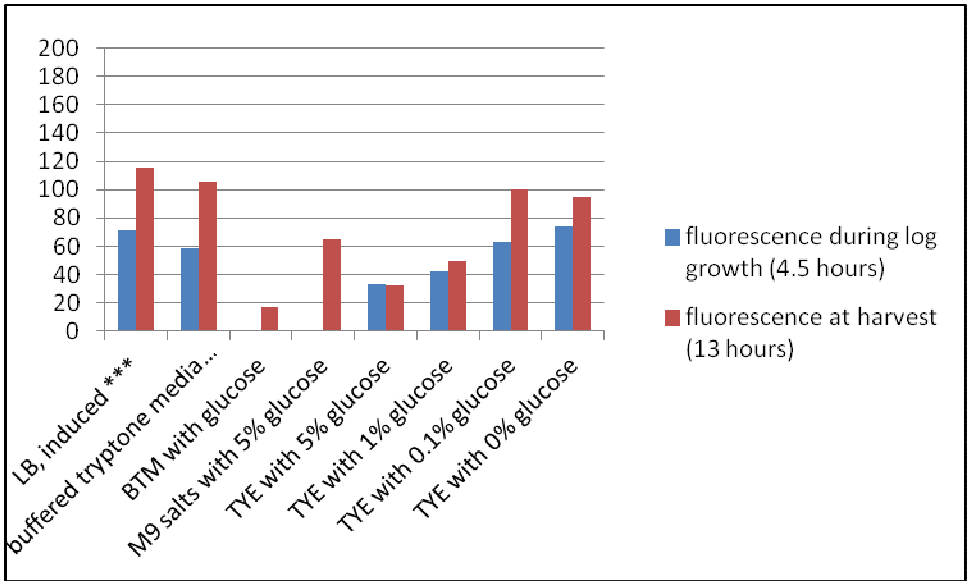


Media were supplemented with .5% arabinose and .2% glucose, except the "LB, induced" culture was induced with arabinose at 4 hours. Fluorescence values are +/- 4.

In Figure 5.3.4, the effects of salt concentration using sodium chloride is demonstrated. The second and third column pairings show significantly lower expression when reduced salt levels were used. As expected, 5% glucose suppressed GFP expression. Surprisingly, glycerol also reduced fluorescence. Since the cells were harvested at 13 hours, it is possible that cells grown on glycerol did not propagate at the same rate as in the other media. However, cell density (OD600) did not vary considerably among all five cultures [data not shown], so the absence of expression is not likely explained by differences in overall cell growth.

Figure 5.3.5 provides the results of different combinations of tryptone and yeast extract, with and without glucose addition. The second column pairing suggests phosphate buffer will have a small but noticeably negative effect on GFP expression. It is probable, but not yet established that this will be offset by the reduced cell growth due to acid produced when glucose is metabolized. The column pairings four through seven give an early indication of glucose concentrations that will give rise to high level protein production. Other than APS media (shown in Figure 5.3.6), other combinations of tryptone and yeast extract provided similar or lower levels of GFP expression [data not shown].

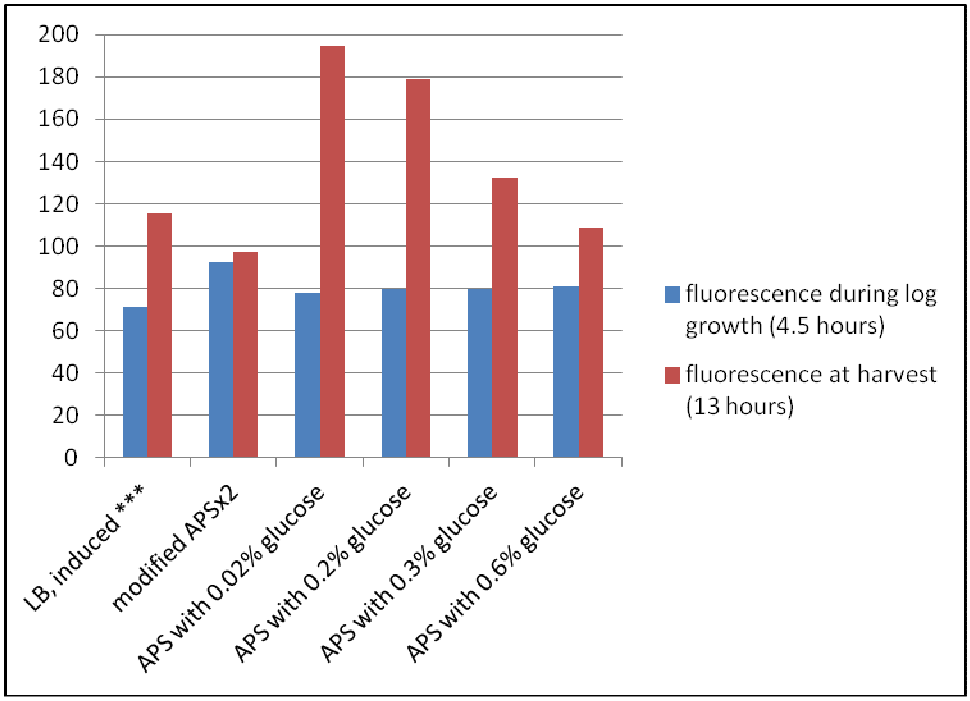
Figure 5.3.5 Evaluation of different tryptone and yeast extract based media



Media were supplemented with .5% arabinose and .2% glucose, except the "LB, induced" culture was induced with arabinose at 4 hours. Fluorescence values are +/- 4.

Figure 5.3.6 provides selected results focused on the APS media with varying levels of glucose. The second column shows a negative impact on GFP expression when the yeast extract was doubled. Highest expression levels were observed with 0.2% glucose, more than 50% higher than observed with the classically induced LB culture. Since all harvesting was conducted at 13 hours, the GFP expression will likely continue for several hours. It is possible 0.2% glucose does not pose the optimal concentration for GFP expression. This was tested and confirmed with cultures grown in a 24 well microplate.

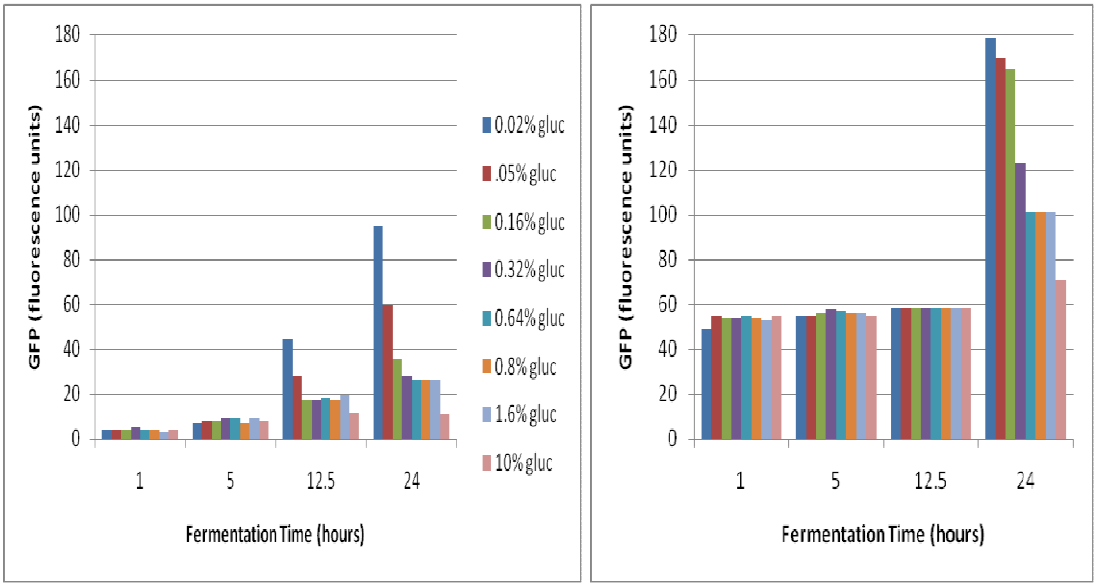
Figure 5.3.6 Evaluation of the APS media with varying glucose concentrations.



Media were supplemented with .5% arabinose and .2% glucose, except the "LB, induced" culture was induced with arabinose at 4 hours. Fluorescence values are +/- 4.

Figure 5.3.7 shows HB101 grown on LB and APS media, respectively, each supplemented with glucose and arabinose. The scale has been maintained to allow direct comparison of the two figures. At 12 hours in the fermentation, cells grown on APS have not yet begun expressing GFP. At harvest time of 24 hours, GFP produced by cells grown on APS is almost double that grown on LB.

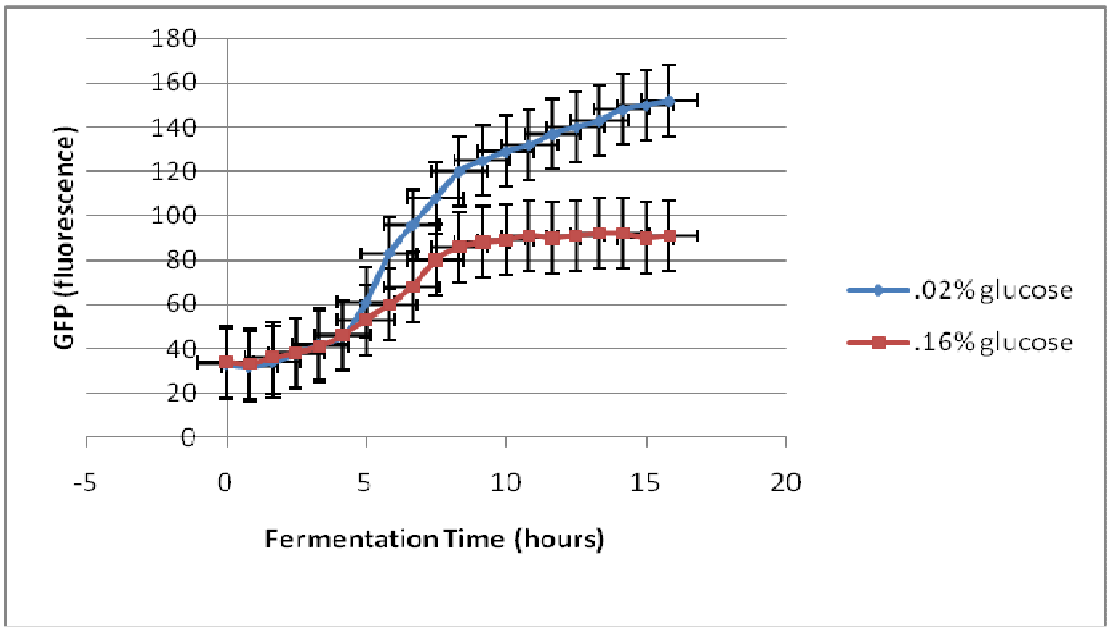
Figure 5.3.7 Relative GFP production by HB101 grown in shaker flasks.



Cultures were grown on LB (left side) and APS media (right side) supplemented with glucose as indicated. Fluorescence values are +/- 16.

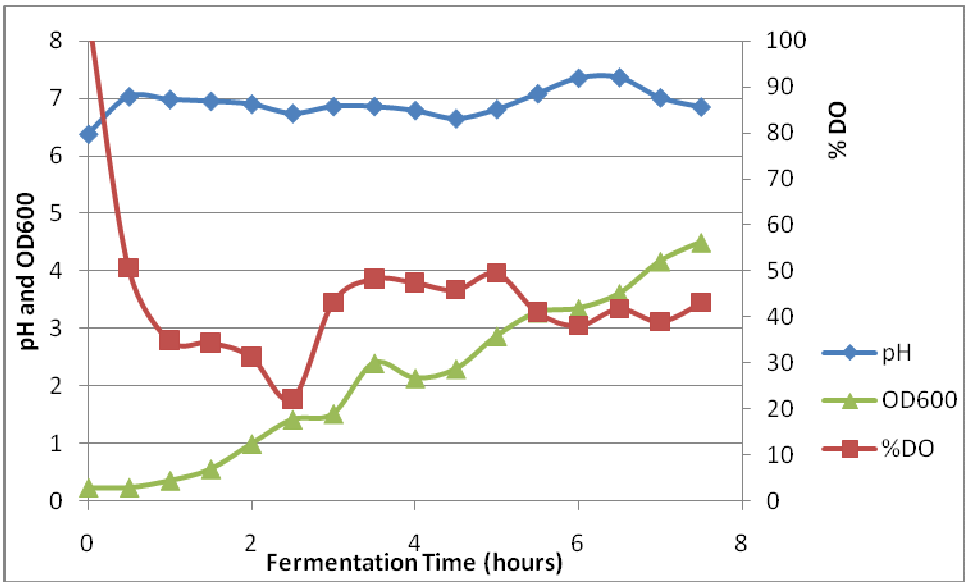
A further evaluation of the lag time and maximum expression for the batch fermentations shown in Figure 5.3.7 is warranted. Two examples are compared in Figure 5.3.8, where the HB101 cells grown with higher starting glucose levels demonstrate an increased time lag for GFP expression as well as an increased duration of the fermentation before maximum GFP expression. Upon closer review of run B1 (0.2% glucose represented by the blue line in Figure 5.3.8 and further evaluated with flow cytometry conducted in the next subsection), the fluorescence has not yet leveled or decreased. Therefore the maximum fluorescence - and correspondingly the maximum GFP expression - may not have occurred yet.

Figure 5.3.8 Effects of glucose on lag time and time to maximum GFP production in HB101.



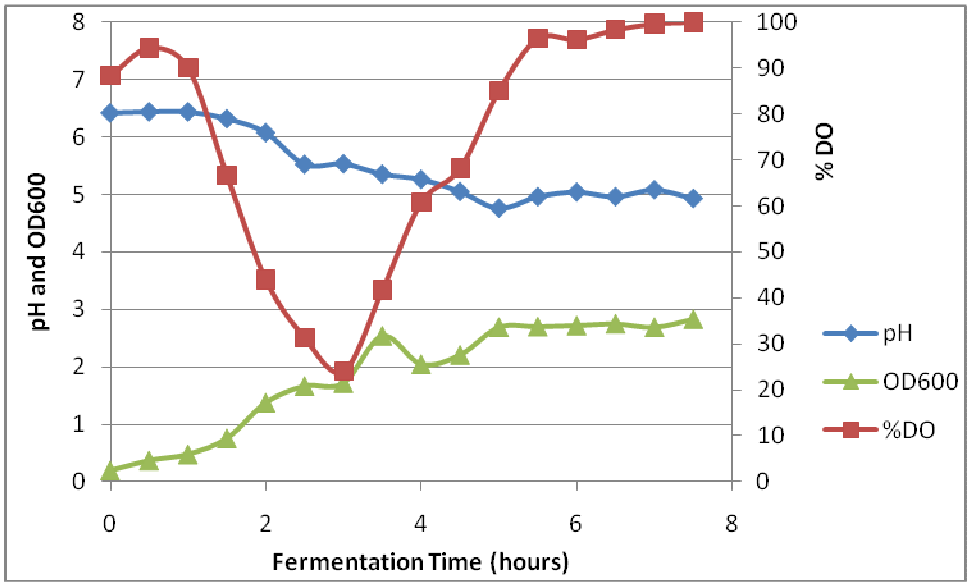
Figures 5.3.9 and 5.3.10 show 10L batch fermentations of HB101 comparing oxygen cascade as a method of providing moderate oxygenation. Cells were grown on APS supplemented with 2% glucose and harvested at 12 hours. The difference between operating conditions of the two fermentations is the first fermentation used an automatic oxygen cascade to 40%, the second fermentation had no cascade. Arabinose was added manually at about 3 hours to induce GFP. Both cultures produced GFP, although the cascaded fermentation produced slightly more (data not shown). Cell growth was almost double in the oxygen cascade than the no cascade fermentation. Under oxygen limiting conditions, *E. coli* shifts to anaerobic respiration or to fermentative metabolism. The metabolism by-products are acidic, causing a drop in pH. *E. coli* grows best in a pH range of 6 to 7, therefore the more acidic pH in the no cascade fermentation likely contributed to reduced cell growth. Oxygen cascade control provides an approximate doubling of cell mass and GFP.

