Recombinant Collagen Production Optimization in *Escherichia coli*

by

Brett A. Whittemore

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Engineering in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 16, 2005

Author . . . .
Department of Electrical Engineering and Computer Science
February 2, 2005

Certified by . . . .
Jean-François P. Hamel
Lecturer
Thesis Supervisor

Certified by . . . .
Karen K. Gleason
Professor
Thesis Supervisor

Accepted by . . . .
Arthur C. Smith
Chairman, Department Committee on Graduate Theses

Recombinant Collagen Production Optimization in

Escherichia coli

by

Brett A. Whittemore

Abstract

An *Escherichia coli*-based collagen-production process was used to investigate several process optimization objectives for use at the industrial scale. The effect of cooling on fermentation growth kinetics was studied, with preliminary results indicating that cooling an early to mid-batch phase fermentation ($OD_{600} = 2$–4) to 10°C for up to 40 hours was not detrimental to the recovery of fermentation growth or the culture's ability to reach high cell density ($OD_{600} > 50$). A cost-effective assay of collagen-like polymer was examined under high-cell density conditions (as opposed to previously-studied low density conditions) and an experimental design for tailoring the assay to high cell density fermentation samples is presented. In addition, a dual-plasmid strain of *E. coli* was designed for use in a novel process for the mass production of collagen-like polymers: one plasmid contains a thermally inducible recombinant collagen gene (CLP3.1-his), and the other contains an arabinose-inducible lytic gene (bacteriophage T4 t-holin) along with the basally expressed T7 lysozyme gene from the pLysS plasmid. A methodology for the optimization of the sequential induction of CLP3.1-his followed by induction of the t-holin is presented; a review of the literature suggests that decreased growth rate is detrimental to lytic efficiency (both the time required for lysis and the degree of lysis). The resulting process will enable lysis to take place in the bioreactor, thus avoiding the extra time and monetary cost of a separate cell homogenization step for *E. coli* disruption and endogenous protein release.

Thesis Supervisor: Jean-François P. Hamel
Title: Lecturer

Thesis Supervisor: Karen K. Gleason
Title: Professor

Acknowledgments

When Jean-François Hamel first approached me on the last day of his bioreactor module in BEH.109 (Laboratory Fundamentals in Biological Engineering), he proposed a summer UROP project in which I would implement an *E. coli* fermentation for the Chemical Engineering department's new course, 10.28 (Biological Engineering Laboratory).

By the end of the summer, I still wasn't completely finished with the project, so I luckily had to stay on in the lab. After two months working under Jean-François' supervision, it was clear what a dream-come-true it would be to have him as a thesis advisor. Thanks to J-F's belief that an electrical engineer had something to offer a chemical engineering lab, in addition to Karen Gleason's sponsorship as a thesis co-advisor, my bioprocess engineering thesis opportunity was born.

Thinking back on the nearly two-year experience working towards the completion of this thesis, there is one overwhelmingly pervasive theme: the flavor of J-F's leadership. Working closely with J-F was an incredible experience: I had not yet had to opportunity to work closely with a brilliant academician who was at the same time so kind, generous, and trusting. The freedom and responsibilities that J-F entrusted to me over the course of my time working in his laboratory have fundamentally changed my approach to work. Following the leadership of someone who's outlook and intentions are fundamentally *good* has shown me that one's capabilities are vastly increased with the confidence instilled by such a thoroughly honest and considerate leader. Thank you, J-F, for this experience and for your support.

My first experience working with an *E. coli* fermentation came through Anil Salgotra, and John Lorusso taught me the practical knowledge of instrumentation and equipment setup that are vital to implementing a working experiment. These two excellent teachers laid a solid foundation for all of the lab work that ensued.

One of the benefits of working in the Hamel Lab was the opportunity to work with chemical engineering masters students from the Institut Químic de Sarrià in Barcelona, Spain. In addition to the colorful language, dancing moves, and out-of-lab

style that they taught me, Jaime "Santi" Compte, Mercè Dalmau, Cristina Ramírez, Jordi Gibert, and Laura Lorenzo-Gonzalez have become great friends. Their unflagging support during the most grueling fermentations was integral to the completion of this thesis and the quality of the undergraduate courses that we taught.

In addition to the members of the lab already mentioned, I would like to emphasize the contributions of Markéta Valterová and Anjali Verghese, whose support, senses of humor, and eagerness to be kind friends greatly humanized my experience in the lab.

As you will read, much of the progress embodied by this thesis is due to the brute force of undergraduates who endured the implementation of our ideas in their courses. The dedication and resilience of the students in the Fall 2003 and 2004 offerings of courses 10.28 (Biological Engineering Laboratory) and 10.26 (Chemical Engineering Projects Laboratory) made my involvement with teaching the courses an awesome experience.

I am also very grateful to the undergraduate UROP students who worked with me to implement parts of this thesis. Their enthusiasm and excellence in contributing original ideas to the project were wonderful. Mike Xiang, Grace Lin, and Bashira Chowdhury made valuable contributions to off-gas analysis, the collagenase assay, and the lytic strain design, respectively.

Contents

List of Figures

11

List of Tables

Chapter 1

Introduction

The field of bioprocess engineering is concerned with the mass production of organic products. A bioprocess is based on the same paradigm as chemical production processes implemented by chemical engineers: chemical reactants are fed into a reactor, in which the reaction takes place, and the reaction products are monitored. The host organisms used in a bioprocess can be thought of as the catalyst for the reaction, which produces a complex organic molecule or enzyme.

The work in this thesis is concerned with the optimization of an *E. coli* process that produces a collagen-like polymer. Collagen is a ubiquitous structural protein that is used for a wide variety of biological purposes, including acting as a drug delivery carrier, a part of artificial skin, and surgical suture strand. There are 19 types of naturally-occurring collagen, but the most ubiquitous (and consequently the most commercially available) are type I and type II, which are present in cells as striated fibrils that serve a structural purpose [50]. Collagen obtained from animal sources is at risk for contamination by infectious agents residing in the host animal (e.g. bovine spongiform encephalopathy, BSE) [67]. These risks, in addition to high maintenance costs and regulations on animal production, make recombinant *E. coli* an attractive alternative to animal sources for obtaining collagen [72].

Gelatin is the denatured, degraded form of collagen, and has applications in the biomedical industry where it is used in matrix implants, drug delivery microspheres, and the production of photographic film. Single-stranded collagen-like polymers are

most similar to gelatin. Collagen-like polymers are defined by the repeating triple-peptide sequence of glycine followed by two variable amino acids (Gly–X–Y). In nature, proline and hydroxyproline (post-translationally modified proline, augments triple helix stability [67]) are the most common X and Y constituents, and collagen exists as a left-handed triple helix. Thus, another motivation for the development of recombinant collagen-like polymers is that applications in which they are used need not be constrained by the properties of naturally-occurring collagen sequences [72]. See Section 3.1 for a complete description of the recombinant product, CLP3.1-his.

Convenience was the main reason for basing the project on a collagen-producing strain rather than using another type of protein product; significant work had already been done with the strains and assay methods by D.I.C. Wang, Jin Yin, and colleagues in the Bioprocess Engineering Center (BPEC) at MIT, and it was straightforward to restart work on this project [72, 73]. However, many of the contributions of this thesis are independent of the identity of the product protein, and can potentially be applied to other *E. coli*-based processes as well.

The Hamel Lab, in which this thesis work was done, serves as both a research and teaching lab. Some of the lab's research is done through liaisons with industry. In particular, several of the projects are focused on the analysis of new instruments which can be used to engineer better, more highly-controlled bioprocesses. Many of the experimental results for these projects came from the *E. coli* processes that were used for this thesis, and are thus presented here.

In accordance with the teaching function of the lab, this thesis work was incorporated into two courses in the Chemical Engineering department in order to teach undergraduates the fundamentals of bioprocess engineering. Through its association with these courses, this thesis work gained the benefit of a large amount of state of the art equipment and many more sets of hands to implement the experiments than I could have done alone.

In order to orient the reader to the context of the work done in this thesis, the next section begins with a brief history of recombinant DNA technology and its importance in the field of bioprocess engineering. The specific motivation for this thesis project

and its contributions are then described, and the chapter concludes with a roadmap for the rest of the thesis.

1.1   Overview of bioprocess engineering

After Watson and Crick discovered DNA in 1953 [65], two decades passed before recombinant DNA technology had advanced sufficiently to make it possible to engineer organisms capable of producing a recombinant protein product. In 1972, Jackson, Symons, and Berg created the first recombinant plasmid, which contained genes from Simian Virus 40, λ-phage genes, and the galactose operon of *E. coli* [26]. In 1973, Cohen, *et al.* created the first genetically engineered organism: *E. coli* transformed with plasmids engineered to confer antibiotic resistance [12]. Because of its genetic simplicity and its rapid growth rate relative to other organisms, *E. coli* became the most popular host organism for the mass production of recombinant proteins. Approved by the FDA in 1982, human insulin produced by recombinant *E. coli* became the first genetically engineered drug to enter the market [15]. Mammalian cells, yeast, and insect cells are also widely used as host organisms. In 1987, human tissue plasminogen activator, made by Genentech, became the first product produced by mammalian cells to be approved by the FDA [71, 15]. (It is important to note that the central dogma of molecular biology—the processes by which DNA is replicated and transcribed then translated into protein—is universal across all organisms; thus, the primary peptide sequence corresponding to any gene sequence can theoretically be manufactured by any organism, provided that the organism is provided with a copy of the gene and an appropriate regulation mechanism. Thus, prokaryotic *E. coli* can be made to produce a eukaryotic peptide sequence.)

The ensuing advances in the fields of genetics and microbiology combined with the progress made by the human genome project have led to an explosion in the number of new recombinant products that are under development and on the market. In general, since these products are most often used therapeutically in humans, they are eukaryotic proteins. For some products, *E. coli*'s lack of the cellular machinery

needed to fold and secrete a eukaryotic product correctly does not pose a problem to protein biochemists who can coax the linear peptide strand from an *E. coli* harvest into the correct correct biologically active conformation. However, *E. coli*-based production is sometimes not feasible for specific products. In addition, concurrent advances in protein purification and analysis technology have shown that the efficacy of many proteins is dependent on the environment in which they were produced. The host organism used is a major environmental factor, but more subtle factors such as growth conditions can have a large effect on a protein's glycosylation state, which can dramatically affect product quality [11]. Consequently, even though *E. coli* processes have the potential to make several orders of magnitude more protein than more slowly growing host organisms (e.g. mammalian cells, insect cells, yeast), product quality considerations related to host organism have come to the forefront. The drive for homogeneous, well-folded, biologically active products of maximum purity has boosted support for processes that use mammalian cells. The rapid growth in the total number of bioprocesses, as well as the increase in mammalian relative to *E. coli* processes is shown in Figure 1.1.

Despite the fact that mammalian cells have replaced *E. coli* as the most common host organism, *E. coli* remains a major workhorse for the bioprocess industry. In a traditional industrial process that uses *E. coli*, a strain is created by inserting the gene of interest into a plasmid, which is then inserted into the *E. coli*. The plasmid usually contains a gene that confers resistance to a specific antibiotic (called the selection marker antibiotic), an origin of replication (so the plasmid is copied during each chromosome replication cycle and persists in successive generations), and an inducible operon which has a specific site into which the product gene can be inserted so that expression of the product can be controlled. The expression systems used in this thesis include the IPTG[1]-inducible *lac* operon, the arabinose-inducible pBAD system, and a thermally-inducible system (cIts). These expression systems are discussed in further detail in Chapters 3 and 6.

The bacterial culture is grown in a vessel called a bioreactor (or "fermenter"),

---

[1]isopropyl-β-D-thiogalactoside.

Approved and manufactured pharmaceuticals
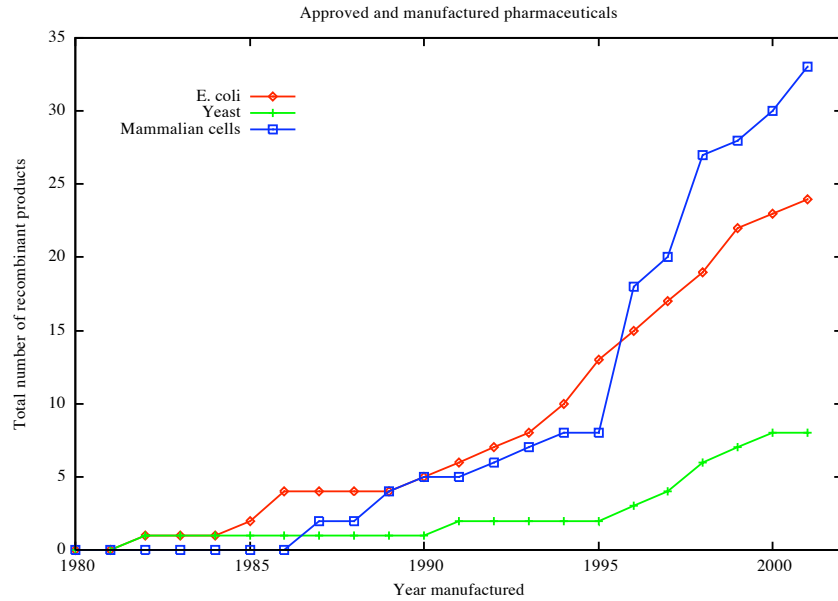
Figure 1-1: Total number of approved and manufactured pharmaceuticals by year for popular host organisms [64].

which contains a nutrient rich culture medium; by convention a process that uses *E. coli* or yeast is called a "fermentation," even if it is aerobic. The vessel is equipped with sensors and controllers that maintain a constant, ideal growth environment in its interior. A typical fermentation designed for production consists of three phases: batch, fed-batch, and induction. Shake flasks (typically 100 mL working volume) are inoculated with a small amount of a frozen stock solution of *E. coli* and incubated until the nutrient supply is depleted. This inoculum is added to a bioreactor containing sterile medium. The first phase of the fermentation is called the "batch" phase because the culture grows until the energy source (usually a carbon source such as glucose) is depleted. Nothing is added or removed from the system during this phase, which allows the culture to adjust to the bioreactor environment and to begin steady growth. To improve process efficiency and net yield, the culture must be grown to a higher cell density than can be reached during the batch phase. Hence, during the fed-batch phase, an energy source (often glucose for *E. coli*) is fed to the culture to achieve higher cell densities. At the end of the fed-batch phase, once the culture has reached the optimum cell density and growth rate for the particular process, expression of

the recombinant gene is induced, and the induction phase begins. At the end of the induction phase, the culture is harvested and the protein product purified.

In this thesis, the fermentation harvest was routinely disrupted by sonication in order to break open the *E. coli* and release the intracellular recombinant product, CLP3.1. By means of a series of centrifugation and endogenous protein precipitation steps, CLP3.1 was purified. The purified product was assayed using a collagenase and ninhydrin[2]-based assay developed by Yin, *et al.* which was a cheap and rapid alternative to a standard immunobased assay [73].

1.2   Motivation

The development of therapeutic products is an expensive process, costing approximately one billion dollars to develop, produce, and market a successful product [20]. In this context, the development of novel bioprocesses, such as the one proposed in this thesis, is a worthwhile approach to keeping production costs as low as possible.

This thesis aims to optimize the cell disruption step of downstream processing by designing a strain of *E. coli* conferred with the recombinant ability to lyse. Achieving cell disruption by inducing lysis would obviate the need for the current industry standard method of cell disruption: high pressure homogenization. Homogenizers are major pieces of equipment and the samples must be passed through the system several times to achieve complete disruption because of the resilience conferred by the tough *E. coli* cell wall.

Any process optimization methodology requires rigorous empirical characterization in order to make the process as reproducible and efficient as possible. In addition to the industrial motivation for such empirical rigor, the use of the processes studied in this thesis in an undergraduate bioprocess engineering course strongly motivates the development of a robust and reproducible process. Thus, the results of an in-depth quantitative analysis of the fermentations performed in this thesis are vital to the design of robust fermentations that will be successful and predictable in a course

---

[2]triketohydrindenhydrate

setting, in which there is a high degree of experimental variability. Furthermore, the limitations on what students may be exposed to during the strictly-scheduled laboratory hours of a course[3] motivate the development of flexible fermentations that may be manipulated to "perform" during lab hours. This thesis explored the use of cooling phases in fermentations to arrest their growth in between laboratory sessions, so that students could view a larger fraction of each process than just the four hour window they would observe in an uninterrupted fermentation.

In order to study the production efficiency of a process, a product-specific assay must be at the process engineer's disposal. Immuno-based assays, which take advantage of the highly specific affinity of particular antibodies for particular proteins, have high sensitivity and specificity. However, these assays are expensive and time consuming. Yin, *et al.* developed a low-cost and relatively fast assay for the detection of collagen-like polymers in *E. coli* crude lysate from shake flask cultures [73]. Use of this assay to analyze samples from high cell density fermentations is desirable in order to transfer the benefits of the low cost and rapidity of the assay to fermentation analysis. As described in this thesis, though, the form of the assay published by Yin, *et al.* is not directly transferable to shake flasks. Hence, there is strong incentive for the study (and improvement) of the assay's performance on high cell density samples.

1.3 Contributions of this thesis

The work done in this thesis comprises several necessary steps towards the implementation of a novel, optimized collagen-production and assay process. These steps are described below:

---

[3]For example, MIT course 10.28 meets twice a week, with four hours per lab session.

1.3.1 Robust *Escherichia coli* fermentation protocol for academic use

Because this research was conducted in close association with MIT courses 10.28 (Biological Engineering Laboratory) and 10.26 (Chemical Engineering Projects Laboratory) in the Chemical Engineering Department at MIT, it was essential to develop a robust high cell density fermentation protocol. The sheer number of fermentations carried out over the course of this project (about 25) enabled the collection of a large amount of data from a variety of conditions. This data has enabled the development of a sturdy fermentation setup and protocol that will be heavily used in future *E. coli* studies motivated by this project.

1.3.2 Analysis of collagenase assay in high cell density conditions

The *E. coli* densities of fermentation induction samples are up to two orders of magnitude greater than the densities studied in the original collagenase assay experiments. In addition, there are other major differences between the shake flask and stirred-tank culture methods. Data from this thesis' experiments imply that the collagenase assay protocol, developed using shake flask experiments [73], is insufficient for the analysis of fermentation samples from bioreactors. Several hypotheses regarding the failure of the current collagenase protocol as well as an experimental plan for the improvement of the assay are proposed. In addition, the prescribed use of the bicinchoninic acid (BCA) assay in conjunction with the collagenase assay to quantify collagen production was based on oversimplistic assumptions about protein reactivity and must be reevaluated.

1.3.3 Bioreactor cooling

Ubiquitously, fermentation processes are run without interruption from inoculation to harvest. High cell density processes are often at least fifteen hours long, mak-

ing the task of monitoring a fermentation grueling. This thesis work experimented with arresting the kinetics of a fermentation by cooling it to 10°C.[4] Further study is required for accurate quantification of the effect of cooling on product production, but initial results suggest that growth kinetics (once the culture has been warmed back up) are unaffected by cooling when cell densities are relatively low. Successful fermentations implemented in course 10.28 that included a period of arrested growth demonstrate the potential of the cooling technique for humanizing the bioprocess engineer's schedule in the classroom and research settings. Use of a cooling phase has enabled the prolongation of course fermentations to increase bioprocess engineering students' exposure time to live processes over noncontiguous laboratory sessions.

1.3.4 Design of lytic strain

A plasmid containing the λ-phage R-gene (a lysozyme: transglycosylase) and used in a published *E. coli* system with both recombinant product and lytic capabilities [74] was tested with BL21(DE3), the host strain optimized for collagen production by Yin, *et al.* [72]. The resultant R-gene-based system yielded unsatisfactory lytic behavior, which lead to the development of the lytic strain design described in this thesis. The final strain design is named BL21(DE3)[pLysS-t][pJHL], containing a thermally-inducible collagen gene (pJHL) and a modified version of the pLysS plasmid (named pLysS-t) which produces a T7 lysin at low basal levels. This strain design will be implemented for use in a process engineering project to optimize its lytic efficiency at high cell densities in MIT course 10.26 during the Spring 2005 term.

1.4 Roadmap

Chapter 2: Literature review summarizes prior efforts in the production of recombinant collagen and collagen-like polymers, summarizes work that has been

---

[4]Wurm, *et al.* obtained positive results experimenting with the cooling of Chinese Hamster Ovary (CHO) cell cultures. When temperature ranges from 10°C to 25°C were used, cooling periods of 3–4 days could be imposed with minimal negative effect on transfection and protein production performance. Lower temperatures (0–4°C) caused a rapid drop in viability [70].

done with mechanical, chemical, and recombinant methods for disrupting *E. coli* to release intracellular products, reviews the current understanding of the bacteriophage lytic mechanism, and gives the theoretical underpinnings of the collagenase and BCA assay mechanisms.

Chapter 3: Experimental setup contains technical information and protocols for recombinant *E. coli* techniques, a description of the bioreactor apparatus, and details of the analytical protein assays.

Chapter 4: Quantitative fermentation results presents experimental fermentation results in order to define a robust and repeatable CLP production process.

Chapter 5: Analysis of quantitative protein assays analyses results of the collagenase and BCA assays in the context of their reaction mechanisms, reevaluates the assumptions governing their use and reliability, and outlines an experimental approach to studying the assays with the intention of changing them to make them more accurate.

Chapter 6: Design of lytic *E. coli* strain motivates the construction of a lytic strain, describes the experiments with λ-phage DNA that prompted the design of a new lytic mechanism, and presents the design for the lytic strain BL21(DE3)[pLysS-t][pJHL].

Chapter 7: Conclusion

Appendix A: Media and reagents contains a detailed description of the *E. coli* growth media, collagenase assay reagents, and lysis buffers for use with strains containing the λ-phage DNA fragment.

Appendix B: Thesis chronology and relation to undergraduate courses describes the chronological and academic contexts for the experiments performed over the course of this thesis.

Appendix C: Microscopic images of *E. coli* contains microscope slide images from shake flask, fermentation, and sonication samples.

Chapter 2

Literature review

This chapter begins with a review of previous work that has been done with the production of recombinant collagen. This chapter then summarizes techniques for releasing intracellular products from *E. coli* and elucidates the need for the studies with lytic phage enzymes at high cell densities discussed in this thesis. In addition, the mechanism by which bacteriophages lyse *E. coli* is presented, and the mechanisms of the collagenase and bicinchoninic (BCA) assays are described.

2.1   Recombinant collagen production

A variety of expression systems have been designed for the production of both natural and synthetic collagens. Table 2.1 presents a summary of the production levels achieved by each system based on information compiled by Yin, *et al.* [72]. It is clear that production levels several orders of magnitude higher can be achieved using expression systems that produce single-chain collagen-like polymers rather than natural triple-helical collagens. For applications in which the triple-helical form is not necessary (discussed briefly in Chapter 1), yeast or bacterial expression systems are the clear choice for maximum production.

The inability of *E. coli* to add post-translational modifications to eukaryotic proteins, the propensity for production of inclusion bodies (solid protein aggregates), and the likelihood of incomplete folding have been cited as disadvantages to *E. coli*

Natural triple-helical collagens (type I-III)		
Host organism	Collagen quantity (mg/L)	Reference No.
Mammalian cells	0.6-20	[72]
Yeasts	15-600	[72]
Insect cells	10-60	[72]
Plants	30 (mg/kg)	[72]

Single-chain collagen-like polymers		
Host organism	CLP quantity (g/L)	Reference No.
Pichia Pastoris (single-copy transformant)	4-5	[72]
Pichia Pastoris (15-copy transformant)	14.8	[72]
Bacillus brevis	0.5	[72]
E. coli	50 (mg/L)	[50]

Table 2.1: Variation in quantity of collagen-like polymers (CLPs) produced with different host strains.

expression systems in general [50]. However, considering that the target product of this thesis (CLP3.1) is intended to be a soluble, single-stranded polypeptide, the post-translational modification, folding, and inclusion body considerations are likely to be irrelevant [72]. In their *E. coli* expression systems, Ruggiero, *et al.* expressed several fragments of a natural collagen gene. One fragment, from the heparin binding domain (contains one glycine every three amino acids), was produced at a level of 0.05 mg per 1 mL of bacterial culture.[1] Other fragments that were less glycine-rich were produced with yields up to 0.5 mg per 1 ml of bacteria. It was theorized that bacterial levels of GlytRNA might be too low to express large amounts of collagen relative to other recombinant proteins [50]. However, based on the results of Yin, *et al.*, where expression levels of glycine-rich CLP3.1 were achieved, lack of GlytRNA does not appear to be a limiting factor in *E. coli* expression systems. Yin, *et al.* claimed that the low production levels in *E. coli* systems seen up to the time of their study were most likely due to the lack of studies optimizing the expression system [72].

Werten, *et al.* achieved 3-6 g/L levels of artificial CLP production using single-copy *Pichia Pastoris* transformants that secreted the CLP into the fermentation broth. They designed their P4 gelatin to be especially hydrophilic relative to natural colla-

---

[1]The OD_{600} value at the time of induction was not given.

gens and to have an isoelectric point similar to that of natural cattle bone gelatin by controlling the number of basic lysine residues. The P4 gelatin was a 36.8 kDa protein [67]. In a previous study, Werten, *et al.* achieved expression levels up to 14.8 g/L of recombinant gelatins based on mouse type I and rat type III collagen sequences [66]. Their 5 copy and 15 copy transformants produced 11.3 g/L, and 14.8 g/L of gelatin, respectively. A disadvantage of the yeast expression system is that proteases can degrade the collagen product (both extra- and intracellular proteases). In the high copy number systems, Werten, *et al.* suppressed some degradation by conducting the fermentation at very low pH (3.0 rather than 5.0). However, based on the results of their P4 system, Werten, *et al.* speculated that arginine residues (which were intentionally limited in the P4 gelatin) increased its susceptibility to proteolytic cleavage (site-directed mutagenesis of the sequence Met–Gly–Pro–Arg to Arg–Gly–Pro–Met resulted in the disappearance of cleavage products).

In another system for the extracellular secretion of collagen-related proteins, Kajino, *et al.* inserted PHI and NEU subunits of human αI collagen into *Bacillus brevis*, obtaining production levels of 0.5 g/L [29]. In recombinant systems where the product gene has a large number of short, repeating sequences, stability of the foreign gene can be a problem. The *Bacillus brevis* system was found to be a less stable system than *E. coli* for expressing such genes [29].

As alluded to above in Ruggiero's theory that lower expression levels of glycine rich fragments were seen because of a lack of glytRNA in *E. coli* systems [50], a protein's amino acid composition and coding sequence can significantly affect the achievable level of protein expression. If significant amounts of rare proteins are required, and furthermore if they are coded for by rare codons, expression levels in *E. coli* systems will suffer. In natural proteins with highly repeated sequences such as structural proteins, rare codons are likely to appear in the gene. Based on these observations, Martin, *et al.* created an *E. coli* system for the expression of a synthetic human tropoelastin (STEL). By modifying the human tropoelasin sequence using a codon adjustment model suitable to *E. coli*, Martin, *et al.* achieved expression levels of 20–30% [36]. An awareness of the amino acid needs and presence of rare codons is

necessary for the design of CLP expression systems in general, despite the fact that Yin, *et al.* did not experience such difficulty with their *E. coli* expression system for CLP3.1 (see below).

In the work leading up to the strains and plamids used in this thesis, Ferrari, *et al.* patented a system for the creation of large polypeptides with repeating sequences. They addressed the problem of decreased stability of repetitive genes by replacing specific codons with alternate codons [19].[2] Cappello and Ferrari subsequently patented designs for high molecular weight, synthetic collagen-like polymers with low proline content ($< 45\%$) and created *E. coli* plasmid vectors for the expression of these genes [7]. Yin, *et al.*, on whose work much of this thesis is based, created the pJHL vector for the thermally-inducible expression of CLP3.1-his from a plasmid described in Cappello and Ferrari's patent. Yin, *et al.* also constructed the pJY-1 and pJY-2 plasmids for the IPTG-inducible expression of CLP3.1-his (see Section 3.1.1 for plasmid information) [72].

2.2   *E. coli* disruption methods

Since *E. coli* rarely secrete recombinant products, protein remains intracellular. Consequently, at some step in the product purification process the *E. coli* must be destroyed. In mechanical and chemical disruption processes, the cells may be washed with a buffer of choice (e.g. Phosphate Buffer Saline) as the first step in product purification prior to lysis. On the other hand, a recombinant system that causes rapid lysis within the bioreactor does not have this advantage, but is more time and energy efficient. The induction of *E. coli* autolysis—lysis due to self-originating lytic processes—has been studied but generally determined to be inefficient. Mechanisms for induced autolysis include osmotic shock, aging, increasing temperature, and addition of organic compounds such as butanol, ethyl acetate, or toluene (this final method of inducing autolysis could also be classified as a chemical disruption method) [61]. Since autolytic methods are not efficient and therefore are not industrially-viable op-

---

[2]Codon alternates exist due the phenomenon of codon redundancy.

tions, the following sections review more competitive techniques for disrupting *E. coli* for the release of intracellular products.

2.2.1 Mechanical disruption

This category of cell disruption techniques is the most widely used in industry because of its practicality. Unlike chemical disruption methods, no additional compounds must be introduced to the *E. coli* harvest which must then be extracted during the downstream purification process. The most common mechanical means of cell disruption are sonication, French Press, and high pressure homogenization. Because the cell wall makes *E. coli* quite difficult to lyse,[3] the disruption processes must be vigorous [37].

Sonication, which is the exposure of cells to high intensity ultrasound, is ideal for use with small sample volumes (i.e. several mL to 50 mL). Thus, it is one the most popular disruption techniques at the laboratory scale. A typical setup consists of a power source connected to a converter (electrical energy → mechanical energy) that drives a disrupter horn, which focuses the ultrasonic vibrational energy at its tip. If the sample volume is small (several mL) the disruptor horn will likely have a microtip that channels the mechanical energy into an even smaller diameter tip that fits easily into the sample tube. Sonication generates large amounts of heat and sometimes foam, so the sample must remain on ice throughout the process. (Refer to Figure C-3 in Appendix C to see images of a sample before and after sonication.)

The French Press technique is also a laboratory-scale disruption method. The operating principle is similar to high pressure homogenization (see below), in which a large pressure differential (7,000–10,000 psi) is applied over a small volume. The sample is put into the French pressure cell which is then pressurized. The sample slowly drains through the small aperture in the cell. Like high pressure homogenization, the sample must be passed through the French Press machine several times to achieve maximum disruption.

---

[3]Relative to *E. coli*, mammalian cells are much more sensitive to disruption techniques and yeast (e.g. *Pichia Pastoris*) are even more difficult to break apart.

High pressure homogenization is most widely used for processing large volumes [16]. Homogenization systems are designed to draw a continuous sample flow. The crux of a homogenizer consists of a large pressure drop within a very small volume. Because of the resilience conferred by its cell wall, an *E. coli* sample must be passed through a homogenizer multiple times to attain thorough disruption. In an experiment with yeast disruption, Dunnill and Lilly found that the amount of protein released, R, after N passes through a homogenizer follows the relation

$$\ln\frac{R_m}{R_m - R} = KN \tag{2.1}$$

where: $K = \kappa p^\alpha$ = first-order rate constant, κ, α depend on the microbe (e.g. yeast, *E. coli*), p is the operating pressure, and R_m is the maximum achievable protein release [14, 3].

2.2.2   Chemical disruption

As an alternative to mechanical disruption, cells may be treated with specific chemicals to disrupt the integrity of their plasma membranes and cell walls in order to release intracellular proteins. A variety of chemical disruption mechanisms are available. In this discussion, "hybrid" methods that require the expression of a recombinant protein as well as the addition of a foreign chemical to effect the action of that protein are classified as chemical disruption techniques (rather than recombinant techniques) because they lose their efficacy without the addition of the extra chemical and they introduce significant quantities of a the chemical to the culture which must then be removed during the downstream purification process.

Chemical disruption techniques are useful when mechanical disruption machines are not available and the protocol for the downstream processing of the lysate accounts for the chemical composition of the lysate (i.e. the addition of lysis-causing chemicals is not detrimental to the experimental results). The Quiagen Plasmid Purification kit used in this thesis (Section 3.1.2) achieved lysis by suspending an *E. coli* pellet in 200 mM NaOH (with 50 mM Tris·Cl, 10 mM EDTA (ethylenediamenetetraacetic

acid), 100 μg/mL RNase A, and 1% SDS (sodium dodecyl sulphate) (w/v)) [23]. Cell wall hydrolyzing enzymes are also used to lyse bacteria. Hen egg lysozyme is the most popular and least effective lysozyme. Mutanolysin (muramidase of *Streptomyces globisporus*) and lysostaphin (peptidase of *Staphylococcus staphylolyticus*) are other commercially-available enzymes (Sigma Chemical Co.). With gram-negative bacteria, these cell wall hydrolases must be used with permeabilizing agents (e.g. EDTA [a chelating agent that is thought to destabilize the bacterial outer membrane by removing divalent cations from it [16]], toluene, chloroform, thymol, other alcohols) that disrupt the outer cell membrane so that the lysozyme may reach its substrate [61, 35].

In general, the energy state of the cell membrane is thought to affect the membrane's susceptibility to hole formation. Addition of cyanide or azide to *Bacillus subtillus* cultures in stationary phase resulted in autolysis; these chemicals dissipate energy gradients in the cell membrane which increases the likelihood of hole formation and the passage of autolytic enzymes through the membrane to the cell wall [61, 28].

Along similar lines, hybrid recombinant and chemical methods have been developed that use specific chemicals to permeabilize the inner bacterial cell membrane in order to allow recombinant lysozyme to pass through the membrane and hydrolyze the cell wall. These methods use a chemical permeabilizing agent to mimic the effect of phage holin proteins. Yin, *et al.* expressed the λ-phage R-gene lysozyme in a poly-β-hydroxybutyrate-producing strain of JM109. Because there was no membrane-permeabilization agent in the growth medium, the culture did not lyse. Complete lysis was achieved after resuspension in a "lysis buffer" containing 50 mmol/L Tris, 2 mmol/L EDTA, and 0.1% Triton X-100 (v/v) [74].[4] In his evaluation of different uses of the phage LL-H holin and lysozyme for bacterial disruption, Vasala described the use of thymol to mimic the membrane-permeabilizing effect of the LL-H *hol* (holin) gene product in order to allow over-expressed LL-H *mur* (lysozyme) protein to reach and degrade the cell wall. This combination of thymol and over-expressed lysozyme caused efficient lysis of *E. coli*. Because inefficient lysis was observed when the LL-H holin and lysozyme were coexpressed (presumably due to low levels of protein ex-

---

[4]These and related experiments were implemented in this thesis; see Section 6.1.

pression at low growth rates), Vasala argued that the thymol method of membrane permeabilization was applicable at all cell densities as long as the *mur* lysozyme had been expressed in sufficient quantity prior to harvest [61].

Based on the representative examples presented in this section, chemicals that permeabilize the cell membrane are often toxic and must be removed during the downstream purification process. The extremely low tolerance for these toxic chemicals lends potential for the downstream purification process to become complicated and risky should the tolerance level be inadvertently exceeded.