Figure 5.3.9 HB101 fermentation with oxygen cascade.



Values for pH are +/- .15, values for OD600 are +/- .16.

Figure 5.3.10 HB101 fermentation with no oxygen cascade.



Values for pH are +/- .15, values for OD600 are +/- .16.

Further discussion

Use of LB media resulted in nutrient limited growth conditions, as shown in Figure 5.4 where the peak fluorescence occurred within 9 hours. The cell growth profiles in each modified LB media exhibited similar patterns of protein expression. Most interestingly, glycerol did not appear to adequately sustain cell propagation, and low salts did not appear to aid in GFP expression. Both the cell density and the total fluorescence of LB variants expressed the lowest fluorescence of the tested media. There is good agreement with the literature when comparing the relative *E. coli* cell density for each tested media. There is also good agreement with the literature regarding the addition of supplemental glucose at a concentration of 0.5 to 1.0% to facilitate catabolite repression in batch culture. It is expected, but not yet demonstrated, that these glucose levels are too low for optimal fed-batch fermentation.

Glucose metabolism is known to form high levels of acetic acid causing a pH drop, especially when oxygen is limited; the phosphate appears to provide an appropriate amount of buffering for batch conditions. Supplementing media with moderately high levels of combined glucose and glycerol resulted in an excessive pH drop. This is in apparent contrast to the auto-induction media formulations for the *lac* operon, which contain both glucose and glycerol as the carbon sources. In the case of the *ara* operon, it is suspected glycerol in the fed-batch fermentation will further lower pH, and possibly defeat the existing buffering capacity. A modified APS media, including tryptone and

yeast extract with phosphate buffer and supplemented with glucose is expected to provide the highest cell density.

At the lowest glucose concentration of 0.1%, the mean GFP expression was also the lowest. Glucose levels of 0.15% to 0.3% provided similar levels of GFP expression and similar distributions of GFP expression across cells. At suitably high arabinose levels, auto-induction with glucose does not appear to result in any appreciable levels of all-or-nothing expression. It was also found that GFP expression at the end of the fermentation (as measured at 36 hours) was similar at all glucose levels, but the time to reach maximum expression increased with increasing glucose levels. The lowest levels of glucose resulted in both the highest overall GFP production and in the shortest time.

As identified in section 4.1, the influence of IPTG was examined. No changes in growth rate (as measured by OD600), total protein production, or fluorescence were observed as compared to controls with no IPTG addition (data not shown). It is likely the recommendation of IPTG addition was a carryover from fermentations using the combined T7lac and ara operon. IPTG addition was not explored further.

In summary, this phase of experimental work set out by selecting an *E. coli* strain and plasmid for GFP expression suitable for auto-induction. Starting with the literature, an evaluation of growth media for auto-induction of *ara* in HB101 commenced. A loose matrix of media components including glucose, yeast extract, tryptone, glycerol, and salt was assessed. Supplementing with glycerol appears to contribute to lower pH as the fermentation proceeds, reducing overall GFP production. High salt concentrations were not identified as precluding high GFP expression. Moderate oxygenation is important to

sustaining high cell density and GFP production. An initial starting concentration of 0.2% glucose was identified as appropriate to keep cAMP levels low and control expression of the *ara* operon. Next cell homogeneity will be briefly examined.

5.4 Cell homogeneity in GFP expression

5.4.1 Fluorescence Microscopy

Using a fluorescence microscope, cell homogeneity can be qualitatively evaluated. Figure 5.4.1 shows HB101 cells without GFP expression; the left hand side is a black and white image, the right hand side is an image of green fluorescence. See Chapter 4 for a description of the equipment, settings, and other materials and methods.

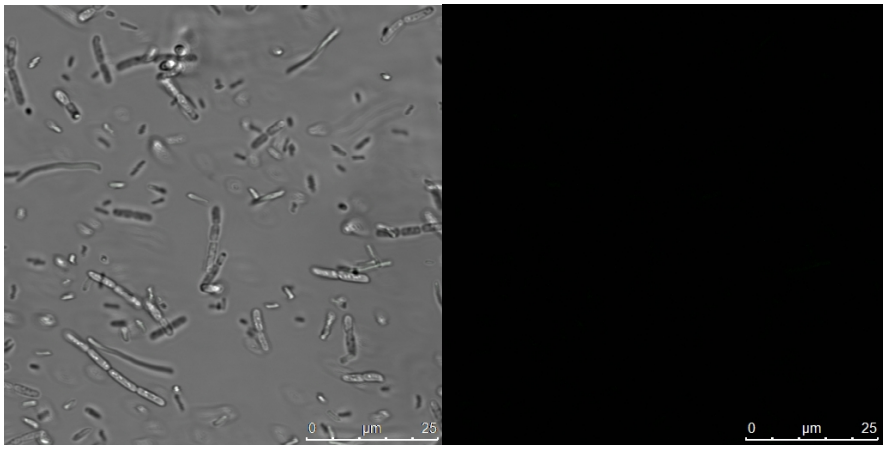


Figure 5.4.1 HB101 cells without GFP expression. Note
The left side photo appropriately appears black because no GFP is expressed.

Figure 5.4.2 shows HB101 cells with GFP expression. Again the left hand side is a black and white image, the right hand side is an image of green fluorescence. The presence of both expressing and non-expressing cells is apparent.

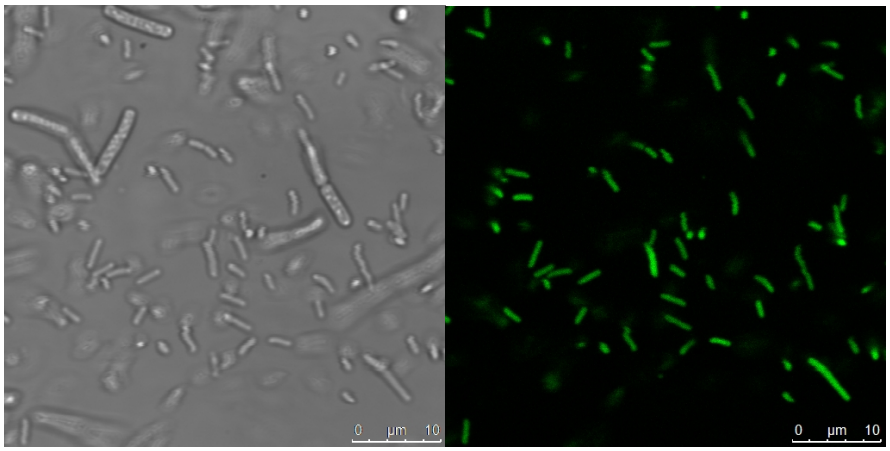


Figure 5.4.2 HB101 cells with GFP expression.

The images can be superimposed to allow immediate characterization of cell homogeneity. Figure 5.4.3 shows under saturating inducer conditions, some cells are present that do not express GFP. In the next section, flow cytometry will be used to quantify GFP expression in individual cells.

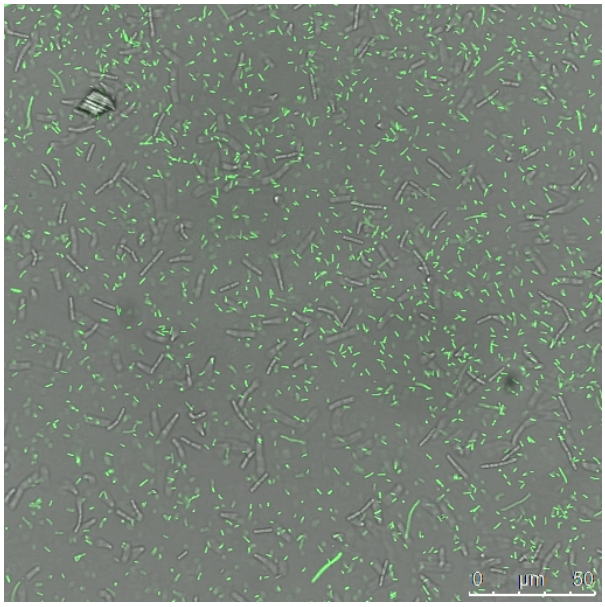


Figure 5.4.3 GFP expression in HB101 saturated with the inducer arabinose.

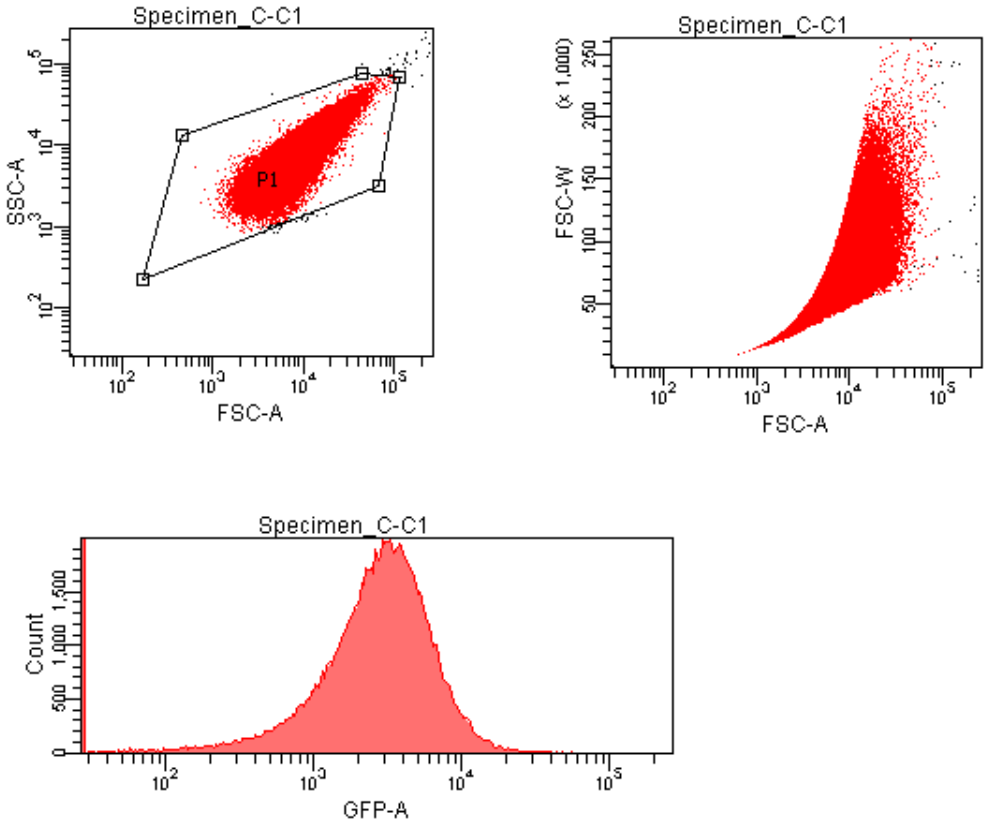
5.4.2 Flow Cytometry

None of the fermentations described thus far utilize subsaturating levels of the inducer arabinose. Flow cytometry was used to evaluate GFP expression profiles at the cellular level rather than the population level. Run "C" consists of several controls: C1 cells have arabinose but no glucose, C2 is an auto-induced culture, C3 is the negative control (no GFP expression), and C4 is a culture producing GFP using the classical induction method. Figure 5.4.5, 7, 7, and 8 show the flow cytometry results for each of these four conditions, respectively. Table 5.4.1 shows the mean and the standard deviation for each condition. The median GFP expression in the auto-induced culture is

80 percent higher than the classically induced culture. Interestingly, the culture with no control over induction (i.e. the inducer arabinose was included in the growth medium) produced slightly higher levels of GFP than the classically induced culture. It is surmised that since cell growth continues beyond induction of *ara*, the need to manually induce the culture at an OD_{600} of 0.5 is obviated. It is also likely the starting OD_{600} (i.e. after addition of inoculums the starting OD_{600} is close to 0.2) is sufficiently close to the preferred cell density of OD_{600} of 0.5 that there would be only minor differences in GFP expression. This is consistent with the observed GFP expression.

Table 5.4.1 Mean and standard deviation of flow cytometry distributions

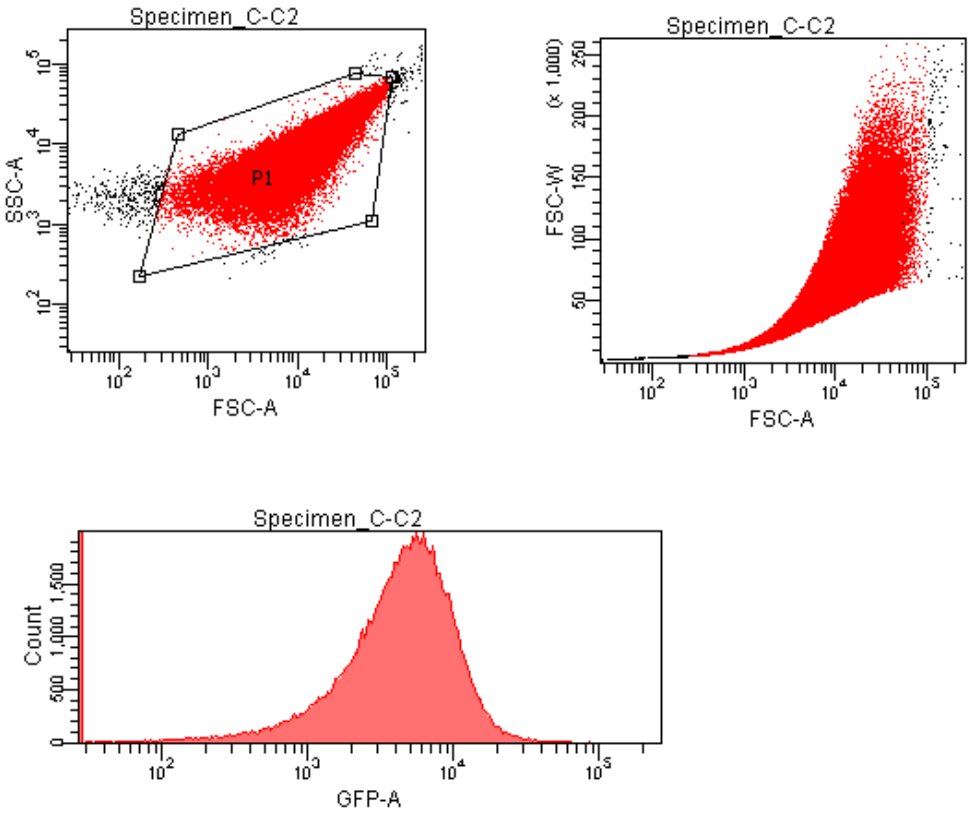
Cell growth condition	C1 (no glucose, induced at start of fermentation)	C2 (auto-induced)	C3 (no GFP expression)	C4 (classical induction, no glucose)
Mean Fluorescence	3,371	5,238	76	2,915
Est. One Standard of Deviation	317	642	83	483



Experiment Name:	Experiment_003
Specimen Name:	Specimen_C
Tube Name:	C1
Record Date:	Feb 18, 2011 11:15:54 AM
$OP:	Administrator
GUID:	9f4bba0c-215d-4c6c-9f05-2ef6a64cecb2

Population	#Events	%Parent	GFP-A Mean
P1	100,103	99.9	3,371

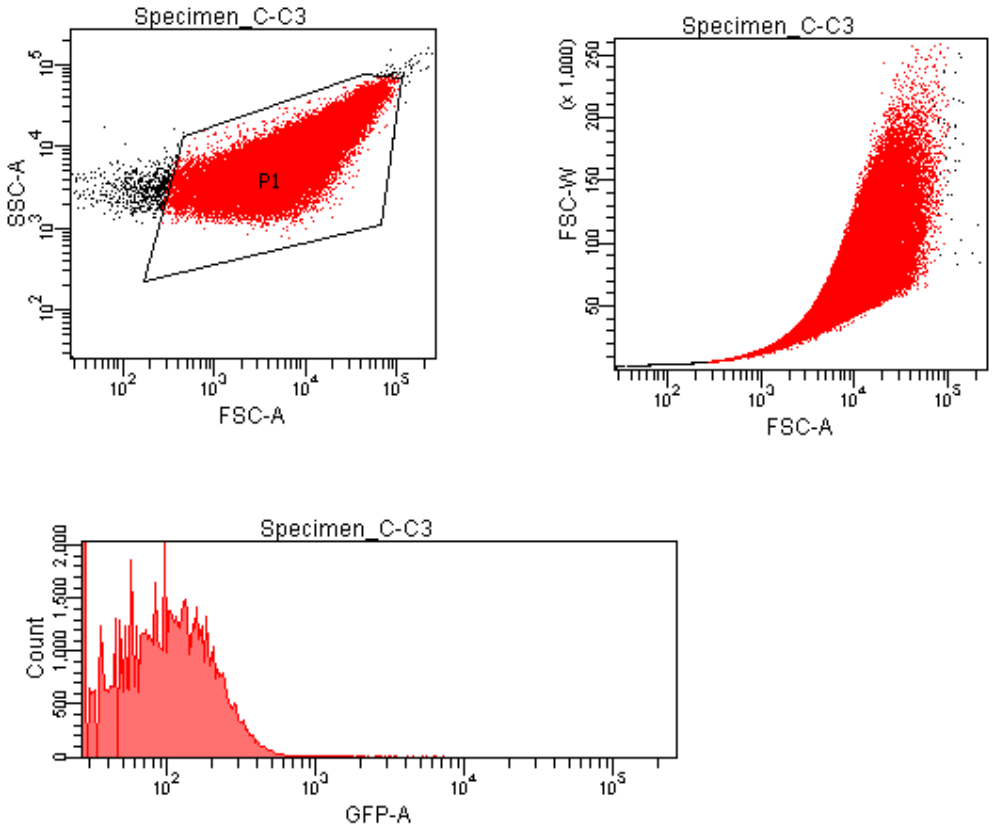
Figure 5.4.4 HB101 cells grown on APS medium supplemented with arabinose; no glucose was added to the medium.



Experiment Name:	Experiment_003
Specimen Name:	Specimen_C
Tube Name:	C2
Record Date:	Feb 18, 2011 11:17:31 AM
$OP:	Administrator
GUID:	1a49f12e-c31f-40bd-b9c2-d7afd29f3be4

Population	#Events	%Parent	GFP-A Mean
P1	101,568	98.7	5,238

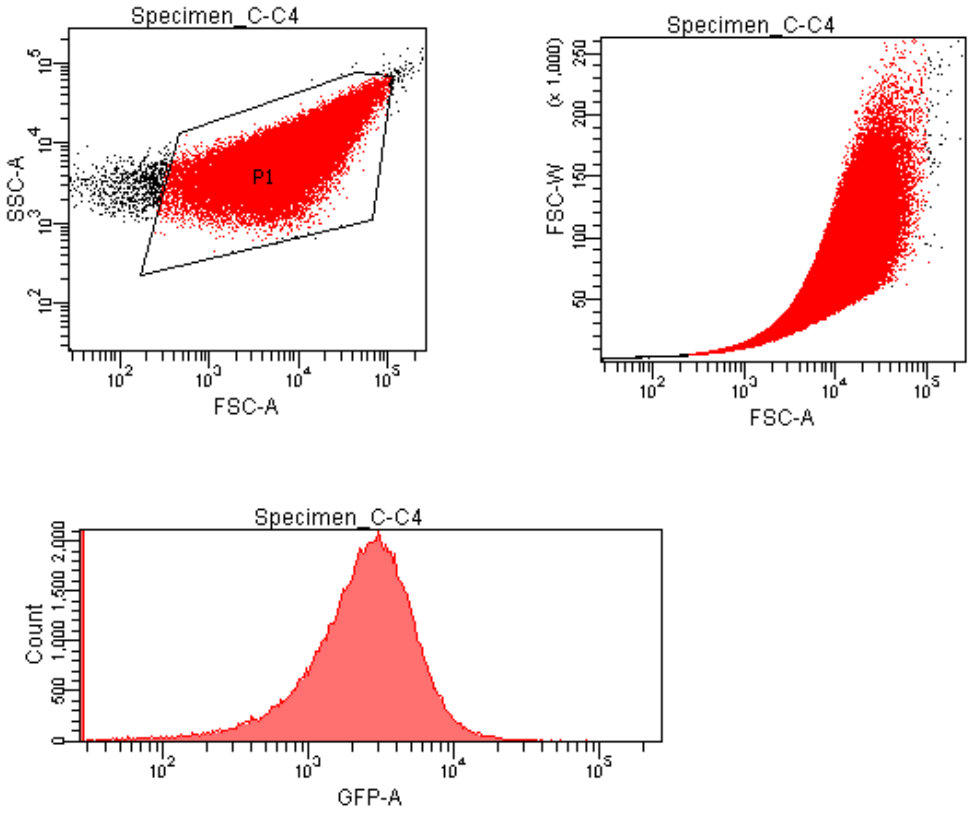
Figure 5.4.5 HB101 cells auto-induced with APS medium supplemented with glucose.



Experiment Name:	Experiment_003
Specimen Name:	Specimen_C
Tube Name:	C3
Record Date:	Feb 18, 2011 11:18:47 AM
$OP:	Administrator
GUID:	f0fd9845-e9da-4187-a409-e6fcdd2793eb

Population	#Events	%Parent	GFP-A Mean
P1	101,640	99.1	76

Figure 5.4.6 HB101 cells grown on APS medium supplemented with 10% glucose and no arabinose.



Experiment Name:	Experiment_003
Specimen Name:	Specimen_C
Tube Name:	C4
Record Date:	Feb 18, 2011 11:19:37 AM
$OP:	Administrator
GUID:	66472eb2-bdbb-487f-b8f6-16840e405b98

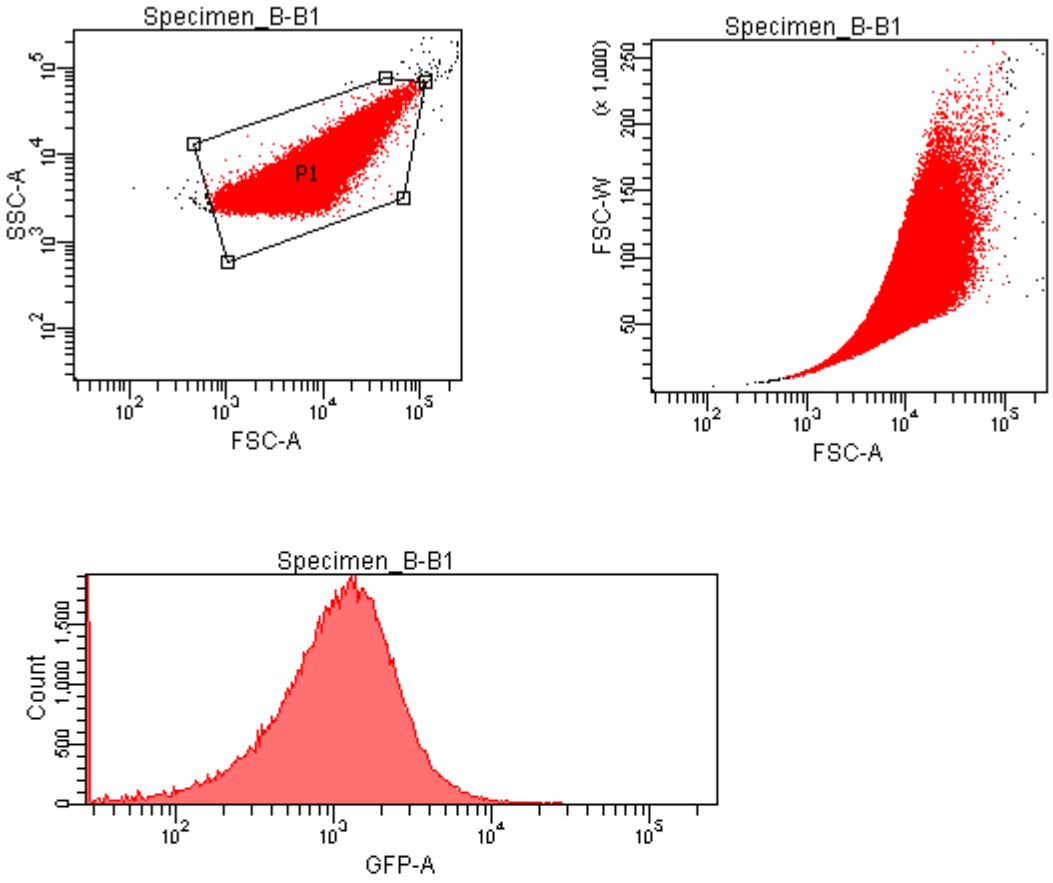
Population	#Events	%Parent	GFP-A Mean
P1	101,679	97.6	2,915

Figure 5.4.7 HB101 cells induced with arabinose at OD_{600} = 0.5. Cells were grown on APS medium with no glucose.

Run series “B” consists of auto-induced cultures with starting glucose concentrations of 0.1, 0.2, 0.25, and 0.3 percent. The flow cytometry for these conditions are shown in Figures 5.4.8, 9, 10, and 11 respectively. Various glucose concentrations are summarized in Table 5.4.2, showing the mean and standard deviation of fluorescence using flow cytometry. The series of runs “B” indicates the GFP expression was three times higher when the media is supplemented with 0.3% glucose compared to 0.1% glucose. One possible explanation is the process of auto-induction may play a role in synchronizing the cell population. Interestingly, the similar GFP expression over a range of starting glucose concentration compares well to the initial 24 well microplate studies, where cell growth increased with increasing starting glucose concentrations, but where the observed GFP expression was highest at low, but non-zero, levels of glucose.

Table 5.4.2 Mean and standard deviation of flow cytometry distributions at varying initial glucose concentrations

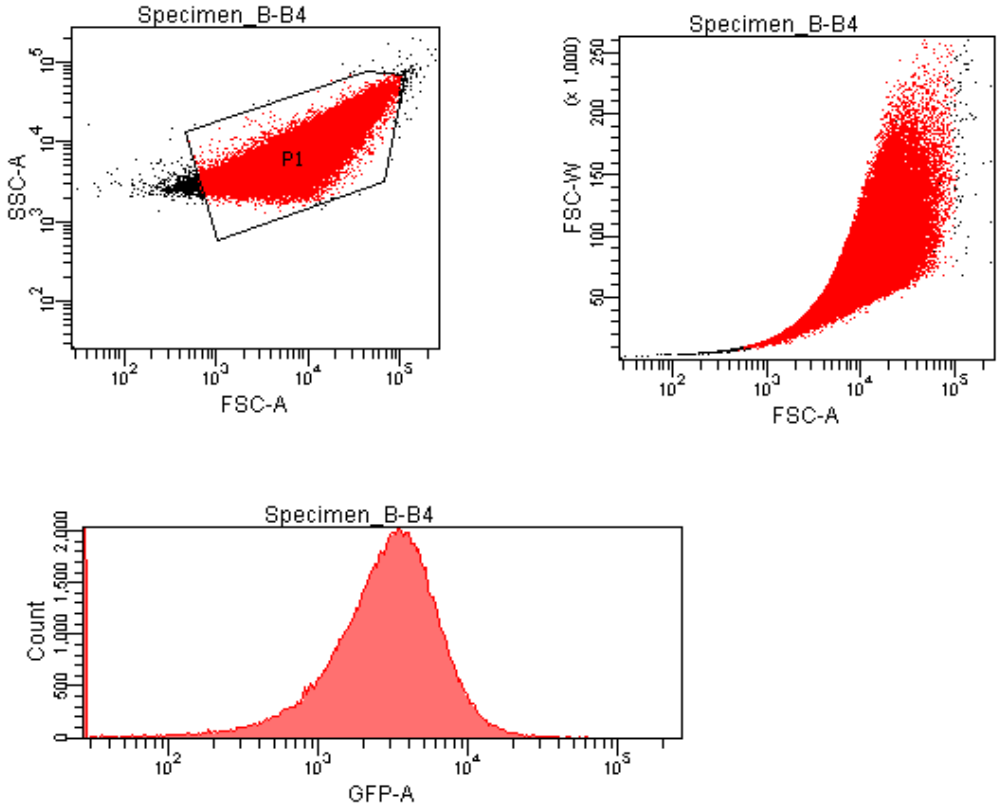
Cell growth condition	B1 (0.1% glucose)	B4 (0.2% glucose)	B5 (0.25% glucose)	B6 (0.3% glucose)
Mean Fluorescence	1,366	3,522	3,136	3,682
Standard Deviation	333	550	567	467



Experiment Name:	Experiment_003
Specimen Name:	Specimen_B
Tube Name:	B1
Record Date:	Feb 18, 2011 11:10:46 AM
$OP:	Administrator
GUID:	7adb1580-eb57-4bee-a88f-55e3fdcbd5fa