However, a relatively non-toxic method of chemical extraction does exist, which uses only chemicals common to the inclusion body solubilization and protein refolding steps of downstream processing. Falconer, *et al.* describe the use of high concentration urea (6 M) and EDTA ($>$ 0.3 mM) to extract intracellular protein from *E. coli*. EDTA acts to destabilize the outer cell membrane, and urea is a chaotropic agent which provides enough hydrogen bonds to compete with those of water, making it a strong solubilizing agent for disruption of the cell wall. The process also describes the use of 15 mM 2-hydroxyethyldisulfide (2-HEDS) to stabilize released inclusion bodies[5] while the sample is centrifuged, thus allowing for the selective purification of product protein inclusion bodies. Levels of protein release comparable to mechanical disruption techniques were attained. The protein extraction incubation times were varied from 30–90 minutes [16, 17, 18, 10].

2.2.3   Recombinant cell lysis

As noted in Section 2.2.2, *E. coli* disruption systems that require the combined action of a recombinant protein and a foreign chemical in order to effect lysis are classified in this discussion as chemical disruption techniques. This section describes recombinant lytic systems that are the most similar to the goals of this thesis: a strain of *E. coli* capable of efficient induced lysis within the bioreactor without the addition of any

---

[5]The product protein in the studies of Falconer, *et al.* was Long-R^3-IGF-I, an insulin-like growth factor.

chemicals other than an inducing agent.[6]

The use of the bacteriophage lytic mechanism is overwhelmingly the most popular and effective means of conferring recombinant lytic capability. Several different well-defined phages have been studied and multiple patents obtained. Auerbach, *et al.* placed the λ-phage S (holin) and R (lysozyme) genes under the control of a operon sensitive to the thermally inactivated cI857 repressor [2]. Raising the culture temperature above 38°C allowed the phage gene products to be expressed, resulting in lysis. Aurbach, *et al.* describe induction of the lytic genes at low cell density ($OD_{650} \approx 0.5$) and suggest concentrating the culture by filtration or centrifugation before inducing lysis. Auerbach, *et al.* did not describe the use of their invention at industrial scales or cell densities; rather, examples of the use of their invention utilized shake flask cultures. They did not include a detailed analysis of the effect of growth rate on the rapidity or efficiency of lysis.

Vasala developed a lytic system based on the *Lactobacillus delbrueckii* subsp. *lactis* bacteriophage LL-H [61]. Experiments with the coexpression of LL-H genes *hol* (membrane permeabilizing holin) and *mur* (cell wall hydrolase) resulted in lysis only at low cell densities (because growth rates and protein synthesis activity are higher at low cell densities); in addition the *mur* protein digests the cell wall inefficiently. Because of inefficient lysis at "high" cell densities ($OD_{600} > 0.5$), Vasala developed a method in which *mur* was accumulated during growth, after which the cell membrane was permeabilized by treatment with thymol. Consequently, Vasala considered the permeabilization method cell density independent.

The lysis mechanism of the lytic strain designed in this thesis was closely based on the work of Tanji and Morita, *et al.*, which used bacteriophage T4 lysis genes as the means of conferring lytic capability to product-producing strains of *E. coli* (see Section 6.2). Tanji, *et al.* created two plamid vectors, one containing T4 gene-e (lysozyme) and one containing gene-t (holin) [58]. They found that growth of *E. coli* JM109 transformed with the gene-e plasmid was not affected by induction of

---

[6]Arabinose is one such inducing agent. The design for the use of the pBAD system (*ara*BAD promoter) to control the expression of the bacteriophage T4 holin is described in Section 6.2.

the lysozyme product, but they observed an immediate halt in cell growth when the strain containing gene-t was induced. They speculated that accumulation of the lysozyme did not affect cell growth because the inner cell membrane prevented the gene-e product from reaching its substrate the murein layer (part of the cell wall). On the other hand, the t-holin was immediately toxic because it disrupted the integrity of the inner cell membrane, making it impossible to tightly control the intracellular environment. Tanji, *et al.* then transformed *E. coli* BL21(DE3)[pLysS] with each plasmid. When gene-e was induced, they observed a decrease in growth rate and a morphological change in the cell shape (cells became "swollen ellipsoids"). A decrease in viability was observed, and resuspension of the harvest in pure water resulted in near-100% lysis. Induction of gene-t during the late-logarithmic growth phase resulted in immediate and complete lysis; 90% β-galactosidase activity was observed in the supernatant 20 minutes after induction began.[7] Induction of gene-t later in the log-phase of growth resulted in larger fragments of cellular debris, and partial cell lysis was observed even with cultures induced in the stationary phase. Tanji, *et al.* suggested that the disruption of inner membrane integrity by the basally-expressed T7 lysozyme (due to the presence of the pLysS plasmid) allowed the lysozyme to reach its substrate and degrade the cell wall. In contrast with the inability of phage ϕX174 to lyse bacteria during the late logarithmic and stationary growth phases, as well as the general decreased efficiency of recombinant lysis mechanisms at lower growth rates described by Vasala [61], the t-holin with T7 lysozyme system appears to be especially potent, giving it promise as a lytic system for high cell density fermentations when growth rate is decreased. In addition, their success with lysing strain BL21(DE3) was an improvement over the inefficient lytic behavior of strain BL21(DE3)[pUC18-S$^-$RRz] created in this thesis (Section 6.1.1).

Based on the work of Tanji, *et al.*, Morita, *et al.* created a dual plasmid *E. coli* system for the production of β-glucuronidase (GUS). The plasmids constructed by

---

[7]The host strains used in these experiments produced β-galactosidase due to the presence of IPTG (isopropyl-β-D-thiogalactoside), a lactose analog. By measuring the amount of β-galactosidase activity due to e/t-gene action relative to the maximum achievable activity obtained by complete disruption with chloroform, Tanji, *et al.* measured the fraction of the culture that was lysed.

Morita, *et al.* [38] combined either the e-lysin or the t-holin bacteriophage T4 genes with a plasmid containing the construction for the pLysS T7 lysozyme. They named the resultant plasmids pLysSE (T7 lysozyme gene + e-gene) and pLysT (T7 lysozyme gene + t-gene). *E. coli* JM109(DE3) was used as the two-plasmid host which was transformed with either pLysSE or pLysT and the plasmid with the GUS gene. Tanji, *et al.*'s results showed that growth rate was decreased by gene-e expression and rapid lysis was achieved by gene-t expression. Cultures were grown in a minimal medium lacking tryptophan so that GUS, regulated by the OPtrp operator, would be expressed constitutively.[8] Induction of gene-t in mid-log phase resulted in 100% lysis within 2 hours, while late-log phase induction required 6 hours to achieve complete lysis.[9]

Rather than isolating and expressing phage lytic genes using a plasmid vector, the direct infection of *E. coli* with specially engineered phages has been investigated. Park, *et al.* used λ-phage infection of *E. coli* to transform the bacteria with a protein product gene (β-galactosidase), which was then thermally induced to enter the lytic cycle, produce recombinant product, and lyse [30, 33, 44]. The system consisted of two bioreactors. In the first bioreactor, *E. coli* were infected with a strain of bacteriophage λ containing the desired product gene. Conditions in the first bioreactor were optimized for cell growth in the lysogenic state ($< 32°$C). When it was time for the cells to be induced, they were transferred to the second bioreactor and thermally induced to enter the lytic cycle ($> 38°$C). During the lytic cycle, the gene product (along with the bacteriophage lytic enzymes) was produced.[10] Once a critical amount

---

[8]This thesis argues that for an industrial scale process, control over product expression by means of a specific induction event is necessary; by delaying product protein expression during the batch and fed-batch phases, higher growth rates and therefore less culture time may be necessary because cellular resources are used solely for growth rather than being shared for product production throughout low cell density growth. However, constitutive expression of GUS in Morita, *et al.*'s experiments made the assay of GUS production throughout the growth and T4 gene induction processes more straightforward.

[9]It must be emphasized that the experiments described here with the T4 phage genes were done with shake flask cultures, at cell densities 1–2 orders of magnitude lower than densities reached in bioreactors. Implementing these studies in bioreactors (as described in Section 6.3), in which the growth rate can be precicely controlled and monitored, will give a precise ranges of growth rates for efficient lysis.

[10]Expression of the product gene was not under the control of an independent expression system; by virtue of its association with the phage's lytic cycle genes, the product gene was expressed when the phage entered the lytic cycle.

of lytic enzymes were manufactured, the cells lysed. This process was optimized and protein levels at 30 times the lysogenic (non-induced, basal) levels were achieved. An advantage of this system is that the recombinant genes were more stable since they were directly inserted into the bacterial chromosome rather than into a separate plasmid, thus avoiding the plasmid stability problem of long-running continuous cultures [52]. However, protein product induction and lysis were entirely coupled, and the only way to increase protein production for a given strain would be to genetically modify the virulence of the lambda phage. In addition, without discussing the level of basal protein expression relative to other production systems, the absolute production magnitude of "30 times the lysogenic levels" may be significantly less than other production systems.

As an alternative to the phage holin lytic mechanism, the use of colicins as membrane spanning proteins to aid in lysis has been studied. Colicins are antibiotic proteins native to *E. coli* that span the cell membrane and form ionpermeable channels which cause a host of problems for bacterial health. The toxicity and instability of colicin clones is so high, that combined with their inefficiency for catalyzing cell wall hydrolysis, colicins are not an efficient and therefore viable option for *E. coli* lysis [61].

2.3   Lytic mechanism of bacteriophage lysis enzymes

"Autolysis" mechanisms, by which *E. coli* are induced to lyse of their own accord, are generally inefficient. In contrast, by nature, bacteriophages cause efficient lysis. This work takes advantage of the lytic capability of the bacteriophage, a bacteria-infecting virus (an obligate bacterial parasite).

In general a phage consists of a protein coat called the "capsid" that contains phage DNA. When a bacterium is infected, the phage attaches to the cell surface and injects its DNA into the bacterium. Synthesis of host DNA and RNA is halted and the manufacture of phage DNA and proteins begins. At the end of this "lytic" cycle, phage lytic enzymes are expressed and fully assembled phages are released from

Phage	Holin	Lysozyme	Lysozyme type
λ	gene S	gene R	transglycosylase
T4	gene T	gene E	muramidase
T7	not included here	gene 15	amidase
LL-H	*hol*	*mur*	*N*-acetylmuramidase[11]

Table 2.2: Phage lysis genes and lysozyme mechanisms.

the bacterium. Some phages can also exist in a second, "lysogenic" life cycle. In this cycle, phage DNA is incorporated into the bacterial chromosome and the phage lies dormant, replicating with the rest of the chromosome during the DNA synthesis phase of each cell cycle. The lytic cycle is usually repressed by suboptimal growth conditions, and initiated in response to a growth-promoting environment (e.g. abundance of nutrients) [35]. (Hence, the BL21(DE3) host strain is called a λDE3 lysogen because DNA from the λ-phage is incorporated into the BL21 chromosome.)

There are two classes of lytic enzymes that work together to lyse a bacterium. The holin enzyme makes the inner cell membrane permeable to the lysozyme (endolysin/lysin), which degrades the cell wall. If the lysozyme is produced but the holin is not present, the lysozyme can not come in contact with its substrate, and consequently can not degrade the cell wall. Table 2.2 lists the holin-lysozyme pairs and lysozyme categories for phages λ, T4, T7, and LL-H [61].

2.4   Quantitative protein assays

An important quantifier of process efficacy is the percentage of total protein comprised by the product of interest. Thus, it is necessary to quantify both CLP and the total protein throughout the course of the process. In the work leading into this project, as well as in the beginning of this project, it was assumed that the results of the collagenase assay for CLP quantification and the bicinchoninic acid (BCA) assay for total protein quantification could be directly related. Based on the mechanisms described below, Chapter 5 argues that the assumptions governing the use of these

---

[11]Effective on cell walls of *Lactobacillus delbrueckii*, *Lb. helveticus*, *Lb. acidophilus*, and *Pediococcus damnosus*. Inefficient for digestion of *E. coli* cell walls [61].

assays are oversimplistic.

2.4.1   Collagenase assay mechanism

Yin, *et al.* developed the collagenase assay studied in this thesis as a cheap, rapid alternative to immuno-based assays or other assays that require extensive purification of the collagen or gelatin prior to performing the assay [73]. The steps of the assay are discussed in detail in Chapter 5 and the reagent compositions are listed in Appendix A, but a brief description of the protocol is given here for reference. After disrupting the *E. coli* harvest by sonication and removing the cell debris by centrifugation, Yin, *et al.* propose the removal of approximately 80% of the excess proteins with a high temperature (80–100°C) precipitation step. The purified sample (containing mostly collagen-like polymer or gelatin standard) is then digested with a bacterial collagenase (see below) which liberates many polypeptide fragments. The digest is incubated with a ninhydrin (triketohydrindenhydrate) reagent that reacts with free amino groups to form a purple complex that is assayed colorimetrically by measuring the absorbance at 570 nm. Thus, despite the presence of interfering factors (whose effect is measured with a blank sample), the liberation of a large amount of amino groups by hydrolysis of the collagen-like substrate is intended to overwhelm the colorimetric signal resulting from the ninhydrin reaction.

There are a variety of collagenases that vary based on substrate and origin. Type VII collagenase is the type used in this protocol, which is derived from *Clostridium histolyticum*, an anaerobic bacterium that invades tissue by using its collagenases to degrade them structurally [56]. The type VII collagenase cleaves the sequence Pro—X—Gly—Pro between the neutral amino acid X (e.g. leucine or alanine, R groups are hydrocarbons) and Glycine, and is inhibited by metal chelating agents (e.g. EDTA, cysteine; bacterial and most mammalian collagenases are metalloproteinases, which require metal ion cofactors such as Ca^{2+} for Type VII collagenase) [54, 31, 32]. Ideal conditions for the collagenase reaction include the presence of Ca^{2+} ions and physiological pH (7.5).

Ninhydrin reacts with most compounds that contain amino groups such as α-

amino acids, imino acids, amino alcohols, primary amides, urea, ammonia, and ammonium ion (because they are secondary amines, proline and hydroxyproline give a yellow color). Relative to free leucine, ammonia results in 60% color yield. Stannous chloride acts as the reducing agent for the reduction of ninhydrin to form the colored complex (a conjugated π-bond system). The ninhydrin reagent is 50% methyl cellosolve (ethylene glycol monomethyl ether), which solubilizes ninhydrin. The ninhydrin binding reaction is carried out at pH 5 and 80–100°C [49, 13].

Since the collagenase assay results are used in conjunction with the results of the BCA assay, the relative reactivities of gelatin and CLP3.1 in the BCA assay are important (see Section 2.4.2 for the assay mechanism and composition of the BSA protein standard). Type IV gelatin from bovine skin is a 46.3 kDa protein with the following selected amino acid compositions (mole %): 3.75% cystine (17 residues), 1.99% tyrosine (9 residues), and 1.10% tryptophan (5 residues) [60]. CLP3.1 does not contain any of these three residues (see Section 3.1.1); this difference between natural gelatin and artificial CLP may be a significan source of error if it is not taken into account by the assay (discussion in Section 5.4).

2.4.2   Bicinchoninic acid (BCA) assay mechanism

Developed in 1985 by Smith, *et al.* [55], the BCA assay is highly sensitive and easily-implemented. In this thesis, a BCA assay kit manufactured by Pierce Biotechnology was used (Part No. 23225).

The BCA assay reaction involves a multi-step mechanism that begins with a biuret reaction. In the first step, 4–6 peptide bonds (from polypeptides of any length greater than 1–2 peptides) form a coordination complex with a Cu^{2+} ion. The complex is blue and absorbs light at 540 nm. In the alkaline reaction mixture, the protein reduces Cu^{2+} to Cu^+, which is then chelated by BCA. Two molecules of BCA react with one Cu^+ ion to form the coordination complex, which absorbs light at 562 nm. Relative to the biuret reagent, the BCA reagent is 100 times more sensitive in the lower limit of detection. Residues that are particularly reactive in the reduction of Cu^{2+} are cystine (R-group: —CH_2SH), tyrosine(R-group: —CH_2–Ph–p-OH), and tryptophan

(R-group: —CH_2–C^2=CH–NH–Ph; ortho single bond between C^2 and Ph); all are basic residues, but experimental evidence suggests that the peptide bond also reduces copper at increased temperatures. Copper chelating and reducing agents will interfere with the assay [45, 59].

Bovine Serum Albumin (BSA) is the protein standard provided with the BCA assay kit. Its molecular weight is approximately 66.4 kDa with the following composition of selected amino acids (mole %): 5.8% cystine (35 residues), 3.5% tyrosine (21 residues), and 0.5% tryptophan (3 residues) [69].

Chapter 3

Experimental setup

The apparati, techniques, and materials used to carry out each part of this project are described in this chapter. The experimental components of this project can be broken up into three general categories: basic recombinant *E. coli* techniques, fermentations, and analytical protein assays (downstream processing, process analysis). It is not the intent of this chapter to give continuity to the experiments and a chronological history of their use in this thesis; it is a reference for protocols and technical specifications. Experimental data is also included in this chapter when it can make the protocol more clear or when it was used to develop or improve the protocol itself.

3.1 Recombinant *E. coli* techniques

The techniques described in this section implement very basic biological concepts that are taught in any introductory biology lab course. The sections below, then, focus on the vital information necessary to reimplement these experiments assuming familiarity with molecular biology and fermentation implementation.[1]

---

[1]For example, MIT courses 7.02 (Introduction to Experimental Biology and Communication) and 10.28 (Biological Engineering Laboratory) would provide useful background knowledge.

3.1.1 Vector design: strain and DNA information

Prior to this project, Yin, *et al.* found the optimal combination of host strain and plasmid expression system for the production of collagen. Their study quantified the expression level of collagen-like polymer (CLP3.1) when the gene was inserted into different plasmid vectors and used with different versions of the BL21 family of *E. coli* [72]. In another project whose results were also studied here, Yin and colleagues in China worked on a lytic process using the JM109 strain of *E. coli* with a λ-phage fragment that contained a lytic gene. Chapter 6 contains more detailed information about the host strains, plasmid vectors, and their operating mechanisms. A concise description of the characteristics of each strain and plasmid as well are listed below.

Host strains

All BL21-derived strains were studied by Yin, *et al.* at MIT [72], and the JM109-derived strain was studied by Yin, *et al.* at Tsinhua University, China [74].

BL21 Base strain used in the construction of other host strains. Analyzed by Yin, *et al.* but low productivity results with the plasmids used. Relevant properties: F^- *ompT hsdS*$_B(r_B^-)$ *gal dcm* (Novagen, [72]). *Not used in this thesis.*

BL21(DE3) A λDE3 lysogen of BL21. The included λ DNA contains a gene for bacteriophage T7 RNA polymerase under the control of the *lacUV5* promoter. Designed to work with pET expression vectors in which the product gene is placed under the control of a T7 promoter. Addition of IPTG induces the expression of T7 RNA polymerase, which transcribes the product gene. There is a small amount of basal ("leakage") expression of T7 RNA polymerase, which leads to basal expression of the product protein [21].

BL21(DE3)[pLysS] The pLysS plasmid confers chloramphenicol resistance and has a T7 lysozyme gene that is always expressed at low basal levels (*tet* promoter). By binding to T7 polymerase, the T7 lysozyme has an inhibitory effect on the activity of T7 RNA polymerase, thus decreasing basal expression of the product

gene. Amidase activity of the T7 lysozyme is retained when the enzyme is bound and inbibiting T7 RNA polymerase [9, 75]. The use of pLysS for the basal expression of T7 lysozyme is a technique for repressing basal expression of the product gene under the control of the T7*lac* system.

JM109 A common strain with which high transformation efficiencies can be obtained.

Structure of the product gene, CLP3.1

In each plasmid vector described below, the product gene sequence CLP3.1 is fused with a 252 bp sequence that codes for a six-histidine tag (from the pET expression vectors, Novagen) at the C-terminus of the protein, making it possible to perform an immunodetection assay to identify the protein. CLP3.1 is a 95 kDa protein made up of 52 repeating peptide sequences with the sequence GAPGAPGSQGAPGLQ [72, 19, 7]. The amino acid composition of CLP3.1 is therefore 33.3% Glycine, 20% Alanine, 20% Proline, 13.3% Glutamine, 6.7% Leucine, and 6.7% Serine. Since CLP3.1 is produced by *E. coli*, the product is a single-stranded protein, contrasted with the triple helix structure of naturally occurring collagen. *CLP3.1 is a non-toxic product; its accumulation in host cells will not hurt the cells.* Gelatin is natural collagen that is partially denatured and degraded, and was considered by Yin, *et al.* to be close enough in structure and composition to be an accurate standard for the quantification of CLP3.1 in the collagenase assay [73]; this claim is considered in more detail in Section 5.4. In the context of this thesis, "CLP3.1" and "collagen" are used interchangeably.

Plasmid vectors

The following information is given for each plasmid: antibiotic resistance, origin of replication, total length. The names pJY-1, pJY-2, and pJHL are nicknames for more descriptive names which are included below. Since this thesis is concerned with the construction of a dual plasmid strain of *E. coli*, it is important to note that when multiple plasmids are present they must replicate by different mechanisms. Origins within

the same compatibility group are not compatible. Compatibility groups relevant to this thesis are: {colE1, pMB1} and {p15A} (groups listings are not exhaustive).

pJY-1 Antibiotic resistance: amp^R, Origin: pMBR, Expression system/operon: *lac*, Full name: pET-14b-CLP3.1-his, Parent vector: pET-14b (Novagen), Length: 7053 bp. **Not used in this thesis.**

pJY-2 Antibiotic resistance: amp^R, Origin: pMBR, Expression system/operon: T7*lac*, Full name: pET-15b-CLP3.1-his, Parent vector: pET-15b (Novagen), Length: 8090 bp.

pJHL Antibiotic resistance: kan^R, Origin: pMB1 (from pBR322), Expression system/operon: cIts, Full name: pPT0296-CLP3.1-his, Parent vector: pPT0296 (Protein Polymer Technologies), Length: 6447 bp [7, 72].

pUC18-S$^-$RRz Antibiotic resistance: amp^R, Origin: pMB1, Expression system/operon: *lac*, Full name: pUC18-S$^-$RRz, Parent vector: pUC18 (GE Healthcare, formerly Amersham Biosciences), Lenth: $\approx$ 4120 bp [62].

3.1.2   Plasmid purification and verification

Plasmids were purified using QIAfilter Plasmid Maxi Kits (QIAgen, Inc.; Cat. No. 12262). To obtain maximum plasmid purity, strict adherence to the steps in the kit's manual was necessary. Following the instructions in the QIAgen kit manual, inoculum tubes containing 3 mL Luria Bertani (LB) medium and the appropriate antibiotic marker were inoculated with a single colony from a streaked plate, grown for 12 hours, and 200 μL used to inoculate a 100 mL working volume LB shake flask which was grown for another 12 hours. The shake flask broth was harvested, pelleted, and the cells were chemically lysed using kit reagents. The plasmid DNA in the lysate was purified using special columns from the kit. The plasmids pJY-1, pJY-2, and pJHL were purified from BL21(DE3) hosts, and the plasmid pUC18-S$^-$RRz was obtained from a JM109 host; all strains were obtained from Jin Yin. The final DNA stocks were dissolved in TE buffer (10 mM Tris, 1 mM EDTA).

Lane 1:   0.5 μg 1 kb DNA ladder
Lane 2:   (Lane 1) $\times$ 2 dilution
Lane 3:   1.25X dilution of stock (repurified) pJHL
Lane 4:   (Lane 3) $\times$ 2 dilution
Lane 5:   (Lane 4) $\times$ 2 dilution
Lane 6:   (Lane 5) $\times$ 2 dilution
Lane 7:   (Lane 6) $\times$ 2 dilution
Lane 8:   (Lane 7) $\times$ 2 dilution

Notes: lanes are numbered from left to right and wells are at the top of the image.

Figure 3-1: Agarose gel used to quantify the repurified pJHL plasmid stock.

The plasmid DNA concentration was quantified by performing agarose gel electrophoresis on the purification eluate (0.8% agarose in 1X TAE buffer; running buffer = 1X TAE; 100 Volts; 35 minute running time; 15-20 minute ethidium bromide stain; 5-10 minute dH$_2$O destain). The gel was then imaged with a Kodak Gel Logic 200 (image capturer and analyzer) under UV light with a 590 nm filter (λ at which ethidium bromide fluoresces), and the digital image was analyzed by the accompanying Kodak image analysis software. The software's quantification algorithm was based on based on comparing the DNA band intensities with those of a DNA standard. The concentration of purified pUC18-S$^-$RRz was found to be 50 μg/mL. The pJY-2 stock was found to be 4 μg/mL[2]. The first pJHL stock was 125 μg/mL, but after a necessary repurification of the stock,[3] the final concentration was found to be 5.4 μg/mL. Figure 3-1 shows the repurified pJHL quantification gel. The length of the DNA fragments is determined by comparison with the DNA ladder in lanes 1 and 2 (New England Biolabs 1kb ladder; Cat. No. N3232S). The brightest band in the ladder is composed of 3kb fragments.

In order to verify that the purified DNA was indeed the plasmid DNA that was expected, each plasmid was each digested with one single-cutting restriction enzyme

---

[2]This low concentration (relative to the other purifications) is probably due to overdrying of the DNA pellet, which made it very difficult to redissolve the pellet in the TE storage buffer.

[3]Even though the stock was sufficiently concentrated, appropriate dilution of the stock resulted in low transformation efficiencies, implying interference by RNA. Large amounts of RNA were also visible in the quantification gel as bright, unexpected bands and smears.
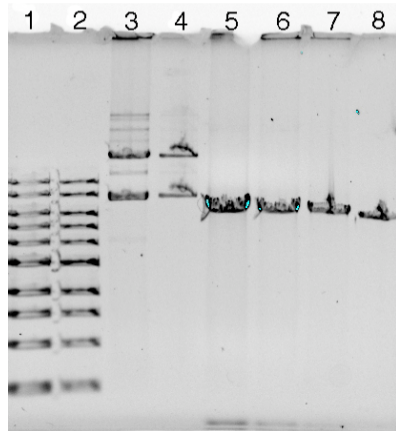
Figure 3-2: Digestion of plasmid pJHL with NcoI results in two fragments of length 6195 bp and 252 bp.

and agarose gel electrophoresis was performed on the digestion mixture using the parameters listed above. Gels were again imaged using the Kodak analyzer. The gel images for the digestions of pUC18-S⁻RRz and pJHL are shown in Figure 3-2. Since the digestion times and restriction enzyme amounts depended on the concentration of the plasmid DNA, the quantification step was carried out first. Restriction endonuclease NcoI (New England Biolabs, Cat. No. R0193S) was used to digest pJHL. As seen in Figure 3-2, the lanes containing the digestion mixture contain two bands of length 6.2 kb and < 0.5 kb (sum ≈ length of pJHL; the < 0.5 kb band is at the edge of the bottom the gel and is very faint). Restriction endonucleases EcoRI (New England Biolabs, Cat. No. R0101S) and HindIII (New England Biolabs, Cat. No. R0104S) were used to digest pUC18-S⁻RRz. As expected, two fragments of length 2.7 kb and 1.4 kb resulted (sum = 4.11 kb ≈ length of pUC18-S⁻RRz). Also notice that lanes 3 and 4 of Figure 3-2 contain multiple bands (> 8 kb) which are not present in the lanes containing digested DNA. These additional bands are due to conglomerations of multiple circular plasmids which are not formed when the plasmids are digested to form linear structures. The disappearance of these bands also suggests the presence of one pure type of plasmid (pJHL).

3.1.3 *E. coli* strains studied

With the exception of BL21(DE3)[pJY-1], all strains that were studied during the course of this thesis were made "freshly" by transforming the appropriate host strain with purified plasmid DNA. More detailed information on the use of and data related to these strains is written in Chapters 4 and 6.

BL21(DE3) plasmids used: pJHL, pJY-1,[4] pJY-2, pUC18-S⁻RRz

BL21(DE3)[pLysS] plasmids used: pJHL, pJY-1

JM109 plasmids used: pUC18-S⁻RRz

3.1.4 Transformation

Competent host strains BL21(DE3) and BL21(DE3)[pLysS] were obtained from Novagen, and competent host JM109 was obtained from Promega. Transformation protocols from the manuals that accompanied the cells were followed as strictly as possible in order to achieve maximum transformation efficiency [22]. Despite the fact that this is a "protocols" chapter, some data from the transformations is discussed here in order to give an experimental basis for successful transformation experiments in the future.

Cells were chemically competent. The most frequently used means of preparing chemically competent cells is by subjecting them to a cold, hypotonic $CaCl_2$ shock then mixing them with plasmid DNA while still cold. The cell and plasmid DNA mixture is then transferred from ice to a hot water bath (42°C) in order to incur a rapid heat shock (30 seconds). The competency "window" for cells is quite short and very sensitive to external conditions, and no DNA at all will be taken up if there are even slight deviations from the optimal conditions [35, 22]. The protocols warned that any warming of the cells before the heat shock would reduce transformation efficiency. It is likely that the early transformation experiments done in this thesis were subject to

---

[4]BL21(DE3)[pJY-1] stock was obtained directly from Jin Yin and plasmid pJY-1 was not repurified and inserted into competent BL21(DE3).

this degradation of competency because the original stocks of BL21(DE3) and JM109 came in stock tubes from which aliquots for individual transformation reactions were to be aliquoted. The pipette tips used were not chilled, which is likely to have warmed up the cells before the DNA was added. In the subsequent transformation experiments that were performed by students in course 10.28, the BL21(DE3) and BL21(DE3)[pLysS] cells came in single-reaction aliquots, which removed the aliquoting step that was previously detrimental to transformation efficiency. The data from the transformation experiments supports the hypothesis that the transformation efficiency was dramatically reduced by the extra handling step: the repurified pJHL plasmid DNA was used in the transformations of both the aliquoted and single use transformation mixtures, and the number of colony forming units increased from 2-3 to several hundred.

Size, purity, and quantity of plasmid DNA also affect transformation efficiency. These effects were also observed over the course of the experiments. The original stock solution of plasmid pJHL was highly concentrated (125 μg/mL), but had significant RNA contamination (note: Figure 3-1 shows the repurified pJHL; the original gel of the contaminated pJHL stock had much darker smears of RNA outside of the band expected of pJHL DNA). Only the competent BL21(DE3) that was transformed with the repurified pJHL DNA resulted in colony forming units when plated on kanamycin-containing plates, implying that the additional RNA contamination interfered with the DNA receptors on the surface of the *E. coli* [35]. In addition, an order of magnitude more colony forming units were obtained when the same cells transformed with the pUC18-S⁻RRz plasmid, which was much more pure. (However, a confounding factor may be that the 4.12 kb pUC18-S⁻RRz plasmid is much shorter than the 6.45 kb pJHL plasmid, and the difference in size effects likelihood of plasmid uptake.)

Finally, the quantification method based on calculating band intensities of plasmid DNA in agarose gels is not highly accurate; it is difficult to obtain an image in which all serial sample dilutions correspond to serial multiplications of the calculated band intensities. In order to get a sense of the degree of quantitative error of this image analysis technique and its effect on the success of transformation protocols which

are based on the concentration values it produces, experiments were done at widely varying plasmid concentrations. For example, the Novagen protocol for transforming BL21(DE3) called for 1-10 ng of plasmid DNA. Using the Kodak analyzer data, 0.5 ng, 5 ng, and 50 ng of pUC18-S$^-$RRz were added to separate transformation mixtures. The 5 ng reaction produced 2 colonies, the other mixtures produced zero. This sparse data implied that the quantification provided by the Kodak analyzer was accurate to within one order of magnitude, which was enough to achieve a successful transformation. Based on this data, transformations in 10.28 (pJHL and pJY-2) were performed under the assumption that the calculated stock concentrations were accurate. For each type of DNA, teams were assigned one of three DNA quantities (4-5 ng $\pm$ 2 ng) with which to perform their transformation. The overwhelming transformation success (hundreds of colony forming units per plate) implies that the Kodak analyzer's accuracy to within roughly one order of magnitude is sufficient for these experimental purposes. Because of the large number of transformants (as well as resultant satellite colonies) on each plate, it was impossible to make quantitative conclusions about the effect of DNA quantity on transformation efficiency.

The experimental data presented here proves the importance of following the transformation protocols and gives a real experimental basis for future experimentalists to design experiments that continue this work.

3.1.5 Cell bank

A cell bank of genetically pure stock of each strain was prepared. A small amount (several μL) of each *E. coli* stock obtained from Jin Yin was streaked on a separate LB-agar-antibiotic plate (LB medium + 1.5% agarose, see Section 3.2.3 for LB composition). For each strain, a 500 mL baffled shake flask with 100 mL of LB medium was inoculated with a single colony to insure genetic purity. The shake flask cultures were grown for 10-12 hours, 220 rpm, 30/37°C until somewhere within the log phase of growth was reached (exact cell density not important, as long as the OD$_{600}$ value is within the range 1-5). The broth was then mixed with pure, sterile glycerol to

make a 50% (v/v) solution[5] which was then distributed to microfuge tubes in 1 mL aliquots and frozen at -80°C. This preparation followed the protocol in the 10.28 lab manual [68].

3.2   Fermentation

The exact parameters for implementing a fermentation vary depending on the strain, operating conditions, and goals of the experiment. The purpose of this section is to give an overview of a typical fermentation in order to define a context for the instrumentation and assays discussed in this chapter. Detailed information on how to set up a bioreactor resides in the 10.28 lab manual [68]. Specific data from a large number of fermentations is presented in Chapter 4.

3.2.1   Bioreactor setup

Fermentations were carried out in the BioFlo 3000 and BioFlo 110 bioreactors (New Brunswick Scientific Co., Edison, NJ). The vessel sizes were 1L and 5L for the BioFlo 3000 reactors, and 0.9L and 10L for the BioFlo 110 units. The major differences between the two models lies in the controllers. The BioFlo 3000 is a console-style controller with a water-jacketed vessel, while the BioFlo 110 controllers are modular boxes that are cabled together on a controller bus, and heat transfer is accomplished with a cooling coil submerged in the culture broth.