Population	#Events	%Parent	GFP-A Mean
P1	100,160	99.9	1,366

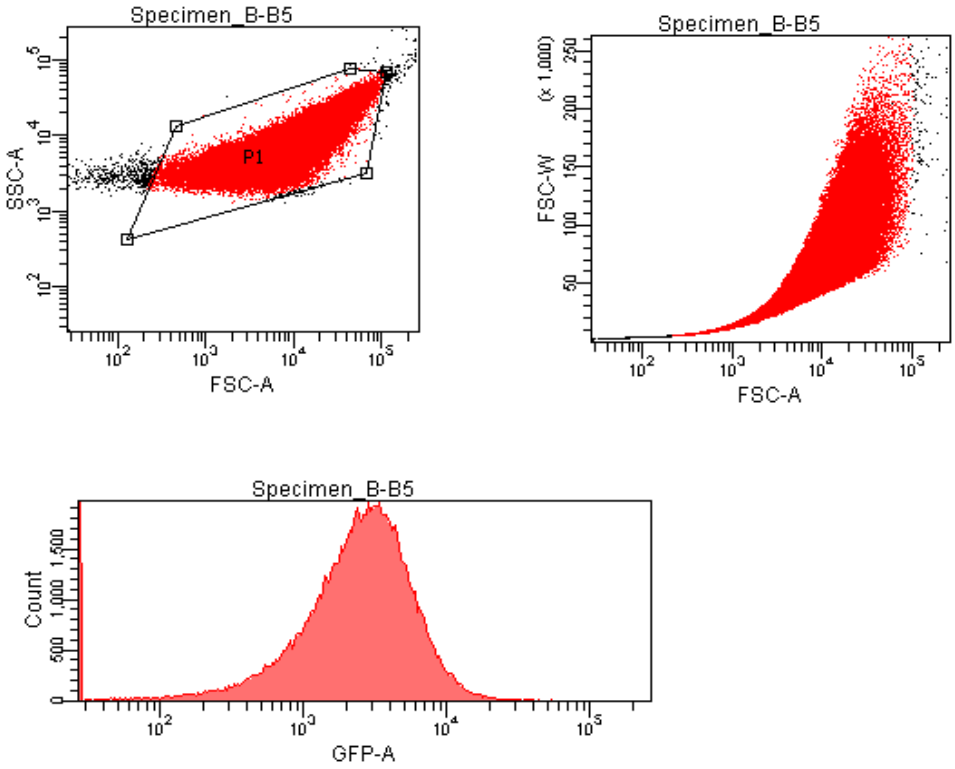
Figure 5.4.8 HB101 cells autoinduced by APS medium supplemented with 0.1% glucose



Experiment Name:	Experiment_003
Specimen Name:	Specimen_B
Tube Name:	B4
Record Date:	Feb 18, 2011 11:11:41 AM
$OP:	Administrator
GUID:	dd415ba1-4a41-41ae-9aa9-3f9a6ffc633f

Population	#Events	%Parent	GFP-A Mean
P1	100,368	99.2	3,522

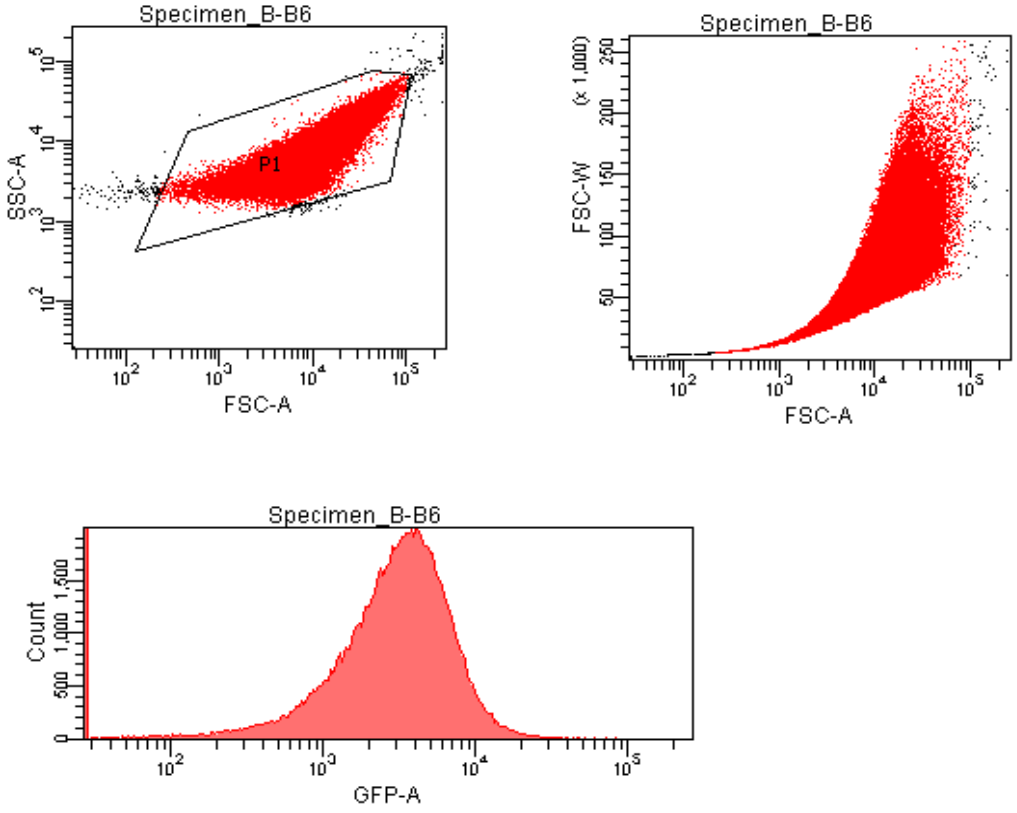
Figure 5.4.9 HB101 cells autoinduced by APS medium supplemented with 0.2% glucose.



Experiment Name:	Experiment_003
Specimen Name:	Specimen_B
Tube Name:	B5
Record Date:	Feb 18, 2011 11:13:29 AM
$OP:	Administrator
GUID:	b5fe752f-2283-48fc-96b9-f895c09fbd28

Population	#Events	%Parent	GFP-A Mean
P1	101,447	98.6	3,136

Figure 5.4.10 HB101 cells autoinduced by APS medium supplemented with 0.25% glucose.



Experiment Name:	Experiment_003
Specimen Name:	Specimen_B
Tube Name:	B6
Record Date:	Feb 18, 2011 11:14:43 AM
$OP:	Administrator
GUID:	986c4a0c-56a1-451e-912c-b72f0e043863

Population	#Events	%Parent	GFP-A Mean
P1	100,705	99.6	3,682

Figure 5.4.11 HB101 cells autoinduced by APS medium supplemented with 0.3% glucose.

Section 5.5 Fed-batch

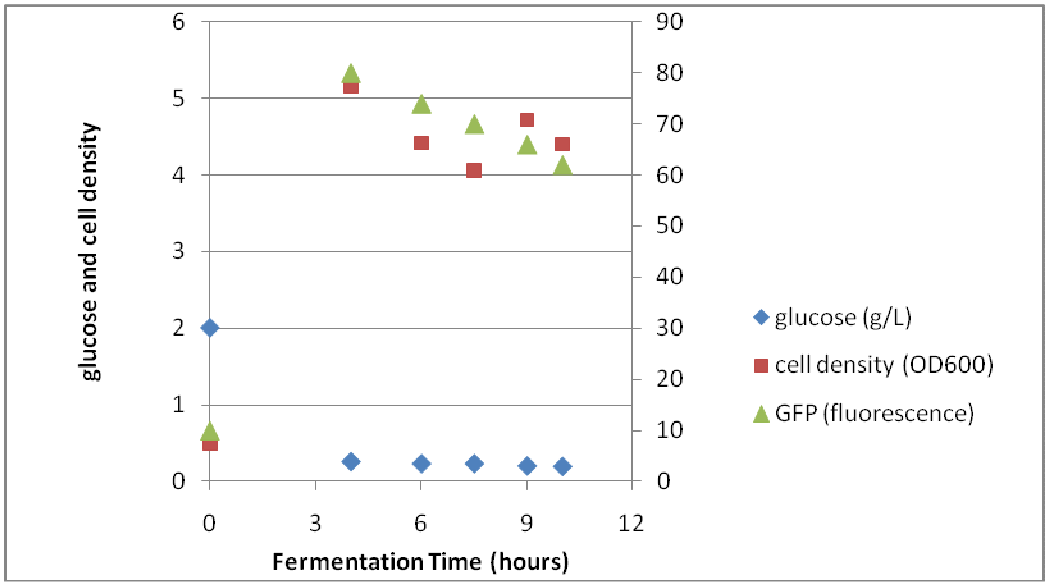
In batch mode, the selected auto-induction media consists of a modified APS media without glycerol and with less than 0.5% glucose. In high cell density culture fed-batch systems, it is assumed significantly more carbohydrates will be necessary. A study of glucose concentrations suggests more than 0.2% glucose will prevent GFP expression. As a first step, APS media was supplemented with glycerol.

Figure 5.5.1 shows cell growth in 250 mL fed-batch fermentation. HB101 cells were grown on modified APS media including 20 g/L glycerol and 2 g/L glucose. 2 g/L sodium chloride and 0.4% arabinose was included in the starting media. After 3 hours, feed commenced at the rate of 0.3 mL per minute (about 7% volume increase each hour) consisting of the same modified APS media.

Initially a good cell density was rapidly obtained (OD_{600} of about 5), however cell growth began to decline after about 5 hours. Glucose levels dropped rapidly as predicted, and the pH remained near neutral, however GFP expression did not occur. It is likely the cells were both oxygen limited in the high density culture and substrate limited.
The experiment was repeated as a cyclic fed-batch and a higher feed rate of 1 mL per min (25% volume increase per hour). Every 4 hours, half of the fermentation volume was harvested (data not shown). Again glucose levels dropped rapidly as predicted, the pH

remained near neutral, and GFP expression did not occur. It is likely the cells were oxygen limited.

Figure 5.5.1 Fed-batch of HB101 using modified APS media for continuous feed.

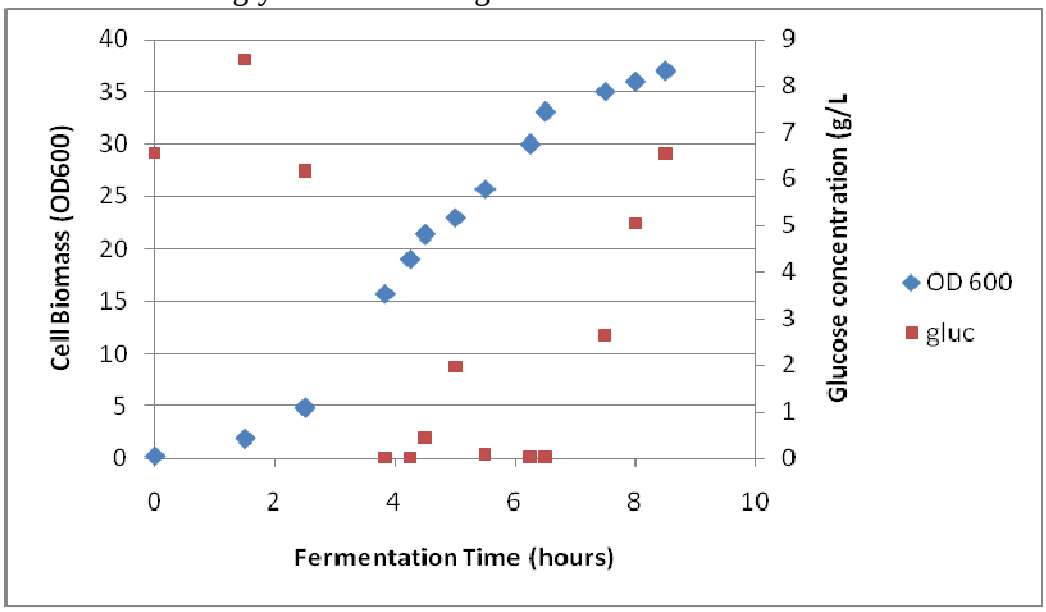


Cell density is +/- 1.7, fluorescence values are +/- 16.

Figure 5.5.2 shows cell growth in fed-batch 100L fermentation with oxygen control cascade to 40%. HB101 cells were grown on APS supplemented with 0.4% glycerol and 2.5% glucose. To provide additional substrate for high density cell growth prior to induction, a five-fold increase in the minimal glucose level necessary for catabolite repression was used. When starting glucose levels dropped to 2 g/L (approximately t= 4 hours), cascade feed of glucose to maintain 2 g/L levels was used to produce a high cell density culture (OD_{600} of at least 20). As just described, maintaining

glucose above 2 g/L will prevent GFP expression. In this fermentation, GFP was manually induced by arabinose in the late log phase of the fermentation, at approximately 6 hours. Cells were harvested within 4 hours of induction. Negligible GFP expression as indicated by fluorescence was measured. It appears the cell growth rate was dropping off such that glucose levels started to increase. The resulting glucose concentration was too high, repressing arabinose induction; GFP fluorescence was negligible for the entire fermentation.

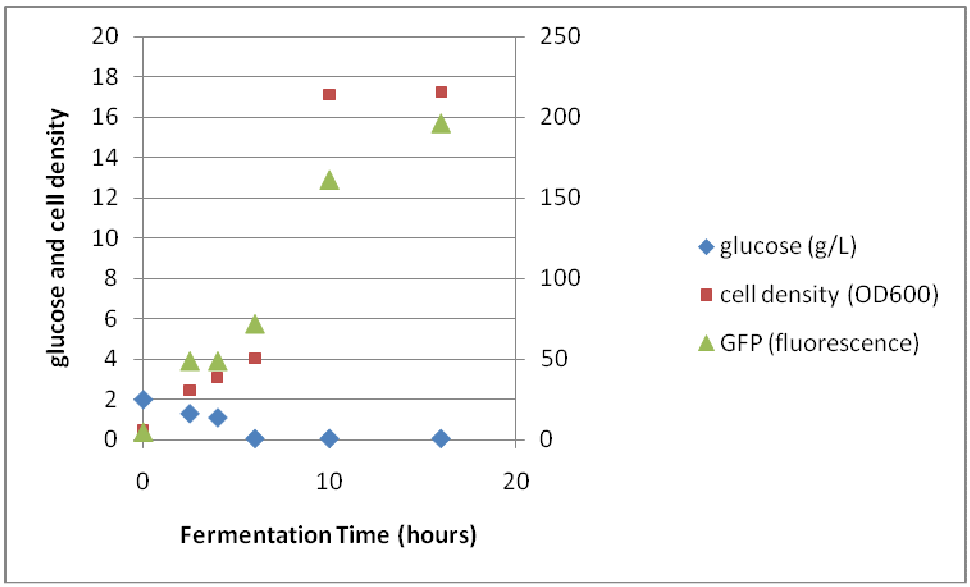
Figure 5.5.2 Fed batch fermentation of HB101 cells grown on APS media supplemented with 0.4% glycerol and 2.5% glucose.



Continuous feed of glucose began at about 4 hours. Arabinose induction occurred at about 6 hours. No fluorescence was measured anytime during the fermentation. Cell density is +/- 1.7.

In Figure 5.5.3, HB101 cells were grown in fed-batch fermentation with modified APS media with 2% glucose and 0.5% glycerol. Oxygen was controlled by cascade. GFP expression is observed once glucose levels drop below 0.1%. Continuous feed of modified APS began at about 6 hours. No glycerol or glucose was contained in the feed. Cell density and GFP expression continued to increase for several hours, to a maximum cell density of OD_{600} = 17. The fermentation was continued for several hours, but no further cell growth was observed, probably due to substrate limitation.

Figure 5.5.3 Fed batch fermentation of HB101 cells auto-induced on APS media supplemented with 0.4% glycerol and 2% glucose.



Cell density is +/- 1.7, fluorescence is +/-16.

This fed-batch fermentation successfully demonstrated auto-induction of the *ara* operon in a high cell density culture. GFP yields were approximately 2.5 times that of the highest batch fermentation identified in Section 5.3.

Chapter 6: Conclusion and Further Work

The most frequently used formulation of auto-induction media is 5052 based media (Studier 2005), consisting of 0.5% glycerol and 0.05% glucose as a supplement to medium consisting of 10g tryptone and 5 g yeast extract. The 5052 media was developed to auto-induce *lac* in *E. coli.* The same auto-induction media was reported to work on a combined *lac-ara*, and in one instance an *ara* gene expression system. Several reports cited in Chapter 3 of this thesis identify those factors influencing the success of auto-induction includes: strain, growth medium, promoter and repressor construction, and the lacI protein concentration. This work begins with a clear demonstration of auto-induction of *ara.* Next an auto-induction media for use *ara* in high cell density *E. coli* cultures was developed. A modified APS media provided the best performance in batch fermentation.

The auto-induction media presented here is not likely an optimized media for all *ara* expression systems. Each plasmid and bacterial strain has slightly different requirements for producing maximum levels of recombinant protein. For example, a construction with reduced lacI protein relieves oxygen sensitivity to auto-induction (Blommel 2007). As discussed in Chapter 2, as a first step an array of media in microwells is an established simple way to test the performance of other plasmid vectors and hosts (Blommel 2007). Other researchers report changing the expression levels by manipulating the auto-induction medium (Gawrylewski 2007). Further, a different target protein will likely pose different needs for optimal production. Ongoing work at NIH is developing different media for each of several dozen auto-induced proteins of interest

(personal communication). Even so, fine tuning a rich media with so many competing factors is “not straight forwards” (Studier 2005). Therefore, further developing media is unlikely to be a further significant contribution.

This work then presents a novel application of auto-induction media and the *ara* operon in a fed-batch fermentation. The combined fed-batch and auto-induction successfully demonstrates a high cell density and high protein expression system that avoids the need for human monitoring and intervention. Yields of GFP were found to exceed 250% of those obtained in batch cultures. The success of auto-induction in fed batch fermentation poses a number of interesting follow-up questions and related topics suitable for further exploration:

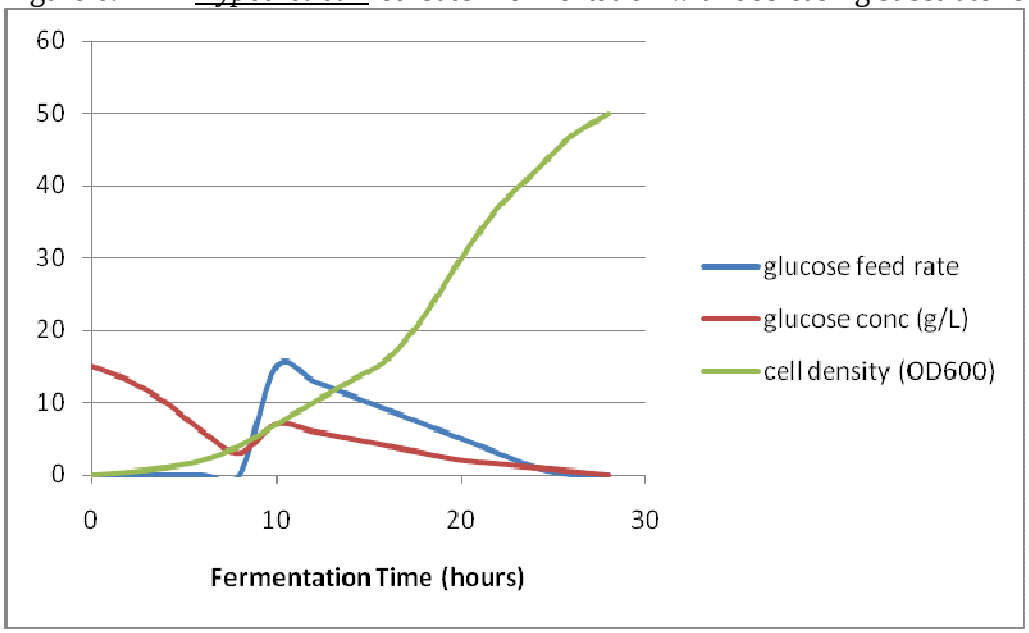
- Will the auto-induction media be sufficient for maximum production in fed-batch high cell density cultures? Fed-batch is more productive with a later induction time, therefore higher initial starting glucose concentrations will be desirable in the auto-induction media. Excess glycerol as the primary carbon source causes an increase in pH, and phosphate is insufficient to buffer the acid produced by metabolism of high glucose levels. Therefore is it reasonable to anticipate further development of the auto-induction media will provide improved recombinant protein production.
- Can fully induced and homogenous expression be sustained in fed-batch cultures at subsaturating levels of arabinose concentrations? An evaluation of the propionate-inducible pPro system by conventional and auto-induction methods (Lee 2008), and deletion of both the ara FGH and the glucose transport genes

show fully induced homogenous cell populations at subsaturating levels of inducer are obtainable. Application of the system to fed-batch fermentation has not been reported.

- Can auto-induction media be further adapted to successfully sustain recombinant protein production in a chemostat? The unstable behavior of plasmids in continuous culture is well documented but not clearly understood. In batch culture, GFP yields of hundreds of milligrams per liter per hour are obtained. Chemostat culture was briefly explored in Chapter 5, and no GFP was observed over the full range of dilution rates tested. In many cases, a plasmid free cell will outgrow a plasmid bearing cell (Dykhuizen 1983). However, when the chemostat feed and siphon pumps are shut off and the culture is operated in batch mode, GFP expression resumes. Similarly, GFP expression is restored by re-plating the harvested culture on agar plates or inoculating shaker flasks. This suggests the plasmid has not been lost or mutated, but more exploration is necessary. A certain type of extended fed-batch - *the cyclic fed-batch culture for fixed volume systems* - refers to periodic harvest of a portion of the culture, and using the remaining culture as the starting point for additional fed-batch process. Once the fermentation reaches a certain stage, (for example, when aerobic conditions cannot be maintained anymore) the culture is removed and the biomass is diluted to the original volume with sterile water or medium (Stanbury 1993). The fixed volume fed batch appears to provide a reasonable next step in exploring the observed limitations of chemostat fermentation.

Finally, this work has led to ongoing development of a novel method for feeding glucose in the fed-batch. Usually, fed-batch feed consists of highly concentrated substrate so as to minimize the increase in fermentation volume. As discussed in Chapter 2, the most commonly used feed strategy is a continuous linear feed, but modern process control allows increasing feed as the cell density increases in order to avoid substrate limited growth. It is anticipated that the new feed strategy would consist of a decreasing addition of glucose. The glucose feed would continue to suppress GFP expression until late in the fermentation. Glucose is the preferred substrate for *E. coli*. Thus glucose feed often produces the highest cell density. Decreased glucose would also reduce the formation of acidic by-products, and may serve to aid in pH control. Figure 6.1 shows the conceptual fermentation for the hypothetical decreasing substrate feed model.

Figure 6.1 *Hypothetical* fed-batch fermentation with decreasing substrate feed.



Certainly there are many approaches to achieve high levels of protein or metabolic products from a microbiological culture. Given the successful methods for auto-induction and fed-batch fermentation, a logical next step would consider models reflecting the various interactions between auto-induction and fed-batch. It is hoped that this work would provide some useful insights for furthering recombinant protein production.

Appendix A: Zeiss Axiophot Settings

Image: Image002

Size: 524.29 KB

File Location: D:\SP5 X Images\Amy\Paul Shriner\040811.lif

Start Time: 4/8/2011 8:13:05 AM.235

End Time: 4/8/2011 8:13:05 AM.235

Total Exposures: 2 (2 channels, 1 frames)

Dimensions

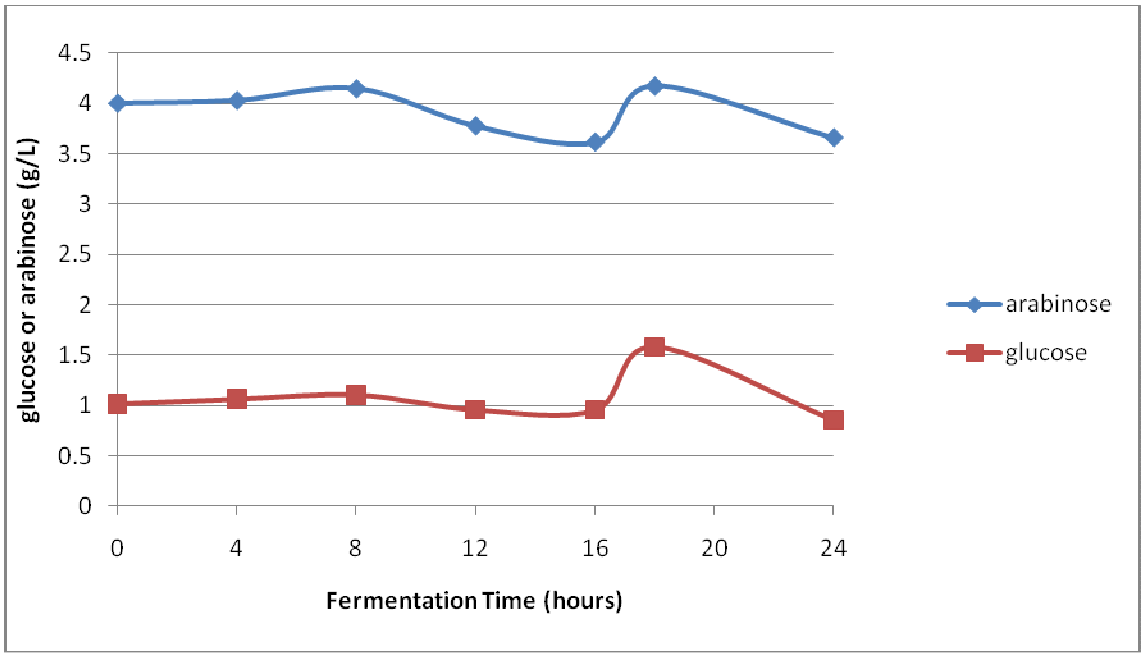
Dimension	Logical Size	Physical Length	Physical Origin
X	512	86.07 µm	42.90 µm
Y	512	86.07 µm	-19.22 µm

ChaserUVShutter	0
ChaserVisibleShutter	0
MP2Shutter	0
MPShutter	0
SupercontinuumVisibleShutter	0
UV405Shutter	0
UVShutter	0
VisibleShutter	1
ScanMode	xyz
Pinhole [m]	111.4 µm
Pinhole [airy]	1.00
Size-Width	86.1 µm
Size-Height	86.1 µm
Size-Depth	0.0
StepSize	0.00 µm
Voxel-Width	168.4 nm
Voxel-Height	168.4 nm
Voxel-Depth	0.0 nm
Voxel-Volume	0 nm^3
Zoom	2.9
Scan-Direction	1

System Type	TCS SP5
RoiScan	0
Is 3D limited ROI scan enable	0
IsSequential	0
StepSize Constant	0
CARS Shutter	0
ChaserUVShutter	0
ChaserVisibleShutter	0
FSOPO Shutter	0
MP2Shutter	0
MPShutter	0
Pulsed 635 (Visible) Shutter	0
Pump Shutter	0
STED 1 Shutter	0
STED 2 Shutter	0
STED 3 Shutter	0
STED 4 Shutter	0
Stokes Shutter	0
SupercontinuumVisibleShutter	0
UV405Shutter	0
UVShutter	0

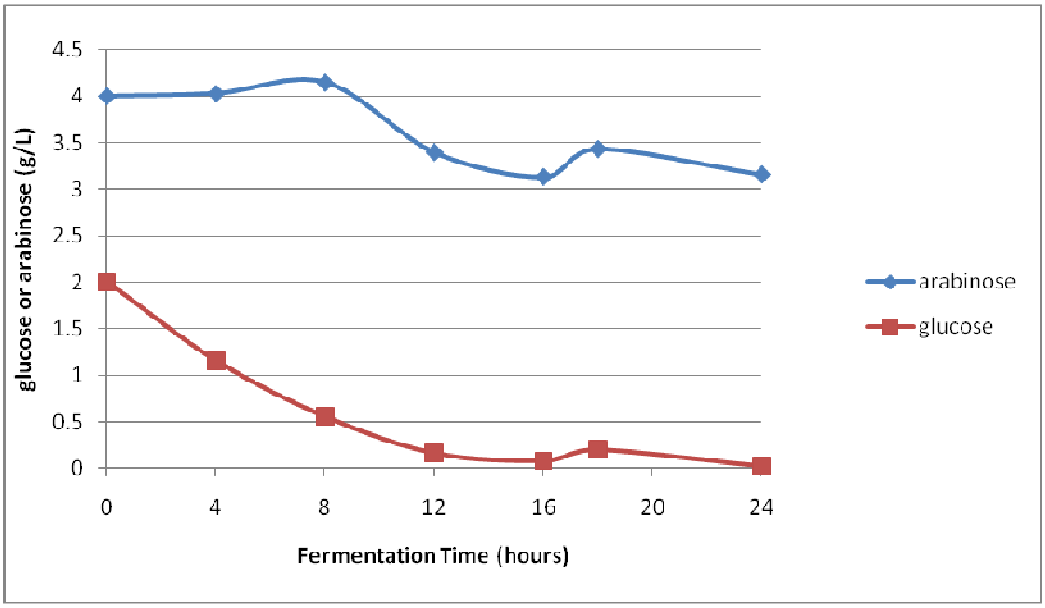
WLL AOTF (489)	0.00 %
WLL AOTF (552)	0.00 %
WLL AOTF (670)	0.00 %
WLL AOTF (631)	0.00 %
WLL AOTF (667)	0.00 %
WLL AOTF (668)	0.00 %
WLL AOTF (659)	0.00 %
WLL AOTF (670)	0.00 %
AOTF (405)	0.00 %
AOTF (458)	0.00 %
AOTF (476)	0.00 %
AOTF (488)	19.00 %
AOTF (496)	0.00 %
AOTF (514)	0.00 %
PMT 1	Inactive
PMT 2	Active
PMT 2 (Offs.)	0.0 %
PMT 2 (HV)	799.7
PMT 2 (HV_Unit)	V
PMT 2 (Preamp)	Direct
PMT 3	Inactive