Either pure air or oxygen-enriched air was filter sterilized (0.2 μm) and sparged through the broth by means of a perforated sparging ring that extended to the bottom of the vessel. The broth was agitated by two Rushton impellers driven by an electric motor connected to the headplate. Inoculum, acid/base, glucose feed, bioreactor samples, etc. were added through sampling/addition lines that passed through the headplate. The headplate also contained a septum port which allowd for a sterile connection to the interior of the vessel to be initiated with a syringe or needle with

---

[5]The formation of the crystal structure of frozen 50% glycerol is less disruptive to the cells than that of a purely aqueous solution.
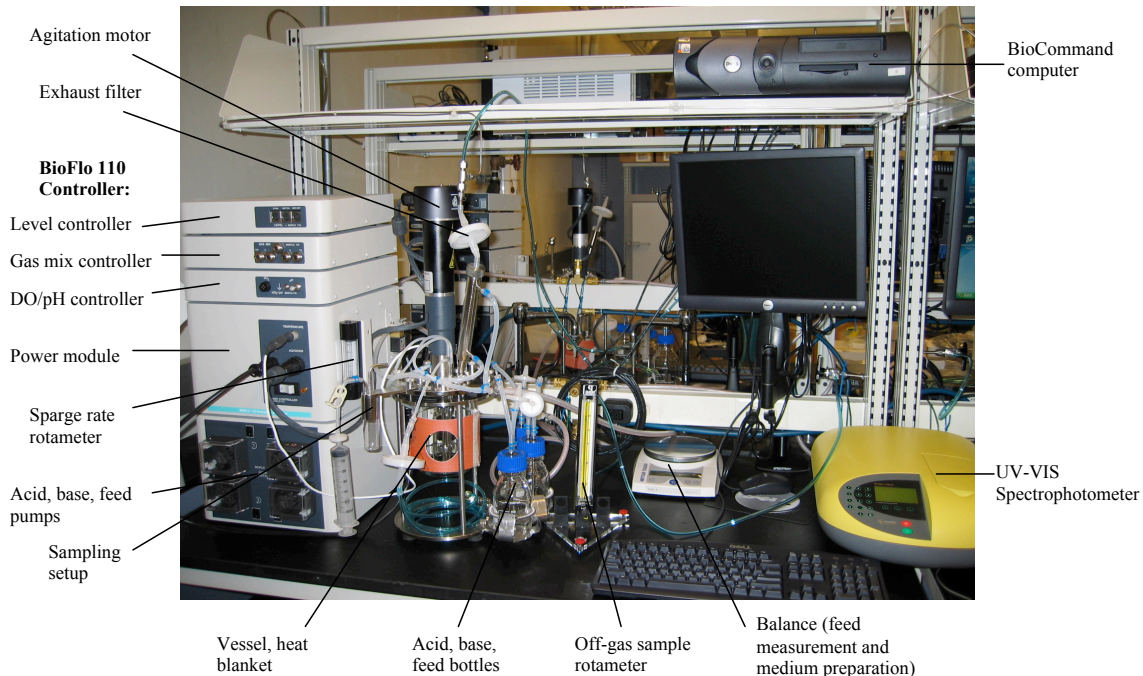
Figure 3-3: Overall view of a typical fermentation setup used in course 10.28 (Fall 2004).

a tubing barb. Tubing generally had an Inner Diameter (ID) of 1/8" or 1/4". A minimum wall thickness of 1/16" was necessary to insure that the peristaltic pumps could completely pinch seal the tubing as they rotated. Silicon tubing was used when it was useful to see the progress of the fluid meniscus, but a major problem with silicon is that it could be permanently pinched closed by a pump or clamp that was left in the same position for too long. For this reason, Pharmed® tubing was generally preferred to silicon for use with pinch clamps and pumps. Barbed tubing connectors were purchased from Cole Parmer; polypropylene and Teflon® (PTFE) were the preferred materials for their composition since they are very inert and can withstand autoclave temperatures (121°C). Tubing pushed onto the barbed connectors was secured with tool-tightened cable ties. A picture of the overall setup for a course 10.28 bench (which contains standard equipment used in all fermentations) is shown in Figure 3-3, and a close-up picture of a 1 L BioFlo 110 headplate is shown in Figure 3-4. Both figures are reproduced from the 10.28 lab manual [68].

A "process loop" is a fermentation parameter under the control of a P-I-D feedback

DO, pH probes
(behind drive shaft)

Agitation drive

Headplate plug
(spare port)

Tri-port: cooling coil,
sparger

Septum

Exhaust condenser

Thermowell

Tri-port: harvest,
sample, extra lines

Tri-port: acid,
base, feed
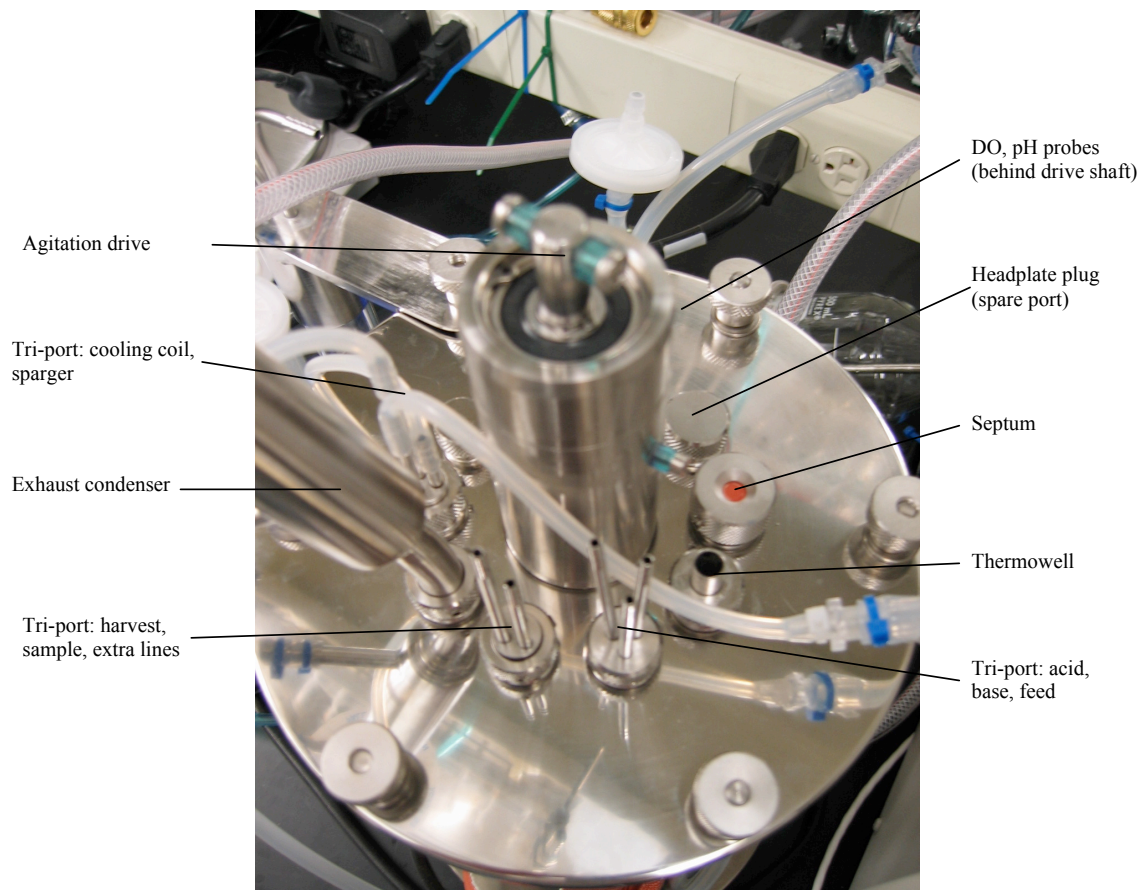
Figure 3-4: Close-up view of a BioFlo 110 headplate.

controller. "Loop" values were recorded by NBS's BioCommand Plus software which was installed on a PC connected to the NBS controller. Following is a list of process loops recorded during each fermentation, with some brief notes on the setup and parameters associated with each loop (the maximum sampling frequency was once every 12 seconds):

Agitation Ranged from 300-800 depending on the oxygen demand of the culture. Could be controlled based on the dissolved oxygen process value, but controllers were often not consistent so manual control of agitation was used.

Temperature Cold water (4–10°) was supplied by VWR recirculating chillers (VWR Part No. 13271-204). BioFlo 3000 consoles had a heating element that could warm the chiller water before circulating it through the SS water jacket. BioFlo 110 controllers warmed the vessel with an electric heating blanket and allowed chilled water to pass through a SS cooling coil. In general, the SS water jackets of the BioFlo 3000 vessels delivered a faster temperature response because there was much more heat exchange surface area and coolant flow. The water pressure was regulated to 15-20 psi for the BioFlo 3000s and 2-4 psi for the BioFlo 110s (silicon tubing was used in the flow path of the BioFlo 110s; high pressure would burst the tubing).

Dissolved oxygen (DO) Measured by a Mettler-Toledo InPro 6900 polarographic probe. Oxygen in the medium diffused through an oxygen permeable membrane was reduced at a platinum cathode, and Silver was oxidized at the anode to form Silver Chloride (AgCl). The probe was autoclaved with the bioreactor and connected to the controller via its cable for at least 6 hours before calibration in order to allow sufficient time for polarization (practically, though, the polarization time did not make a noticeable difference). The probe was calibrated at the operating temperature: the zero-point was calibrated by saturating the medium with nitrogen, and the 100% point was defined as the saturation point when air was sparged.

pH Measured by a Mettler-Toledo probe (Order No. 10 405 4480). Calibrated before bioreactor sterilization and autoclaved with the bioreactor. pH was controlled by a P-I-D controller that could add acid or base. Since the fermentations usually produced acidic biproducts, the acid pump was turned off to eliminate the possibility of controller oscillation and over addition of acid and base.

Acid/Base/Feed pumps These pumps were two state (on or off) peristaltic pumps. The controller achieved a variable flow rate by varying the percentage of the time that the pump was on. The pumps could be controlled manually, by P-I-D controllers, or by software programs created within BioCommand Plus. The maximum pump rate achievable using 1/8" ID tubing was 3.13 mL/min.

Inlet gas composition The percentage enrichment of air with pure oxygen could be controlled. Air came from the building's air compressors, Oxygen came from high pressure tanks, and Nitrogen came from the headspace of a liquid N_2 dewar. All gases were regulated to 3-10 psi and delivered to controller manifolds via 1/4" braided PVC (or braided rubber) tubing. Connections were made with Coil Hose 1500 Series brass quick release connectors (New Hampshire Tool Outlet, Salem, NH).

Foam probe This probe was unreliable and there was no effective way of pumping an appropriate amount of viscous antifoam (polypropylene glycol) into the vessel. The BioFlo setup was equiped with this functionality, but it was not used.

Fermentation exhaust gas composition was measured by a Perkin Elmer mass spectrometer and/or an ARI Laser Gas Analyzer. Exhaust gas passed through a 10°C condenser before traveling through Stainless Steel tubing to a rotary valve that controlled which gas stream was sampled (see Section 3.2.2 for details on these instruments). Cole Parmer variable area (ball and tube) rotameters were used to monitor the flow through the sampling lines. Because SS is very inert and permeates very little gas, as much of the length of the sampling system as possible was constructed using stainless steel and Swagelock® ferrule-based connectors. Silicon tubing is very

permeable to oxygen, so these lengths were kept to a minimum. Teflon® (DuPont) tubing is permeable to water, but this effect is less pronounced than the Si/O_2 effect [46].

3.2.2 Supporting instrumentation

Mass spectrometer This is the gold-standard instrument for measuring the composition of bioreactor off-gas. The exhaust gas from each bioreactor was piped through a separate length of 1/4" Stainless Steel (SS) tubing to a rotary valve (Hamilton-Sundstrand) which could be configured to automatically swith between lines. The rotary valve had one outlet which delivered the gas flow to the mass spec. The residence time for gas in the tubing system varied from approximately 40 seconds to 2 min, 40 sec, depending on the distance from the bioreactor to the rotary valve (between 15 and 40 feet). It was important to keep the fermenter's exhaust condenser cold at all times in order to prevent condensation from building up in the gas lines and putting the mass spec in danger of getting wet (water destroys the system). Requires a flow rate of $\approx$ 100–120 mL/min.

Laser Gas Analyzer The first prototype manufactured by Atmosphere Recovery, Inc. (Plymouth, MN) for use with bioprocesses (ARI's instruments are usually used in industrial manufacturing plants). Used in one fermentation so far (BL21(DE3)[pJY-1], October 4-6, 2004). Based on the principal of Raman spectroscopy, specially configured to measure eight gases that are biologically relevant: N_2, O_2, CO_2, H_2O, NH_3, H_2, CH_4, and C_2H_5OH. The instrument has only been used to measure N_2, O_2, and CO_2 in experiments so far. Gas transport system is setup the same as the Perkin Elmer, with 1/8" SS tubing and a rotary valve.

Noninvasive biomass probe BugEye prototype manufactured by BugLabs, LLC. (Danville, CA). Used on one *E. coli* fermentation so far (October 4-6, 2004). The device is strapped to the outside of the glass vessel so that it has an

unobstructed view of the broth inside. Measures the degree of scattering of infrared laser light to calculate the biomass concentration.

UV-VIS spectrophotometer Models 1100pro and 2100pro were used to measure OD_{600} (cell density), OD_{570} (collagenase assay), and OD562 (BCA protein assay). Manufactured by Amersham Biosciences (Piscataway, NJ).

3.2.3 Fermentation medium and feed components

A complete list of media and reagents, as well as notes on their preparation, is included in Appendix A. All shake flask cultures were grown in Luria-Bertani (LB) medium (10 g/L NaCl, 10 g/L Tryptone, 5 g/L Yeast extract). MR medium was used for all fermentations (note that the optimal minimal medium varies with strain; since the host strain used in all fermentations was BL21(DE3), no other minimal medium was necessary). A note on media design: the media composition for optimal growth varies with host strain. MR medium, designed for the BL21 family, does not work with JM109; in a fermentation performed with JM109[pUC18-S⁻RRz] and MR medium, the culture never grew past an OD_{600} of 2.

3.2.4 Typical fermentation chronology

After preparing the bioreactor for use according to the description in Section 3.2.1, MR medium is added to the bioreactor. The bioreactor is then sterilized in the autoclave (121°C, 45 min). After autoclaving, nitrogen is slowly sparged until just before inoculation to deter growth of possible aerobic contaminants. Just before inoculation, glucose and antibiotic are added through the septum and the DO probe is calibrated. Table 3.1 describes the components that are added before and at inoculation (reproduced from the 10.28 lab manual [68]). The divergence of the pH probe is evaluated by measuring a sample with an external pH meter. Throughout the process, samples are taken every 1-3 hours, and the assays described in Section 3.3 are performed. The BCA and collagenase assays are only performed on induction samples, though.

The initial inoculum volume for a fermentation is 10% of the total working volume.

Medium component	Volume (/L)	Final concentration
"Partial" 1X MR medium	**883.3 mL**	**1X**
Antifoam (PPG)	**0.4 mL**	**0.04% (v/v)**
25 mg/ml ampicillin	2 mL	50 μg/mL
700 g/L glucose with 14 g/L MgSO$_4 \bullet$ 7H$_2$O	14.3 mL	10 g/L
Seed culture	100 mL	10% (v/v)
Total	1000 mL	-

Table 3.1: Components of initial bioreactor medium per liter of final bioreactor working volume. Components in bold are added before autoclaving, the remainder are added before and at inoculation (10.28, Fall 2004).

The night before inoculation, an appropriate number of 500 mL baffled shake flasks with 100 mL of LB medium are inoculated with 100 μL of glycerol stock and grown for 12-16 hours until the end of the log growth phase. Once the bioreactor has been completely prepared, the inoculum is added through a sterile bottle connected to tubing with a needle that pierces the septum. The BioCommand Plus software is set to begin recording process values and the batch phase begins. The batch phase lasts approximately 6 hours, and the end is signified by a sudden spike in the DO%: when the glucose is completely consumed aerobic respiration can no longer occur resulting in a sudden decrease in the consumption of oxygen. Fermentations in which bioreactor cooling is done are most successful when the cultures are cooled midway through the batch phase, when cell density is sufficiently low. The length of the "warm time" batch phase for a culture that is cooled is usually only slightly shorter than the lengh of the batch phase of an uncooled culture.

During the fed-batch phase, glucose/MgSO$_4$ solution is fed to the culture until a sufficiently high cell density is reached ($>$ 40 OD$_{600}$; $>$ 15 g/L DCW). A primary goal during this phase is to maintain the culture in a feed-limited state, rather than an oxygen-limited state. Physical limits on the fermentation system make it impossible to sustain the ever-increasing oxygen transfer rates (OTR) necessary to maintain high growth rates at the high cell densities of the fed-batch phase. Overfeeding of glucose and the consequent tendency towards an oxygen-limited state can lead to the accumulation of acetate, which can be detrimental to product production [27].

One strategy to maintain a feed-limited state is called DO-stat feeding. In this

strategy, a small amount of feed is added when the DO level rises, which immediately causes the DO to drop. If the DO remains floored at zero, there is most likely an excess of glucose. DO-stat feeding can be frustrating because a steady state is hard to achieve and the DO-stat feeding algorithms that the NBS equipment is capable of can be easily confounded by the delay of the DO probe and response time of the controller, leading to erratic feeding. Section 4.1.2 describes the use of vigilant DO-stat feeding to determine a range of acceptable feeding rates for a given set of process values, which are then applied to subsequent processes in a prescribed, rather than reactive feeding strategy. During the fed-batch phase, when relatively high cell densities are reached, foaming is prone to occur. Foam is dealt with by using a syringe to add about 5 drops of antifoam at a time through the septum port.

Once a sufficiently high cell density has been reached and while the culture is still actively growing (usually $OD_{600} \approx 50$, DCW $\approx$ 15-20 g/L), the culture is induced to express the product gene. In the *lac* system, IPTG is added to a working concentration of 1 mM, and in the thermally-inducible system, the temperature is increased from 30°C to 41°C. In most cases, it is ideal to have a high growth rate at the start of induction. In a high-growth state, a high level of transcriptional and translational machinery is present in a cell to support its growth activitity. Suddenly inducing the culture can cause a major shift in the application of that machinery from the synthesis of proteins required for cell replication to the synthesis of the product protein. Consequently, a marked decrease in growth rate may be observed 30–45 minutes after the start of induction [20]. After a 5 hour induction period, the process is stopped and a final harvest sample is taken. Unused cells are killed with bleach. The BCA and collagenase assays are used to analyze the induction samples for intracellular protein content. *E. coli* is lysed by sonication (Branson Ultrasonics Sonifier 250D: 1/8" microtip, 70% intensity, 70% duty cycle, 1 second period, sample cooled in ice bath) after washing twice with Phosphate Buffer Saline (PBS).

3.3 Analytical assays and fermentation measurements

Over the course of the fermentation, the assays listed below were performed. All protocols for the following assays are described in detail in the 10.28 lab manual [68]. In addition, the theoretical background and techniques for measuring the volumetric oxygen mass transfer coefficient ($k_L a$) are given in Section 3.3.2.

3.3.1 Analytical assays

OD_{600} In order to quantify the cell density within a few minutes of sample-taking, a sample was diluted in a 1.5 mL cuvette to within the linear range of the UV-VIS instrument (between 0.1 and 1) and the absorbance at 600 nm was measured. Since *E. coli* are particulate matter they scatter light. As a result, the measured absorbance is affected both by scattering and absorbance of the sample. This means that UV-VIS instruments with different physical setups will measure different absorbances for the same sample: OD_{600} measurements should not be used as the gold standard to compare different fermentations with which different instruments may have been used. To be rigorous, optical density measurements must be normalized for the interference due to medium absorbance. However, relative to the OD_{600} values for most cell-containing measurements, the medium's absorbance is minimal: for LB medium $OD_{600} \approx$ 0.014, and for MR medium, the value is even less.

Dry Cell Weight (DCW) 1.5 mL samples were washed twice with deionized water and dried in an 80°C oven to a constant mass ($\approx$ 2 days). In contrast with the OD_{600} measurement, DCW *is* an assay that can be used to compare cell densities across fermentations and laboratories. Even with a balance with a sensitivity of $\pm$ 0.1 mg, the signal from the dry mass of the sample only began to dominate the noise of the instrument when the DCW reached $\approx$ 20 g/L, and in the analysis of the BL21(DE3)[pJY-1] data from the 10/4/04 fermentation, only samples with

$OD_{600} > 40$ were considered in the quantitation of the relationship between OD_{600} and DCW (Section 4.1.2). In the future, larger sample volumes should be used to measure the DCW of lower cell densities. Major sources of error are introduced in the washing and drying steps. If the samples are left in the oven for too long, organic material will sublimate before the sample masses are measured [20]. Sublimation will result in the underestimation of the DCW associated with a specific OD_{600}. In the washing step, if cell pellets are not tight, cell mass may be lost when the supernatant is poured off.

Glucose A Yellow Springs Instruments (Yellow Springs, Ohio) 2700 Select enzymatic probe was used to measure glucose concentration in the bioreactor samples. To avoid rapid clogging of the membrane, samples were centrifuged for 1 minute at 3000 rpm before the YSI took an aliquot.

pH A Mettler-Toledo benchtop SevenEasy pH meter was calibrated daily and the pH of the fermentation broth measured every few hours to monitor the drift of the bioreactor pH probe. If the pH measured by the bioreactor probe and the Mettler probe disagreed, the pH setpoint was changed to compensate for the bioreactor probe drift.

Collagenase assay Developed by Yin, *et al.* [73] as a cheap, simple alternative to an immunodetection assay. Protocol involves precipitation of insoluble proteins and subsequent digestion of the remaining (soluble) collagen with collagenase. The excess freed N- and C-terminus react with ninhydrin to form a purple complex that absorbs light at 570 nm, allowing for a colorimetric assay using a UV-VIS spectrophotometer. Gelatin is used to make the standard curve.

BCA assay Used BCA Protein Assay kit (Pierce Biotechnology, Cat. No. 23225) to measure total intracellular protein in lysed samples. Assay uses Bovine Serum Albumin (BSA) as protein standard. Mechanism involves the reduction of Cu^{2+} to Cu^{1+} which forms a complex with BCA and results in purple color. Specific amino acid residues and the peptide bond act as reducing agents [69].

HPLC An Agilent 1100 Series High Performance Liquid Chromatography instrument was set up and run by Judy Yeh to assay carbohydrates and organic acids from the course 10.28 fermentation samples. The HPLC can detect concentrations as low as 1 mg/L, which made it useful for assaying fermentation products such as acetate.

3.3.2 Oxygen mass transfer theory and experimental approach

Many calculations can be carried out on the process loop and growth kinetic data to further analyze the process. These calculations will be introduced and carried out in Section 4.3, where the actual fermentation data is presented. There is one system parameter, though, whose measurement requires special instrumentation and technique, the details of which are most appropriately described in this section. That parameter is the volumetric oxygen mass transfer coefficient (k_La, units = hr^{-1}). The k_La is a quantification of the fermentation system's ability to encourage the transfer of oxygen from the gas to the liquid phase. The higher the value of the k_La, the greater the Oxygen Uptake Rate (OUR) of the culture can be. The k_La is the product of the liquid side oxygen mass transfer coefficient, k_L, and the gas-liquid interfacial area, a. k_L is dependent on the physical properties of the liquid. In this thesis, a is primarily dependent on agitation (vigorous stirring means smaller bubbles and a larger surface to area ratio) and sparging rate, while k_L is effected by ionic strength and broth composition (see Section 4.3.4 for a discussion of k_La data from this thesis in the context of literature analyses).

The k_La appears in the oxygen mass transfer equation:

$$\frac{dC_L}{dt} = k_l a(C^* - C_L) - OUR \qquad (3.1)$$

C_L is the concentration of dissolved oxygen in the broth $[\frac{mol}{L}]$, C^* is the saturation concentration of oxygen in the broth at the given operating temperature, and OUR is the oxygen uptake rate $[\frac{mol}{L \cdot hr}]$.

Two methods were used to measure the k_La: the dynamic method and the off-gas

method. Both of these methods are based on achieving conditions in the bioreactor that allow for the simplification of the oxygen mass transfer equation so that it can be easily rearranged and the $k_L a$ solved for.

Because of the **dynamic method**'s dependence on the dynamic response of the polarographic dissolved oxygen (DO) probe, results of this method are most accurate at low cell densities (< 1 g/L, $OD_{600} \leq 1 - -2$), when the DO probe can more accurately track the relatively slower dynamics of the DO response. The first step in this method is to turn off the bioreactor's air supply when the DO is at a high value (preferably at least 50–60% of air saturation). Assuming the experiment is accomplished within a small fraction of the doubling time of the culture, the dissolved oxygen will be consumed at a constant rate equal to the OUR. When the DO approaches zero, the air flow is stepped back to its original rate and the recovery of the DO is plotted until the DO stabilizes at some steady-state value. The $k_L a$ can be calculated in two ways from this data: either by using the OUR (calculated above) along with the steady-state driving force, as seen in Equation 3.2, or by plotting $\frac{dC_L}{dt}$ versus $(C^* - C_L)$ and calculating the slope ($k_L a$ will be the negative slope of the plot).

$$\text{In steady state}: \quad \frac{dC_L}{dt} = 0 \quad \Rightarrow \quad k_l a = \frac{OUR}{(C^* - C_L)} \tag{3.2}$$

Because the dynamic method is only useful for calculating the $k_L a$ for a small fraction of the total fermentation time, the **off-gas method** of $k_L a$ analysis is a particularly useful technique, applicable at any time there is a measurable amount of oxygen consumption. The off-gas method is most accurate when a steady state is achieved: a constant, measurable difference between the inlet and exhaust oxygen percent compositions ($\%O_2^{in}$, $\%O_2^{out}$) is used to calculate the OUR (given the flow rate of sparged gas), and the value of C_L is used to calculate the driving force. The resultant values can be plugged into Equation 3.2 to calculate the $k_L a$. If there is water vapor present in the exhaust gas, the calculated OUR will be artificially high because the percent composition of $\%O_2^{out}$ is reduced because of the presence of water vapor in the exhaust. To compensate for this effect, a normalization factor using inert

nitrogen is introduced to the equation for the percentage OUR ($OUR\%$), as shown in Equation 3.3.

$$OUR\% = \%O_2^{in} - \left(\frac{\%N_2^{in}}{\%N_2^{out}} \right) \%O_2^{out} \tag{3.3}$$

It must be noted that emperical correlations between $k_L a$ and system parameters such as agitation, aeration, and physical properties of the impellers, etc. are not examined in this thesis; in the experimental analysis, more effort was devoted to developing other aspects of the project.

Chapter 4

Quantitative fermentation results

The overarching goal of this chapter is to present the data and analysis necessary to implement repeatable *E. coli* fermentations. One of the unique contributions of this thesis that defines the context for these analyses is the development of a cooling phase to arrest the growth of fermentations for up to several days. To accomplish this goal, the following study analyzes data spanning the topics of fermentation growth kinetics, cooled fermentations (10–12°C), and the oxygen mass transfer coefficient ($k_L a$), and carbon mass balance.

4.1 *E. coli* growth kinetics

Because of their relevance to an organism's metabolic state, which directly impacts the organism's capacity for recombinant product production, a detailed study of growth kinetics is essential to the definition of a viable bioprocess. In general, high growth rates at the time of induction are desirable because they necessitate the buildup of large amounts of protein-making machinery required for growth, which can be used for product production at the time of induction. High cell densities are desirable because product quantity is usually related to the sheer amount of biomass accumulated (*E. coli* densities of up to 100 g/L [$OD_{600} \approx 250$] have been reported) [8]. The following sections report the kinetic data for *E. coli* shake flask cultures and bioreactor fermentations. In addition, high time-resolution data from an on-line, noninvasive

biomass sensor is presented.

4.1.1 Shake flask culture

In the context of fermentations, shake flasks are primarily used to multiply an *E. coli* stock to generate an inoculum culture, hence the criteria for successful shake flask growth processes are less stringent than those for fermentations. In general, *E. coli* densities in shake flasks follow a logistic growth curve with four phases: lag (adjustment to new conditions just after inoculation, low growth rate), log (once adjusted to culture conditions, a constant growth rate is attained, leading to an exponential increase in cell density while nutrients are plentiful), stationary (remaining nutrients are not sufficient to maintain exponential growth), and death (cells die due to the depletion of nutrients and accumulation of harmful waste products). By the end of the log phase the available nutrients per cell has dramatically decreased, acetate waste has accumulated,[1] and the cells begin to lose their propensity for exponential growth. Thus, in preparation for a fermentation in which rapid accumulation of biomass is desired, the ideal inoculum must contain a high density of cells that are still in the log phase of growth.

Shake flask cultures were grown in 500 mL baffled flasks with 100 mL of LB medium and an appropriate antibiotic, and were inoculated with 100 μL of glycerol stock (see Section 3.1.5). Using the lab's Amersham 1100pro UV-VIS spectrophotometers, the maximum OD_{600} values attained in shake flasks permitted to reach the stationary phase were in the range 5–6. The target optical density range for inoculum shake flasks was 4–5. Strains containing an IPTG inducible CLP3.1 gene (e.g. BL21(DE3)[pJY-1/pJY-2]) that were grown at 37°C, 220 rpm, reached an OD_{600} in the range 4–5 within 10–12 hours. Thermally-inducible (BL21(DE3)[pJHL]), grown

---

[1]In a set of experiments for a separate, industry-related *E. coli* fermentation project (12/17/03), shake flasks were inoculated with 50 mL bioreactor samples from different times in the fermentation. Shake flasks inoculated with broth from the bioreactor with an initial $OD_{600} \approx 15$ reached a final pH in the range 5.1–5.5, while shake flasks with an initial $OD_{600} \approx 40$ reached a final pH in the range 4.7–5.0. (Because oxygen mass transfer is much more limited in the shake flask than in the bioreactor, the biomass of these high cell density cultures did not significantly increase in the shake flasks.)

at 30°C, 220 rpm, reached an OD_{600} in the targeted range within 11–12 hours. Since shake flask growth can vary significantly, it is best to measure the OD_{600} every hour after 10 hours has elapsed.

Before testing the effect of cooling on a bioreactor culture, cooling experiments were done with shake flask cultures (8/4/03–8/5/03). Shake flasks containing *E. coli* strains BL21(DE3)[pJHL] and BL21(DE3)[pJY-2] were inoculated with one colony each into 100 mL LB with an appropriate antibiotic (see Appendix A.1 for medium and antibiotic compositions). The cultures were grown to a low cell density in the beginning of the log phase, put into a refrigerator overnight, and returned to the incubator the following day. The growth curves for the two strains are plotted in Figure 4-1. When shake flasks are inoculated with a colony from an agar plate, the lag phase may vary greatly (up to 8 hours) because of large differences in the number of cells that are initially dislodged from the inoculation loop in the medium. The growth curves of the two cultures in Figure 4-1 are synchronized; the time between inoculation and the first data point for BL21(DE3)[pJHL] was 9.5 hours, and 2 hours for BL21(DE3)[pJY-2]. The time associated with each OD_{600} value is the effective elapsed time that the cultures spent in the incubator; the 6.3 hour period that the samples spent in the refrigerator was effectively removed when the data was plotted. It is evident that both cultures recovered from the cooling period and resumed exponential growth. Also note that the OD_{600} doubled while the cells were in the refrigerator. It is likely that the time it took the cells to cool to 4°C while in the refrigerator accounts for a significant amount of the additional growth, so that the majority of growth that lead to the doubling of OD_{600} occured in the very beginning of the time that the cells spent in the refrigerator. Based on the encouraging results of this experiment, further cooling experiments were performed with bioreactors (see Section 4.2).

A discussion of the effects of phage lytic enzymes on bacteria growth is also relevant to this thesis. Inconsistent results and low growth rate and growth extent were sometimes seen in shake flask cultures with strain BL21(DE3)[pUC18-S⁻RRz]. See Section 6.1 for more detailed results and Chapter 6 for a general discussion of the
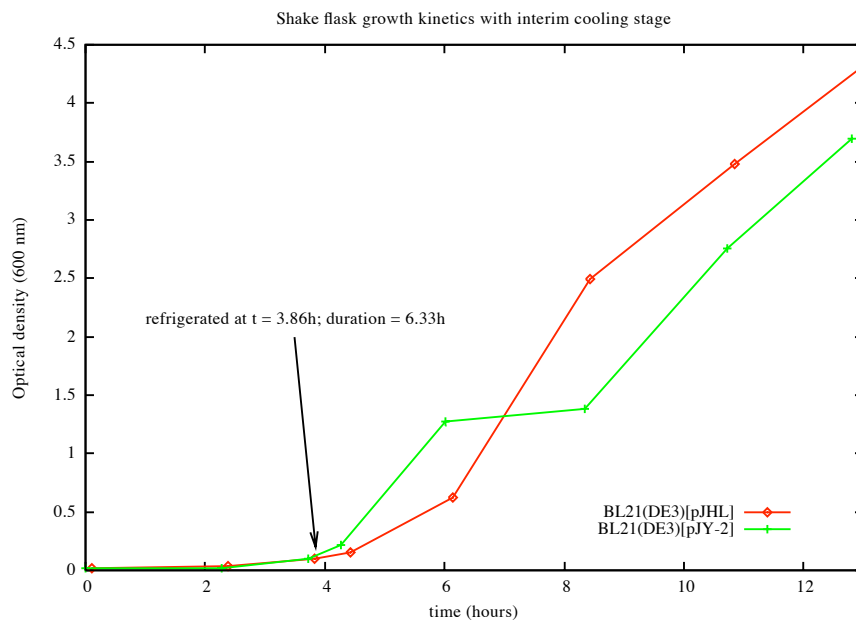
Figure 4-1: Growth kinetics of shake flask cultures refrigerated at low cell densities.

effect of basal expression of bacteriophage lysozyme on cell growth.

4.1.2 Fermentation

The purpose of this section is to describe the general fermentation growth characteristics of the BL21(DE3)/CLP3.1-producing family of strains. Detailed calculations are performed on the data from the BL21(DE3)[pJHL] fermentation on 3/25/04, since this is the primary fermentation that the design of the course 10.28 (Fall 2004) fermentations was based on. Data from other fermentations is presented where applicable. Data from the course 10.28 fermentations (Fall 2003 and Fall 2004) is primarily used to analyze the efficacy of different cooling strategies (Section 4.2).

Data for several fermentations on the length of the batch, fed-batch, and induction phases, as well as the OD_{600} at the end of each phase are listed in Table 4.1. In the 3/25/04 fermentation, the oxygen enrichment level was increased from 0% to 4.5% over the course of the fed-batch phase and kept at 11% during the induction phase. At the beginning of the fed-batch phase, the culture was fed in bursts until the feed bottle and pump system were set up correctly; the culture alternately fed and starved for about 1.5 hours while the feeding system was repaired, and oxygen enrichment

| | | Final OD_{600} and phase length | | |
Date	Strain	Batch	Fed-batch	Induction
3/25/04	BL21(DE3)[pJHL]	14.88 (7.5 hr)	55.8 (7.1 hr)	70.9 (4 hr)
6/24/04	BL21(DE3)[pJHL]	36.0^3 (5.8)	54 (4.6)	102 (5.0)
7/25/04 (1)	BL21(DE3)[pJY-2]	10.4 (4.7)	37.6 (4.3)	52.3 (4.75)
7/25/04 (2)	BL21(DE3)[pJY-2]	10.9 (4.7)	42.9 (4.3)	65 (4.75)
10/4/04	BL21(DE3)[pJY-1]	9.1 (16)	46.4 (6)	96.7 (5.5)

Table 4.1: Summary of fermentation kinetics for various strains and fermentation runs: phase length and final OD_{600} measurement for batch, fed-batch, and induction phases.

was not begun until the feeding system was repaired.[2]

In contrast, to the 3/25/04 run, the fermentation conducted on 6/24/04 was immediately fed and oxygen-enriched when the batch phase glucose was depleted. During the fed-batch phase the oxygen enrichment level was steadily increased to 16% and held at that level throughout induction. Most likely, because of the less stressful conditions at the beginning of batch phase the 6/24/04 fermentation grew more vigorously and consistently reached higher optical densities in less time than the 3/25/04 fermentation. Since only one fermentation was completed with strain BL21(DE3)[pJY-1], the reproducibility of long lag period during the batch phase is unknown. Because of the accumulation of waste and consumption of unreplaced nutrients, the longevity of a batch style culture is expected to be finite (in contrast to a continuous culture, in which spent medium is continually replaced by fresh medium). In most fermentations conducted during this project, the growth rate had dramatically slowed or stopped by the end of the induction phase and the cultures were especially prone to foaming.

Since the fermentations performed on 3/25/04 and 6/24/04 were experimental

---

[2]When the feeding was started for the first time and tubing from the glucose feed bottle was put into the peristaltic pump, the tubing was not unclamped downstream of the pump, so the tubing eventually burst. In the 1.5 hour period that it took to prepare and sterilize another feed bottle assembly, the broken feed tubing was patched and glucose periodically drained into the bioreactor. Because there is slight back pressure in the bioreactor due to the restriction on the exhaust imposed by the exhaust filter, the air supply had to be shut off in order for the bioreactor pressure to dissipate and for the draining setup to work. Consequently, in addition to starving periods, there were periods of oxygen deprivation.

[3]The final OD_{600} value reported for the batch phase of this fermentation ($OD_{600}= 36.0$) is likely to be an error; two hours into the fed-batch phase the OD_{600} was 30.9 and the subsequent measurements were most consistent with the 30.9 value.

fermentations whose results were used to develop a defined fermentation strategy for course 10.28, a vigilant, care-intensive DO-stat feeding strategy was implemented (see Section 3.2.4 for a general description of DO-stat feeding). This strategy involved making minute adjustments to the feed addition rate as the culture grew while gradually increasing the level of oxygen enrichment. The primary goal was to keep the culture feed-limited at all times so that glucose would not accumulate. Table 4.2 summarizes the relationship between biomass increase and quantity of glucose added for several fermentations. These ratios ($\frac{\text{(g glucose)/(L broth)}}{\text{OD}_{600} \text{ unit}}$) were calculated by plotting the cumulative amount of glucose added (i.e. the absolute mass of glucose, not the mass of the glucose feed) versus OD_{600}, and taking the slope of the linear regression on this data. Table 4.2 includes the R^2 values for each linear regression. Note that the BL21(DE3)[pJHL] fermentations (working volume for both fermentations was 5 L) resulted in similar quantities of glucose consumed per unit of biomass increase, while the BL21(DE3)[pJY-1] fermentation (10 L working volume) required slightly more glucose. The increase in working volume over time due to feed addition was not taken into account. A result of this simplification in the calculation is that the amount of biomass increase per unit volume is overestimated, especially as the elapsed fermentation time inceases. If a mean volume of 11 L during the fed-batch and induction phases was used for the BL21(DE3)[pJY-1] fermentation, the glucose consumption ratio would have been 1.46 $\frac{\text{g/(L broth)}}{\text{OD}_{600} \text{ unit}}$ (significantly closer to the values for the BL21(DE3)[pJHL] fermentations). The data used to calculate the values in Table 4.2 remained linear (rather than beginning to plateau), even at high OD_{600} values. This result implies that the energy requirements for growth far outweigh the requirement for cellular maintenance (otherwise, increasing cell density would have required an increasing fraction of the glucose feed to be used for maintenance, rather than for biomass increase). Therefore, assuming that the variability of the data in Table 4.2 is purely experimental, the mean mass of glucose consumed per unit increase is approximately $1.37 \pm 0.40 \left[\frac{\text{(g glucose)/(L broth)}}{\text{OD}_{600} \text{ unit}} \right]$.

Since a culture's growth rate is never constant, designing the feeding and oxygen enrichment strategies based on the rate of biomass increase is more straightforward;

Date	Strain	Glucose consumed per OD_{600} unit	R^2
3/25/04	BL21(DE3)[pJHL]	1.19 $[\frac{\text{g/(L broth)}}{OD_{600} \text{ unit}}]$	0.98
6/24/04	BL21(DE3)[pJHL]	1.30	0.95
10/4/04	BL21(DE3)[pJY-1]	1.61	0.98

Table 4.2: Specific glucose consumed $[\frac{\text{g}}{\text{L broth}}]$ per unit OD_{600} increase in biomass; obtained from slope of the linear regression on glucose added vs. OD_{600} data during fed-batch and induction phases of various fermentations.

if a critical minimum growth rate for maximal induction is determined in the future, then that value of μ can be used to calculate an acceptable rate of biomass increase that satisfies range in μ necessary for effective induction, which can then be used to develop a feeding strategy based on the results presented in this section.[4] Thus, using the results of the 3/25/04 and 6/24/04 fermentations, non-dynamic feeding strategies were constructed for the course 10.28 (Fall 2004) fermentations. Implementing these strategies involved calibrating the feed tubing with the controller pump to determine the maxim flow rate, then making a program with the BioCommand Plus software (New Brunswick Scientific Co.) that would increase the feeding rate linearly with time. (The strategy for oxygen enrichment was to use the DO% process value to manually respond to the change in oxygen demand caused by a change in feed rate by changing the enrichment level about one percent at a time until a new equilibrium was reached where the DO% was non-zero; often during the later phases of induction, the feed demand would decrease, allowing the O_2 enrichment to be concordantly decreased.) Results of the Fall 2004 course 10.28 fermentations are described in Section 4.2.

As a side note, an explanation of why we used the methodology described above to develop fixed feeding programs rather than "simply" using DO-stat feeding control programs for course 10.28 is relevant. Work concurrent to but unrelated to this thesis was done with several *E. coli* fermentations in an attempt to implement a DO-stat feeding program for an unmanned implementation of the fed-batch phase. A combination of the DO probe time delay and delay in the BioFlo controller's response

---

[4]The growth rate at the time of induction of the bacteriophage T4 t-holin protein is critical to the efficiency and rapidity of lysis. Sections 2.2.3 and 6.3 discuss the effect of growth rate on induced lysis in more detail.
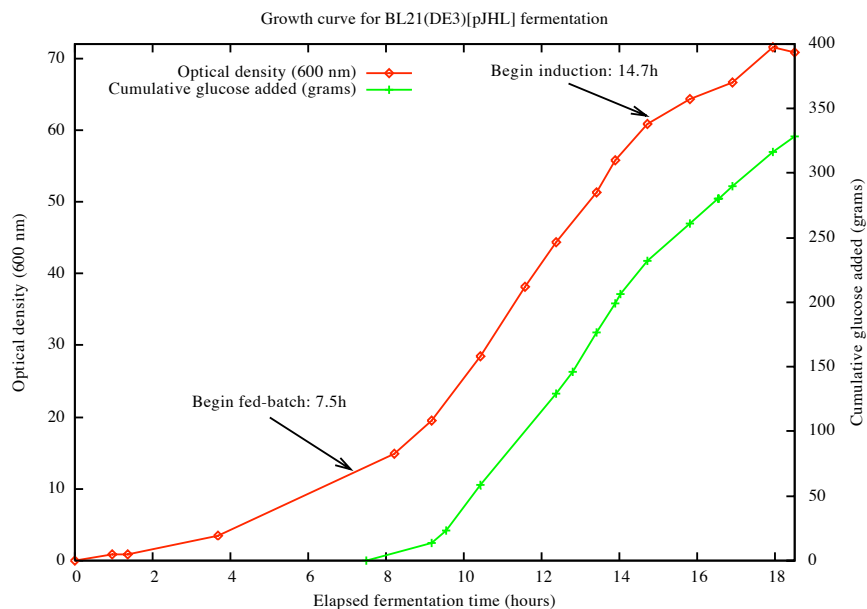
Figure 4-2: Growth curve for BL21(DE3)[pJHL] fermentation with cumulative glucose added during the fed-batch and induction phases.

time caused the resultant "feed back controller" to feed in erratic patterns, resulting in prolonged periods of starvation and glucose glut. We decided that developing fixed feeding strategies based on data from several labor-intensive implementations of glucose-limited feeding would allow us to develop fermentations with more predictable growth kinetics, consequently causing them to be more robust.

In order to provide a detailed characterization of the relationship between growth rate, feeding rate, and biomass accumulation rate, the 3/25/04 BL21(DE3)[pJHL] fermentation is analyzed below. Figure 4-2 shows the growth curve for the 3/25/04 fermentation along with the cumulative amount of glucose added during the fed-batch and induction phases.[5] The trend in OD_{600} is qualitatively correlated with the amount of glucose that has been fed; periods of relatively rapid increase in OD_{600} are accompanied by periods of relatively rapid increase in the amount of glucose that has been added.

The trends in Figure 4-3 are quantitative representations of this noticeable correla-

---

[5]The feed contained 700 g/L glucose with a density of almost exactly 1 g/mL. Thus, 0.7 g of glucose was added per 1 mL of glucose feed. All plots of glucose-related data are in terms of glucose mass, not the mass of glucose feed.

tion between feeding and growth. Figure 4-3(a) shows the time courses of the growth rate, μ, and the feed addition rate.[6] Based on this data (unfortunately the data was limited in this part of the fermentation), the highest growth rate was achieved in the middle of the batch phase and steadily decreased as the fermentation progressed. The decrease in growth rate was primarily due to limitations on the magnitude of the OTR that can be achieved in the bioreactor; as the bacterial population increases, the amount of oxygen available per cell decreases. This trend can be countered using oxygen enrichment, but judging from the consistent decrease in growth rate with time for all fermentations, an exponential growth rate is not sustainable. Figure 4-3(a) demonstrates that feeding rate and growth rate are not analogous quantities (as are the rate of biomass increase and the feeding rate). The maximum feed demand from 13–14 hours is the combined result of a relatively high biomass with a high metabolic demand.

Figure 4-3(b) shows the time courses of the change in biomass and the feed addition rate. As expected, the shapes of the trends closely track one another. Since the culture was glucose-limited (i.e. there was no glucose accumulation), the feed was converted to biomass. The close correlation of these trends also implies that a relatively constant fraction of the feed was converted to biomass (see Table 4.2).

The analysis of cell growth kinetics treats the OD_{600} measurement as a gold standard for biomass concentration in shake flasks and bioreactors.[7] A brief discussion of the validity of the assumption that OD_{600} and biomass are linearly related is presented here. DCW measurements were performed on the BL21(DE3)[pJY-1] fermentation (10/4/04). Because a sample volume of 1 mL was used for the assay, the differential mass measurements ($mass_{tube+cells} - mass_{tube}$) for samples with $OD_{600} < 40$ were not used because there was much more variability among these samples than among the higher cell density samples. A linear regression on the five "usable"

---

[6]Ideal exponential growth satisfies the differential equation $\frac{d(OD_{600})}{dt} = \mu \cdot OD_{600}$. The growth rate, μ, between each pair of OD_{600} data points was estimated by dividing the rate of biomass increase ($\frac{dOD_{600}}{dt}$) by the linearly-interpolated biomass (OD_{600}) at the intermediate time between OD_{600} data points.

[7]Since OD_{600} is significantly less work-intensive than dry cell weight (DCW) measurement, establishment of the validity of OD_{600} is convenient in order to obviate the need for DCW measurements in all cultures.
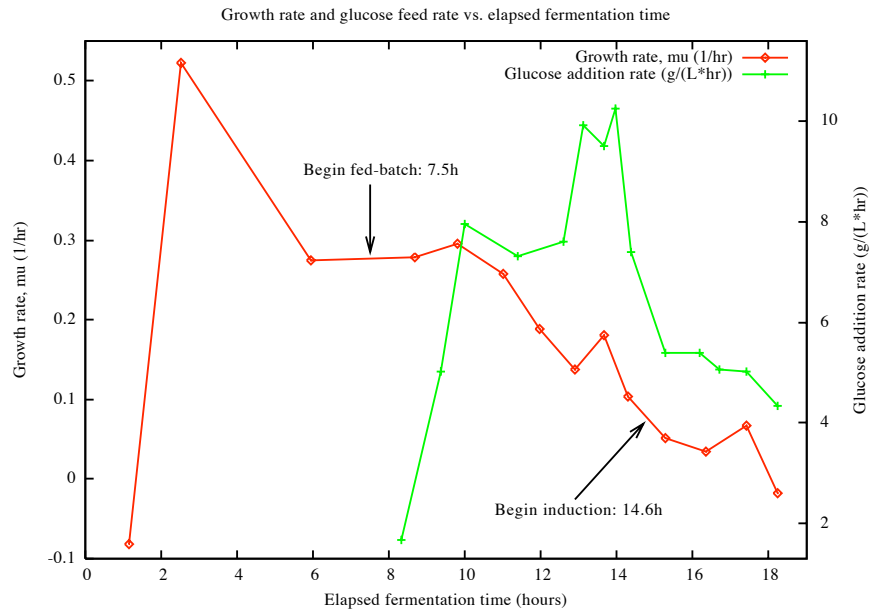
(a) Growth rate (μ, hr^{-1}) and rate of glucose addition ($\frac{\text{g}}{\text{L}\cdot\text{hr}}$) over time.
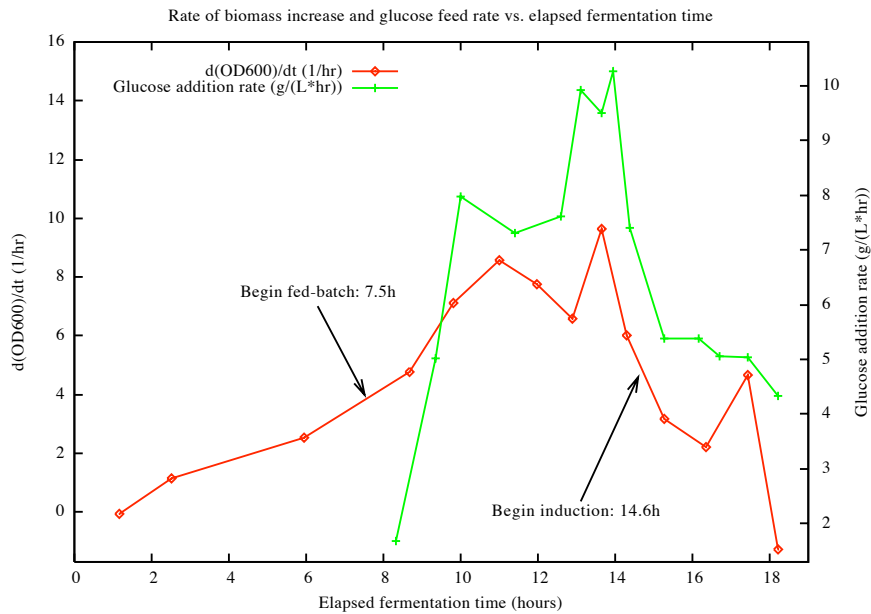
(b) Rate of biomass change ($\frac{\text{dOD}_{600}}{\text{dt}}$, hr^{-1}) and rate of glucose addition ($\frac{\text{g}}{\text{L}\cdot\text{hr}}$) over time.

Figure 4-3: Growth rate, rate of biomass change, and rate of glucose addition over time for BL21(DE3)[pJHL] fermentation.
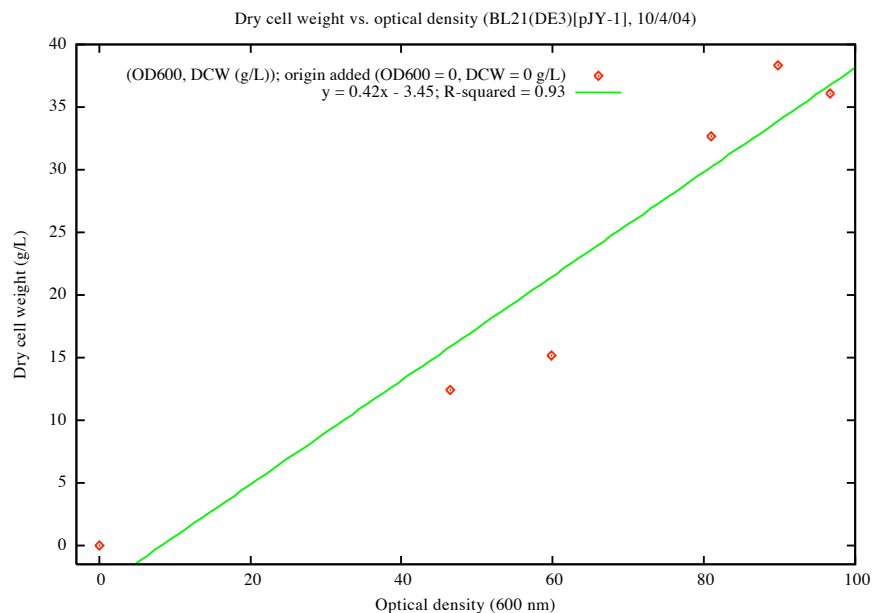
Figure 4-4: Optical density (600 nm) vs. dry cell weight (g/L)

(DCW (g/L), OD_{600}) ordered pairs from the BL21(DE3)[pJY-1] fermentation gave a regression equation with $R^2 = 0.93$. See Figure 4-4 for a plot of the ordered pairs, the regression line, and the equation of the regression line. Considering that there were only six sample points (five samples + the origin), the linearity of the DCW vs. OD_{600} relationship ($R^2 = 0.93$) is strong, implying that OD_{600} is a good biomass quantification method. Based on the slope of the linear regression model in Figure 4-4, one may convert from OD_{600} to DCW by multiplying by $0.42 \; \frac{\text{(g DCW)/(L broth)}}{OD_{600} \; \text{unit}}$.

4.1.3 Non-invasive biomass probe results

Measurement of the optical density at a visible wavelength (e.g. 600 nm) is a standard method of obtaining a relative measure of the cell density of microbial cultures. The method is quick, easy, and sensitive (relative to dry cell weight, which requires large volumes of sample at low cell densities; see Section 3.3.1 for notes on the OD_{600} assay). On the other hand, this cell density assay requires several mililiters of sample for every data point that is taken. With small bioreactors, such as those used in research settings (e.g. 1 L working volume), taking too much volume out of the bioreactor in order to keep track of the growth kinetics is undesirable. For example, lowering the

surface of the liquid eventually exposes an impeller, resulting in excessive foaming. Consequently, the total number of data points that can be taken to measure cell density may be limited.

As an alternative to limited "invasive" methods such as the OD_{600} assay, biomass probes that provide real-time, on-line measurements have been developed. A fixed path length absorbance probe passes through the headplate and is submerged in the fermentation broth. The end of the probe has an emitter and a dectector with a notch between the windows of the two diodes that the fermentation broth can flow through. As the cell density increases, the emitted light is absorbed and scattered by the medium contents and cells, decreasing the amount of light that reaches the detector. optek-Danulat, Inc. (www.optek.com) is a manufacturer of this type of probe.

The BugEye noninvasive biomass probe (BugLab, Inc.) is another instrument that can measure cell density in real-time. The BugEye is based on the principle that particles in the fermentation medium will scatter light. The major particulate component of the medium is cell mass; the more cells there are, the more light is scattered. The BugEye shines infrared light into the fermentation vessel and quantifies the amount of light that is scattered. Based on the scattering that it detects, the BugEye outputs a number in BugEye units. Like OD_{600} readings, the numbers output by the BugEye do not mean anything (beyond value for relative comparisons) without being calibrated to a universally meaningful measurement such as DCW.

Since the BugEye output is simply based on scattering of light caused by the contents of the medium, the bacteria are not the only medium components that affect the BugEye measurement. The two major factors that affect measurements are sparging and agitation, which both affect the size and amount of bubbles in the medium, which also scatter the light the BugEye emits. Consequently, the BugEye data will have sharp changes due to changes in agitation and/or sparging. When interpreting and calibrating data, this effect must be taken into account.

The BugEye was used to monitor the fermentation conducted from 10/4/04–10/6/04 with strain BL21(DE3)[pJY-1] grown in a 10 L BioFlo 110. Figure 4-5

76

shows both the BugEye and OD_{600} data for the entire fermentation.[8] The BugEye data is "messy" in several sections because of its sensitivity to changes in agitation and aeration. At several points in the fermentation these parameters were changed and the BugEye was even taken off of the bioreactor for several minutes. These disturbances cause the unexpected artifacts in the BugEye data; the BugEye instrument recorded data every 20 seconds, while the OD_{600} was recorded every couple of hours, so the OD_{600} data artificially appears much more smooth.

Depending on which trend in Figure 4-5 is considered correct, during the later stages of the fermentation (time > 20 hours) it could be concluded that the growth rate increased or decreased, depending on which dataset is more accurate. The most likely explanation is that the linearity of one of the measurement methods changes at higher cell densities. Since neither the feed rate nor the oxygen enrichment levels changed over the course of the late fed-batch and induction phases, it is expected that as biomass increased, the ratio of feed energy used for growth to that used for culture maintenance metabolism (i.e. just keeping the culture alive) would decrease, thus resulting in a lower growth rate as the biomass accumulated. The BugEye data reflects this predicted gradual decrease in growth rate while this behavior in the OD_{600} curve is not readily apparent. But, these differences are subtle and more data must be generated to determine their significance. The near-constant rate of biomass accumulation from the late fed-batch phase through the remainder of the fermentation is consistent with the fact that the feed rate remained constant and did not change significantly.

Because of the noise introduced by the non-uniformity of the bubble mixture in the broth, the BugEye employs a smoothing algorithm that takes the running average of the previous n points ($n \approx 5$). Even with smoothing, noise is still visible in the data. However, the instrument still appears to be very sensitive to subtle changes in the growth of the culture. For example, during the fed-batch and induction phases of the fermentation the feed bottle repeatedly had to be refilled with the glucose feed

---

[8]The following data analysis considers the raw BugEye data and OD_{600} values without worrying about calibrating these data to a physically significant value such as DCW. The discussion focuses on relative differences in the data and does not make conclusions about absolute quantities of biomass.
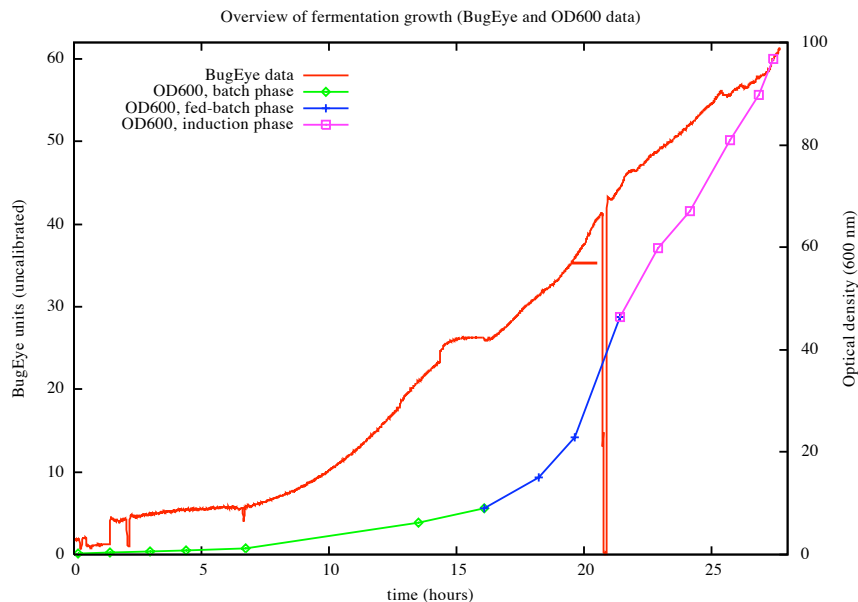
Figure 4-5: BugEye and OD_{600} data for the entire BL21(DE3)[pJY-1] fermentation.

solution. Sometimes, due to complications with the fermentation setup (e.g. excessive foaming, need to change exhaust filter, etc.), the feed bottle remained empty for periods up to about 15 minutes before it could be refilled. Without the addition of the carbon source, the culture's biomass can not increase, and the culture will theoretically stop growing. Halts in growth at times when the culture starved followed by a recovery of growth once the feeding was resumed are reflected in the BugEye data. Figure 4-6 shows such behavior over the course of three bottle refills. The first and last bottle refill occurred after the culture had starved, while the intermediate refill was done before the feed bottle was empty and no change in growth kinetics was observed. The fluctuation in BugEye measurements in the hour leading up to the third feed refill (hours 25–26) was probably due to a halt in air/O_2 sparging due to a clogged exhaust filter. The implications of using such finely-tuned monitoring of bioreactor growth kinetics for process definition and control are far-reaching, and further experiments with probes for the on-line monitoring of biomass must be done. A first step will be to establish the repeatability of these results and to calibrate the probe's BugEye units to OD_{600} and DCW data.
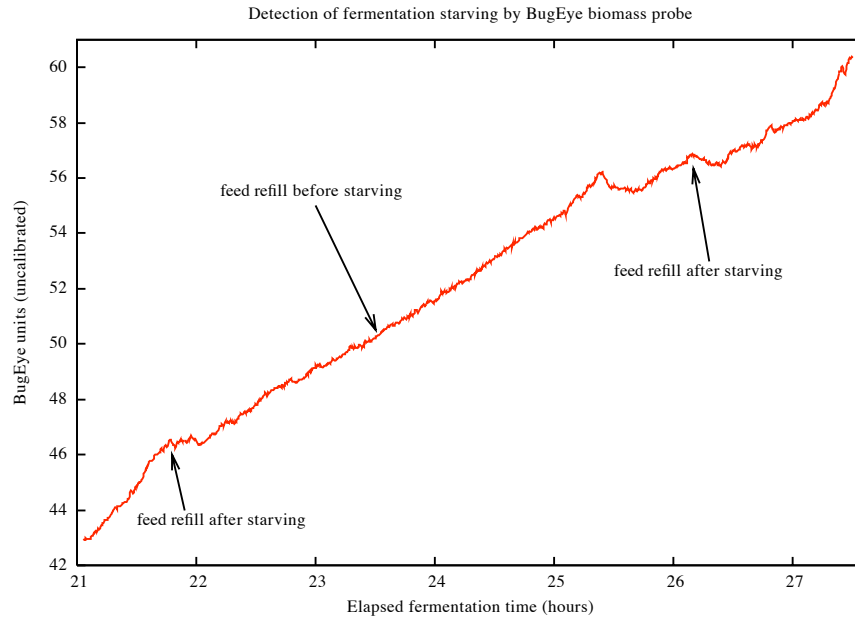
Figure 4-6: Change in growth rate due to starving as detected by BugEye noninvasive biomass probe.

4.2 Fermentation cooling

Based on the successful shake flask cooling experiments described in Section 4.1.1, fermentations implemented in course 10.28 (Fall 2003 and Fall 2004) included cooling phases. The original purpose of the cooling phases was to allow students to inoculate their bioreactors on one lab day then delay the post-batch phases of the fermentation until the following lab day (40 hours later). (See Appendix B for a description of the 10.28 course structure and a brief summary of the fermentations that were conducted.) The initially positive results of these experiments suggest that cooling can be used to make a fermentation's schedule more flexible for researchers (and possibly industry) without affecting the quality of a fermentation run.

4.2.1 First experimental iteration: course 10.28, Fall 2003

The fermentations of the Fall 2004 offering of course 10.28 were the first to be cooled. In order to design the cooling methodology, it was hypothesized that the best time to cool the bioreactors was after lag phase had been completed and the culture was grow-

ing vigorously, but before the cell density had increased greatly. The more healthy the cells were at the time of cooling, the more likely they would be to recover quickly when re-warmed. But, if the cell density was allowed to increase to a high enough level, the agitation and aeration of the broth would have to be kept at a near-process value in order to prevent the live but cold culture from being oxygen-limited, which is a potentially unhealthy condition. Since the cooling phase was to last from 24–48 hours, it was desirable to decrease the agitation and aeration as much as possible in order to reduce the likelihood of foaming, the resultant complications, and therefore the need for constant monitoring of the process.

During the Fall 2004 offering of 10.28, four fermentations were inoculated: two with BL21(DE3)[pJHL] (both at 1 L working volume) and two with BL21(DE3)[pJY-2] (one at 4 L and one at 10 L working volume. After about 3–4 hours of growth, the bioreactors were cooled to $12°C$.[9] While cooled, one of the 1 L fermentation controllers failed and the fermentation was lost. About 36–40 hours after cooling, the remaining three fermentations were warmed to their appropriate temperatures, and growth resumed. The combination of an ineffective gas mixing system and an inconsistent implementation of DO-stat feeding (see Section 3.2.4, page 57) limited the OD_{600} of the 10 L culture to about 20. However, the 5 L culture grew vigorously and reached an OD_{600} of approximately 55. The results from this set of fermentations proved that full recovery (from a growth kinetics point of view) from a 36–40 hour cooling phase is possible.

4.2.2 Cooling during batch and fed-batch phases

The design of the cooling phases of the Fall 2004 fermentations was based on the Fall 2003 results. In order to maximize the amount of time that the students spent with the bioreactors and to minimize the amount of work that the staff did to monitor the bioreactors outside of class, the first set of 10 fermentations was designed with

---

[9]Because of the lack of insulation of the cool water lines leading from the chiller to the bioreactor controller and the fact that the chiller could not cool below $0°C$, $12°C$ (and not lower) was a reasonable bioreactor temperature for the setup to sustain.

| Team No. | Strain | OD$_{600}$ at cooling phase | | Final OD$_{600}$ |
		Batch cool	Fed-batch cool	(growth time)
G1	BL21(DE3)[pJY-2]	7.93	16.1	23.2 (5 hr)
H1	BL21(DE3)[pJY-2]	7.74	12.5	19.3 (4–5 hr)
J1	BL21(DE3)[pLysS][pJY-2]	5.58	7.225	8.42 (4–5 hr)
L1	BL21(DE3)[pLysS][pJY-2]	6.08	7.54	11.3 (3 hr)
G2	BL21(DE3)[pJY-2]	(data missing)	15.9	20.6 (3 hr)
H2	BL21(DE3)[pJY-2]	5.80	14.4	15.1 (3 hr)

Table 4.3: Growth of fermentations subjected to cooling phase interuptions (10°C) of the batch and fed-batch phases.

two cooling stages. Students inoculated bioreactors on a Tuesday and Wednesday, the bioreactors were cooled to 10°C, all bioreactors were warmed back up on Thursday, reached the end of the fed-batch phase during the lab session, were cooled down near the end of the lab session, and were warmed back up on Friday to be induced. In short, this set of 10 student fermentations was not successful: high cell densities were not achieved, and after the second cooling phase (which arrested the cells in intermediate fed-batch phase), only low growth rates were achieved in the most healthy fermentations. Table 4.3 summarizes the optical densities at different stages of some of the module 1 fermentations. Each cell density is the OD$_{600}$ value at the time of the respective cooling phase. The growth time (in parenthesis) is the time it took the culture to grow from the listed OD$_{600}$ at the fed-batch cooling phase to the final OD$_{600}$, when the fermentation was terminated.

In general, it was observed during the cooling experiments that as the temperature decreased, the oxygen demand of the culture decreased, thus decreasing the necessary driving force to sustain the lower oxygen uptake rate (OUR), which resulted in an increase in dissolved oxygen concentration within the bioreactor. If the cell density was low enough when the cooling phase was performed, a significant increase in the dissolved oxygen concentration was seen when the culture was cooled. However, during the second cooling phase, the dissolved oxygen content of many cultures consistently remained at 0%. Because of the higher cell density, the cultures' oxygen demand was larger than cultures cooled at lower densities. In addition, as mentioned in Section 4.2.1, it is desirable to decrease the agitation and aeration during the cool-

ing phase in order to minimize the likelihood of a disaster due to foaming while the fermentation is left unattended. Before being cooled for a second time the agitation ranges of the module 1 cultures were in the range 400-600 rpm, and the aeration was uniformly at 1 vvm (0.9–1 lpm). During both cooling phases, these parameters were arbitrarily reduced to 150 rpm and 0.5 vvm.[10] It is because of the increased oxygen demand of the higher density cultures and the imposed conditions causing lower oxygen transfer rates (OTR) that the dissolved oxygen concentration was consistently at 0%. It is hypothesized that by imposing conditions that decreased the sustainable OTR during the fed-batch cooling phase, the cultures were severely oxygen limited and consequently damaged during this phase.

Based on the poor growth in the module 1 fermentations (Table 4.3) and the fact that the second, higher cell density cooling phase was the primary culprit for causing the poor growth, the module 2 fermentations were designed to have a single low cell density cooling phase to try to reproduce the more successful results of the 10.28 Fall 2003 experiments.[11] Consequently, the Fall 2003 fermentations were mimiced as closely as possible, with one cooling phase ($\approx$ 36 hours) inserted in the middle of the batch phase. After inoculation, the bioreactor cultures were grown for 3.5–4 hours ($OD_{600} \approx$ 2–4) before cooling to 10°C. About 36 hours later, the bioreactors were warmed to their process temperatures and reached the end of the batch phase after 1.5–2 hours. Because the fed-batch phase started in the morning (about 8:00 am), before the students got to lab (1:00 pm), there were not enough staff present to take measurements on the bioreactors during the initial part of fed-batch phase. Consequently, limited data is available on early fed-batch phase. The growth results for selected module 2 fermentations are shown in Table 4.4. To give an idea of the growth rate during the later stages of the fermentations, the OD_{600} measurement when the students entered lab (end of fed-batch phase) and several hours later (near the beginning of induction) are given along with the growth time for this change in

---

[10]Such a reduction had worked succesfully for the Fall 2003 10.28 fermentations, so without further data, the same reduction was used.

[11]Variability in culture behavior due to strain identity is unlikely since successful cooling experiments were performed with both strains in module 2 of the 10.28 Fall 2004 offering.

Team No.	Strain	OD_{600} change (growth time)
A1	BL21(DE3)[pJY-2]	44.0 → 59.1 (2.0 hr)
B1	BL21(DE3)[pJY-2]	41.5 → 53.4 (1.5 hr)
C1	BL21(DE3)[pJY-2]	48.4 → 70.0 (2.75 hr)
D1	BL21(DE3)[pLysS][pJY-2]	7.15 → 9.25 (1.25 hr)
E1	BL21(DE3)[pLysS][pJY-2]	21.9 → 27.0 (3.0 hr)
F1	BL21(DE3)[pLysS][pJY-2]	28.3 → 37.6 (2.25 hr)
A2	BL21(DE3)[pJHL]	53.7 → 80.4 (1.2 hr)
B2	BL21(DE3)[pHL]	18.0 → 26.0 (3.1 hr)
C2	BL21(DE3)[pLysS][pJHL]	32.7 → 42.3 (1.5 hr)
D2	BL21(DE3)[pLysS][pJHL]	26.1 → 29.7 (3.0 hr)

Table 4.4: Growth of course 10.28 module 2 fermentations subjected to a single cooling phase interuption (10°C) during the batch phase.

cell density. Most of the fermentations saw little growth during the induction phase.

The generally more vigorous growth of the module 2 bioreactors leads to the conclusion that with the current experimental setup, a single cooling phase at low cell density during the batch phase is the most robust design. Cooling at higher cell densities or for longer periods of time could possibly be done if the bioreactors are cooled to less than 10°C, further slowing bacterial metabolism.

4.2.3 Glucose consumption during the cooling phase

A danger of allowing the fermentation to progress too far through the batch phase before cooling is that there might not be enough glucose in the bioreactor to sustain the culture until the fermentation is re-warmed. Even though *E. coli* metabolism is slowed down greatly at 10°C relative to a normal process temperature, thus drastically reducing glucose consumption, 36–40 hours (the typical length of the cooling phase required by the courses) is 2–3 times longer than the average process length; for cooling phases of significant length, then, it is likely that glucose limitation becomes a critical factor that must be addressed by the process design. For the first set of cooled fermentations (course 10.28, Fall 2003), the cooling phase was arbitrarily initiated after approximately 3 hours of fermentation had elapsed. Since the first fermentations were successful, this strategy was not changed for the fermentations

implemented in the Fall 2004 offering of course 10.28; all resultant fermentations that were healthy at the time of batch-phase cooling recovered upon re-warming and reached the fed-batch phase.