Appendix B: Example Chemostat Fermentation Carbohydrates Analysis Using HPAE-PAD Chromatography – Summary of Results

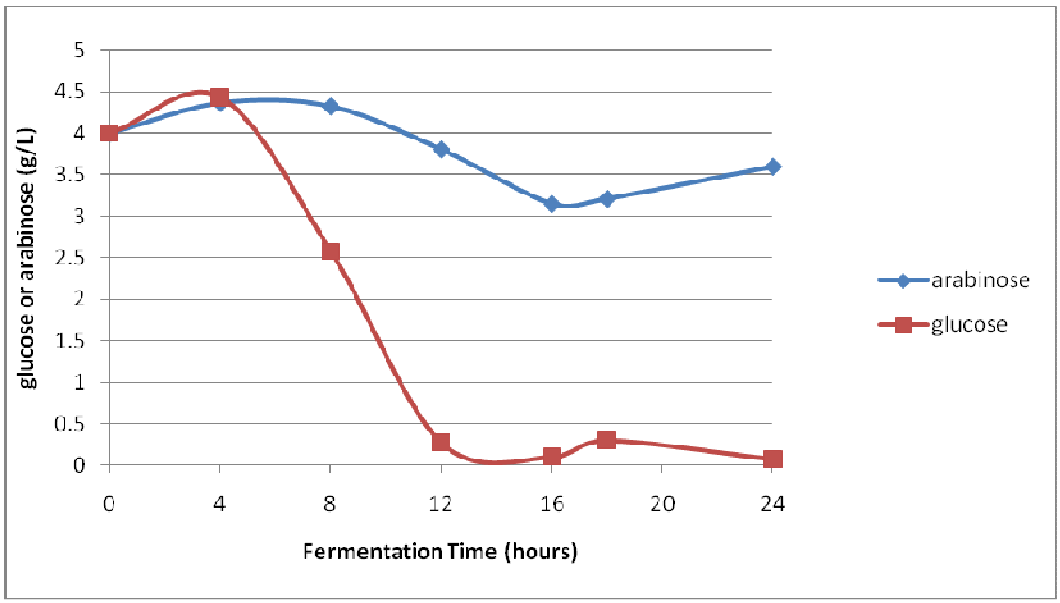


HPAE-PAD analysis of a chemostat fermentation of HB101. Results plotted as glucose and arabinose concentrations in g/L. Starting arabinose concentration of 4/g/L and glucose concentration of 1 g/L. This chromatography uses galactose as an internal standard. The plot shows the mean of triplicate samples. One standard of deviation: 0.24 arabinose and 0.23 glucose.

Appendix C: Example Batch Fermentation Carbohydrates Analysis Using HPAE-PAD Chromatography – Summary of Results

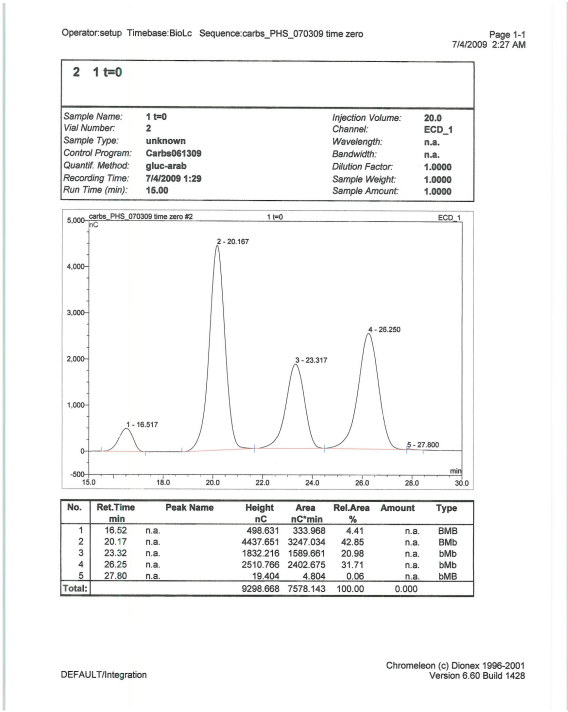


HPAE-PAD analysis of a batch fermentation of HB101. Results plotted as glucose and arabinose concentrations in g/L. Starting arabinose concentration of 4/g/L and glucose concentration of 2 g/L. This chromatography uses galactose as an internal standard. The plot shows the mean of triplicate samples. One standard of deviation: 0.49 arabinose and 0.23 glucose.



HPAE-PAD analysis of a batch fermentation of HB101. Results plotted as glucose and arabinose concentrations in g/L. Starting arabinose concentration of 4/g/L and glucose concentration of 4 g/L. This chromatography uses galactose as an internal standard. One standard of deviation: 0.49 arabinose and 0.23 glucose.

Appendix D: Example Batch Fermentation Carbohydrates Analysis Using HPAE-PAD Chromatography – Individual Sample Analysis



Operator:setup Timebase:BioLc Sequence:carbs_PHS_070309 time zero

Page 1-1
7/4/2009 2:27 AM

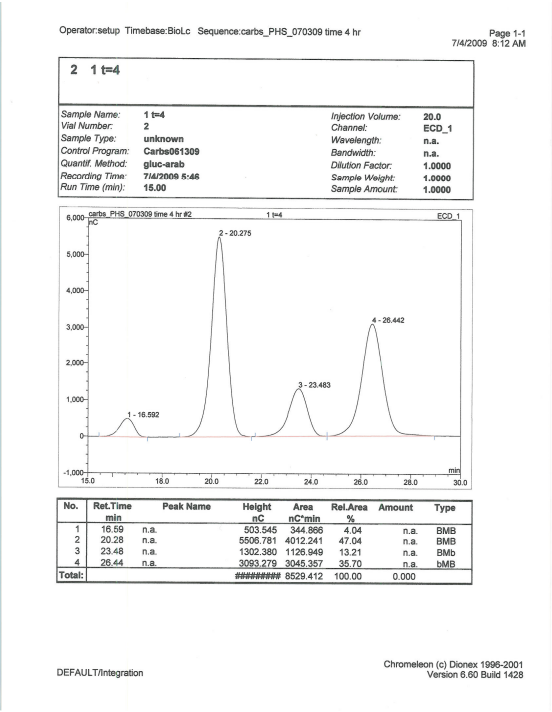
2 1 t=0

Sample Name:	**1 t=0**	*Injection Volume:*	**20.0**
Vial Number:	**2**	*Channel:*	**ECD_1**
Sample Type:	**unknown**	*Wavelength:*	**n.a.**
Control Program:	**Carbs061309**	*Bandwidth:*	**n.a.**
Quantif. Method:	**gluc-arab**	*Dilution Factor:*	**1.0000**
Recording Time:	**7/4/2009 1:29**	*Sample Weight:*	**1.0000**
Run Time (min):	**15.00**	*Sample Amount:*	**1.0000**

No.	Ret.Time min	Peak Name	Height nC	Area nC*min	Rel.Area %	Amount	Type
1	16.52	n.a.	498.631	333.968	4.41	n.a.	BMB
2	20.17	n.a.	4437.651	3247.034	42.85	n.a.	BMb
3	23.32	n.a.	1832.216	1589.661	20.98	n.a.	bMb
4	26.25	n.a.	2510.766	2402.675	31.71	n.a.	bMb
5	27.80	n.a.	19.404	4.804	0.06	n.a.	bMB
Total:			9298.668	7578.143	100.00	0.000	

DEFAULT/Integration

Chromeleon (c) Dionex 1996-2001
Version 6.60 Build 1428



Operator:setup Timebase:BioLc Sequence:carbs_PHS_070309 time 4 hr

Page 1-1
7/4/2009 8:12 AM

2 1 t=4

Sample Name:	**1 t=4**	*Injection Volume:*	**20.0**
Vial Number:	**2**	*Channel:*	**ECD_1**
Sample Type:	**unknown**	*Wavelength:*	**n.a.**
Control Program:	**Carbs061309**	*Bandwidth:*	**n.a.**
Quantif. Method:	**gluc-arab**	*Dilution Factor:*	**1.0000**
Recording Time:	**7/4/2009 8:46**	*Sample Weight:*	**1.0000**
Run Time (min):	**15.00**	*Sample Amount:*	**1.0000**

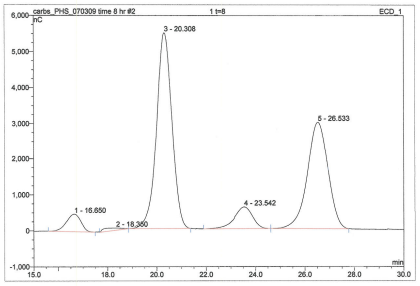
No.	Ret.Time min	Peak Name	Height nC	Area nC*min	Rel.Area %	Amount	Type
1	16.59	n.a.	503.545	344.866	4.04	n.a.	BMB
2	20.28	n.a.	5506.781	4012.241	47.04	n.a.	BMb
3	23.48	n.a.	1302.380	1126.949	13.21	n.a.	BMb
4	26.44	n.a.	3093.279	3045.357	35.70	n.a.	bMB
Total:			#########	8529.412	100.00	0.000	

DEFAULT/Integration

Chromeleon (c) Dionex 1996-2001
Version 6.60 Build 1428

2 1 t=8

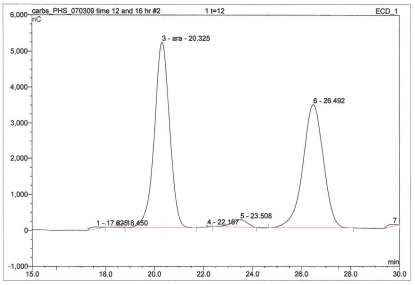
Sample Name:	**1 t=8**	*Injection Volume:*	**20.0**
Vial Number:	**2**	*Channel:*	**ECD_1**
Sample Type:	**unknown**	*Wavelength:*	**n.a.**
Control Program:	**Carbs061309**	*Bandwidth:*	**n.a.**
Quantif. Method:	**gluc-arab**	*Dilution Factor:*	**1.0000**
Recording Time:	**7/4/2009 9:57**	*Sample Weight:*	**1.0000**
Run Time (min):	**15.00**	*Sample Amount:*	**1.0000**



No.	Ret.Time min	Peak Name	Height nC	Area nC*min	Rel.Area %	Amount	Type
1	16.65	n.a.	485.078	332.135	4.28	n.a.	BMB
2	18.35	n.a.	56.280	59.379	0.77	n.a.	BMb
3	20.31	n.a.	5460.306	3965.328	51.15	n.a.	bMB
4	23.54	n.a.	603.435	515.460	6.65	n.a.	BMb
5	26.53	n.a.	2980.067	2880.166	37.15	n.a.	bMB
Total:			9585.167	7752.469	100.00	0.000	

2 1 t=12

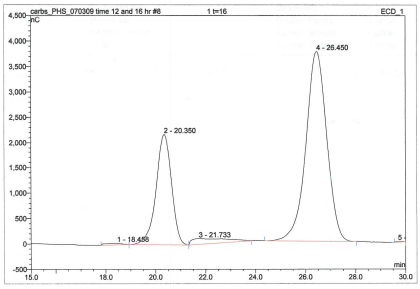
Sample Name:	**1 t=12**	*Injection Volume:*	**20.0**
Vial Number:	**2**	*Channel:*	**ECD_1**
Sample Type:	**unknown**	*Wavelength:*	**n.a.**
Control Program:	**Carbs061309**	*Bandwidth:*	**n.a.**
Quantif. Method:	**gluc-arab**	*Dilution Factor:*	**1.0000**
Recording Time:	**7/4/2009 14:11**	*Sample Weight:*	**1.0000**
Run Time (min):	**15.00**	*Sample Amount:*	**1.0000**



No.	Ret.Time min	Peak Name	Height nC	Area nC*min	Rel.Area %	Amount	Type
1	17.63	n.a.	53.134	28.008	0.38	n.a.	BMB
2	18.45	n.a.	12.825	6.410	0.09	n.a.	bMb
3	20.33	ara	5171.026	3732.282	51.25	n.a.	bMB
4	22.17	n.a.	33.746	3.251	0.04	n.a.	BMb
5	23.51	n.a.	205.947	152.837	2.10	n.a.	bMB
6	26.49	n.a.	3449.733	3333.733	45.78	n.a.	BMB
7	29.73	n.a.	52.700	25.611	0.35	n.a.	BMB
Total:			8979.111	7282.131	100.00	0.000	

8 1 t=16

Sample Name:	**1 t=16**	*Injection Volume:*	**20.0**
Vial Number:	**7**	*Channel:*	**ECD_1**
Sample Type:	**unknown**	*Wavelength:*	**n.a.**
Control Program:	**Carbs061309**	*Bandwidth:*	**n.a.**
Quantif. Method:	**gluc-arab**	*Dilution Factor:*	**1.0000**
Recording Time:	**7/4/2009 20:26**	*Sample Weight:*	**1.0000**
Run Time (min):	**15.00**	*Sample Amount:*	**1.0000**

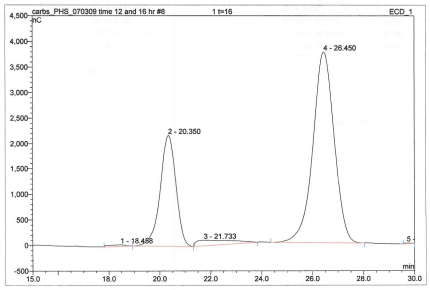


No.	Ret.Time min	Peak Name	Height nC	Area nC*min	Rel.Area %	Amount	Type
1	18.46	n.a.	28.140	24.957	0.46	n.a	BMh
2	20.35	n.a.	2178.172	1557.827	28.96	n.a.	bMb
3	21.73	n.a.	107.820	167.653	3.12	n.a.	bMB
4	26.45	n.a.	3746.950	3625.175	67.39	n.a.	BMB
5	29.70	n.a.	14.276	3.991	0.07	n.a.	BMB
Total:			6075.359	5379.602	100.00	0.000	

Operator:setup Timebase:BioLc Sequence:carbs_PHS_070309 time 12 and 16 hr
Page 1-1
7/5/2009 9:56 AM

8 1 t=16

Sample Name:	1 t=16	Injection Volume:	20.0
Vial Number:	7	Channel:	ECD_1
Sample Type:	unknown	Wavelength:	n.a.
Control Program:	Carbs061309	Bandwidth:	n.a.
Quantif. Method:	gluc-arab	Dilution Factor:	1.0000
Recording Time:	7/4/2009 20:26	Sample Weight:	1.0000
Run Time (min):	15.00	Sample Amount:	1.0000