Because of the time constraints on the course, limitations on staff time outside of course lab hours, and periods of intermittent chaos, complete sets of glucose measurements bounding the start and end of the cooling phase were not taken for any of the bioreactors during both module 1 and 2! As a result, the following information is the best reconstruction of the experiments from the limited data available. In general, batch phases with strains BL21(DE3)[pJHL] and BL21(DE3)[pJY-2] are about 5–7 hours long (see Section 4.1.2). The Fall 2004 course 10.28 fermentations began with 10 g/L glucose at inoculation and were grown for approximately 3.5–4 hours before cooling to 10°C. Glucose concentrations at the beginning of the cooling phase were approximately 5 g/L $\pm$ 1 g/L. After the 36–40 hour period at 10°C, glucose concentrations were in the approximate range 1.5–3 g/L. After warming back up to the process temperature, batch phases lasted another 1–2 hours before the glucose was completely consumed. Based on this approximate data, it appears that a measurable amount of glucose *was* consumed. However, the length of the batch phase does not seem to have been greatly affected. More useful conclusions, though, can not be made without more rigorous experimentation. In addition to more thoroughly defining the glucose consumption curve through the interrupted batch phase, it would be interesting to prolong the cooling phase for several days to verify whether the glucose consumption characteristics of a 36–40 hour cooling phase match those of an extended phase; the idea being that glucose consumption over a longer period of time at cold temperatures would be especially accurate, since the culture has more time to consume glucose and establish a measurable glucose differential over time.
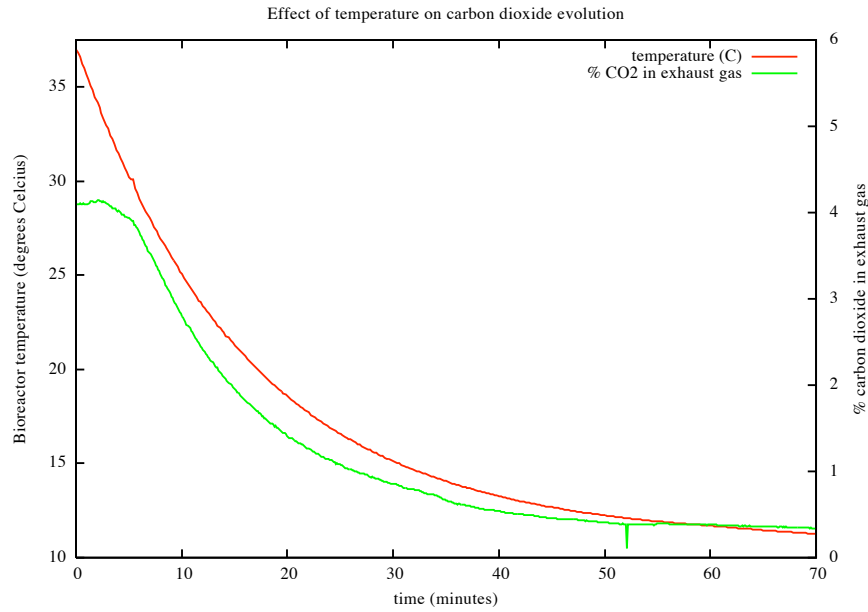
4.2.4   Effect of cooling on carbon dioxide evolution

In order to study the effect of temperature on culture metabolism, the response of dissolved oxygen and carbon dioxide evolution due to a steady decrease in temperature were measured. During module 2 of course 10.28, after the induction phase was com-

plete, the temperature setpoint of group C1's fermentation was changed from 37°C to 10°C. In this fermentation, strain BL21(DE3)[pJY-2] was grown in a BioFlo 110 vessel with a working volume of 0.9 L. The culture reached an OD_{600} of 70 (30 g/L) midway through the induction phase and remained at that constant cell density throughout the rest of the experiment.[12] At the time this experiment was performed, the agitation was 700 rpm, sparged air was enriched with pure oxygen 20% of the time, and the gas flow rate was 1 vvm (0.9 lpm). Since the temperature in a BioFlo 110 is decreased by flowing cool water through a cooling coil submerged in the broth, the temperature response due to a step change in the setpoint is a long and smooth decrease ($\frac{dT}{dt} \approx -k(T - T_{cooling\ coil})$). Figure 4-7 shows the response of carbon dioxide in the exhaust gas and dissolved oxygen due to the gradual temperature decrease. Note that since the culture's growth rate was near-zero at the time of this experiment, the oxygen demand was much lower than that of a vigorously-growing culture. This can be seen by the fact that at full temperature, the maximum level of carbon dioxide in the exhaust gas was 4%, while near-10% values have been measured for rapidly-growing cultures near that cell density.

One concern in the analysis of this data is the magnitude of time lags in the data that is presented and the effect of these lags on the conclusions that can be made. It should be explicitly noted that there was a two minute delay to the off-gas measurements caused by the dead space in the stainless steel sampling lines (other bioreactors had longer and shorter delays depending on their proximity to the mass spectrometer). In the data processing steps leading up to the generation of Figure 4-7, this delay was taken into account. The time between a change of temperature at the probe and the registering of that change by the BioCommand Plus software was not more than 10–20 seconds, and that delay combined with the inherent 30–40 second DO probe measurment delay was not more than a minute. Based on the plotted data, these delays relative to the rate of change of temperature are negligible. However, the relationship between the temperature and the metabolism of a culture is more subtle.

---

[12]Since the cell density was constant throughout the cooling experiment described below, it is assumed that the effect of changes in biomass on the culture's oxygen consumption during the experiment were negligible.

(a) Off-gas %CO$_2$ and temperature vs. time.
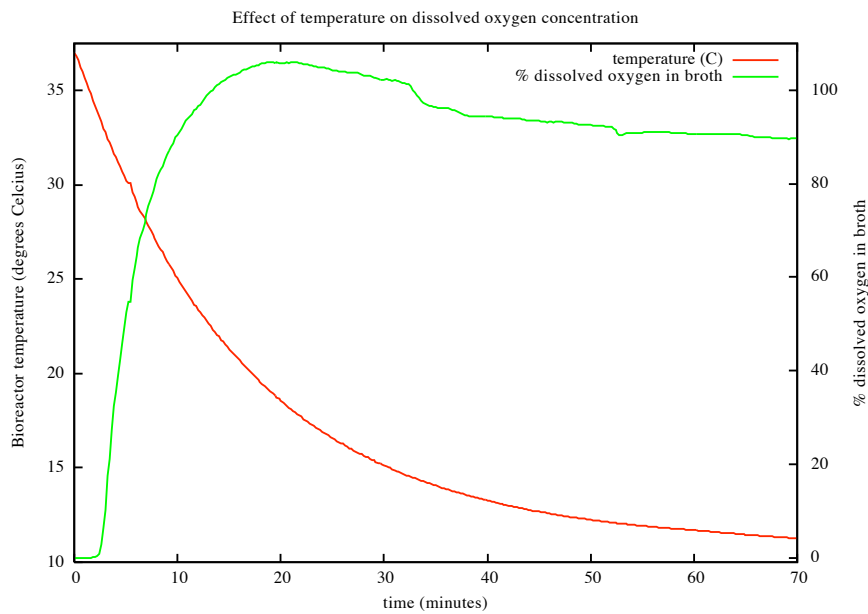
(b) Dissolved oxygen percent and temperature vs. time.

Figure 4-7: Effect of cooling a high cell density culture on carbon dioxide evolution in the off-gas and dissolved oxygen concentration in fermentation broth.

It is a given that reducing the temperature affects the rate of metabolic processes from a basic reaction kinetics standpoint. However, less is understood about longer term responses of *E. coli* to cold temperatures. If the *E. coli* were allowed to stabilize by fixing the temperature and not continually decreasing it, a shift in metabolic state evidenced by an additional slower dynamic in the DO and $\%CO_2$ may have been seen. In other words depending on the rate of the temperature drop, the DO and $\%CO_2$ responses that are measured may be dominated by the fast dynamics of reaction kinetics or slower changes in bacterial state in response to the temperature change. To study this further, a culture should be fixed at a range of below-process-setpoint temperatures and the DO and $\%CO_2$ measured, plotted, and compared with the shape of the curves obtained in the experiment implemented by this thesis.

Based on the data, the culture's oxygen demand dramatically decreases with temperature. It is interesting that the shape of the $\%CO_2$ curve closely tracks the decreasing temperature, but that the correlation between temperature and $\% CO_2$ is not exactly linear; the rate of change of the $\%CO_2$ is slightly greater than that of the temperature. This nonlinearity may be due to an unaccounted for mechanism involved in the measurements, or it may be due to the effect of a biological process that is activated by the falling temperature and augments the decline in CO_2 production due to slowed reaction kinetics. The $\%CO_2$ stabilizes at 0.32% as the temperature asymptotes in the 11°C range. Further experiments should be done with cultures of varying growth vigor at the time of cooling to see if their carbon dioxide production ends up stabilizing at the same low value. In addition, it would be useful to see the effect of cooling the culture even further on the CO_2 output.

It is also noteworthy that the percent dissolved oxygen rises to its maximum value[13] much sooner than the $\%CO_2$ reaches its minimum. Then the %DO begins to decrease again, and as the $\%CO_2$ stabilizes at 0.32% (with a very slow rate of continued decrease, mirroring the asymptotically decreasing temperature), the %DO continues to fall slowly. The discrepancy between a falling %DO (implying increasing

---

[13]The maximum %DO is greater than 100% because the probe was calibrated with pure air, but at the point in the fermentation when this experiment was done, air enriched with pure oxygen (20% of the time) was being sparged, thus increasing the saturation concentration of O_2.

oxygen consumption because of the increasing driving force) and the falling $\%CO_2$ in the off-gas (implying decreased aerobic metabolism) can not be explained by the simple assumptions made thus far about the relationship between aerobic metabolism and carbon dioxide evolution. This experiment should be repeated to determine whether the results are reproducible.

4.3 Volumetric oxygen mass transfer coefficient (k_La)

Because oxygen mass transfer is most often the rate-limiting process in aerobic fermentations, precise quantification of a process' oxygen demand and the fermentation system's ability to supply that demand is important. This topic is especially relevant to scale-up methodologies, which take processes from the laboratory "bench" scale (0.2–100 L) to the industrial scale (10,000–100,000 L). Effective scale-up is not simply a matter of scaling the dimensions of bioreactor components (e.g. impellers, baffles) and the volume of gas per volume of liquid (vvm) sparged through the culture broth; the relationship between dimensions and mass transfer at different scales is nonlinear and difficult to predict. The correct scaled parameters are often result in immoderate dimensions which must be engineered to within acceptable ranges; because of the complexity of the relationship between dimensions and mass transfer at different scales, this engineering process may not ensure mass transfer sufficient for a particular process [4]. In contrast to the physical dimension approach to scale-up, the volumetric oxygen mass transfer coefficient (k_La), which appears in the oxygen mass transfer equation (Equation 4.1), may be used to encapsulate a fermentation system's ability to sustain a particular oxygen transfer rate. The paradigm of k_La is fundamentally scalable because it allows for variation in bioreactor design in order to accomodate the varying constraints of different scales. Thus, maintenance of a particular value of k_La at different scales is more likely to insure sufficient oxygen transfer rates than a purely dimensional-based methodology of scaleup.

The calculations below are based on the theory described in Section 3.3.2. For convenience, the oxygen mass transfer equation and it's steady state simplifications

are reproduced here as Equations 4.1 to 4.3. For a detailed description of the variable definitions and discussion of these equations, see Section 3.3.2 (pages 61–63).[14]

$$\frac{dc_L}{dt} = k_L a(c^* - c_L) - OUR \tag{4.1}$$

$$\text{In steady state}: \quad \frac{dc_L}{dt} = 0 \implies k_L a = \frac{OUR}{(c^* - c_L)} \tag{4.2}$$

$$\%OUR = \%O_2^{in} - \left(\frac{\%N_2^{in}}{\%N_2^{out}}\right)\%O_2^{out} \tag{4.3}$$

4.3.1 Determining the oxygen saturation concentration, c*

Judging from Equations 4.1 and 4.2, the accuracy of the $k_L a$ calculation depends on the accuracy of the value of the oxygen solubility at saturation. Under fixed conditions, c* is linearly proportional to the partial pressure of oxygen in the sparged gas and headspace. This relationship is described by Henry's law in Equation 4.4 (k_H is the Henry's law constant, c* is the saturation concentration of oxygen for a fixed set of conditions, and p_g is the partial pressure of oxygen in the sparged gas).

$$k_H \overset{def}{=} \frac{c^*}{p_g} \tag{4.4}$$

The primary factors affecting k_H under fermentation conditions are temperature, ionic strength, and composition of the fermentation medium [3].[15] Thus, a different value of k_H must be used for 30°C and 37°C processes. Because of the dependence of the constant on broth composition, the $k_L a$ will also vary as the fermentation progresses: cell density will increase and solute composition changes as cells consume/produce substrates and lyse. Variation in k_H over a range of conditions is shown in Table 4.5. Based on this data, temperature and/or medium composition differences can cause about 10–20% variation in k_H depending on the conditions chosen. For the 37°C

---

[14]c_L is the concentration of dissolved oxygen in the fermentation broth $[\frac{mol}{L}]$, c* is the saturation concentration of oxygen in the broth at the given operating temperature, and OUR is the oxygen uptake rate $[\frac{mol}{L\cdot hr}]$. Variables $\%O_2^{in}$, $\%O_2^{out}$, $\%N_2^{in}$, and $\%N_2^{out}$ are the percentages of the respective gas in the inlet or outlet gas.

[15]Oxygen solubility decreases with increasing ionic strength and temperature.

Aqueous solution	Henry's constant	Reference
Pure water (15°C)	1.54 $\left[\frac{\text{mmol}}{\text{atm·L}}\right]$	MIT BE.360 handout, [42]
Pure water (25°C)	1.3	Sander, NIST, [53], [39]
Pure water (35°C)	1.09	MIT BE.360 handout, [42]
Pure water (35°C)	1.11	Parente, et al., [43]
0.25 M NaSO$_3$ (35°C)	0.92	Parente, et al., [43]
1 M NaCl (25°C)	0.89	MIT BE.360 handout, [42]
Cell culture medium (37°C)	0.93	MIT BE.360 handout, [42]

Table 4.5: Examples of variations in the Henry's law constant (k_H) for oxygen gas for different temperatures and liquid compositions.

fermentation analyzed below, the value $k_H = 0.926 \frac{\text{mmol}}{\text{atm·L}}$ is used [42].

Depending on the particular experimental setup, it was sometimes not convenient to set up a system for sampling the bioreactor inlet gas to directly measure the inlet gas composition. Early in a fermentation batch phase, the lack of inlet gas measurement was not an issue because the lab air could be sampled before or after the fermentation. Later in the fermentation, though, when the sparged lab air was enriched with pure oxygen, the gas composition at the inlet had to be approximated. The New Brunswick BioFlo controllers mixed air and oxygen by alternately turning on and off valves connected to these gas supplies. The controller defined the level of oxygen enrichment by the percent of the time that pure oxygen, rather than air, was flowing (the duty cycle of the system was on the order of one second). With these parameters in mind, Equation 4.5 can be used to estimate the partial pressure of oxygen in the inlet gas ($\%O_2^{in}$, this is in terms of a volume percentage). $\%O_2^{air}$ is the percent oxygen in air (20.95%), $\%O_2^{enrich}$ is the percent of the time that pure oxygen flows to the bioreactor, and $\%O_2^{enrich}$ is the percent of the time that the controller sparges pure oxygen rather than air through the bioreactor.

$$\%O_2^{in} = \left(1 - \frac{\%O_2^{enrich}}{100}\right)\%O_2^{air} + \%O_2^{enrich} \tag{4.5}$$

So, when calculating the $k_L a$ during times when oxygen enrichment was used, $\%O_2^{in}$ from Equation 4.5 must be used as the partial pressure of oxygen used to calculate c* rather than $\%O_2^{air}$(see Equations 4.4 and 4.6).

4.3.2 Dynamic method

The dynamic method analysis here will focus on the three dynamic method experiments that were performed using the 10 L fermentation with strain BL21(DE3)[pJY-1] (10/4/04–10/6/04, process temperature = 37°C). The experiments were performed early in the fermentation when the OD_{600} ranged from 0.45 to 0.68. Three experiments were implemented by varying the agitation rate in the following order: 400 (run time = 1.3h, $OD_{600} \approx 0.4$), 300 (run time = 1.7h, $OD_{600} \approx 0.5$), and 500 rpm (t = 3.1h, $OD_{600} \approx 0.7$). The air flow rate was kept constant at 0.5 vvm. Because the dynamic method experiments were performed early in the fermentation, oxygen enrichment was not used. Therefore, the partial pressure of oxygen was that of the atmosphere, and the saturation concentration of oxygen can be calculated as:

$$c^* = p_g \cdot k_H = (0.2095 \text{ atm}) \left(0.926 \, \frac{\text{mmol}}{\text{atm} \cdot \text{L}} \right) = 0.194 \times 10^{-3} \, \frac{\text{mol}}{\text{L}}. \qquad (4.6)$$

In Figure 4-8 (400 rpm experiment), the slope of the OUR linear regression is $-16.9 \, \frac{\%\text{DO}}{\text{min}}$. Convert to units of molarity per hour as follows:

$$\frac{\text{mol O}_2}{(\text{L of broth}) \cdot \text{hr}} = \left(\frac{\%\text{DO}}{\text{min}} \right) \left(\frac{60 \text{ min}}{1\text{hr}} \right) \left(\frac{c^* \, [\frac{\text{mol O}_2}{\text{L}}]}{100\% \text{ saturation}} \right) \qquad (4.7)$$

$$OUR = 2.15 \times 10^{-3} \text{ M/hr} \qquad (4.8)$$

The values of OUR for the 300 and 500 rpm experiments were calculated using the same method, and the results of the three dynamic method experiments are summarized in Table 4.6. The calculated values of OUR did not change significantly over the course of the experiments; this trend is validated by the near constant growth curve at the beginning of the fermentation (Figure 4-5). In Table 4.6, "OUR $[\frac{\text{mol}}{\text{L} \cdot \text{hr}}]$ (regression)" refers to the value of OUR calculated from the slope of a linear regression on the initial decreasing trend in %DO (Figure 4-8(a)); "k_La $[\frac{1}{\text{hr}}]$ (regression)" refers to the value of k_La calculated from the slope of the regression on the plot of $\frac{d(DO)}{dt}$ vs. DO (Figure 4-8(b)); "k_La $[\frac{1}{\text{hr}}]$ (steady-state)" refers to the value of k_La calculated using the OUR (regression) and the steady state value that the %DO reaches after air
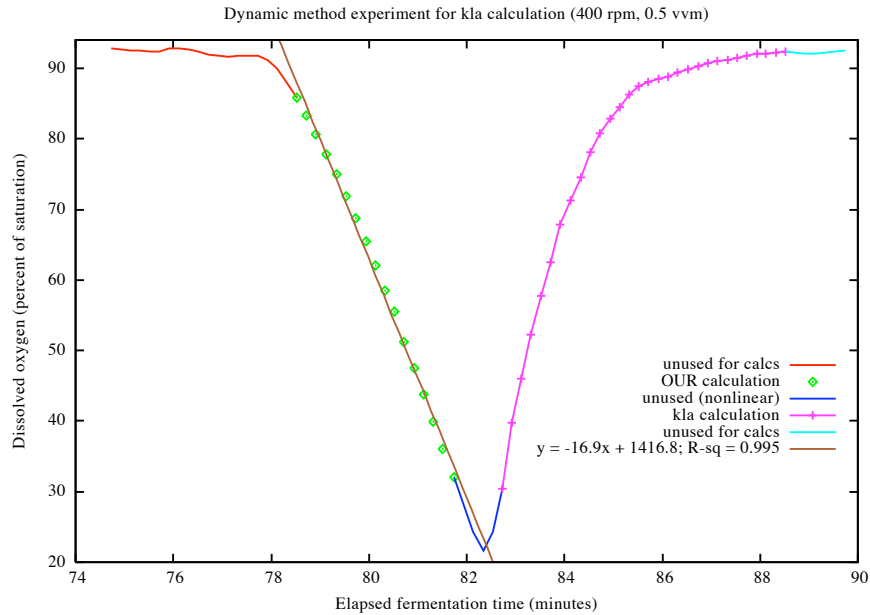
(a) Categorization of data used in dynamic method calculations with OUR regression plotted.
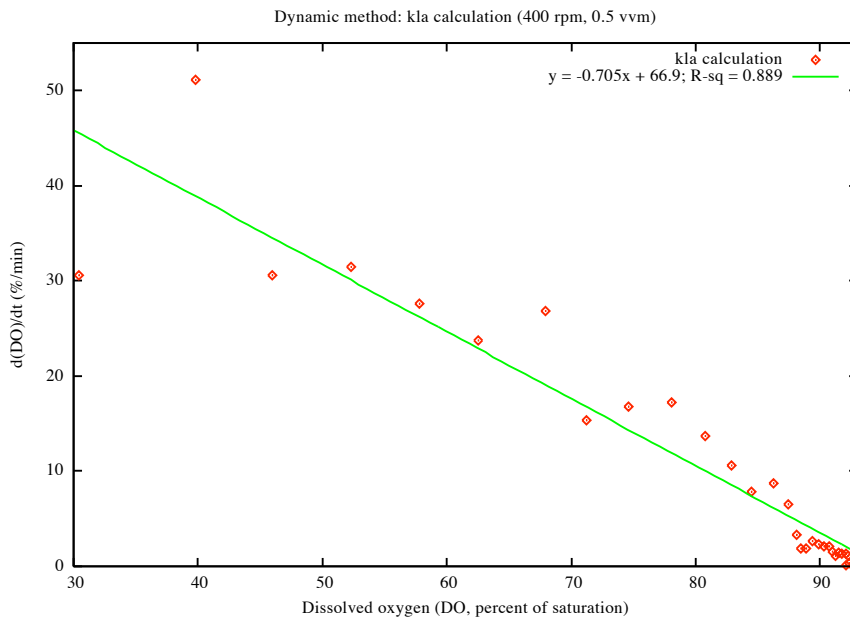
(b) Plot of $\frac{d(DO)}{dt}$ vs. DO; the k_La is the magnitude of the slope of this regression.

Figure 4-8: Dynamic method calculations of oxygen uptake rate (OUR) and volumetric oxygen mass transfer coefficient (k_La) at 400 rpm, 0.5 vvm.

Agitation (rpm)	OUR $[\frac{\text{mol}}{\text{L}\cdot\text{hr}}]$ (regression)	k_La $[\frac{1}{\text{hr}}]$ (regression)	k_La $[\frac{1}{\text{hr}}]$ (steady-state)
300	$2.15 \cdot 10^{-3}$ ($R^2 = 0.999$)	20 ($R^2 = 0.611$)	37
400	$1.97 \cdot 10^{-3}$ (0.995)	42 (0.89)	145
500	$2.22 \cdot 10^{-3}$ (0.999)	63 (0.94)	286

Table 4.6: Dependence of k_La on agitation; OUR and k_La were calculated by linear regression and steady-state analysis of dynamic method data.

sparging is recommenced (using Equation 4.2). Considering the fact that the experiments were performed in the order 400, 300, and 500 rpm, and that the cell density increased over that time period, the trend of increasing OUR with time is expected and contributes to the level of confidence in the accuracy of these measurements.

In performing this dynamic analysis, nonlinear data bounding the critical sections of the curves used for the calculations was intentionally not used because it is not accounted for by the simplistic models described in Section 3.3.2. These models do not take the dynamics of the DO probe into account, which are the cause of the nonlinearities. First consider the data used for the linear regression to calculate the OUR (i.e. the semi-linear decrease in DO% when the airflow was turned off, Figure 4-8(a)). When the airflow was stopped at 78 minutes, the magnitude of the slope gradually increased until it reached the magnitude of the slope of the regression line. Near the "bottom" of this curve, where the DO% decreased to 25%, the magnitude of the slope then began to decrease with time (i.e. the slope became more shallow). The drawing in Figure 4-8(a) makes it clear that these nonlinear sections were not used in the OUR regression calculation. Likewise, there is a brief lag in the onset of the rapid rise in DO% when the airflow is turned on at 82.3 minutes, which is not accounted for by the solution to the oxygen mass transfer (differential) equation which has a single exponential term (Equation 4.1). Neither this nonlinear data at the onset of the rise in DO% or the data when the DO% reaches steady-state after 88.5 minutes was used in the calculation of the k_La. Including more data points of $\frac{d(DO)}{dt}$ vs. DO would only add more points to the cluster of points at (92%, 1%/min) in Figure 4-8(b), which would minimally affect the slope of the regression line. It is the points near the onset of the DO rise that most affect the slope, so the choice of

the time at which to begin using the data in the $k_L a$ regression calculation is very important.

As mentioned above, the nonlinearities not accounted for by the model used for the calculations are caused by the nonideality of the DO probe. Because the DO probe has an oxygen permeable membrane separating the electrode from the fermentation broth, oxygen must diffuse through the membrane to the electrode. For diffusion to take place, a concentration gradient must exist. This means that while the broth concentration is changing, the concentration of oxygen at the probe electrode will never equal that of the broth. In other words, there is a delay between the time that the probe is presented with a specific broth concentration and the time at which the probe's signal reflects that concentration. Because the probe delay, τ_{probe}, is governed by diffusion, it can be modeled as a first order differential equation (Equation 4.10). Incorporating the probe delay into the original oxygen mass transfer model results in a set of coupled differential equations which can be solved to give the probe concentration as a function of time (Equations 4.9 and 4.10). The solution will not be presented here, but it is important to recognize that the general form of the solution has two exponential terms: the time constant of one term is the DO probe delay time constant, τ_{probe} (this term dominates at the onset of the rise in the DO% when the airflow is turned on) and the other term's time constant, $\tau_{k_L a}$, is the inverse of the $k_L a$ (this term dominates after the initial rise in DO).

$$\frac{dc_L}{dt} = \frac{1}{\tau_{k_L a}}(c^* - c_L) - OUR; \tau_{k_L a} = \frac{1}{k_L a} \tag{4.9}$$

$$\frac{dc_{probe}}{dt} = \frac{1}{\tau_{probe}}(c_L - c_{probe}) \tag{4.10}$$

If $\tau_{probe} \gg \tau_{k_L a}$, the oxygen concentration will reach steady state much faster than the DO probe reading will change, and the resultant response of the DO probe will approach the probe response that would be measured if there was a step change in dissolved oxygen within the bioreactor. Thus, for values of $k_L a$ that are too large, the method of performing a linear regression on the $\frac{d(DO)}{dt}$ vs. DO points would not

work; the calculated $k_L a$ would closely approximate $\frac{1}{\tau_{probe}}$. The time constant, τ_{probe}, of the polarographic DO probe used in these $k_L a$ experiments was approximately $40\ s^{-1}$, or $90\ hr^{-1}$. Therefore, the dynamic method is only suitable for measuring the OUR and $k_L a$ at low cell densities ($< 0.5\ g/L$) and low values of $k_L a$ ($\ll 90\ hr^{-1}$), when the dissolved oxygen dynamics induced by the experiment have time constants significantly less than the probe delay. Reconsidering the OUR and $k_L a$ data in Table 4.6, the values of OUR calculated from the regression are likely to be accurate considering that the drop in DO spanned a period of 3.2 minutes ($\gg \tau_{probe} = 0.7$ min). On the other hand, the values of $k_L a$ calculated from the regression on $\frac{d(DO)}{dt}$ vs. DO for the 400 and 500 rpm experiments were significant fractions of $90\ hr^{-1}$. The $k_L a$ values calculated using the OUR and the steady-state concentration differential are likely to be much more accurate for the values of $k_L a$ corresponding to the 400 and 500 rpm experiments; see Section 4.3.4 for a comparison with published values. (The two methods of $k_L a$ calcultion for the 300 rpm experiment resulted in values that were at least on the same order of magnitude, which can not be claimed for the higher rpm experiments. The discrepancy in the $k_L a$ values for the 300 rpm experiment is in part due to the lack of abundant data; since the DO probe data was recorded once every 12 seconds and most of the rise in DO took place over 1.5 minutes, there were only 7–8 data points in the range that had the most influence on the $k_L a$ regression slope—see Figure 4-8(b).)

4.3.3   Off-gas method

The off-gas approach to calculating $k_L a$ is based on the same steady-state simplification of the oxygen mass transfer equation as that used to calculate the $k_L a$ from dynamic method data (Equation 4.2). Once a steady-state in the oxygen mass transfer of the system is achieved,[16] the OUR $[\frac{M}{hr}]$ is calculated by first measuring the difference in oxygen composition of the inlet and exhaust gas flows to calculate the %OUR [units = %O_2 (volume)] (Equation 4.3) then converting the result to the cor-

---

[16]That is, constant %DO, constant OUR, constant process values; this is only a pseudo steady-state since the culture is growing all of the time

rect units using Equation 4.11.

$$OUR \ \left[\frac{\text{M}}{\text{hr}}\right] = \frac{\left(\frac{\%O_2^{in}}{100\%}\right)\left(Q_{sparge} \ \left[\frac{\text{L}}{\text{min}}\right]\right)\left(V_{gas \ law} \ \left[\frac{\text{L}}{\text{mol}}\right]\right)^{-1}\left(\frac{60 \ \text{min}}{1 \ \text{hr}}\right)}{\text{Broth volume [L]}} \qquad (4.11)$$

The value Q_{sparge} is the gas flow rate through the bioreactor in L/min, and $V_{gas \ law} = 24.5$ L/mol: the volume in liters occupied by 1 mol of ideal gas at room temperature and 1 atmosphere of pressure.[17]

The off-gas method of determing the OUR is ideal for high cell density fermentations; if the oxygen supply to the bioreactor were suddenly stopped (i.e. a step change in the driving force, $c^* - c_L$; the partial pressure of O_2 effectively goes to zero $\Rightarrow c^* = 0$) the OUR of the culture would be large enough to deplete the dissolved oxygen in the broth within a small fraction of the DO probe time delay ($\tau_{probe} \approx 40$ sec).[18] Consequently, the dynamics of the depletion of the dissolved oxygen if the gas supply were turned off would be much too fast to be measured accurately by the DO probe. An added benefit of the off-gas method is that there is no need to stress the culture by limiting its oxygen supply in order to measure the $k_L a$, as in the dynamic method. However, in order to use the off-gas data to measure the OUR, the cell density must be high enough to produce a measurable difference in oxygen composition between the inlet and outlet. The following analysis concerns the fermentation from 10/4/04–10/6/04 with strain BL21(DE3)[pJY-1] in order that the results may be compared with the dynamic method calculations of OUR and $k_L a$ in Section 4.3.2.

One hour after inoculation, when the OD_{600} was approximately 0.55 (0.2 g/L),

---

[17]After passing through the exhaust condenser, the off-gas passes through about 2 feet of PharMed® tubing before the flow rate is measured. Thus, the off-gas is at approximately room temperature when the flow rate is measured. Depending on the state of the exhaust filter, which may be moist or clogged, the pressure within the bioreactor may be greater than 1 atm, which is the pressure assumed for all calculations performed here.

[18]The solubility of glucose is > 700 g/L, which is much greater than the solubility of oxygen ($c^* = 0.93 \ \frac{\text{mmol } O_2}{\text{atm} \cdot \text{L}} = 6.2 \ \frac{\text{mg } O_2}{\text{atm} \cdot \text{L}}$). The general, simplified equation for aerobic respiration is $C_6H_{12}O_6 + 6 \cdot O_2 \rightarrow 6 \cdot H_2O + 6 \cdot CO_2$. The stoichiometry of oxygen glucose consumption is 6 mol O_2 consumed per 1 mol glucose, so O_2 is severely limited. The fact that the respiratory quotient (RQ = (volume of CO_2 generated)/(Volume of O_2 consumed)) may not equal 1 does not affect this conclusion of oxygen limitation.

there was a measurable %OUR of about $0.14\% \pm 0.03\%$ (measured using the Perkin Elmer mass spectrometer). Since the working volume was 10 L, the inlet gas flow rate of 0.5 vvm corresponds to 5 lpm. Plugging these values into Equation 4.11 gives an OUR of about $3.4 \cdot 10^{-3}$ $\frac{M}{hr}$. Considering that the OUR is calculated from a %OUR at the limit of detection of the mass spectrometer, this value is quite close to the OUR range calculated using the dynamic method at approximately the same time in the fermentation ($2.0 \cdot 10^{-3}$ M/hr, Table 4.6).

Next, consider data at later times in the fermentation, when the dynamic method for calculating $k_L a$ was unusable because of high OUR values. Off-gas OUR calculations, the resultant $k_L a$, and relevant process values at different times in the fermentation are listed in Table 4.7. The column "$\%O_2^{enrich}$" describes the amount of time that pure oxygen, rather than air, was flowing into the bioreactor (see Section 4.3.1, page 90 for a description of how oxygen enrichment works with a BioFlo controller). The column "c^*" describes the saturation concentration of oxygen at 37°C after taking the increase in partial pressure due to oxygen enrichment into account (Equation 4.5). Normalization of the measured oxygen in the off-gas to account for the effect of water vapor on the outlet gas compositions was not carried out (Equation 4.3); in the BL21(DE3)[pJHL] fermentation on 6/24/04, the difference between $\%N_2^{in}$ and $\%N_2^{out}$ was consistently $< 0.5\%$ of $\%N_2^{in}$ (the 10°C exhaust condenser removed most of the vapor that came off of the warm fermentation broth.). Thus, in the calculations done here, the effect of water vapor in the off-gas on the kla and OUR calculations is considered negligible relative to the other measurment error (e.g. sparging rate, which was accurate to within 0.2 lpm, $\approx 4\%$ of the sparging rate). When calculating the %OUR before calculating the OUR, the gas composition measurments of both the mass spectrometer and the laser gas analyzer were compared; composition measurements given by each instrument were within 1% (volume) for each particular gas.

The calculated values of $k_L a$ in Table 4.7 are quite variable. It is clear that the $k_L a$ increases as the agitation increases up to 800 rpm, but the value of 101 hr^{-1} is anomalous considering that it is nearly one-third of the value calculated using

OD$_{600}$	Agitation	%Dissolved O$_2$	%O$_2^{enrich}$	c*(at 37°C)	OUR	k$_L$a
0.55	400 (rpm)	92%	0%	0.194 [mM]	3.4 $\left[\frac{\text{mM}}{\text{hr}}\right]$	219 [hr^{-1}]
6.14	500	19	0	0.194	15.8	101
46.4	800	0	40	0.487	132	272
81.0	800	0	40	0.487	162	332

Table 4.7: Off-gas method of OUR calculation and k$_L$a values at high cell densities (fermentation with BL21(DE3)[pJY-1], 10/4/04).

the dynamic method (286 hr^{-1}, Table 4.6) and considering the relatively high k$_L$a magnitudes calculated by the off-gas method at other points in the fermentation.

4.3.4 Discussion of results in the context of literature

Depending on whether the "regression" or "steady-state" method of calculating k$_L$a from the dynamic method data was used (Section 4.3.2), the dependence of k$_L$a on agitation varied significantly (Table 4.6). In order to evaluate the accuracy of these two discordant sets of results, an emperical correlation developed by Van't Riet based on a large number of experiments in different vessels will be considered. The correlation fits emperical data to within 20–40% (Equation 4.12) [3, 48]. The correlation applies to noncoalescing fluids, which is typical of fermentation broths [1, 34].

$$k_L a = 2.0 \times 10^{-3} \left(\frac{P}{V}\right)^{0.7} (u_{gs})^{0.2} \text{ [s}^{-1}] \tag{4.12}$$

u_{gs} is the superficial gas velocity,[19] P/V is the volumetric power input, and k$_L$a is the volumetric mass transfer coefficient (units = s^{-1}). Equation 4.12 is applicable under the conditions: $V \leq 2600$ L; $500 < P/V < 10,000$ W/m^2. A correlation for the volumetric power input for agitated vessels, P/V, is shown in Equation 4.13 [3, 51].

$$\frac{P}{V} \approx \text{const} \times \frac{\rho_l N_i^3 D_i^5}{D_i^3} \text{ for } \text{Re}_i \geq 10^4 \tag{4.13}$$

The subscript i denotes the relation of the variables to the impeller (i.e. rather than another bioreactor component). ρ_i is the density of the liquid, N_i is the impeller

---

[19]$u_{gs} = \frac{\text{(gas feed volumetric flow rate)}}{\text{(vessel cross section area)}\cdot\text{(gas holdup)}}$

rotation rate (revolutions per time), D_i is the diameter of the impeller, and "const" is a proportionality constant dependent on other system parameters, including broth viscosity.

The combination the Equations 4.12 and 4.13 suggests the following proportionality:

$$k_L a \propto N_i^{2.1} \tag{4.14}$$

Considering the results of the "regression" method of calculating the k_La from the dynamic method data (Table 4.6, page 93),[20] k_La and agitation appear to be linearly related. Since the regression method depends on the dynamic response of the dissolved oxygen probe, the accuracy of the k_La calculation depends on the ability of the DO probe to track the true value of oxygen concentration in the broth. As discussed in Section 4.3.2, the limits of the probe delay decrease the accuracy of k_La calculations as the k_La approaches 90 hr^{-1}. This method will not result in a calculated k_La greater than 90 hr^{-1}, no matter what the true value of k_La is. Considering that the values of k_La at 400 and 500 rpm (42 and 63 hr^{-1} respectively) are large fractions of the 90 hr^{-1} limit, it is likely that these values are underestimates of the true k_La.

On the other hand, when k_La is calculated using the "steady-state" approach to analyzing dynamic method data,[21] only the accuracy of the OUR calculation depends on the probe response being much faster than the change in oxygen concentration. For low enough cell densities ($OD_{600} < 1$–2), this constraint on the change in oxygen concentration due to consumption is satisfied. Even though the correlation between k_La and agitation (N_i) appears to be governed by a polynomial of degree greater than 2, the fact that the increase in k_La with agitation is more rapid than a linear increase implies that the "steady-state" approach to the dynamic method gives more physical results with typical fermentation setups (i.e. high agitation and aeration).

Relative to the values of k_La calculated with the dynamic method, the off-gas

---

[20]In the "regression" method, k_La is equal to the slope of the linear regression on the plot of $\frac{d(DO)}{dt}$ vs. DO after the oxygen supply was reintroduced to the culture.

[21]In the "steady-state" method, the k_La is equal to the quotient of the OUR and driving force ($c^* - c_L$), where c_L is the oxygen concentration when there is oxygen mass transfer equilibrium, after the experiment to determine OUR has been performed.

method calculations gave results that were less predictable (Table 4.7, page 98). The most notable abberations are the low k_La value calculated at 500 rpm (101 hr^{-1}) and the relatively large value calculated at 400 rpm (219 hr^{-1}). Assuming the validity of the correlations in Equations 4.12 and 4.13, and disregarding the effect of change in broth composition over time, one would expect the k_La values at 400 and 800 rpm to differ by a factor of approximately 4.[22] However, the k_La values at 400 and 800 rpm only differ by a factor of approximately 1.5, implying a measurement error or insufficiency in the correlation. The fact that both k_La measurements at 800 rpm deviate from the mean value (302 hr^{-1}) by about 10% suggests that these values are relatively accurate. It is likely that because the steady-state dissolved oxygen concentration (92%) was near saturation during the 400 rpm k_La calculation, error in the DO probe measurement was amplified since it is the difference $(c^* - c_L)$ that is used in the calculation.

If the k_La values resulting from the dynamic method and off-gas calculations are at all relatable, the similarity of the magnitude of the dynamic method calculation at 500 rpm (286 hr^{-1}) and the off-gas calculations at 800 rpm (mean = 300 hr^{-1}) is unexpected considering the correlation predictions; one would expect the values at 800 rpm to be much higher than the value at 500 rpm, but instead they are quite similar. As mentioned before, a major source of variability not taken into account by the k_La correlation (Equation 4.12) is the effect of broth composition on k_La. In a *Bacillus subtilis* bioreactor culture agitated at 700 rpm (6-blade Rushton impeller) and 2 vvm, Richard, *et al.* observed the k_La increase to 154 hr^{-1} during exponential growth and decrease to 134 hr^{-1} as the maximum viscosity was reached. k_La was measured using the dynamic gas-out/gas-in method [47].

At the lowest agitation rate (300 rpm), the k_La values calculated using the "regression" (20 hr^{-1}) and "steady-state" (37 hr^{-1}) methods of analyzing dynamic method data were at their highest degree of agreement. These values are somewhat corroborated by a study done by Nielsen, *et al.* [41], who found that there was a sharp increase in k_La when the agitation was increased from 200 rpm to 400 rpm, the k_La

---

[22]From Equation 4.14, $\frac{k_L a \text{ at } 800 \text{ rpm}}{k_L a \text{ at } 400 \text{ rpm}} = \frac{800^{2.1}}{400^{2.1}} = 2^{2.1} \approx 4$

at 200 rpm was approximately 20 hr^{-1}, and the maximum k$_L$a attained was approximately 180 hr^{-1} (800 rpm, 2.0 vvm).

In general, the values of k$_L$a calculated using the "steady-state" dynamic method approach and the off-gas method were about 100 hr^{-1} higher than the maximum values obtained by Richard, *et al.* and Nielsen, *et al.*. Without a more complete set of data, though, it is difficult to say why this discrepancy exists, and whether it is due to measurement error or differences in the systems that were studied (i.e. both physical parameters of the bioreactor and the broth composition).

4.4 Carbon mass balance

The carbon mass balance technique is used to ensure that all process substrates and products are accounted for. Performing a mass balance serves as a check to make sure that the assumptions about a fermentation's reaction processes that govern the transfer of carbon from one form to another (e.g. glucose to dry cell weight or CO_2) are correct and that the measurement techniques for process characterization are sound. A carbon mass balance on the BL21(DE3)[pJY-1] fermentation from 10/4/04–10/6/04 is performed here. The entire fermentation is considered, from time zero (inoculation) to harvest. Spanning the entire fermentation with the mass balance maximizes the error in mass balance closure, thus amplifying any cumulative error in the calculation.

Table 4.8 lists the carbon-containing substrates added to the bioreactor from the time of inoculation to the end of the fermentation, and Table 4.9 lists the reaction products that were produced during the fermentation. It is assumed that citric acid is fully consumed and converted to "products" over the course of the reaction. High Performance Liquid Chromatrography (HPLC) analysis of sample supernatant (performed by Judy Yeh) did not detect acetate in any of the samples.

Carbon dioxide output was measured with both the Perkin Elmer mass spectrometer and the ARI laser gas analyzer. Both instruments agreed on the CO_2 composition to within 1% (volume %), so the measurements from the ARI were used for this mass

balance. The differential amount of CO_2 evolved between each successive set of time points (sample frequency was 1 per 15–16 seconds) was calculated for the entire time course of the fermentation using Equation 4.15 (note similarity to Equation 4.11 on page 96).

$$CO_2^{evolved} \text{ [g]} = \sum_{dt} \left(\frac{\%CO_2^{out}}{100\%} \right) (Q_{sparge})(V_{gas\ law})^{-1} \left(\frac{44.01 \text{ g}}{1 \text{ mol } CO_2} \right) (dt \text{ [min]}) \quad (4.15)$$

$CO_2^{evolved}$ is the total amount of carbon dioxide that exited the bioreactor over the time course of the fermentation in grams, $\%CO_2^{out}$ is the percent of CO_2 in the outlet gas (volume %), Q_{sparge} is the gas flow rate through the bioreactor in L/min,[23] $V_{gas\ law} = 24.5$ L/mol: the volume in liters occupied by 1 mol of ideal gas at room temperature and 1 atmosphere of pressure, and dt is the differential amount of time between off-gas samples (approximately 15–16 seconds). Using this methodology, the total amount of evolved CO_2 was calculated to be 689 g.

Thus, based on the calculations summarized in Tables 4.8 and 4.9, the carbon mass balance was closed to within 78%. This is highly incomplete closure of the mass balance. Possible sources of error include the reliance of the final DCW on a single OD_{600} measurement,[24] the estimation of the flow rates over time based on limited data from the fermentation log sheet, and the fact that cell debris and released cell contents due to cell lysis were not taken into account. In addition, normalization of the % CO_2 in the off-gas to adjust for the effect of water vapor (analogous to Equation 4.3 for O_2) was not done, which could have led to an underestimate of the CO_2 evolved, expecially at points when the filter was clogged and wet.

The remaining discussion (unrelated to the 10/4/04 BL21(DE3)[pJY-1] fermentation) pertains to mass balance in general, and considers interesting data on acetate production experienced by several course fermentations. During course 10.28 module

---

[23]Q_{sparge} varied with time because the exhaust filter would periodically get wet from condensation and clogged with cells, causing the flow rate through it to decrease below 5 lpm. The change in flow rate was taken into account based on the data recorded in the fermentation log sheet.

[24]The DCW measurement of 36.1 g/L for this time point would have resulted in even more incomplete closure.

Substrate	Substrate quantity (g)	% carbon	Carbon mass (g)
Inoculum dry cell weight	1.0^a	55%	0.55
Glucose (batch)	98^b	40%	39.2
Glucose (fed-batch)	1332^c	40%	533
Citric acid	8.5^d	38%	3.23
Total carbon in:			576 g

^{a}(4.73 OD_{600} units) $\left(0.42 \frac{\text{g/L DCW}}{OD_{600} \text{ unit}}\right)$ (0.5 L shake flask inoculum) = 1.0 g DCW

^{b}(9.8 $\frac{\text{g}}{\text{L}}$) (10 L bioreactor volume) = 98 g

^{c}(1903 g feed) $\left(1 \frac{\text{mL feed}}{\text{g feed}}\right) \left(0.7 \frac{\text{g glucose}}{\text{mL feed}}\right)$ = 1332 g glucose

^{d}(0.85 $\frac{\text{g}}{\text{L}}$) (10 L bioreactor volume) = 8.5 g

Table 4.8: Mass of carbon contained in fermentation substrates (cumulative from inoculation to harvest).

Product	Substrate quantity (g)	% carbon	Carbon mass (g)
Dry cell weight (in bioreactor)	447^a	55%	246
Dry cell weight (samples, foam)	27^b	55%	15
CO_2	689	27.3%	188
Acetate	0	40%	0
Total carbon out:			449 g

^{a}(96.7 OD_{600} units) $\left(0.42 \frac{\text{g/L DCW}}{OD_{600} \text{ unit}}\right)$ (11 L bioreactor volume) = 447 g DCW

^{b}(71 OD_{600} units) $\left(0.42 \frac{\text{g/L DCW}}{OD_{600} \text{ unit}}\right)$ (0.9 L bioreactor volume removed) = 27 g DCW; During induction approximately 0.9 L of cumulative bioreactor volume was either sampled, discarded, or lost due to foam. The approximate average OD_{600} value during this period was 71.

Table 4.9: Mass of carbon contained in fermentation products (cumulative from inoculation to harvest).

Induction time (hours)	OD_{600}	Acetate concentration (g/L)
0.0	26.1	0.31
0.8	27.0	0.78
1.5	26.2	0.68
2.3	23.4	1.17
3.0	29.7	1.42
4.5	18.3	1.89

Table 4.10: Acetate production during anaerobic fermentation.

2 (Fall 2004), group D2 saw relatively high levels of acetate production in their fermentation (*E. coli* BL21(DE3)[pLysS][pJHL]). When the culture recovered from its cooling phase (early-mid batch phase) and the fed-batch phase began, the acetate concentration was measured to be zero. For unknown reasons, the culture did not metabolize glucose as quickly as other cultures during the fed-batch phase and glucose accumulated to approximately 20 g/L by the beginning of induction phase ($OD_{600} =$ 26.1). Also near the beginning of the induction phase, the exhaust filter was flooded, clogging it and stopping air flow through the reactor during the remainder of the induction phase. As seen in Table 4.10, the culture stopped growing, and acetate steadily accumulated, implying that the culture entered anaerobic metabolism which resulted in acetate production.

In contrast to the acetate production of group D2, group A2's fermentation (*E. coli* BL21(DE3)[pJHL]) reached the highest cell density of all course fermentations ($OD_{600} = 80.4$) while appearing to remain glucose-limited throughout the fed-batch and induction phases. HPLC analysis of group A2's samples measured zero acetate production.

Chapter 5

Analysis of quantitative protein assays

When quantifying the production of a specific protein such as a collagen-like polymer (CLP), it is ideal to have an assay at one's disposal that is specific to that protein. To elucidate the motivation for a specific CLP protein assay, first consider an alternative method of CLP quantification. One approach could be to simply measure the concentration of total protein before and after a purification step designed to remove endogenous (non-CLP) protein, as done by Werten, *et al.* with the differential acetone method [67]. However, this method assumes that interfering components in the purified CLP mixture are minimal, that the measurement component due to CLP in the purified mixture overwhelms the component due to interfering factors, and that the purification step removes 100% of the endogenous protein. A more ideal system would use a specific assay (e.g. the collagenase assay) whose measurement magnitude is non-linearly amplified by CLP in the purified sample (relative to the remaining impurities), in order to increase the signal to noise ratio of the measurement and thus its accuracy.

This chapter analyzes experimental data from the collagenase and bicinchoninic acid (BCA) assays in the context of their chemical reaction mechanisms (described in Section 2.4.1). The purpose of this approach is to develop a case for their insufficiency (in their current form) for product quantification. Rather than presenting all of the

results of the fermentation sample assays here, only specific data that is useful for analyzing the assays is shown. In addition, an experimental approach is outlined in each relevant section for determining how the collagenase assay must be changed to obtain useful results. The protocols discussed in this chapter are based on Yin, *et al.* and the 10.28 manual (Fall 2004) [73, 68]. A list of the assay reagents is included in Appendix A.2.

The use of the collagenase assay in course 10.28 resulted in a wide range of experimental conditions that are useful in evaluating the assay. Unfortunately, because of time constraints on the course, students were not able to perform duplicate experiments on their samples, preventing each data set from carrying statistical information on measurement variation. However, the existing data was useful in designing experiments to improve the assay.

5.1   Efficacy of protein precipitation step

After cell debris are removed from the cell lysate, the samples contain soluble intracellular proteins released during sonication. Since the assay is based on the binding of ninhydrin to the exposed N- and C- termini of the proteins, the assay is more accurate when non-collagen-related peptides are removed from the sample. The assay calls for a 80–100°C incubation for 15-20 minutes. The extreme heat denatures many of the proteins which causes them to precipitate, much like a raw egg that becomes "hard" when it is boiled. Collagen-like polymers remain soluble at high temperatures.

There was a large amount of variation among 10.28 teams' results in the amount of protein that was precipitated during the heating step. After pelleting the precipitate by centrifugation, some groups observed large pellets that reached the 0.1 ml graduation on their 1.5 ml microfuge tubes. Other groups saw very faint pellets that were less than 1 mm in diameter. This variation occurred despite the fact that some of the teams were analyzing identical samples.[1] In addition, one group was asked to

---

[1]In the first module of the Fall 2004 offering of 10.28, students did not analyze their own samples because their fermentations did not reach high cell densities. Instead, they analyzed samples from the 10 L BL21(DE3)[pJY-1] fermentation (10/4/2004–10/6/2004).

perform the BCA total protein assay on fermentation samples before and after the protein precipitation step. They observed an intermediate to low level of precipitation after the heating step, but no significant difference in the BCA data due to the precipitation.

These results draw attention to the fact that there is no rigorously quantitative data associated with the reduction in non-CLP protein by the heating purification step. Yin, *et al.* mention that "more than 80% of non-specific [CLP assay] reactions can be eliminated" by thermal treatment [73]. Quantitative data on the amount of protein removed during the heating step would lead to a more thorough understanding of the degree of interference due to components that remain in solution throughout the assay (see Section 5.3 for a discussion of other interfering factors).

In addition, it would be useful to evaluate a second method for removing endogenous proteins from the sonication lysis supernatant. Werten, *et al.* describe a method of differential acetone precipitation, in which cold acetone is added to the sample to a concentration of 40% (v/v), endogenous proteins are pelleted, the acetone concentration is increased to 80% (v/v), and the resultant CLP precipitate is pelleted and washed once more with 80% (v/v) acetone. Werten, *et al.* achieved $> 96\%$ purity of artificial P4 gelatin using acetone method to treat the *Pichia Pastoris* fermentation supernatant. Like the protein precipitation method, this acetone-based technique relies on the relatively high solubility of gelatin due to its highly polar amino acid composition [67].

5.2 Length of incubation

The collagen assay protocol as described by Yin, *et al.* calls for a 5 hour incubation of the samples with collagenase. During course 10.28, different sets of standards and samples were incubated for varying lengths of time to determine whether reductions in the incubation time have an effect on the linearity of the standard curve. Representative standard curves from incubation times of 2h:30m and 4h:14m are plotted in Figure 5-1.
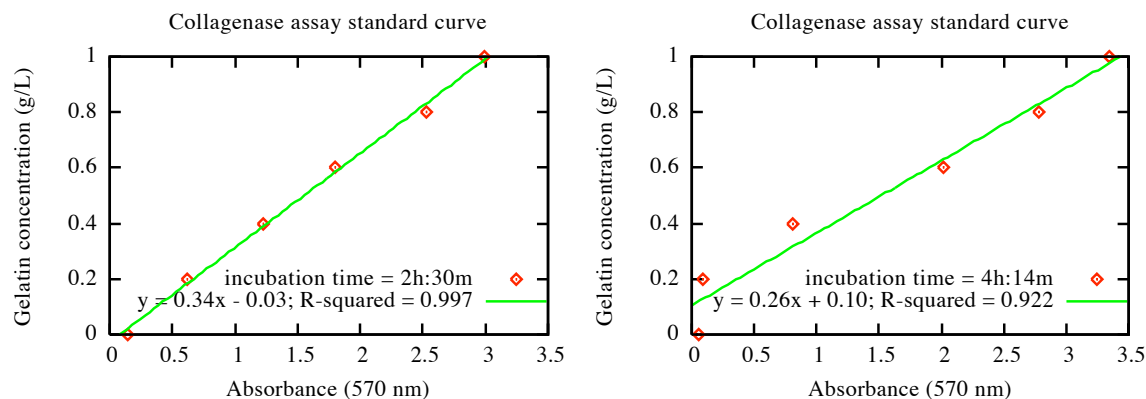
Figure 5-1: Collagenase assay gelatin standard curves from incubations of varying length.

The consistent linearity of the standard curves despite drastically different incubation times suggests that times much shorter than 5 hours are sufficient for achieving the same relative amounts of digested sample as in longer incubations, thus preserving the linearity of the standard curve. This data supports a reduction of the incubation time called for by the protocol, but should not be done until more accurate information about the concentration of CLP3.1 in the samples is found out.

Even if a shorter incubation time is sufficient for complete digestion of the standards, the effect of very high amounts of CLP3.1 on the accuracy of the assay must be considered. The CLP3.1 concentration in samples from a successful fermentation is likely to be much higher than that of the gelatin in the standards (maximum standard gelatin concentration = 1 g/L). Consequently, the standards are likely to be more fully digested than the fermentation samples in a limited-time incubation. The use of the linear relation between the fermentation and standard sample measurements in order to quantify the amount of CLP3.1 in the fermentation samples assumes that the amount of gelatin/CLP3.1 digested is proportional to the amount of substrate present in the sample.

Additional considerations related to very short incubation times ($\ll$ 2 hrs) are concerned with determining whether the enzyme–substrate reaction kinetics are in the saturated or linear ranges (as described by the Michaelis-Menten analysis). When stopping the digestion reaction early, the linearity assumption discussed above is

based on the assumption that the digestion kinetics are in the same range for all samples. At high CLP concentrations, this may not be the case; CLP concentrations may not allow the reaction to be in the linear range, causing the measured CLP values to saturate past a certain real CLP concentration for a given incubation time.

One way to circumvent the error-introducing possibility of these considerations is to allow enough time for samples at a maximum concentration to be fully digested. After several experiments with a particular strain, it will be possible to estimate the level of CLP in a fermentation sample (based on cell density and expected % expression level) and to consequently dilute it to within a specified range for the incubation time corresponding to a given set of incubation conditions. One way to experimentally determine the critical incubation time for a specific CLP concentration would be to make a set of identical samples of gelatin standard (concentration = highest possible standard concentration, most likely 1-2 g/L). The samples would be incubated for varying lengths of time, stopping the reaction by heating in boiling water. The point at which the colorimetric assay data stops increasing is the point at which the sample is fully digested.

5.3 Interference in the collagenase assay

To calculate the absorbance due to the added colorimetric binding of ninhydrin to the liberated peptides of digested CLP/gelatin, the difference between the OD_{570} values of the samples incubated with collagenase ("test samples") and the undigested samples incubated with the enzyme solvent ("control samples") was calculated (see Appendix A.2 for reagent details). However, the absolute magnitude of the absorbance of each sample directly determines the significance of this difference. For example, if the difference between two measurements remains constant as the magnitude of the measurements increases, the size of the difference between the measurements is likely to approach the same order of magnitude as the noise of the measurement, resulting in the difference between the two measurements becoming statistically insignificant.

The magnitude of the collagenase assay OD_{570} measurements varied greatly de-

Sample type	OD$_{570}$ range	Uncertainty
Gelatin standard	0.1–5.0	± 0.5
Washed *E. coli* samples (2 times with PBS)	10–20	± 1–1.5
Washed *E. coli* samples (1 time with PBS)	20–50	± 4–5
E. coli samples lysed in spent medium (no washing)	50–85	± 5–10

Table 5.1: Effect of collagenase assay sample composition on magnitude of OD$_{570}$ signal.

pending on the methods used to prepare the samples for the assay. General ranges and estimates of the variability in the measurements are shown in Table 5.1. As the degree of "cleanliness" of the sample increases due to either washing or to the baseline purity of the standards, the OD$_{570}$ of the samples as well as the magnitude of the measurement variability decrease.[2] The most likely cause of this data trend is the presence of interfering factors both in the medium and within the cells. Here, and interfering factor is a ninhydrin binding reaction with any non-CLP-derived species.

Before continuing the discussion further, the motivation for testing the collagenase assay without washing the cells with PBS must be discussed. In light of one of the objectives of this thesis—to create a lytic strain that is induced to lyse at the end of the fermentation—it would be ideal to be able to use the collagenase assay (or some variant, to be determined) to assay CLP in the lysate. Since the cells are lysed in the bioreactor, without the option of a washing step, any assay for CLP must be performed on the maximally impure broth mixture. Thus, if the pre-assay purification steps and the collagenase assay itself can be made to work with samples that start off with a composition similar to the broth, the collagenase assay would remain a cheap and rapid alternative to immunobased methods of assaying specific proteins.

In their collagenase assay study, Yin, *et al.* grew BL21(DE3)[pPT0296][3] in shake flasks containing MR medium. Based on the size of their absorbance increases, which ranged from 0 to 0.8, as well as on the data in Table 5.1, their absolute absorbance

---

[2]Because the relationship between absorbance and concentration is only linear at relatively low concentrations as described by Beer's law ($A = \log\left(\frac{I_0}{I}\right) = \ln\epsilon cl$), concentrated samples must be diluted to the linear range of the UV-VIS spectrophotometer (0.1–1) and the resultant raw absorbance must then be multiplied by the dilution factor to get the estimated absorbance of the undiluted sample.

[3]The precursor to the pJHL plasmid, thermally inducible expression of CLP3.1. In the creation of pJHL, the 6-his tag was added to the C-terminus of CLP3.1.

magnitudes were probably near 10. Based on experimental results obtained in this thesis by performing the assay on fermentation samples, reproducible, confidence-inspiring results have not yet been achieved. Even with fermentation samples washed twice with PBS, the magnitude and variability of the OD_{570} measurements are large. A thorough set of experiments must be performed to quantify the effects of interfering factors on the assay results.

Despite the fact that MR medium was used in the experiments of both Yin, *et al.* and of this thesis, the fact that Yin, *et al.* analyzed samples of *E. coli* grown in shake flasks while this thesis' samples came from high cell density fermentations is a major experimental difference. The conditions in shake flasks are left uncontrolled throughout the growth process, while fermentations receive base and feed solutions throughout the process which is deliberately prolonged. Ammonium hydroxide is the primary component of the based used to adjust pH; NH_4^+ interferes with the ninhydrin binding reaction. In addition, large amounts of foam accumulate in the bioreactor in the latter stages of a fermentation. Foam often arises because of decreased surface tension due to intracellular proteins released by dead, lysed cells.

Another major difference between shake flask and fermentation experiments is the initial cell density of the samples when they are resuspended just before sonication. Based on his results from the development of the collagenase assay, Jin Yin recommended that the cells from a 10 mL bioreactor sample at a cell density of about 15 g/L Dry Cell Weight be resuspended in 2.5 mL of Buffer A (20 mM Tris, 5 mM $CaCl_2$). The resultant mixture of proteins and CLP is likely to be much more concentrated than if the samples came from shake flasks. The high concentrations of intracellular components in the lysate is probably a major source of the high OD_{570} values. In future experiments with the collagenase assay, it will be necessary to find an optimal resuspended cell density to dilute the bioreactor samples to. It is probable that the cell density will be on the order of the density of a shake flask sample.

In addition, a range of assays on different medium formulations would be useful to determine the effect of various medium components on the assay results. For example, the assay can be performed on Buffer A alone (see above), MR medium, MR medium

and additional ammonium salt (e.g. extra $(NH_4)_2HPO_4$), and spent medium from a bioreactor and from a shake flask culture that was done in MR medium. It would also be useful to reproduce the shake flask results of Yin, *et al.* with one of the CLP3.1-producing strains as a baseline for the fermentation sample optimization experiments.

5.4 Use of gelatin and Bovine Serum Albumin (BSA) as assay standards

The results of both the BCA total protein assay and the collagenase assay depend on the standards used. The following sections analyze the validity of the use of BSA and gelatin as standards for the assays and make suggestions for improving the protocols and data interpretation.

5.4.1 Total protein assay standards

In their paper on the collagenase assay, Yin, *et al.* used the results of the BCA assay (Pierce) to maintain the concentration of total protein in the crude cell lysate of a BL21(DE3) shake flask culture constant at 1.20 g protein/L. The gelatin concentration was varied from 0 to 0.835 g/L to simulate levels of CLP3.1 expression from 0% to 41% of the total protein. While this is an effective approach to studying the interference due to non-CLP proteins, the paper did not suggest methods of measuring total protein in the crude cell lysate of a CLP expression system. Without accurate data on the amount of total protein in the lysate (including CLP), a meaningful CLP production level as a percent of total protein can not be measured. Based on the discussion of the BCA and collagenase mechanisms (Sections 2.4.1 and 2.4.2), the results of the assays are expected to vary depending on the relative amounts of the amino acid constituents of the proteins being measured. In optimized CLP expression systems such as those developed by Yin, *et al.* [72], CLP is expected to be expressed at levels near 40-50% of the total protein. Therefore, the results of an ideal total protein assay must not be skewed by differences in the CLP versus total protein composition.
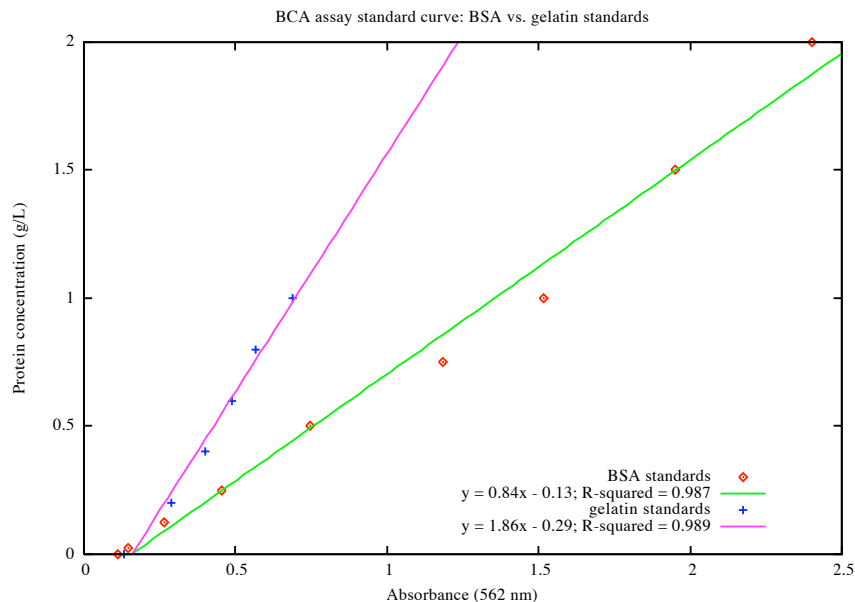
Figure 5-2: BCA assay standard curves using two different standards: Bovine Serum Albumin (BSA) and gelatin.

Stated in another way, the protein standard used for the BCA assay should give the same assay results as if CLP were used for the assay.[4]

Based on the data in Figure 5-2, it is clearly not the case that BSA and gelatin are equivalent standards for the BCA assay.[5] Relative to the results with BSA as a standard, equivalent gelatin concentrations result in significantly lower OD_{562} measurements. The slope of the curve from gelatin standards $(1.86 \; \frac{g/L}{OD_{562} \; unit})$ is approximately twice the slope of the BSA standard curve $(0.84 \; \frac{g/L}{OD_{562} \; unit})$. This implies that the amino acid residues of natural gelatin caused much less reduction of Cu^{2+} to Cu^+ than the residues of BSA. These results are consistent with the fact that gelatin has less residues that are especially reactive in the assay than BSA (cystine, tyrosine,

---

[4]It is recognized that CLP is not the only protein that reacts significantly differently than BSA in the BCA assay. The reasoning here is based on the assumption that the variation in protein concentration across samples from a given experiment is constant, and that the large proportion of the total protein comprised by BCA will dominate the effect of mis-measurement of any other specific proteins due to the use of BSA as a standard.

[5]Note that in Figure 5-2 the standard curves do not pass through the origin. This is because the data was not normalized for the non-zero absorbance caused by the reagents present in the blank sample (which are present in the same quantity in the other samples). Whether or not the standard curve is normalized does not matter, as long as the normalization is performed at some point in the data analysis process.

tryptophan) and CLP3.1 does not have any of these residues (see BSA and gelatin compositions in Section 2.4, and CLP3.1 composition in Section 3.1.1). Thus, if BSA was used as a standard for the measurement of a mixture of gelatin/CLP and common bacterial proteins, the result would be an underestimate of the total protein content and consequently an overestimate of the amount of CLP produced by the culture.

A method of grossly adjusting for the difference in reactivities of CLP and BSA, without calling for an additional assay method, would be to take the difference in reactivities into account in the total protein calculation. Rather than simply dividing the result of the collagenase assay (CLP concentration) by the result of the BCA assay (total protein concentration) to get percent CLP production, a second term must be added to the measured total protein concentration. Equation 5.1 is the new expression for the percent of total protein comprised by CLP.

$$CLP \text{ \% total protein} = \frac{[CLP]}{[BCA] + (1 - m)[CLP]} \times 100\% \qquad (5.1)$$

In Equation 5.1, $[CLP]$ is the concentration of CLP obtained from the collagenase assay, $[BCA]$ is the total protein concentration from the BCA assay (using BSA as a standard), and m is the normalization factor accounting for the difference in reactivities of gelatin and CLP, calculated by dividing the slope of the BCA assay standard curve from a BSA standard by the slope of the curve from a gelatin standard. The factor m is the effective fraction of CLP that was detected by the BCA assay using BSA as the standard. The value $(1 - m)$, then, is the fraction of CLP "undetected" by the BCA assay. Based on the information in Figure 5-2, $m \approx \frac{0.84}{1.86} = 0.45$, implying that the BCA assay with a BSA standard accounts for slightly less than half of the gelatin present in a sample.

Even though the complexity added by Equation 5.1 is likely to improve the accuracy of the assay, the efficacy of this equation is based on the assumption that the recombinant CLP and gelatin have similar reactivities when involved in the mechanism of the BCA assay. However, this may not be a valid assumption depending on the amino acid composition of the CLP. The ideal method of making a CLP standard

114

for the BCA assay in order to calculate an accurate value of m would be to use the CLP itself as a standard rather than gelatin. However, because it would be costly and time consuming to obtain a purified, quantified sample of the CLP, this method may not be feasible. Another approach to validating Equation 5.1 would be to use theory to predict the differences in reactivity between the CLP and gelatin. Without quantitative validation, though, this approach would render the results merely qualitative. The implications of a non-analogous gelatin standard are also discussed in Section 5.4.2.

Finally, the most readily-implementable method of validating Equation 5.1 would be to use a second total protein assay based on as different a mechanism as possible. For example the experiments described above to determine the m-factor could be performed with a Coomassie dye based protein assay and the results of the calculations based on Equation 5.1 compared. The Coomassie dye based assay is primarily based on the binding of basic amino acid residues (e.g. arginine, lysine, and histidine) to the dye to produce a colored complex detectable using colorimetric methods. The dye does not bind to short peptides or small proteins less than 3 kDa [45]. Thus, an additional set of experiments aimed at determining the level of CLP in a sample could be based on measuring the difference in dye binding between lysate samples both undigested and digested with collagenase. Fully digested CLP should result in zero contribution to the protein concentration values obtained by the Coomassie dye. Results from such experiments performed on both pre- and post-protein precipitation step samples would be useful.

5.4.2 Collagenase assay standards

The specificity of the particular type of collagenase used for the digestion of a particular collagen-like polymer must also be considered if other types of CLP are to be measured with the assay. Yin, *et al.* "confidently predict that this collagenase assay can be applied in the direct measurement of other types of CLP in *E. coli* cells." [73]. However, depending on the peptide sequence of the CLP, the activity of the collagenase used may vary. One potential way to change collagenase activity would be to

change the peptide sequence of the CLP substrate by substituting a charged amino acid in place of the uncharged residue that normally flanks the cleavage site (see mechanism details in Section 2.4.1). Decreased CLP relative to gelatin hydrolysis would result in an underestimate of the amount of CLP present in a sample if gelatin were used as the assay standard. In the current CLP3.1 system, this effect does not seem to be a problem: in their paper on the assay, Yin, *et al.* showed that the measurements of CLP3.1-his by the immunodetection assay closely matched the results of the collagenase assay, proving that the use of gelatin as a standard for CLP3.1-his digested with type III collagenase produces accurate results. But, the assay must be revalidated if significant changes to the CLP sequence are made.

Chapter 6

Design of lytic *E. coli* strain

This chapter describes the progress towards creating the lytic strain motivated by the research in Section 2.2.3. Based on experimental results, use of the λ-phage genes in the lytic strain design was rejected in favor of a system combining the T4 phage t-holin and the T7 lysozyme. The experimental methodology for creating the final strain is presented, followed by the primary considerations for optimizing a fermentation that uses the final lytic strain.

6.1 Experiments with λ-phage fragment S⁻RRz

Because of its ready availability from Jin Yin of MIT's BPEC and the demonstrated lytic efficacy with host strain JM109, the design of the dual plasmid lytic strain originally called for use of the λ phage fragment S⁻RRz.[1] However, experiments using the pUC18-S⁻RRz plasmid with host strain BL21(DE3) resulted in unsatisfactory lytic behavior which motivated the new strain design described in Section 6.2. A summary of the experiments follows.

---

[1]Refer to Section 2.3 for an explanation of the functions and mechanism of the S-gene (holin) and R-gene (lysozyme). The Rz-gene product is unknown; it may be an endopeptidase [74]. The "−" superscript in the λ fragment name means that the S-gene has been inactivated, so that it has no activity as a membrane-permeablizing agent.