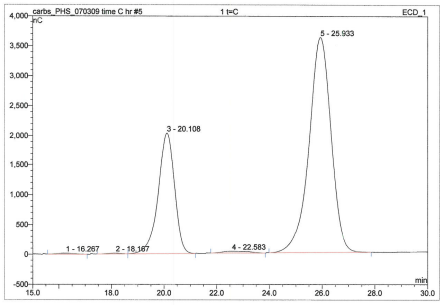


No.	Ret.Time min	Peak Name	Height nC	Area nC*min	Rel.Area %	Amount	Type
1	18.46	n.a.	28.140	24.957	0.46	n.a	BMh
2	20.35	n.a.	2178.172	1557.827	28.96	n.a.	bMb
3	21.73	n.a.	107.820	167.653	3.12	n.a.	bMB
4	26.45	n.a.	3746.950	3625.175	67.39	n.a.	BMB
5	29.70	n.a.	14.276	3.991	0.07	n.a.	BMB
Total:			6075.359	5379.602	100.00	0.000	

DEFAULT/Integration
Chromeleon (c) Dionex 1996-2001
Version 6.60 Build 1428

5 1 t=C

Sample Name:	**1 t=C**	*Injection Volume:*	**20.0**
Vial Number:	**5**	*Channel:*	**ECD_1**
Sample Type:	**unknown**	*Wavelength:*	**n.a.**
Control Program:	**Carbs061309**	*Bandwidth:*	**n.a.**
Quantif. Method:	**gluc-arab**	*Dilution Factor:*	**1.0000**
Recording Time:	**7/12/2009 13:27**	*Sample Weight:*	**1.0000**
Run Time (min):	**15.00**	*Sample Amount:*	**1.0000**



No.	Ret.Time min	Peak Name	Height nC	Area nC*min	Rel.Area %	Amount	Type
1	16.27	n.a.	17.100	11.469	0.23	n.a.	BMB
2	18.17	n.a.	13.286	6.715	0.13	n.a.	BMB
3	20.11	n.a.	2032.493	1462.765	29.16	n.a.	BMB
4	22.58	n.a.	30.028	34.407	0.69	n.a.	BMB
5	25.93	n.a.	3611.730	3500.652	69.79	n.a.	BMB
Total:			5704.637	5016.007	100.00	0.000	

Appendix E: HPAE-PAD Run Settings

REAGENTS AND STANDARDS

Sodium hydroxide, 50% (w/w) (Fisher Scientific and J. T. Baker)
Deionized water, 18 MΩ-cm resistance or higher (Fisher)
Sodium acetate, anhydrous (Fisher)

CONDITIONS
Columns: MA1 Analytical (Dionex)
Flow Rates: 0.5 mL/min
Eluent: A: Water
B: 480 nM NaOH, kept under a nitrogen blanket at 5-8 psi
C: 1 M sodium acetate
Injection Volume: 10 µL
Temperature: 30 °C
Detection (ED50): Integrated pulsed amperometry, disposable Au working electrodes; reference electrode in pH mode

Waveform for the ED50:

Time(s)	*Potential (V)*	*Integration*
0.00	+0.13	
0.04	+0.13	
0.05	+0.33	
0.21	+0.33	Begin
0.22	+0.55	
0.46	+0.55	
0.47	+0.33	
0.56	+0.33	End
0.57	–1.67	
0.58	–1.67	
0.59	+0.93	
0.60	+0.13	

Appendix F: Saturation Density and GFP at harvest for different concentrations of supplemented APS media

All cultures supplemented with .5% arabinose and .2% glucose. Cell density is +/_ 1.6, fluorescence is +/- 32, pH is +/- 0.15.

Table F.1 Sodium chloride

Sodium Chloride (g/L)	Saturation Density (OD600)	pH	GFP (fluorescence)
0	3.5	6.04	95
1.5	3.2	5.98	97
3.0	3.5	5.92	97
4.5	3.4	5.81	98
6.0	3.1	5.74	96
7.5	4.7	5.62	109
9.0	3.1	5.69	98

Table F.2 Tryptone

Tryptone (g/L)	Saturation Density (OD600)	pH	GFP (fluorescence)
0	3.5	6.04	95
5	4.8	5.74	108
10	5.1	5.88	108
15	5.5	6.04	109

Table F.3 Martone

Martone (g/L)	Saturation Density (OD600)	pH	GFP (fluorescence)
7	3.9	-	110
15	3.5	6.04	95

Table F.4 Glycerol

Glycerol (% w/v)	Saturation Density (OD600)	pH	GFP (fluorescence)
0	3.5	6.04	95
.4	4.5	5.50	103
.6	4.2	5.53	102
.8	4.9	5.68	106
1.5	4.2	5.88	100

Table F.5 Phosphate

Phosphate (mM)	Saturation Density (OD600)	pH	GFP (fluorescence)
20	3.2	6.07	93
35	2.8	5.93	89
50	3.5	6.04	95
75	3.0	5.97	93
100	3.0	6.00	97

Table F.6 Aeration

“aeration”	Saturation Density (OD600)	pH	GFP (fluorescence)
very high	8.3	5.39	141
high	7.5	5.39	97
moderate	6.1	5.40	67
poor	2.5	5.72	29

“Aeration” conditions created by filling baffled shaker flask at 10, 20, 40, and 100% of the working volume.

Bibliography

Agrawal P., G. Koshy, and M. Ramseier. “An algorithm for Operating a Fed-Batch Fermentor at Optimum Specific Growth Rate.” *Biotechnology and Bioengineering* 33 (1989): 115 -125.

Ahmed, H. *Principles and reactions of protein extraction, purification, and characterization.* CRC Press, Boca Raton, Florida, 2005.

Albano, C., L. Randers-Eichhorn, Q. Chang, W. Bentley, and G. Rao. “Quantitative measurement of green fluorescent protein expression.” *Biotechnology Techniques* 10 (1996): 953-958.

Application Note 27. “Rapid Measurement of Glucose during Fermentation in Bioethanol Production.” YSI Life Sciences.

Application Note 122. “The Determination of Carbohydrates, Alcohols, and Glycols in Fermentation Broths.” Dionex.

Atkinson B., and F. Mavituna. *Biochemical Engineering and Biotechnology Handbook.* Stockton Press, 1991.

Attrill H., P.J. Harding, E. Smith, S. Ross, A. Watts. “Improved yield of a ligand-binding GPCR expressed in E. coli for structural studies.” *Protein Expression and Purification* 64, no. 1(2009):32-8.

Bailey, J., and D. Ollis. *Biochemical Engineering Fundamentals, 2nd ed.* McGraw Hill: New York, 1986.

Bernard, A., and M. Payton. “Fermentation and Growth of *Escherichia coli* for Optimal Protein Production.” *Current Protocols in Protein Science.* Wiley (2001). http://onlinelibrary.wiley.com/doi/10.1002/0471140864.ps0503s00/pdf

Beverin, S., D. Sheppard, and S. Park. “d-Fucose as a Gratuitous Inducer of the l-Arabinose Operon in Strains of *Escherichia coli* B/r Mutant in Gene *araC.” Journal of Bacteriology* 107, no. 1 (1971): 79-86.

Berrow, N., K. Büssow, B. Coutard, J. Diprose, M. Ekberg, G. Folkers, N. Levy, V. Lieu, R. Owens, Y. Peleg, C. Pinaglia, S. Quevillon-Cheruel, L. Salim, C. Scheich, R. Vincentelli , D. Busso. “Recombinant protein expression and solubility screening in Escherichia coli: a comparative study.” *Acta Crystallography Section D: Biological Crystallography* 62, no. 10 (2006):1218-26.

Biotek. “Quantitation of GFP in Microplates using the FL600.” Application Notes.

Blommel, P., K. Becker, P. Duvnjak, and B. Fox. “Enhanced Bacterial Protein Expression During Auto-induction Obtained by Alteration of Lac Repressor Dosage and Medium Composition.” *Biotechnology Progress* 23, no. 3 (2007): 585–598.

Box, G.,V. Hunter, and J.S. Hunter. *Statistics for Experimenters*. Wiley, New York, 1978.

Brautaset T., R. Lale, and S. Valla.“Positively regulated bacterial expression systems.” *Microbial Biotechnology* 2 (2008):15-30.

Campbell, N. *Biology:Concepts & Connections*. Pearson Education. San Francisco, 2003.

Chalfie, M., Y. Tu, G. Euskirchen, W. Ward, and D. Prasher. “GFP as a Marker for Gene Expression.” *Science* 263 (1994):802-805.

Connell, N. “Expression systems for use in actinomycetes and related organisms.” *Current Opinions in Biotechnology* 12 (2001):446–449.
Deacon S., P. Roach, V. Postis, G. Wright, X. Xia, S. Phillips, J. Knox, P. Henderson, M. McPherson, and S. Baldwin. “Reliable scale-up of membrane protein over-expression by bacterial auto-induction: from microwell plates to pilot scale fermentations.” *Molecular Membrane Biology* 25, no. 8 (2008):588-598.

Debanu D., S. Qian, J. Lee, I. Ankoudinova, C. Huang, Y. Lou, A. DeGiovanni, R. Kim, and S. Kim. “Crystal structure of the multidrug efflux transporter AcrB at 3.1 Å resolution reveals the N-terminal region with conserved amino acids.” *Journal Structural Biology* 158, no. 3 (2007): 494-502.

Document 031824-08. “CarboPac Combined Product Manual.” Dionex.

Don, S. “Optimal Conditions for the Growth of E. Coli.” EEI paper (2008).

Donovan, R., C. Robinson, and B. Glick. “Optimizing expression of a monoclonal antibody fragment under the transcriptional control of the *E. coli* lac promoter.” *Canadian Journal of Microbiology* 46 (2000): 532-541.

Dykhuizen, D. and D. Hartl. “Selection in chemostats.” Microbiology Reviews 47, no. 2 (1983): 150-168.

Epstein, W., L. Rothman-Denes, and J. Hesse. “Adenosine 3”:5”-cyclic monophosphate as mediator of catabolite repression in *E. coli*.” *Proceedings of the National Academy of Science* 72, no. 6 (1975): 2300-2304.

Ferrer-Miralles, N., J. Domingo-Espin , J. Corchero, E. Vazquez, and A. Villaverde. “Microbial factories for recombinant pharmaceuticals.” *Microbial Cell Factories* 8 (2009): 17.
Ferriera, G. “Review on Fed-Batch Fermentations: Mathematical Modelling, Parameters.” Accessed 2010. http://www.scribd.com/doc/27561494/Review-on-Fed-Batch-Fermentations-Mathematical-Modelling-Parameters

Fox, B. and P. Blommel. “Autoinduction of Protein Expression.” *Current Protocols in Protein Science* 56, no. 5 (2009): 1-18.

Gawrylewski, A. “Auto-induction protein production.” *The Scientist* 21, no. 11 (2007): 77.

Geertsma, E., G. van den Bogaard, and B. Poolman. “Maximizing Functional Transport Portein Expression Controlled by the AraC/P_{BAD} System in *E. coli*.” Dissertation, 2005.

Gendron, R. and D. Sheppard. “Mutations in the l-Arabinose Operon of *Escherichia coli* B/r That Result in Hypersensitivity to Catabolite Repression.” *Journal of Bacteriology* 117, no. 2 (1974): 417–421.

Gilbert, S. *Developmental Biology, Eighth Edition*. Sinauer Associates Inc, 2006.

Giomarelli B., K. Schumacher, T. Taylor, R. Sowder 2, J. Hartle, J. McMahon, and T. Mori . “Recombinant production of anti-HIV protein, griffithsin, by auto-induction in a fermentor culture.” *Protein Expression and Purification* 47, no. 1, (2006): 194-202.

Glick, B., J. Pasternak, and C. Patten. *Molecular Biotechnology, 4th ed.* American Society for Microbiology: Washington DC, 2010.

Gosset, G. “Recombinant protein production in cultures of an E. coli trp2 strain.” *Applications in Microbiology Biotechnology* 39 (1993): 541-546.

Grabski, A., M. Mehler, D. Drott. “The overnight express autoinduction system: high-density cell growth and protein expression while you sleep.” *Nature Methods* 2, no. 2 (2005):233.

Graumann K., and Premstaller A."Manufacturing of recombinant therapeutic proteins in microbial systems". *Biotechnology Journal* 1, no. 2 (2006): 164–86.

Greenblatt, J. and R. Schleif. “Arabinose C Protein: Regulation of the Arabinose Operon in vitro.” *Nature* 233 (1971): 166-170.

Gregory, M. *Control of Gene Expression* (2006). Biology Web.
Guzman, L., D. Belin, M. Carson, and J. Beckwith. “Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter.” *Journal of Bacteriology* 177, no. 14 (1995): 4121-30.

Hafiz, A. Principles and Reactions of Protein Extraction, Purification, and Characterization. CRC Press: New York, 2005.

Hammill, J., S. Stoner, J. Hazen, J. Jackson, and R. Mehl. “Preparation of site-specifically labeled fluorinated proteins for 19F-NMR structural characterization.” *Nature Protocols* 2, no. 10 (2007) 2601-2607.

Hannig, G. and S. Makrides. “Strategies for optimizing heterologous protein expression in *Escherichia coli.” Trends in Biotechnology 16 (1998): 54-60.*

Hedge, J. and B. Hofreiter. *Carbohydrate Chemistry, 17.* Eds. Whistler R. and J. Miller. Academic Press: New York, 1962.

Hellens, P., E. Edwards, N. Leyland, S. Bean and P. Mullineaux. “pGreen: a versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation.” Plant Molecular Biology 42, no.6 (2000): 819-832.

Hewitt, C., and A. Nienow. “The scale-up of microbial batch and fed-batch fermentation processes.” *Advances in Applied Microbiology* 62 (2007): 105-135.

Hoskisson, P. and G. Hobbs. “Continuous culture – making a comeback?” *Microbiology* 151 (2005): 3153-3159.

Ingolia, N., and A. Murray. “Positive-feedback loops as a flexible biological module.” *Current Biology* 17, no. 8 (2007): 668–677.

Jeong, K., and S. Lee. “High-Level Production of Human Leptin by Fed-Batch Cultivation of Recombinant *Escherichia coli* and Its Purification.” *Applied Environmental Microbiology* 65, no. 7(1999): 3027-3032.

Johnson M., and J. Lucey. "Major technological advances and trends in cheese". *Journal of Dairy Science* 89, no. 4 (2006) : 1174–8.

Johnson C., and R. Schleif. “In vivo induction kinetics of the arabinose promoters in E. coli.” *Journal of Bacteriology* 177, no. 12 (1995): 3438-3442.

Keasling J. “Gene-expression tools for the metabolic engineering of bacteria.” *Trends in Biotechnology* 17, no.11 (1990): 452-60.

Khlebnikov, A., O. Risa, T. Skaug, T. Carrier, and J. Keasling. “Regulatable arabinose-inducible gene expression system with consistent control in all cells of a culture.” *Journal of Bacteriology* 182, no. 24 (2000): 7029-7034.

Kolodrubetz, D., and R. Schleif. “ L-arabinose transport systems in E. coli K-12.” *Journal Bacteriology* 148 (1981): 472-479.