6.1.1 Comparison of JM109 and BL21(DE3) lysis

Plasmids were purified from host strains JM109[pUC18-S$^-$RRz] and BL21(DE3)[pJHL], then the plasmids were reinserted into competent BL21(DE3) cells (see Section 3.1 for protocol details). The goal of the experiments with these strains was to determine the dependence of lytic behavior on both host strain and whether the inducer (IPTG) is added. The protocol was based on the course 10.28 lab manual (Fall 2003, Part A). Appendix A.3 contains details on lysis Buffer A and control Buffer B.

In one experiment, two shake flasks were inoculated with JM109[pUC18-S$^-$RRz]. IPTG (1 mM) was added to one of the flasks 5 hours before the end of the experiment, and both cultures were harvested at an OD_{600} of 4.87. Each harvest was split into two tubes, pelleted, and the supernatant was discarded. One pellet was resuspended in lysis Buffer A (50 mM Tris, 2mM EDTA) and the other was resuspended in control Buffer B (50 mM Tris). The OD_{600} of the tubes was measured every 20 minutes for 100 minutes. The experiment was repeated with strain BL21(DE3)[pUC18-S$^-$RRz] using BL21(DE3)[pJHL] as a control. The following OD_{600} values were reached before harvesting the cultures: 5.68 (BL21(DE3)[pUC18-S$^-$RRz], induced with IPTG), 5.53 (BL21(DE3)[pUC18-S$^-$RRz], not induced with IPTG), and 6.78 (BL21(DE3)[pJHL]). The results of these experiments are plotted in Figures 6-1 (JM109 cells) and 6-2 (BL21(DE3) cells). The fact that the mean OD_{600} values differ among samples is not important; the value of the figures lies in the OD_{600} trends they show. The variation in the mean values is due to variability in the initial cell densities, which were resuspended in the same buffer volumes (21 mL). This assay is based on the assumption that as the cells lyse, the resulting debris is unable to scatter as much 600 nm light as a suspension of completely intact cells.

The pUC18 host vector employs the *lac* operon (IPTG-inducible) to control expression of the λ-phage fragment. Jin Yin, *et al.* found that the addition of IPTG had no effect on the fraction of host JM109 cells that were lysed, implying that high basal levels of the R-gene lysozyme were produced regardless of the presence of the inducer [74]. This result was verified by the experimental results in Figure 6-1. It
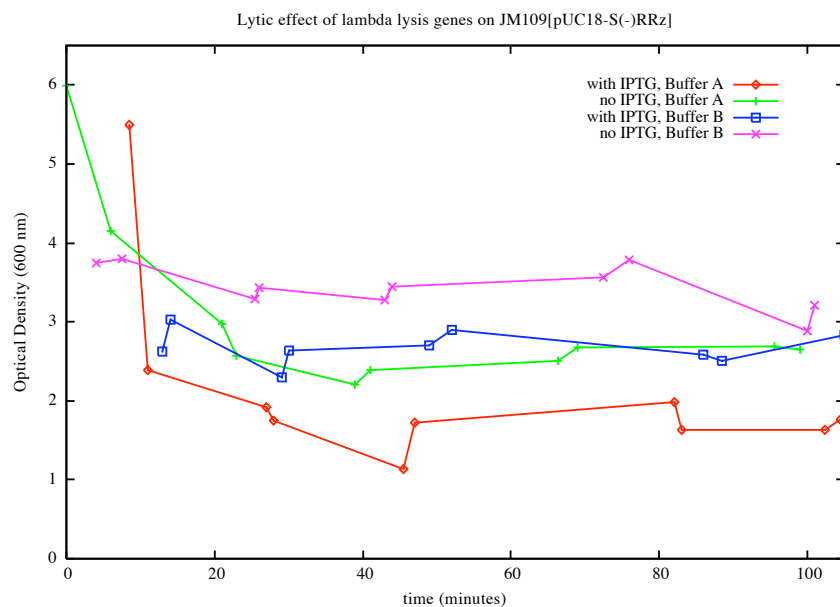
Figure 6-1: Efficacy of the λ phage genes for lysis of *E. coli* strain JM109[pUC18-S$^-$RRz].
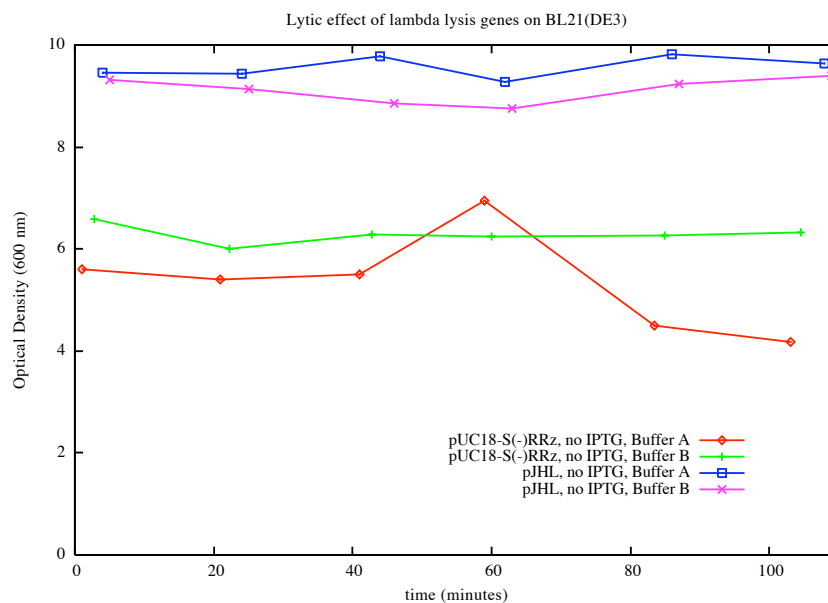

Figure 6-2: Efficacy of the λ phage genes for lysis of *E. coli* strain BL21(DE3)[pUC18-S$^-$RRz].

was because of these results demonstrating the lack of effect of IPTG on lysis that no IPTG was used in the experiments with strain BL21(DE3)[pUC18-S$^-$RRz] (Figure 6-2). The rapid drop in OD_{600} is consistent with data from Yin's study that showed lysis within 35 minutes of resuspension [74].

It is clear from Figure 6-2 that the BL21(DE3) cells are not lysed effectively; the rapid drop in OD_{600} seen with JM109 is not present, and the trends are similar, regardless of plasmid type. In addition, resuspensions of JM109[pUC18-S$^-$RRz] in lysis Buffer A were consistently more viscous than any of the other resuspensions, including the lysis Buffer A resuspensions of BL21(DE3)[pUC18-S$^-$RRz]. The viscosity is primarily due to DNA released from lysed cells. The data is also consistent with the increased difficulty of achieving BL21(DE3) lysis that Vasala, *et al.* experienced with their experiments: treatment of BL21(DE3)[pLysS] and BL21(DE3)[pLysE] with thymol, which was expected to make the inner membrane permeable to the T7 lysozyme, did not result in lysis [61].

6.1.2   Effect of temperature on *E. coli* autolytic tendency

Even though the autolytic capability of *E. coli* is not as potent as the action of bacteriophages [61], the fact that the expression of CLP3.1 with plasmid pJHL requires a temperature increase from 30°C to 41°C was cause for concern; temperature increase can be an inducer of autolysis [61]. In addition, it was considered possible that the presence of the R-gene product could augment the autolytic potency. To test this hypothesis, 4 flasks were inoculated with BL21(DE3)[pUC18-S$^-$RRz] and grown at 30°C for 11 hours to $OD_{600} \approx 3$. Since it was concluded from the results of Section 6.1.1 that the R-gene lysozyme was produced at significant basal levels, IPTG was not added to the flasks. The incubation temperature was then increased to 42°C and the OD_{600} was measured every hour for 3 hours. Over the 42°C period, the OD_{600} increased slightly to ≈ 4. No significant decrease in cell density was observed. To complete the experiment, the contents of two flasks were resuspended separately in Lysis Buffer A, and the contents of the other two flasks in Buffer B (buffer contents listed in Appendix A.3). As in the experiments in Section 6.1.1, the OD_{600}

of the resuspensions did not decrease dramatically, as in the resuspensions of strain JM109[pUC18-S⁻RRz]. It was concluded that the effect of the temperature increase required for use of plasmid pJHL did not contribute to significant autolysis.

6.1.3 Limited fermentation growth

Four fermentations were run in parallel with a set of four DasGip® FedBatchPro® bioreactors.[2] The working volume of MR medium in each bioreactor was 0.5 L, and the fermentations were implemented according to the protocols described in Section 3.2 and the course 10.28 lab manual (Fall 2004) [68]. Two vessels were inoculated with strain BL21(DE3)[pJY-2], and two with BL21(DE3)[pUC18-S⁻RRz]. All other fermentation parameters were held constant. (Because of the slower growth in the BL21(DE3)[pUC18-S⁻RRz] fermentations, glucose was consumed less quickly, and the start of the fed-batch phase came later.) Growth curves of the cultures are shown in Figure 6-3.

An advantage of using this miniaturized fermentation system is that duplicate fermentations can be carried out with a minimal amount of extra work. The fact that identical strains grew similarly implies that the results are reproducible. Thus, considering the growth curves shown in Figure 6-3, the presence of the pUC18-S⁻RRz plasmid impeded growth significantly more than the pJY-2 plasmid. As mentioned in Section 6.1.1, Yin, *et al.* speculated that the pUC18/*lac* system lends to high basal expression of the R-gene product [74]. It is likely that the high intracellular levels of the λ-phage R-lysozyme prevented the vigorous growth seen with strain BL21(DE3)[pJY-2]. This speculation is corroborated by the fact that the presence of the pLysE plasmid in *E. coli*, which is explicitly engineered to express high basal levels of the T7 phage lysozyme, causes significantly decreased growth rates [21].

---

[2]Conditions within these vessels were monitored and controlled in much the same way as those in the stirred tank bioreactors described in Section 3.2.1. The results of numerous fermentations performed in the lab with the DasGip® system have shown that high cell densities analogous to traditional STR fermentations can be consistently achieved. The experiments described in this section were performed on 7/25/04.
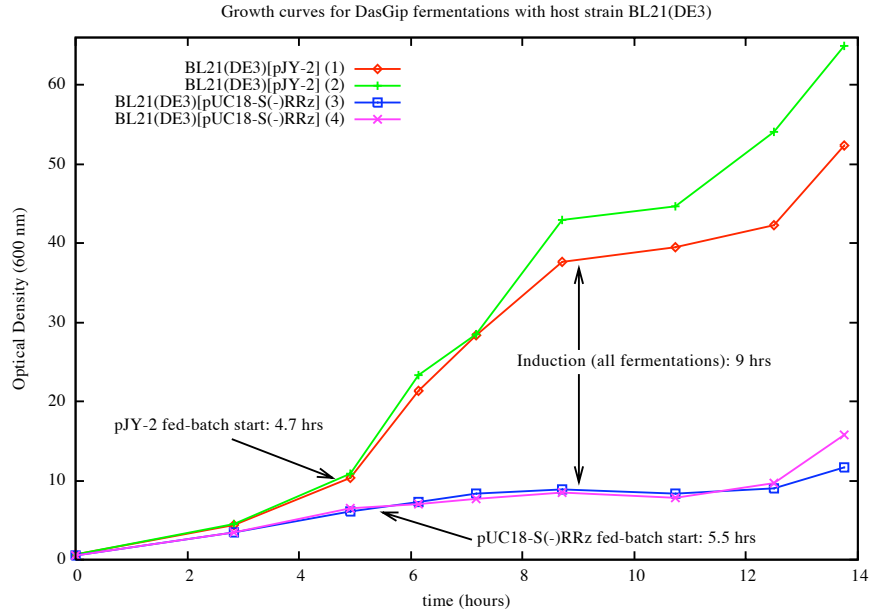
Figure 6-3: Comparison of growth of host strain BL21(DE3) containing either plasmid pUC18-S⁻RRz or pJY-2 in DasGip$^{\tiny\textregistered}$ fermentation vessels.

6.2 Lytic strain design

Since suboptimal results for the lysis of BL21(DE3) were obtained using the λ phage lysin, an alternate design was developed. In order to have maximum confidence that the lytic mechanism would be effective, previous work with successful results was considered. Also in the theme of making the lystic mechanism as reliable as possible, the new mechanism was designed to be more closely akin to the natural mechanism of phage-induced lysis, which requires the presence of both the holin and lysozyme enzymes. The study performed by Morita, *et al.* [38] was particularly effective and in line with the goals of this thesis, so the design of the new lytic strain was based on the constructions in their paper. Refer to Section 2.2.3 for an explanation of their study.

In order to maximize the likelihood of a successful high cell density process, the intermediate strains and experiments were designed with the following characteristics in mind:

High cell density capability The final strain must be capable of growing to high cell densities (30-40 g/L), expressing the product protein, and lysing efficiently.

122

The behavior of a strain at high cell density is directly effected by its metabolic state, which is a direct result of the fermentation design. If high cell densities are not achievable, the process will not be industrially viable because absolute production levels will be to low.

Lysis enzymes effective with host strain Efficient, rapid, and complete lysis must be attained to make this system viable. Efficacy of particular lytic enzymes may vary with host strain.

Stringent t-holin suppression The t-holin is extrememly toxic and even slight overexpression is detrimental to cell growth, since it causes the inner membrane to leak cellular constituents, adversely affecting cell health [61, 58]. Because of its harmful effect, expression of the t-holin causes an immediate halt in cell growth and lysis if a lysozyme is present [38, 58]. Thus, even low basal levels of t-holin are unacceptable, motivating the need for the stringent suppression of gene-t expression. Considering commercially available, stringent expression systems with mechanisms independent of the *lac* or thermally-inducible operons, the arabinose-inducible pBAD system (Invitrogen) is a good candidate because its induction mechanism is different from that of the *lac* (pJY-1, pJY-2) or cIts (pJHL) systems [25].

Independent induction of CLP3.1 and t-holin The induction of the two recombinant proteins must be entirely separate to make the process maximally optimizable. As was seen in Section 2.2.3, process design was fundamentally limited when the expression of the lytic and product proteins could not be independently timed. In light of this requirement, the pBAD system continues to be a good candidate.

The experiments outlined below for the design and validation of the lytic strain are based on an iterative approach. This process begins with a set of simple, "working" strains that are predictable and whose behavior is understood. Each successive experimental iteration adds functionality to the strains characterized in the previous step,

and the set of new strains is once again evaluated. This process is repeated until the final target strain, BL21(DE3)[pLysS-t][pJHL], is achieved. The motivation for this experimental approach, rather than creating and evaluating BL21(DE3)[pLysS-t][pJHL] at the outset, is that the effect of various design components on the overall process mechanism can be isolated and studied in a set of intermediate strains. Detailed data on individual design components will help in deciding how to alter the design, should it not work at all, or how to tweak the design or process to further optimize it.

The following subsections describe the successive experimental iterations and the experimental data expected from each iteration. The title of each subsection describes the design goal its experiments are intended to accomplish. In general, once a strain is created, a minimal amount of validation will be performed with shake flask cultures. If encouraging results are produced by a set of shake flask experiments, a bioreactor will be inoculated and the strain's high cell density behavior characterized.

6.2.1   Growth kinetics of multiple plasmid strains

The first step of the lytic strain design is to evaluate the behavior of single plasmid strains containing pJHL alone, pLysS alone, and both pJHL and pLysS. The copy number of the pLysS plasmid (*ori* p15A) is expected to be generally lower than that of the pJHL origin (*ori* pMB1) [5]. One possible effect on process growth kinetics is that higher plasmid copy number could cause a noticeable increase in metabolic load resulting in a decreased growth rate. However, the presence of a selection antibiotic for which a particular plasmid has resistance can increase copy number [5]. Nevertheless, a dual-plasmid system is expected to impose more metabolic load than either single-plasmid system. In shake flasks, Morita, *et al.* achieved specific growth rates of 0.62 and 0.53 [h^{-1}] with their control (no plasmids) and dual plasmid strains respectively [38]. These growth rates are as large as the maximum growth rate achieved in the oyxgen-enriched fermentation with *E. coli* BL21(DE3)[pJHL] on 3/24/04 (Section 4.1.2: $\mu = 0.52$ hr^{-1}), implying that multiple plasmids will not be a limiting factor for fermentation growth rate. However, the effect of additional plasmids on growth in the context of high cell density fermentations must be quantified: a set of

initial high cell density control experiments will serve as both a reference point for subsequent experiments as well as a relative study of the effect of plasmid type and numbers on growth kinetics.

Initial fermentations with the BL21(DE3) host containing the following plasmids have already been carried out in course 10.28 (Fall 2004): pJHL, pLysS + pJHL, pJY-2, and pLysS + pJY-2. Even though the pJY-2 CLP3.1 expression system will not be used in the final collagen-producing strain, the fermentation results are still useful for general process quantification and validation. See Chapter 4 for the results.

6.2.2 Choice of host strain

The experimental results of this thesis (Section 6.1) combined with Vasala's unsuccessful attempt to use the T7 lysozyme to lyse BL21(DE3) [61] imply that BL21(DE3) is "hardier" than other strains, which would make it an undesirable host strain for the process designed in this thesis. On the other hand, Morita, *et al.* reported efficient BL21(DE3) lysis using the T7 lysozyme and T4 t-holin [58] and found efficient means of chemical lysis using T7 as the lytic agent. Since Yin, *et al.* has already optimized CLP3.1 production with BL21(DE3) [72], BL21(DE3) would be ideal to use as the host for the final lytic strain designed in this paper (thus removing host strain as a source of experimental variability which could effect product production). However, in case BL21(DE3) does not meet the requirements of the host strain listed at the beginning of Section 6.2, JM109(DE3) should also be considered a potential host strain in the experiments described below. The selection of JM109(DE3) as another possible host strain is based on the lytic successes of Tanji, *et al.* and Morita, *et al.* using JM109(DE3) with the T4 lysis genes [58, 38].

6.2.3 Insertion of bacteriophage T4 t-holin gene into the pBAD system

The pBAD expression system was chosen because of its commercial availability (Invitrogen, Inc.), the stringent suppression capability of the *ara*BAD promoter, and its

induction mechanism which is entirely unrelated to the *lac* (pJY-1, pJY-2) or cIts (pJHL) systems. Many variants of the pBAD vector are offered by Invitrogen that add amino acid sequences to the C- and N-termini of the product protein to aid in protein detection and purification. Another variant of the vector insures directional insertion of the product gene, thereby obviating the screening step to determine insertion orientation. One consideration introduced by the extra DNA sequences that are not necessary for product function but are intended to be "helpful" is that the resulting extra peptide sequences added to the protein will interfere with the protein's active site. Thus, depending on the variant of pBAD vector that is chosen, the biological activity of the t-holin product may vary. A call to Invitrogen technical support (September 22, 2004) verified that there is no dependable way of predicting the effect of the additional peptide sequences, and that indeed, interference with product activity had been observed before.[3] In order to minimize the chance of a negative outcome, the pBAD TOPO® TA Expression Kit (Invitrogen, Part No. K4300-01) was chosen because it had the minimal amount of extra sequence relative to other pBAD systems.[4]

The Polymerase Chain Reaction (PCR) will be used to amplify the t-holin from T4 phage DNA (Nippon Gene Co./Wako Chemicals USA, Inc.; T4GT7 DNA, Part. No. 318-037971) using the primers listed in Table 6.1 (Operon, Inc.) and *Taq* DNA Polymerase. The amplified fragment will be 654 base pairs long. As described in the pBAD TOPO® TA manual, the primers must be designed to preserve the N-terminal reading frame; *Taq* DNA Polymerase adds a 3' A-overhang to the PCR product that binds to the sticky T-overhangs of the pBAD vector. The reverse primer includes the native stop codon to prevent the inclusion of the V5 epitope and six-histidine tag. The resultant plasmid, pBAD-t, will be 4780 base pairs long. Following the Invitrogen protocol, pBAD-t will be screened, amplified with the competent *E. coli*

---

[3]Note that this consideration is not important with the CLP3.1 product because there is no biological activity of the protein that is assayed.

[4]The t-gene product is 219 amino acids long. The pBAD TOPO® TA vector would add 14 amino acids to the N-terminus and 28 amino acids to the C-terminus of t-holin—making the resultant protein about 20% larger than the natural form of the t-holin [40, 24]. Retaining the native stop codon in the reverse primer will cause the 28 C-terminus amino acids to be left out.

Orientation	Promoter sequence	Length (bp)	% GC
Forward	5' GCAGCACCTAGAATATCATTTTCGCCCTCT 3'	30	47%
Reverse	5' TTATTTAGCCCTTCCTAATATTCTGGCCGC 3'	30	43%

Table 6.1: Primers for cloning the bacteriophage T4 t-holin gene.

provided with the kit, purified, and quantified (See Section 3.1).

Host strains BL21(DE3) and BL21(DE3)[pLysS], as well as JM109(DE3) and JM109(DE3)[pLysS] will then be transformed with pBAD-t.[5] In order to increase plasmid stability, the antibiotics for the selection marker included in each vector (pLysS-t: Chloramphenicol, pJHL: Kanamycin, pJY-1/2: Ampicillin) must be added to every step of the transformation and growth process for the remainder of the experiments described in this chapter. The shake flask experiments described in Section 6.1 will be performed on these strains to determine the degree of difference between the lytic suceptibility of the two potential host strains (see Section 6.2.2). If both strains lyse equally well, BL21(DE3) will be chosen as the host since CLP3.1 expression in this strain has been proven. If JM109(DE3) is lysed more efficiently, experiments must be performed to evaluate the expression of CLP3.1 in JM109(DE3) relative to BL21(DE3). Ideally, one of the two strains will be a clear choice for the future experiments. If effective lysis is not achieved with either strain, pLysS should be replaced with pLysE and the experiments repeated. Plasmid pLysE is identical to pLysS except that the T7 lysozyme gene is inserted opposite to its orientation in the pLysS plasmid, resulting in higher levels of basal expression (in pLysS, the lysozyme gene is inserted in the antisense direction relative to the *tet* promoter) [57].

6.2.4 Combining the pBAD-t and pLysS: creation of pLysS-t

Once the lysozyme/holin lytic mechanism has been proven viable, the constituents of the lytic system will be condensed into a single plasmid. This is motivated both by the increased simplicity of having the lytic capibility incorporated into one rather than two plasmids and by the predicted decrease in metabolic load as total amount

---

[5]Note that pLysS and pBAD-t are compatible because they contain the p15A and pMB1 origins of replication respectively.

of DNA decreases. In order to combine the plasmids, PCR will be used to copy the pBAD fragment containing araC and the *ara*BAD promoter + t-gene. The araC gene must be copied because its product is a repressor of the *ara*BAD system; strict regulation of the t-gene must be maintained. The fragment will then be inserted into the pLysS plasmid (or pLysE, if dictated by the results of Section 6.2.3). Note that the *ara*BAD–t-gene fragment must be inserted into pLysS because the pMB1 origin of replication of the pBAD vector is incompatible with the same origin in the CLP3.1 vector plasmids (see Section 3.1.1).

6.2.5 Addition of capability to produce CLP3.1

Once pLysS-t is validated as permissive of high cell density growth and an effective lytic system at high cell density, a plasmid containing the CLP3.1 gene will be inserted into the strain. This will be achieved by purchasing competent BL21(DE3) or JM109(DE3) cells (Promega) and transforming them with a mixture of pLysS-t and the CLP3.1 plasmid. Based on the current data, pJHL and pJY-1 are the most promising candidates for the final lytic strain. Yin, *et al.* found that CLP3.1 production levels are drastically reduced from 45% to 5% when pJY-2 (T7*lac*) is used with a pLysS system rather than pJY-1 (*lac*) or pJHL (cIts) [72]. Plasmid pJY-1 achieves expression levels of near 40% regardless of whether or not pLysS is present. However, in a fermentation experiment with BL21(DE3)[pJY-1], the strain's lag phase was about 8 hours longer than in fermentations using BL21(DE3)[pJHL] (see Section 4.1.2, Table 4.1). The significant increase in process length was undesirable, but may be a worthwhile tradeoff for the nearly four-fold increase in CLP3.1 expression using pJY-1 (40% of total protein) versus pJHL (12% of total protein). Additional experiments must be performed with plasmid pJY-1 to verify the long lag phase.

6.3 High cell density process design

This section builds on the strain design ideas of Section 6.2 to develop a methodology for the optimization of a high cell density fermentation that uses the product-

producing, lytic strain of this chapter. The issues addressed in this section focus on optimizing the growth and induction phases of the CLP3.1 and t-holin proteins.

6.3.1   Effect of growth rate on lytic efficiency

It is clear from a review of previous work that the growth rate at the time of t-holin induction greatly effects lytic efficiency; both the completeness (% of culture that is lysed) and the required length of a lytic phase (time needed to achieve lysis in a certain fraction of the culture) are affected. In shake flask cultures, Vasala described decreased lytic efficiency in late-logarithmic and stationary growth phases when bacteriophage LL-H *mur* (lysozyme) and *hol* (holin) genes were coexpressed [61]. Bläsi, *et al.* found that efficient lysis of *E. coli* due to expression of phage ϕX174 holin was only achieved at cell densities below 5×10^{10} cells/mL [6]. Also in shake flask cultures, Tanji, *et al.* and Morita, *et al.* found that in BL21(DE3)[pLysS] host strains, induction of the phage T4 t-holin in mid-log phase ($OD_{600} = 0.5$) resulted in complete lysis within 2 hours of induction, while late-log phase induction required 6 hours to reach lysis in 100% of the culture. Induction of the t-holin during the stationary phase resulted in partial lysis and larger cell debris than cultures lysed at higher growth rates. [58].

In order to optimize the fermentation process, then, it will be necessary to quantify the relationship between growth rate and lytic efficiency to a greater degree than has been done by the prior work described above. The bench-top scale stirred-tank bioreactors used thus far in this thesis combined with the continuous and precise measurement capability of an on-line biomass probe comprise an ideal system for pursuing this study. This experiment proposes the use of both on-line biomass probes available to the laboratory:[6] the BugEye® non-invasive biomass probe (BugLab, LLC.), and a fixed path length absorbance probe that must be submerged in the culture broth (Optek, Inc.). Once enough calibration data is collected to correlate the probe measurements to OD_{600} and DCW values, accounting for variation due to

---

[6]MIT Chemical Engineering Department Undergraduate Teaching Laboratory/Hamel Laboratory.

changes in agitation and aeration, the online data can be used to measure sensitive changes in the growth rate of the culture. Section 4.1.3 describes preliminary results that demonstrate the sensitivity of the BugEye$^{\circledR}$ instrument. Based on these results, the implications of the emperical power for culture growth characterization conferred by these on-line probes are very positive.

Correlating high resolution growth rate curves calculated from the on-line biomass probe data (analogous to the one in Figure 4-3(a) which used OD_{600} measurements) with the efficiency of lytic t-holin induction phases initiated at a variety of times in various fermentations will give a precise relationship between growth rate and lytic efficiency. Such a curve would present a significant improvement over the current qualitative knowledge of this correlation. In addition, the curve would be essential for analalyzing the tradeoff between the time of t-holin induction (earlier induction causes more rapid and complete lysis) and the length of the CLP3.1 induction phase (a longer phase allows more time to accumulate the product protein). Since the culture's growth rate diminishes as a fermentation progresses (Figure 4-3(a), page 74), it is conceivably worthwhile to boost the growth rate just before the t-holin induction (e.g. 30 minutes to 1 hour prior to induction). Assuming that oxygen transfer rate from the gas to liquid phases is the primary factor limiting growth rate at high cell densities, sharply increasing the level of oxygen enrichment in the sparged gas may be a viable way of prolonging the CLP3.1 induction phase while still being able to boost the growth rate into a range appropriate for efficient lysis.

In order to simplify the molecular biology of this experiment and to closely model the proven dual-plasmid systems of Tanji and Morita, *et al.*, *E. coli* BL21(DE3)[pLysS] may be directly transformed with pBAD-t (Section 6.2.3) rather than transforming *E. coli* BL21(DE3) with pLysS-t (Section 6.2.4). Plasmids pLysS and pLysS-t have compatible origins of replication. However, whether pLysS-t or pBAD-t is used for this study is not critical.

To measure the degree of culture lysis, a total protein assay as well as a turbidity assay should be used. For example, the 100% intracellular protein baseline can be established by sonicating a duplicate sample (i.e. an appropriate sonication protocol is

assumed to be a gold standard cell disruption technique) and comparing the results of a total protein assay (e.g. BCA assay) on the sonicated and induced-to-lyse duplicate samples. The OD_{600} assay is a sufficient turbidity assay, and was used by Tanji and Morita, *et al.* to quantify rates of lysis [58, 38].

6.3.2 Competition for gene expression machinery upon co-expression of CLP3.1 and t-holin products

A limitation on the universality of the conclusions that can be drawn from the experiments proposed in Section 6.3.1 is the fact that the t-holin is the only recombinant product produced in the experimental fermentations. However, once a plasmid vector with the CLP3.1 gene is introduced to the system, induction of both recombinant genes at the same time[7] may result in competition for available transcription and translation machinery, which could decrease the amount of t-holin produced during the lytic phase. Such a decrease in t-holin production would result in a decrease in lytic efficiency. Therefore, even though the results of the Section 6.3.1 fermentations will be essential for designing the lytic phase of the CLP3.1 production process, it may be the case that the lytic phase efficiency predictions (both length and degree of lysis) of these fermentations will be over-optimistic relative to a double-recombinant strain.

---

[7]Once the CLP3.1 gene has been induced with IPTG or a temperature change, it would be difficult if not impossible to turn off its expression.

Chapter 7

Conclusion

This thesis comprises several necessary steps towards the implementation of a novel, optimized collagen-like polymer (CLP3.1) production process. In addition, detailed quantification of the process and its use in three bioprocess engineering course offerings (MIT courses 10.28, Biological Engineering Laboratory, Fall 2003 and Fall 2004, and 10.26, Chemical Engineering Projects Laboratory, Spring 2004) allowed the development of a robust, schedule-friendly process for academic use.

7.1 Advancements made by this thesis

Robust fermentation protocol: Three CLP3.1-producing strains were characterized in high cell density fed-batch fermentations (*E. coli* BL21(DE3) with plasmids pJHL, pJY-1, and pJY-2). The quantitative results presented in this thesis are necessary for the design of predictable fermentations required by courses with strict schedules. With *E. coli* BL21(DE3)[pJY-1], a peak maximum growth rate of 0.53 hr^{-1} during mid batch phase was observed which steadily decreased to approximately 0.50 hr^{-1} at the end of induction. Combining the results of three fermentations, the factor for the conversion of glucose consumption to biomass increase was estimated to be $1.37 \pm 0.40 \; \frac{\text{(g glucose)/(L broth)}}{\text{OD}_{600} \; \text{unit}}$. The OD$_{600}$ to dry cell weight conversion factor was estimated to be $0.42 \; \frac{\text{(g DCW)/(L broth)}}{\text{OD}_{600} \; \text{unit}}$. Dynamic method calculation of the oxygen uptake rate (OUR) early in the

fermentation ($OD_{600} < 1$) resulted in values of the volumetric mass transfer co-efficient (k_La) of $37\,hr^{-1}$ (300 rpm), $145\,hr^{-1}$ (400 rpm), and $286\,hr^{-1}$ (500 rpm). OUR calculations based on off-gas analysis resulted in k_La values in the range 272–$332\,hr^{-1}$ at 800 rpm ($OD_{600} = 46.4$–81.0).

Analysis of collagenase assay with high cell density fermentation samples: The collagenase assay for recombinant collagen-like polymers developed by Yin, *et al.* [73] in shake flasks can not be used for analysis of fermentation samples without modification. Depending on the amount of washing of the *E. coli* with Phosphate Buffer Saline (PBS), the magnitude of the OD_{570} measurements in this colorimetric assay ranged from 4 to 10 times the measurement values for pure gelatin standards ($OD_{570} \approx 0.1$–5.0). Cell samples washed twice with PBS resulted in high absorbance values with significant variation (10–20 ± 1–1.5), and samples resuspended in spent medium resulted in OD_{570} values of 50–85 $\pm\ 5$–10. Calibration curves for the bicinchoninic acid assay differed by a factor of two depending on whether gelatin or Bovine Serum Albumin standards were used. A methodology for taking differences in amino acid composition and consequent protein reactivity in the total protein assay into account was presented.

Fermentation cooling phase: Motivated by the desire to prolong *E. coli* fermentations (which last 15–25 hours without interruption) so that students could be exposed to more parts of the same fermentation during course lab hours, fermentation cooling strategies were implemented in the Fall 2003 and Fall 2004 offerings of MIT course 10.28 (Biological Engineering Laboratory). Fermentations that were cooled to 10°C in early to mid batch phase ($OD_{600} = 2$–4) recovered from 36–40 hour cooling phases to reach high cell densities (up to $OD_{600} = 80$) during the induction phase. Fermentations that were cooled once in early to mid batch phase and again in early fed-batch phase ($OD_{600} = 14$–16) only saw up to a 50% increase in biomass after the second cooling phase, regardless of growth time. Consequently, the most successful cooled fermentations

were those that were cooled once during early batch phase.

Design of lytic strain: A two plasmid strain of *E. coli* capable of collagen-like polymer production followed by inducible lysis in the bioreactor was designed based on the results of Tanji and Morita, *et al.* [58, 38]. While CLP3.1 was the recombinant product of interest in this thesis, the design ideals of this lytic process is intended to be applied to other recombinant products. In the dual plasmid strain, one plasmid contains the CLP3.1 gene (either an IPTG-inducible pET vector [pJY-1] or a thermally-inducible cIts system [pJHL]), and the other plasmid contains the lytic genes. To create this lytic plasmid, the bacteriophage T4 t-gene (a holin enzyme) will be inserted into the pBAD vector (Invitrogen). The entire *araBAD* operon with t-gene and the *araC* gene will be cloned by PCR and inserted into plasmid pLysS, which is engineered for low basal expression of phage T7 lysozyme. The presence of the resultant plasmid in an *E. coli* host will result in the accumulation of T7 lysozyme, which will be permitted to reach (and consequently digest) the bacterial cell wall by the expression of the inner cell membrane-disrupting t-holin.

7.2   Future work

Modification of the collagenase assay for high cell density sample analysis: Because of the inapplicability of the current form of the collagenase assay to the reliable measurement of CLP3.1 in high cell density fermentation samples, this rapid and inexpensive assay will be modified. Based on the results in Section 5, a study will be implemented to decrease the signal to noise ratio observed with samples originating from high cell density fermentations.

Evaluation of lytic strains of *E. coli*: Section 6.2 describes an iterative approach to the construction of a dual-plasmid *E. coli* strain capable of induced lysis within the bioreactor. In order to assess the effect of each successive complication of the design as it converges to the final strain, the growth kinetics, product

production (CLP3.1) levels, and lytic behavior of each strain will be evaluated in shake flasks (increased experimental simplicity) before their introduction into bioreactors.

Effect of growth rate on lytic efficiency: Induction of recombinant lytic protein in shake flask cultures at low growth rates during the late logarithmic and stationary phases results in incomplete culture lysis (about 20% in stationary phase) and prolonged lytic periods (up to 6 hours). On the other hand, induction at the higher growth rates of mid log phase ($OD_{600} = 0.5$) results in rapid and complete lysis (100% lysis within 1 hour) [58, 38, 61]. In order to engineer an optimal, reproducible production process ending with efficient lysis, a study that provides detailed quantification of the correlation between growth rate and lytic efficiency must be performed. *E. coli* strain BL21(DE3)[pLysS][pBAD-t] will be constructed and its growth in bioreactors monitored by on-line biomass probes (manufactured by BugLab, LLC. and optek-Danulat, Inc.) which afford high time granularity for precise measurement of growth rate (preliminary results with the BugEye$^{®}$ sensor are discussed in Section 4.1.3). Detailed considerations for high cell density lytic process design are discussed in Section 6.3. The results of this project will be combined with the gross strain behaviors observed in the "Evaluation of lytic strains of *E. coli*" future work project to optimize the lytic phase of the final collagen-like polymer production process.

Appendix A

Media and reagents

This appendix contains detailed information on the composition of all media and reagents used in a fermentation and other assays. What follows is a near-exact reproduction of the lab manual written by Whittemore, *et al.* for course 10.28 [68]. Buffers for the lysis of strain JM109[pUC18-S⁻RRz] come from the lab manual for the Fall 2003 offering of 10.28 [62]. NOTE: All solutions are prepared with deionized water (dH$_2$O).

A.1 Fermentation

LB medium (/L) (for shake flask cultures)

NaCl	10 g
Tryptone	10 g
Yeast extract	5 g

- Tryptone varies considerably between manufacturer and even lot number. The tryptone used in this thesis was manufactured by BD Biosciences: Bacto$^{\text{TM}}$ Tryptone, pancreatic digest of caesin, part no. 211705.

- pH is adjusted to 7.5 by 1 M NaOH (M$_\text{w}$ = 40), added 1 ml at a time.

- 100 ml LB / 500ml flask, autoclaved at 121°C, 30 min (no dry time).

- Pre-pH adjustment, pH = 6.8–6.9; after adding about 1.2 mL of 1M NaOH,

pH increases to 7.5.

Trace metal solution (TE) (/L) (for MR medium)

$ZnSO_4 \cdot 7H_2O$	2.2 g
$MnCl_2 \cdot 4H_2O$	0.5 g
$CuSO_4 \cdot 5H_2O$	1.0 g
$Na_2MoO_4 \cdot 2H_2O$	0.1 g
$Na_2B_4O_7 \cdot 10H_2O$	0.02 g

- pH is adjusted to 2.0 by 1 M HCl ($M_w = 36.5$).

Ferrous sulfate solution (/L) (for MR medium)

$FeSO_4 \cdot 7H_2O$	10 g

- pH is adjusted to 2.0 by 1 M HCl.

Calcium chloride solution (/L) (for MR medium)

$CaCl_2 \cdot 2H_2O$	20 g

1X MR medium (/L) (for bioreactors)

KH_2PO_4	13.5 g
$(NH_4)_2HPO_4$	4.0 g
$MgSO_4 \cdot 7H_2O$	0.7 g
Citric acid	0.85 g
Ferrous sulfate solution	10 ml
Trace metal solution (TE)	10 ml
Calcium chloride solution	1 ml

- All salts are soluble after mixing and before pH adjustment (pH = 5.2)

- pH is adjusted after autoclaving; do not adjust pH initially. (Note that increasing pH causes precipitation of magnesium salts, e.g. $MgSO_4$ and $MgPO_4$, this is expected).

Antifoam (to prevent bioreactor foam)

pure polypropylene glycol, PPG-2000

- Autoclaved at 121°C, for 30 min.

Base solution (for pH adjustment during fermentation)

4 M (NH_4OH:NaOH:KOH = 2M:1M:1M)

- Sterilization is not necessary.

Phosphoric acid solution (for pH adjustment during fermentation)

1 M H_3PO_4

- Sterilized by filtration with 0.2 μm filter: in a biosafety cabinet, a vacuum pump is used to draw a vacuum in a sterile bottle with the filter screwed to the bottle mouth. The vacuum speeds up the filtration process significantly.

Glucose-Magnesium Sulfate Solution (/L) (fermentation carbon source)

Glucose   700 g

$MgSO_4 \cdot 7H_2O$ 14 g

- Sterilized with 0.2 μm vacuum filtration while still very warm.

Ampicillin (sodium salt) (selection marker, resistance conferred by pJY-1 and pJY-2 plasmids)

25 mg/ml in deionized water

- Plan to use about 2 mL of ampicillin stock per 1 L of final bioreactor volume and 200 μL for each 100 mL shake flask (final concentration is always 50 μg/mL).

- Sterilize by filtration with 0.2 μm filter. Store at -20°C. Use at 50 μg/ml.

Kanamycin (selection marker, resistance conferred by pJHL plasmid)

15 mg/ml in deionized water

- Plan to use about 2 mL of kanamycin stock per 1 L of final bioreactor volume and 200 μL for each 100 mL shake flask (final concentration is always 30 μg/mL).

- Sterilize by filtration with 0.2 μm filter. Store at -20°C. Use at 30 μg/ml.

Chloramphenicol (selection marker, resistance conferred by pLysS plasmid)

10 mg/ml in deionized water

- Plan to use about 2 mL of chloramphenicol stock per 1 L of final bioreactor volume and 200 μL for each 100 mL shake flask (final concentration is always 50 μg/mL).

- Sterilize by filtration with 0.2 μm filter. Store at 4°C. Use at 20 μg/ml.

- Chloramphenicol will precipitate at room temperature, so reheat in a boiling water bath before taking aliquots.

IPTG (isopropyl-β-D-thiogalactoside, *lac* operon inducing agent)

1 M stock solution in deionized water

- Prepare IPTG as a filter-sterilized (0.2 μm) 1 M stock solution (M_w = 238.3 g/mol). Will use 1 mL for every 1 L of working volume.

- Stored at -20°C until use.

- IPTG is used at 1 mM in the bioreactors.

100% Glycerol (for cell bank E. coli stock)

- Autoclaved at 121°C, for 30 min.

A.2 Collagenase assay

Collagenase assay reagents from collagenase assay protocol [73, 68]:

- **Gelatin**: Sigma (P-0959, 2 mg/ml in 0.9% NaCl containing 0.05% sodium azide).

- **Buffer A**: 20 mM Tris (adjust pH to 7.5 by HCl) with 5 mM $CaCl_2$.

- **Buffer B**: 0.5 M acetic acid, 0.1M citric acid, pH was adjusted to 5.0 by 1 M NaOH.

- **Ninhydrin (triketohydrindenhydrate) reagent**: 20 mg $SnCl_2$ and 200 mg ninhydrin are mixed with 10 ml methyl cellosolve (ethylene glycol monomethyl ether), followed by the addition of 10 ml Buffer B. This mixture is prepared freshly each time.

- **50% n-Propanol**: n-Propanol / water = 1 : 1 (v/v).

- **Collagenase type VII**: (Sigma Cat. NO. C 0773, Lot 77H8609. Collagenase digestion activity: 1600 units/mg solid. FALGPA hydrolysis activity: 10 units/mg solid. 1.0 mg solid). Dissolved into 0.5 ml Buffer A and 0.5 ml glycerol. Stored at -20°C.

A.3 JM109[pUC18-S$^-$RRz] lysis buffers

Adjust pH of all buffers to 8.0 with 1 M HCl.

Lysis Buffer A 50 mM Tris, 2 mM EDTA, pH 8.0

Control Buffer B 50 mM Tris, pH 8.0

TE Buffer 10 mM Tris, 1 mM EDTA, pH 8.0

Appendix B

Thesis chronology and relation to undergraduate courses

This Appendix provides a chronological list of the experiments carried out over the course of this thesis, organized by seasonal unit. Much of the work of this thesis was carried out in conjunction with courses offered by the MIT Chemical Engineering department, so descriptions of the logistical structure of the courses is also presented in order to give a complete context for the experiments.

B.1 Summer 2003

Work with *E. coli* strains was begun in order to implement the fermentation that would be used by the fermentation module of course 10.28 (Part B, Fall 2003).

B.2 Fall 2003

Course 10.28 (Biological Engineering Laboratory) was offered for the first time. The course was composed of a microbiology module that lasted two-thirds of the term, and a fermentation module that lasted for the final one-third of the term. There were 16 students in total. In the fermentation module, students were divided into four teams (four students per team). The fermentations carried out by each team are

Team Number	*E. coli* strain	Fermenter
1	BL21(DE3)[pJHL]	1 L BioFlo 3000
2	BL21(DE3)[pJHL]	1 L BioFlo 3000
3	BL21(DE3)[pJY-2]	5 L BioFlo 3000
4	BL21(DE3)[pJY-2]	10 L BioFlo 110

Table B.1: Descriptions of fermentations implemented in the Fall 2003 offering of course 10.28.

summarized in Table B.1.

B.3   Spring 2004

Course 10.26 (Independent Projects Laboratory) offered a fermentation engineering project to a group of four students: Bashira Chowdhury, Johanna Salazar-Lazaro, Terence Tai-Weng Sio, and Jessica Zaman. The project asked the students to evaluate the dynamic and off-gas methods for evaluating the volumetric oxygen mass transfer coefficient, $k_L a$. One successful *E. coli* fermentation was carried out with strain BL21(DE3)[pJHL] on 3/25/04.

B.4   Fall 2004

The Fall 2004 offering of 10.28 was divided into two modules: cell culture and fermentation, and had 39 students. Half of the students were in a module at a time, and they switched half way through the term. In each module, there were two sections: Tuesday/Thursday and Wednesday/Friday. The Tuesday/Thursday section had six teams with two students per team. The Wednesday/Friday section had four teams with two students per team. This structure allowed for a total of 20 fermentations to be implemented throughout the course (10 fermentations per module). Table B.2 describes the fermentations carried out in this course. Teams G–L participated in the module 1 fermentation module, and teams A–F participated in the module 2 fermentations. Team names are composed of a letter and section number (1 for Tuesday/Thursday, 2 for Wednesday/Friday). All teams used 0.9 L BioFlo 110 fermenters,

Strain	Team Numbers
BL21(DE3)[pJY-2]	G1–I1, G2–H2, A1–C1
BL21(DE3)[pLysS][pJY-2]	J1–L1, I2–J2, D1–F1
BL21(DE3)[pJHL]	A2–B2
BL21(DE3)[pLysS][pJHL]	C2–D2

Table B.2: Descriptions of fermentations implemented in the Fall 2004 offering of course 10.28.

except for teams K1, L1, E1, and F1, who used 1 L BioFlo 3000 units.

Appendix C

Microscope images of *E. coli*

Looking at samples through a microscope is often the most direct way of determining whether a culture is healthy, whether contamination is present, or whether a particular sample treatment is working. The following figures show images of *E. coli* at 400× magnification. *E. coli* are rod shaped and are approximately 1 μm in length.

Many individual bacteria are visible in Figure C-1 at 1000× magnification. The samples came from a BL21(DE3)[pJHL] fermentation's shake flask inoculum (3/25/04), and were at an OD_{600} of 3.88. The 1000× lens combination is an oil-immersion setup; for these images no oil was available, so the images are slightly fuzzy. The long filaments are daughter *E. coli* that did not quite complete division. The filaments only grow at the ends. Breaking a filament kills the entire chain because the intracellular spaces are continuous. Filaments form more often with recombinant strains than with non-recombinant strains [20, 63].

Pictures in Figure C-2 were taken over the course of the BL21(DE3)[pJHL] fermentation of 3/25/04. Samples were not diluted prior to putting them on the slide. *E. coli* can secrete polysaccharides under stressful conditions that cause them to stick together, forming conglomerations [20]. The healthy growth throughout this fermentation (discussion of fermentation results interspersed throughout Chapter 4) implies that the cells were not as "sick" as the images shown here may lead one to believe. It is possible that the cell densities in the samples were simply much too high for them to *not* conglomerate because of their proximity, and without the influence of
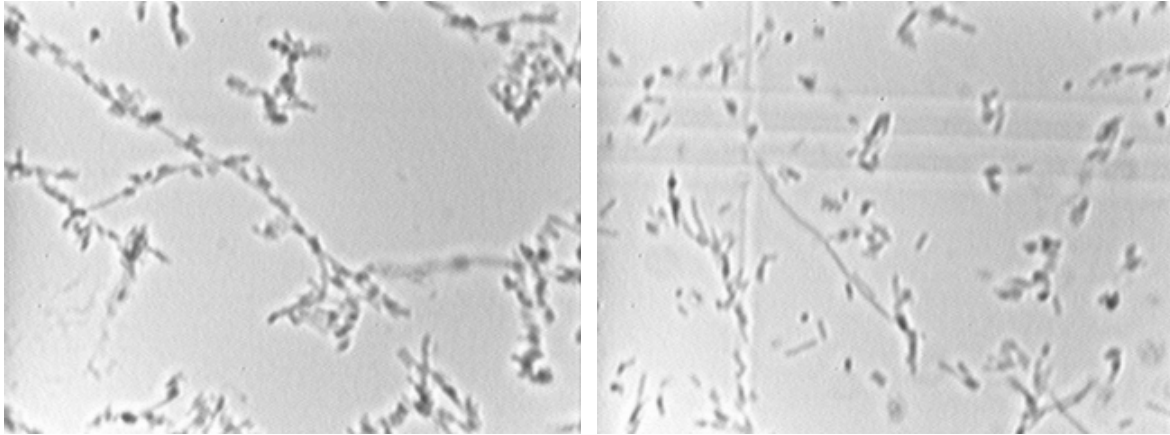
(a) OD$_{600}$ = 3.88, 1000× magnification   (b) OD$_{600}$ = 3.88, 1000× magnification

Figure C-1: *E. coli* from shake flask inoculum (BL21(DE3)[pJHL], 3/25/04).

the vigorously agitated environment within the bioreactor, they rapidly stuck together. Notice the increase in size of the conglomerations as OD$_{600}$ increases. The presence of filaments in the OD$_{600}$ = 3.45 sample but apparent absence in the higher cell density samples may be due to the reduced shear due to lower agitation (500 rpm for OD$_{600}$ = 3.45; 700 rpm for OD$_{600}$ = 51.3, 66.7).

Figure C-3 shows a sample before and after sonication with a Branson Ultrasonics sonifier (5 min; 70%, 1 sec duty cycle; 70% intensity). The samples originated from the course 10.28, module 1, group I1 fermentation with strain BL21(DE3)[pJY-2]. Many intact cells are visible in Figure C-3(a) while Figure C-3(b) shows primarily cell debris, indicating effective disruption.
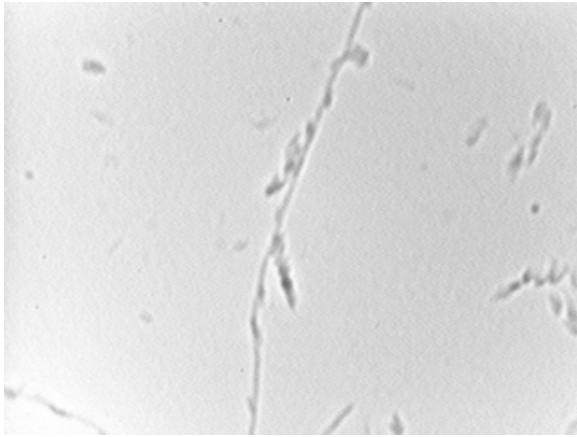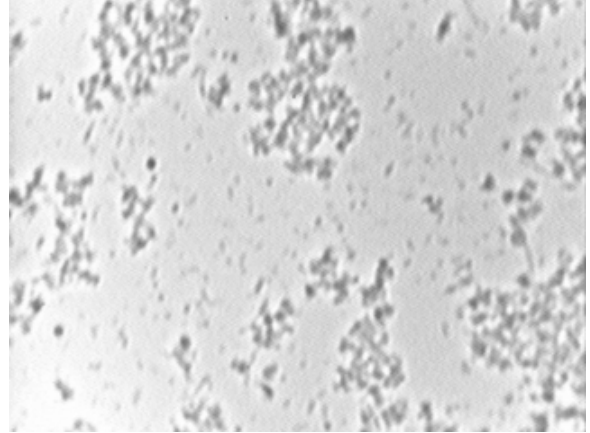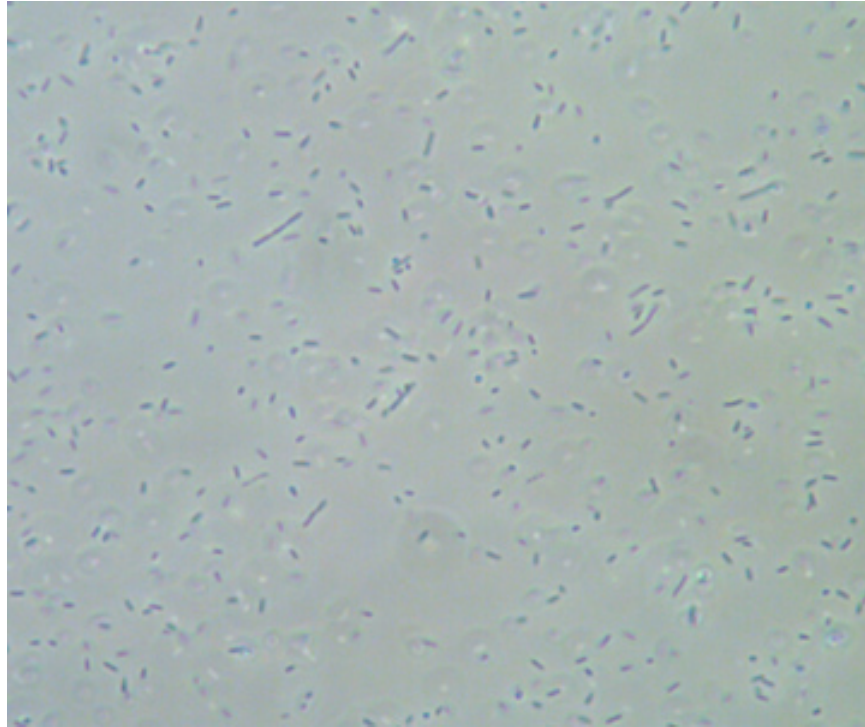
(a) $OD_{600} = 3.45$
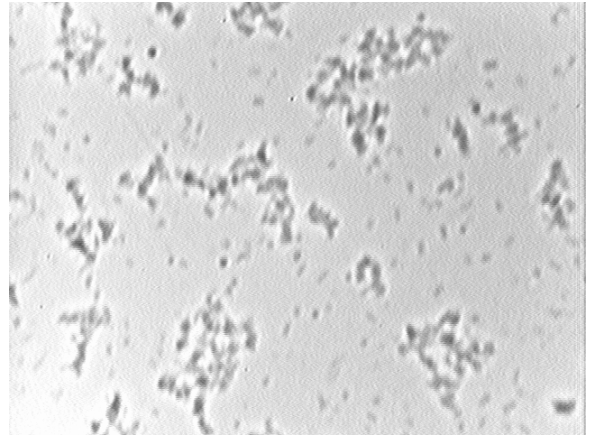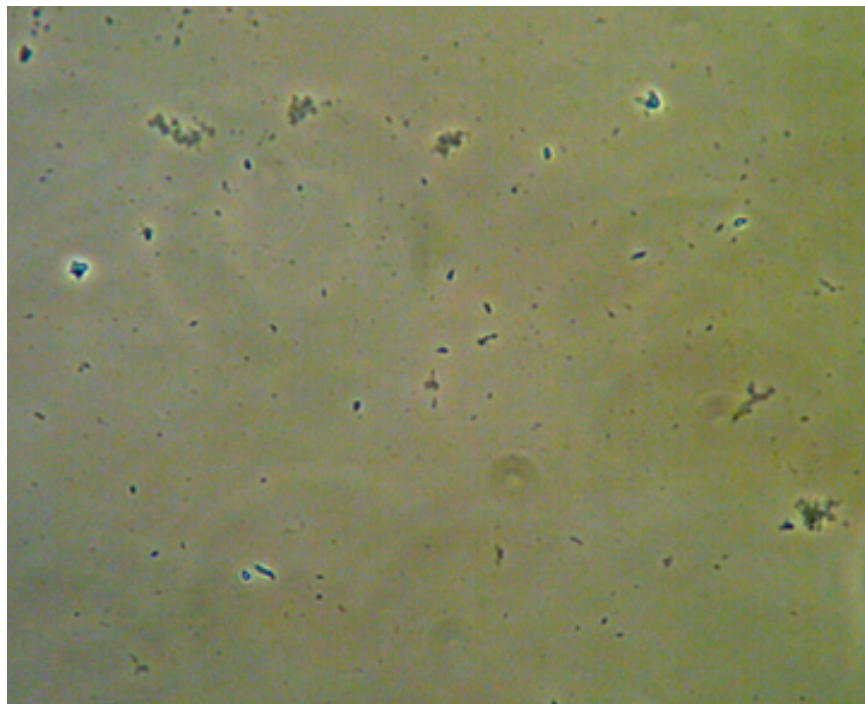
(b) $OD_{600} = 51.3$
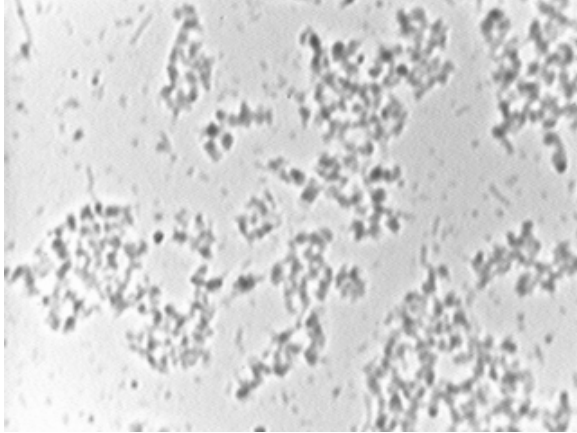
(c) $OD_{600} = 66.7$

(d) $OD_{600} = 66.7$

Figure C-2: *E. coli* at varying densities over the course of a fermentation.

(a) Before sonication, 400×, phase contrast.

(b) After sonication, 400×, phase contrast.

Figure C-3: *E. coli* before and after sonication.

Bibliography

[1] A. Amanullah, B. C. Buckland, and A. W. Nienow. Mixing in fermentation and cell culture industries. In E. L. Paul, V. A. Atiemo-Obeng, and S. M. Kresta, editors, *Handbook of industrial mixing*, pages 1071–1170. John Wiley and Sons, Inc., Published online, 2004.

[2] J. I. Auerbach and M. Rosenberg. Externalization of products of bacteria, January 1987. U.S. Patent No. 4,637,980.

[3] James E. Bailey and David F. Ollis. *Biochemical Engineering Fundamentals*. McGraw Hill, New York, New York, second edition, 1986.

[4] G. T. Benz. Optimize power consumption in aerobic fermenters. *CEP*, 99(5):32–35, May 2003.

[5] EMD Biosciences. *pRSF and pCDF Vectors*. EMD Biosciences, 2003. Plasmid vector manual.

[6] U. Bläsi, G. Halfmann, and W. Lubitz. Induction of autolysis or *escherichia coli* by ϕX174 gene E product. In C. Nombela, editor, *Microbial Cell Wall Synthesis and Autolysis*, pages 213–218. Elsevier Science Publishers, New York, 1984.

[7] J. Cappello and F. A. Ferrari. High molecular weight collagen-like protein polymers, June 1998. U.S. Patent No. 5,773,249.

[8] W. Chen, C. Graham, and R. B. Ciccarelli. Automated fed-batch fermentation with feed-back controls based on dissolved oxygen (DO) and pH for production of

DNA vaccines. *Journal of Industrial Microbiology and Biotechnology*, 18:43–48, 1997.

[9] X. Cheng, X. Zhang, J. W. Pflugrath, and F. W. Studier. The structure of bacteriophage T7 lysozyme, a zinc amidase and an inhibitor of t7 rna polymerase. *Proceedings of the National Academy of Science*, 91:4034–4038, 1994.

[10] W. Choe and A. P. J. Middelberg. Direct chemical extraction of a recombinant viral coat protein from *escherichia coli* at high cell density. *Biotechnology and Bioengineering*, 75:451–455, 2001.

[11] John Chon. Group Leader, BioEngineering, Genzyme Corporation; Conversation on October 20, 2004.

[12] S. N. Cohen, A. C. Y. Chang, H. W. Boyer, and R. B. Helling. Construction of biologically functional bacterial plasmids in vitro. *Proceedings of the National Academy of Sciences*, 70(11):3240–3244, November 1973.

[13] E. Doi, D. Shibata, and T. Matoba. Modified colorimetric ninhydrin methods for peptidase assay. *Analytical Biochemistry*, 118:173–184, 1981.

[14] Dunnill and Lilly. Protein extraction from microbial cells. In S. Tannenbaum and D. I. C. Wang, editors, *Single cell protein II*. MIT Press, Cambridge, Mass., 1975.

[15] Access Excellence. Biotechnology: 1977-present: the dawn of biotech. National Health Museum, 1999. Website: http://www.accessexcellence.org/RC/AB/BC/1977-Present.html.

[16] R. J. Falconer, B. K. O'Neill, and A. P. J. Middelberg. Chemical treatment of *Escherichia coli*: 1. Extraction of intracellular protein from uninduced cells. *Biotechnology and Bioengineering*, 53(5):453–458, 1997.

[17] R. J. Falconer, B. K. O'Neill, and A. P. J. Middelberg. Chemical treatment of *Escherichia coli*: II. Direct extraction of recombinant protein from cytoplasmic

inclusion bodies in intact cells. *Biotechnology and Bioengineering*, 57(4):381–386, 1998.

[18] R. J. Falconer, B. K. O'Neill, and A. P. J. Middelberg. Chemical treatment of *Escherichia coli*: 3. Selective extraction of a recombinant protein from cytoplasmic inclusion bodies in intact cells. *Biotechnology and Bioengineering*, 62(4):455–460, 1999.

[19] Ferrari, *et al.* Construction of synthetic DNA and its use in large polypeptide synthesis, September 1993. U.S. Patent No. 5243038.

[20] J.-F. Hamel. Private communication.

[21] Novagen, Inc. *pET System Manual*, July 2002.

[22] Novagen, Inc. *Veggie^{TM} Singles^{TM} Competent Cells.* Novagen, Inc., 2003. BL21(DE3) transformation manual.

[23] Qiagen, Inc. *QIAGEN$^{®}$ Plasmid Purification Handbook*, August 2003.

[24] Multiple cloning site data sheet for the pBAD TOPO$^{®}$ TA vector (Invitrogen, Inc.).

[25] Invitrogen. *Tightly controlled bacterial protein expression*, 2002. Brochure for pBAD vector expression systems.

[26] D. A. Jackson, R. H. Symons, and P. Berg. Biochemical method for inserting new genetic information into DNA of Simian Virus 40: circular SV40 DNA molecules containing λ-phage genes and the galactose operon of *E. coli*. *Proceedings of the National Academy of Sciences*, 69:2904–2909, 10 1972.

[27] E. Bech Jensen and S. Carlsen. Production of recombinant human growth hormone in *escherichia coli*: Expression of different precursors and physiological effects of glucose, acetate, and salts. *Biotechnology and Bioengineering*, 36(1):1–11, 1989.

[28] L. K. Joliffe, R. J. Doyle, and U. N. Streips. The energized membrane and cellular autolysis in *Bacillus subtilis. Cell*, 25(3):753–763, September 1981.

[29] T. Kajino, H. Takahashi, M. Hirai, and Y. Yamada. Efficient production of artificially designed gelatins with a *bacillus brevis* system. *Applied and Environmental Microbiology*, 66(1):304–309, 2000.

[30] T. S. Kim and T. H. Park. Optimization of bacteriophage λ q$^-$-containing recombinant *escherichia coli* fermentation process. *Bioprocess Engineering*, 23:187–190, 2000.

[31] Website: http://www.eurobio.fr/.

[32] D. V. Lim, R. J. Jackson, and C. M. Pull-VonGruenigen. Purification and assay of bacterial collagenases. *Journal of Microbological Methods*, 18:241–253, 1993.

[33] C.-S. Lin, B.-Y. Chen, T. H. Park, and H. C. Lim. Characterization of bacteriophage λ q$^-$ mutant for stable and efficient production of recombinant protein in *escherichia coli* system. *Biotechnology and Bioengineering*, 57:529–535, 1998.

[34] Mercè Dalmau Mallorquí. Optimization of the fed-batch fermentation of an antibody fragment-producing *Pichia pastoris* culture. Master's thesis, Institut Químic de Sarrià, September 2004.

[35] Stanley R. Maloy, Jr. John E. Cronan, and David Freifelder. *Microbial Genetics*. Jones and Bartlett, Sudbury, MA, second edition, 1994.

[36] S. L. Martin, B. Vrhovski, and A. S. Weiss. Total synthesis and expression in *escherichia coli* of a gene encoding human tropoelastin. *Gene*, 154:159–166, 1995.

[37] M. Moo-Young and Y. Chisti. Biochemical engineering in biotechnology. *Pure and Applied Chemistry*, 66(1):117–136, 1994.

[38] M. Morita, K. Asami, Y. Tanji, and H. Unno. Programmed *escherichia coli* cell lysis by expression of cloned T4 phage lysis genes. *Biotechnology Progress*, 17:573–576, 2001.

[39] National Institute of Standards and Technology. Website: http://webbook.nist.gov/chemistry/name-ser.html.

[40] Enterobacteria phage T4, complete genome summary (www.ncbi.nlm.nih.gov, Accession: NC_000866).

[41] D. R. Nielsen, A. J. Daugulis, and P. J. McLellan. A novel method of simulating oxygen mass transfer in two-phase partitioning bioreactors. *Biotechnology and Bioengineering*, 83(6):735–742, 2003.

[42] Massachusetts Institute of Technology. Phase equilibria: how to calculate oxygen solubility in cell culture medium. MIT course BE.360J/10.449J supplementary handout, Fall 2002.

[43] E. Parente, P. Piraino, M. Fidaleo, and M. Moresi. Overall volumetric oxygen transfer coefficient in an aerated bench-top stirred fermenter in aqueous dispersions of sodium alginate. *Biotechnology and Applied Biochemistry*, 40:133–143, 2004.

[44] S. H. Park and T. H. Park. Analysis of two-stage continuous operation of *escherichia coli* containing bacteriophage λ vector. *Bioprocess Engineering*, 23:557–563, 2000.

[45] Protein assays. From Pierce, Inc. technical reference website, www.piercenet.com.

[46] Ronald R. Rich. President, Atmosphere Recovery, Inc.; Conversation on August 25, 2004.

[47] A. Richard and A. Margaritis. Rheology, oxygen transfer, and molecular weight characteristics of poly(glutamic acid) fermentation by *bacillus subtilis*. *Biotechnology and bioengineering*, 82(3):299–305, 2003.

[48] K. Van't Riet. Mass transfer in fermentation. *Trends in Biotechnology*, 1(4):113, 1983.

[49] H. Rosen. A modified ninhydrin colorimetric analysis for amino acids. *Archives of Biochemistry and Biophysics*, 67:10–15, 1957.

[50] F. Ruggiero, H. Chanut, and A. Fichard. Production of recombinant collagen for biomedical devices. *Biopharm*, 13(5):32–36, May 2000.

[51] J. H. Rushton, E. W. Costich, and H. J. Everett. Power characteristics of mixing impellers. *Chemical Engineering Progress*, 46:467, 1950.

[52] A. M. Sandén, I. Prytz, and I. Tubulekas. Limiting factors in *escherichia coli* fed-batch production of recombinant proteins. *Biotechnology and Bioengineering*, 81:158–166, 2003.

[53] R. Sander. Henry's law constants. Website: http://www.mpch-mainz.mpg.de/ sander/res/henry.html.

[54] Information page on product number C0773, collagenase, from www.sigmaaldrich.com.

[55] P. K. Smith, R. I. Krohn, G. T. Hermanson, A. K. Mallia, F. H. Gartner, M. D. Provenzano, E. K. Fujimoto, N. M. Goeke, B. J. Olson, and D. C. Klenk. Measurement of protein using bicinchoninic acid. *Analytical Biochemistry*, 150(1):76–85, October 1985.

[56] D. L. Stevens. Necrotizing soft tissue infections. *Current Science Inc.*, 2:359–368, 2000.

[57] F. W. Studier. Use of bacteriophage T7 to improve an inducible T7 expression system. *Journal of Molecular Biology*, 219(1):37–44, 1991.

[58] Y. Tanji, K. Asami, X.-H. Xing, and H. Unno. Controlled expression of lysis genes encoded in T4 phage for the gentle disruption of *escherichia coli* cells. *Journal of Fermentation and Bioengineering*, 85(1):74–78, 1998.

[59] The BCA Assay. Website: http://www-class.unl.edu/biochem/protein_assay/bca_assay.htm.

[60] Website: www.ag.unr.edu/Reitz/bch303/001/BSAandGelatinComp.htm.

[61] Antti Vasala. *Characterization of Lactobacillus bacteriophage LL-H genes and proteins having biotechnological interest.* PhD dissertation, Oulu University, Finland, November 1998.

[62] A. Verghese, J.-F. Hamel, and J. Yin. *10.28 Biological Engineering Laboratory: Part A: Molecular biology.* M.I.T., Cambridge, MA, September-December 2003.

[63] D. I. C. Wang. Private communication.

[64] D.I.C. Wang. Lecture slide presented in M.I.T. summer course.

[65] J. D. Watson and F. H. C. Crick. Molecular structure of nucleic acids. *Nature*, 171:737–738, 1953.

[66] M. W. T. Werten, T. J. van den Bosch, R. D. Wind, H. Mooibroek, and Frits A. de Wolf. High-yield secretion of recombinant gelatins by *pichia pastoris. Yeast*, 15:1087–1096, 1999.

[67] M. W. T. Werten, W. H. Wisselink, T. J. Jansen van den Bosch, E. C. de Bruin, and F. A. de Wolf. Secreted production of a custom-designed, highly hydrophilic gelatin in *pichia pastoris. Protein Engineering*, 17(6):447–454, 2001.

[68] B. Whittemore, J.-F. Hamel, and J. Yin. *10.28 Biological Engineering Laboratory: Fermentation module lab manual.* M.I.T., Cambridge, MA, September-December 2004.

[69] K. J. Wiechelman, R. D. Braun, and J. D. Fitzpatrick. Investigation of the bicinchonic acid protein assay: Identification of the groups responsible for color formation. *Analytical Biochemistry*, 175(1):231–237, 1988.

[70] F. Wurm, L. Hunt, D. Hacker, M. Jordan, and M. de Jesus. Pausing of CHO cells. Talk given at the Cell Culture Engineering IX conference in Cáncun, México, March 2004.

[71] F. M. Wurm. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nature Biotechnology*, 22(11):1393–1398, November 2004.

[72] J. Yin, J.-H. Lin, W.-T. Li, and D.I.C. Wang. Evaluation of different promoters and host strains for the high-level expression of collagen-like polymer in *Escherichia coli. Journal of Biotechnology*, 100:181–191, 2003.

[73] J. Yin, L. Tomycz, G. Bonner, and D.I.C. Wang. A simple and rapid assay of collagen-like polymer in crude lysate from *Escherichia coli. Journal of Microbiological Methods*, 49:321–323, 2002.

[74] J. Yin, Y. Xu, H. Li, P. Zhou, and Z. Shen. Overexpressions of lambda phage lysis genes and biosynthetic genes of poly-β-hydroxybutyrate in recombinant *E. coli. Tsingua Science and Technology*, 4(2), 1999.

[75] X. Zhang and F. W. Studier. Mechanism of inhibition of bacteriophage T7 RNA polymerase by T7 lysozyme. *Journal of Molecular Biology*, 269:10–27, 1997.