Larry, E. and R. Block. “Mechanism of initiation and repression of *in vitro* transcription of the *lac* operon of *E. coli*.” *Proceedings of the National Academy of Science* 68 (1971): 1828-1832.

Lee, S. 1996. “High cell-density culture of E. coli.” *Trends Biotechnology* 14 (1996): 98-105.

Lee, S., and H. Chang. “High cell density cultivation of *Escherichia coli* W using sucrose as a carbon source.” *Biotechnology Letters* 15, no. 9 (1993): 971-974.

Lee, S. and J. Keasling. “Heterologous protein production in *Escherichia coli* using the propionate-inducible pPro system by conventional and auto-induction methods.” *Protein Expression and Purification* 61, no. 2. (2008): 197-203.

Lengeler, J., G. Drews, and H. Schlegel. *Biology of the prokaryotes*. New York: Blackwell, 1999.

Life Science Technology. pGLO Bacterial Transformation Kit and Student Manual. BioRad: CA.

Lim, H., K. Jung, D. Park, and S. Chung. “Production characteristics of interferon-□ using an L-arabinose promoter system in a high-cell-density culture.” *Applied Microbiol and Biotechnol* 53 (2000) 201-208.

Lim, H. and K. Jung. 1998. “Improvement of heterologous protein productivity in controlling post induction specific growth rate in recombinant E Coli under control of the pL promoter.” *Biotechnology Progress* 14 (1998): 548-553.

Lobell, R. and R. Schleif. “AraC-DNA looping: orientation and distance-dependent loop breaking by the cyclic AMP receptor protein.” *Journal of Molecular Biology* 218 (1991): 45-54.

Longobardi, G. “Fed-Batch versus Batch Fermentation.” Bioprocess Engineering 10 (1994):185-194.

Lu, C., R. Albano, W. Bentley, and G. Rao. “Differential rates of gene expression monitored by green fluorescent protein.” *Biotechnology and Bioengineering* 79, no. 4 (2002): 429-437.

Lukjancenko, O., T. Wassenaar, and D. Ussery. "Comparison of 61 sequenced Escherichia coli genomes." *Microbiological Ecology* 60, no. 4 (2010): 708–720.

Matthew J., A. Gentile, B. Lovitt, N. Berkley, C. Gunderson, and M. Surber. “Toxic protein expression in Escherichia coli using a rhamnose-based tightly regulated and tunable promoter system.” *BioTechniques* 40, no. 3 (2006): 355–364.

Mathews, C. *Biochemistry*. Benjamin Cummings Publishing Company: California, 1996.
McNeil B. and L. M. Harvey. *Fermentation, a Practical Approach.* IRL Press, Tokyo, 1990.

McNeil, B. and Harvey, L. *Practical Fermentation Technology*. Wiley, 1990.

Megerle, J., G. Fritz, U. Gerland, K. Jung, and J. Radler. “Timing and dynamics of single cell gene expression in the arabinose utilization system. “ *Biophysical Journal* 95 (2008) 2103-2115.

Meszaros A. and V. Bales. “A Contribution to Optimal Control of Fed-Batch Biochemical Processes.” *Bioprocess Engineering* 7 (1992): 363 – 367.

Miller, J. Standards in Fluorescence Spectrometry. Chapman: New York, 1981.

Monteith, C., B. Bowzard, P. Ranjan, and R. Kahn. “Arf and Arl Protein Expression and Purification Using Auto-Induction of BL21 Bacteria.” Web poster session obtained from Emory School of Medicine. (2004).

Morgan-Kiss, R., C. Wadler, and J. Cronan, Jr. “Long term and homogenous regulation of the *E. coli* araBAD promoter by use of a lactose transporter of relaxed specificity.” *Proceedings of the National Academy of Science* 99, no. 11 (2002): 7373-7377.

Mosher, R. “Using pGLO to demonstrate the effects of catabolite repression on gene expression in *E coli*.” *Bioscene* 28, no. 3 (2002).

Neerathilingam M. and J. Markley. “Auto-induction medium containing glyphosate for high-level incorporation of unusual aromatic amino acids into proteins.” *Biotechniques* 49, no. 3 (2010): 659–661.

Nelson, D. and M. Cox. Lehninger Principles of Biochemistry, 5th ed. Freeman, 2008.

Nielson, B., V. Willis, and C. Lin. "Western blot analysis to illustrate relative control levels of the lac and ara promoters in E. coli." *Biochemistry and Molecular Biology Education* 35, no. 2 (2007): 133-137.

Nishi, K., I. Kim, and S. Ma. "Expression of the human soluble epoxide hydrolase in *Escherichia coli* by auto-induction for the study of high-throughput inhibition assays." *Protein Expression and Purification* 69 (2010): 34-38.

NIST/SEMATECH e-Handbook of Statistical Methods, http://www.itl.nist.gov/div898/handbook/, 2011 (2006).

Novick A., and M. Weiner. "Enzyme Induction as an all or nothing phenomenon." *Proceedings of the National Academy of Science* 43, no. 7(1957): 553-66.

Novotny C., and E. Englesberg. " The L-arabinose permease system in E. coli." *Biochim Biophys Acta* 117 (1996):217-230.

Ogden, S., D. Haggerty, C. Stoner, D. Kolodrubetz, and R. Schleif. "The E. coli L-arabinose operon: binding sites of the regulatory proteins and a mechanism of positive and nedgative regulation." *Proceedings of the National Academy of Science* 77 (1980): 3346-3350.

Ozbudak E., M. Thattai, H. Lim, B. Shraiman, and A. van Oudenaarden. "Multi-stability in the lactose utilization network of Escherichia coli." *Nature* 427 (2004):737–740.

Paalme T., K. Tiisma, A. Kahru, K. Vanatalu, and R. Vilu. "Glucose-limited fed-batch cultivation of Escherichia coli with computer-controlled fixed growth rate." *Biotechnology and Bioengineering* 35 (1990): 312-319.

Ramirez D. and W. Bentley. "Fed-Batch Feeding and Induction Policies that Improve Foreign Protein Synthesis and Stability by Avoiding Stress Responses." *Biotechnology and Bioengineering*, 47 (1995): 596-608.

Remans T., P. Schenk, J. Manners, C. Grof , and A. Elliott. "A protocol for the fluorometric quantification of mGFP5-ER and sGFP(S65T) in transgenic plants." *Plant Molecular Biology Reports* 17 (2000): 385–395.

Richards, H., M. Halfhill, R. Millwood, and C. Stewart. "Quantitative GFP fluorescence as an indicator of recombinant protein synthesis in transgenic plants. *Plant Cell Reports* 22 (2003):117–121.

Reisenber, D. and R. Guthke. “High cell density cultivation of microorganisms.” *Applications in Microbiological Biotechnology* 51 (1999): 442-430.

Riesenberg, D., V. Schulz, W. Knorre, H. Pohl, D. Korz, E. Sanders, A. Ross, and W. Deckwer. “High cell density cultivation of E. coli at controlled specific growth rate.” *Journal of Biotechnology* 20, no. 1 (1991):17-27.

Rowe, L., M. Ensor, R. Mehl, and S. Daunert. “Modulating the bioluminescence emission of photoproteins by *in Vivo* site-directed incorporation of non-natural aminao acids.” *ACS Chemical Biology* 5, no. 5 (2010): 455-460.

Rozkov, A., C. Avignonerossa, P. Ertl, P. Jones, R. Okennedy, and J. Smith. “Fed batch culture with declining specific growth rate for high-yielding production of a plasmid containing a gene therapy sequence in Escherichia coli DH1.” *Enzyme and Microbial Technology* 39 (2006): 47-50.

Savageau, M. “Demand Theory of Gene Regulation. II. Quantitative Application to the Lactose and Maltose Operons of *E. coli*.” *Genetics* 149 (1998): 1677-1691.

Schleif, R. “AraC protein: a love-hate relationship.” *BioEssays* 25 (2003):274-282.

Schleif, R. “DNA Looping.” *Annual Reviews Biochemistry* 61 (1992): 199-223.

Schleif, R. “DNA Looping.” *Science* 241 (1988):127.

Schleif, R. *Genetics and Molecular Biology, 2nd ed*. The Johns Hopkins University Press, 1993.

Schleif, R. “Regulation of the L-arabinose operon of E. coli.” *Trends in Genetics* 16 (2000): 559-565.

Schallmey, M., A. Singh, and O. Ward. “Developments in the use of Bacillus species for industrial production.” *Canadian Journal of Microbiology* 50 (2004):1-17.

Shuler, M. and F. Kargi. *Bioprocess Engineering Basic Concepts, 2nd ed*. Prentice Hall: Upper Saddle River, NJ (2002).

Siegele, D., and J. Hu. “Gene expression from plasmids containing the *araBAD* promoter at subsaturating inducer concentrations represents mixed populations.” *Proceedings of the National Academy of Science* 94, no. 15 (1997): 8168-817.

Sigma Chemicals. Material Safety Data Sheet: IPTG. (2010).

Simons, R., F. Houman, and N. Kleckner. “Improved single and multicopy lac-based cloning vectors for protein and operon fusions.” *Gene* 53, no. 1(1987): 85-96.

Sorensen, H., and K. Mortensen. “Advanced genetic strategies for recombinant protein expression in *Escherichia coli.” Journal of Bacteriology* 115 (2005): 113-128.

Sreenath H., C. Bingman, B. Buchan, K. Seder, B. Burns, H. Geetha, W. Jeon, F. Vojtik, D. Aceti, R. Frederick, G. Phillips Jr, B. Fox. “Protocols for production of selenomethionine-labeled proteins in 2-L polyethylene terephthalate bottles using auto-induction medium.” *Protein Expression and Purification* 40, no. 2, (2005):256-67.

Stanbury P., A. Whitaker, and S. Hall. *Principles of Fermentation Technology*. Pergamon, 1993.

Steinmetz, E. “Expresso Cloning and Expression Systems: Expressioneering Technology streamlines recombinant protein expression.” *Nature Methods* 8 (2011).

Studier, F. “Protein production by auto-induction in high-density shaking cultures.” *Protein expression and Purification* 41 (2005): 207-234.

Studier, F., J. Sutherland, L. Miller, and L. Yang. “High-throughput production and analyses of purified proteins.” *Technology Development and Use* (2006). http://genomicscience.energy.gov/pubs/2006abstracts/2006gtlabstractsweb_protein.pdf

Sung K., H. Chou, B. Pfleger, J. Newman, Y. Yoshikuni, and J. Keasling. “Directed Evolution of AraC for Improved Compatibility of Arabinose- and Lactose-Inducible Promoters.” *Appl Environ Microbiol*. 73, no. 18 (2007) : 5711–5715.

Svensäter, G., O. Björnsson, and I. Hamilton. “Effect of carbon starvation and proteolytic activity on stationary-phase acid tolerance of *Streptococcus mutans.” Microbiology* 147, no. 11 (2001): 2971-2979.

Technical Note 21. “Optimal Settings for Pulsed Amperometric Detection of Carbohydrates Using the Dionex ED40 Electrochemical Detector.” Dionex.

Technical Notes. “Bacterial E. coli growth media.” Expression Technologies (2003).

Terpe, K. “Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems.” *Applied Microbiology and Biotechnology* 72, no. 2 (2006): 211-22. Epub 2006 Jun 22.

Timmes, A., M. Rodgers, and R. Schleif. “Biochemical and Physiological Properties of the DNA Binding Domain of AraC Protein.” *Journal of Molecular Biology* 340 (2004): 731-738.

Tyler, R., R. Tyler, H. Sreenath, S. Singh, D. Aceti, C. Bingman, J. Markley, and B. Fox. “Auto-induction medium for the production of [U-15N]- and [U-13C, U-15N]-labeled proteins for NMR screening and structure determination.” *Protein Expression and Purification* 40, no. 2 (2005): 268-278.

U.S. Patent no. 4,952,496. The pET system. Genetics Institute, Inc., 87 Cambridge Park Drive, Cambridge, MA 02140.

U.S. Patent pending 20100068758. An L-arabinose inducible expression system comprising a mutant arabinose promoter. E.I. Dupont Nemours.

Venter, J. C. “Summary of Escherichia coli, Strain K12, version 12.1.” SRI International, Marine Biological Laboratory, DoubleTwist Inc., The Institute for Genomic Research, J. Craig Venter Institute, University of California at San Diego, UNAM, and Macquarie University (2008). http://biocyc.org/ECOLI/organism-summary?object=ECOLI

Watson, J., R. Myers, A. Caudy, and J. Witkowski. *Recombinant DNA, 3rd ed.* Freeman, 2006.

Watson, J., T. Baker, S. Bell, A. Gann, M. Levine, R. Losick. *Molecular Biology of the Gene, 6th ed.* Benjamin Cummings, 2007.

Weast, R. *CRC Handbook of Chemistry and Physics* (62nd edition). Boca Raton, FL: CRC Press, 1981.

Wei. T., K. Cai, X. Gao, L. Xiao, R. Chen, J. Shi, H. liu, X. Hou, Q. Wang, and H. Wang. “Improved production of holotoxin Stx2 with biological activities by using a single-promoter vector and an auto-induction expression system.” *Protein Expression and Purification* 67 (2009): 169-174.

Wei Y. and D. Fu. “Selective Metal Binding to a Membrane-embedded Aspartate in the *Escherichia coli* Metal Transporter YiiP (FieF).” *The Journal of Biological Chemistry*, 280 (2005).

White, D. *The physiology and biochemistry of prokaryotes, 2nd ed*. New York: Oxford University Press, 1999.

Wilcox, G. “The Interaction of l-Arabinose and d-Fucose with *AraC* Protein.” *The Journal of Biological Chemistry* 249 (1974): 6892-6894.

Wiseman, A. “Enzyme induction in microbial organisms.” *Basic Life Science* 6 (1975):1-26.

Wrolstad, R., E. Decker, S. Schwartz, and P. Sporns. Handbook of Food Analytical Chemistry, Water, Proteins, Enzymes, Lipids, and Carbohydrates. Wiley, 2005.

Xia X., V. Postis, M. Rahman, G. Wright, P. Roach, S. Deacon, J. Ingram, P. Henderson, J. Findlay, S. Phillip, M. McPherson, S. Baldwin. “Investigation of the structure and function of a Shewanella oneidensis arsenical-resistance family transporter.” *Molecular Membrane Biology* 25, no. 8 (2008):691-705.

Yee, L., and H. Blanch. “Recombinant protein expression in high cell density fed-batch cultures of E. coli .” *Nature Biotechnology* 10(1992): 1550-6.

Yoon, S., W. Kang, and T. Park. “Fed-batch operation of recombinant Escherichia coli containing trp promoter with controlled specific growth rate.” *Biotechnology and Bioengineering* 43 (1994): 995-999.

Zhang, X., and R. Schleif. “Catabolite gene activator protein mutations affecting activity of the araBAD promoter.” *Journal of Bacteriology* 180, no. 2 (1998): 195-200.

Zhou, Y. and S. Gottesman. “Regulation of Proteolysis of the Stationary-Phase Sigma Factor RpoS.” *Journal of Bacteriology* 180, no. 5 (1998): 1154-1158.