Optimizing Recombinant Microbial Fermentation Processes

An Integrated Approach

Monika Cserjan-Puschmann, Reingard Grabherr, Gerald Striedner, Franz Clementschitsch, and Karl Bayer

A new strategy for controlling recombinant gene expression improves efficiency, maximizes host vector exploitation, reduces costs, improves product consistency, and accelerates product development. Continuous feeds of limited amounts of inducer proportional to biomass growth grant optimal control over the ratio between gene expression and host cell metabolism, providing stable, prolonged recombinant protein production.

Monika Cserjan-Puschmann *is a postdoctoral fellow,* ***Reingard Grabherr*** *is an associate professor,* ***Gerald Striedner*** *and* ***Franz Clementschitsch*** *are PhD students, and corresponding author* ***Karl Bayer*** *is a professor and group head of microbial fermentation at the University of Agricultural Sciences, Institute of Applied Microbiology, Muthgasse 18, A-1190 Vienna, Austria, +43.1.36006.6220, fax +43.1.369.7615, bayer@mail.boku.ac.at, www.boku.ac.at/iam/ferm and www.boku.ac.at/iam/acbt.*

Process development of bacterial host vector systems for recombinant DNA therapeutic products is expensive, time consuming, and stringently regulated. Those factors affect a product's time-to-market and the producing company's competitiveness. The average capitalized cost of developing a human pharmaceutical product is about $300 million. Development time for therapeutic biopharmaceuticals increased continuously over the past four decades, from less than four years in 1960 to more than 10 years since 1989 (1). Consequently, the average patent protection time is reduced.

Development is further complicated because each part of the process must be carried out in a defined, sequential order. Each development phase must meet certain criteria before the next phase can start. Procedures have to be defined early in the process because later steps, such as downstream processing and clinical trials, depend on the availability of sufficient recombinant protein of reproducible quality. Therefore, short-term set-up of an optimized production process is a key goal in bioprocess development.

Difficulties in Process Optimizing

Process optimization for biopharmaceutical production has traditionally focused on genetics-based solutions. Strong expression systems (such as the T7 system) have typically been used to increase yields of recombinant protein (2). However, such systems often override host cell metabolism causing cell damage or death (3). Because a host cell's protein synthesis can be maintained for only a short time, the production phase is shortened and maximum yield is unattainable.

Host cell metabolic capabilities. Process optimization needs to adapt recombinant gene expression rates to the host cell metabolic capabilities. If an equilibrium can be established between recombinant and cellular protein biosynthesis, then stable and prolonged recombinant protein production is possible. Extending recombinant protein formation is the key to increased product yields. Five periods of doubling are needed to achieve maximum specific content. The specific content numbers can be derived from the kinetics of the recombinant protein production (4). Host cells, by providing the protein synthesis machinery, building blocks, and energy for the production of recombinant proteins, are critical to the process. Those factors have not always been fully considered in the past.

An Integrated Approach

Applying an integrated system that includes aspects of host metabolism, process engineering, and recombinant vector modification can increase the efficiency (optimize) and accelerate the development process. The strategy we employ uses strong and regulated host vector systems, which permit continuous tuning of the recombinant gene expression rate in relation to the metabolic potential of the host cell. We show the feasibility and efficiency of this approach by producing a model system of recombinant human superoxide dismutase using an *E. coli* T7 host vector system.

The key to maximum cell factory exploitation is in determining the highest ratio between cellular and recombinant gene expression. Also important are the specific solutions that prevent runaway plasmid replication at high gene expression rates and the strategies that allow tuning of the recombinant gene expression. We established analytical methods to monitor the metabolic load exerted by recombinant gene expression to enable recombinant protein production within tolerable load limits set by the host cell metabolism.

Undesirable increases in plasmid copy numbers (PCNs) at high recombinant gene

expression rates are prevented by modifying the replication system. To circumvent complex and tedious analytical procedures for monitoring PCNs, we used a mathematical model. We established strategies to regulate recombinant gene expression by modulating the transcription rates with a continuous supply of limited amounts of inducer.

Material and Methods

Our **bacterial host strain** was *E. coli* (HMS174(DE3) from the Institute of Applied Microbiology collection). We also used plasmids (pET11ahSOD and pET11ahSODclone3) from the same collection (4–6). The coding region of recombinant human Cu/Zn superoxide dismutase (rhSOD) is equivalent to the nucleotide sequence of human complementary DNA (cDNA).

Growth conditions and media. We used a 20-L (12-L net volume) computer-controlled bioreactor (MBR Wetzikon, Switzerland) equipped with standard measurement and control units, including pH, pO_2, temperature, and rpm. We maintained pH control at setpoint 7.0 by adding potassium hydroxide (10% w/v), foam formation control by adding 5% antifoam suspension (Glanapon 2000, Bussetti, Vienna), and maintained cultivation temperature at 37 °C. Process control and visualization used Lookout software, version 3.8 (National Instruments Inc., www.ni.com). The fermentation methods we used and the nutrient media are described in the "Fermentation Media" sidebar.

Analytical Methods

We measured optical density (OD) at 600 nm and determined bacterial dry matter (BDM) by centrifuging a 10-mL cell suspension, which was then resuspended in distilled water, centrifuged again, dried at 105 °C for 24 hours, and reweighed. We determined guanosine tetraphosphate (ppGpp) using ion-pair high performance liquid chromatography (HPLC) to quantify the metabolic load on the host cells (7). We derived PCNs from the ratio of the amount of genomic DNA to that of plasmid DNA, which we quantified by capillary electrophoresis (8). We determined the content of the rhSOD by enzyme-linked immunosorbent assay (ELISA) (9,10).

Colony forming units (CFUs) were used to determine the number of plasmid-containing cells after 24 hours of cultivation on nutrient broth (NB) agar plates and on plates containing 100 μg/mL of ampicillin. To confirm an overproduction of rhSOD in the plasmid-carrying cells, we counted CFUs after induction by isopropylthiogalactoside (IPTG) (at 200 μg/ml) on plates containing ampicillin. We considered nonviable cells to have overexpressed rhSOD and, therefore, to have lost the ability to form colonies.

Monitoring the Metabolic Load

Metabolic load is a complex variable resulting from a number of dynamic biochemical reactions, which makes it

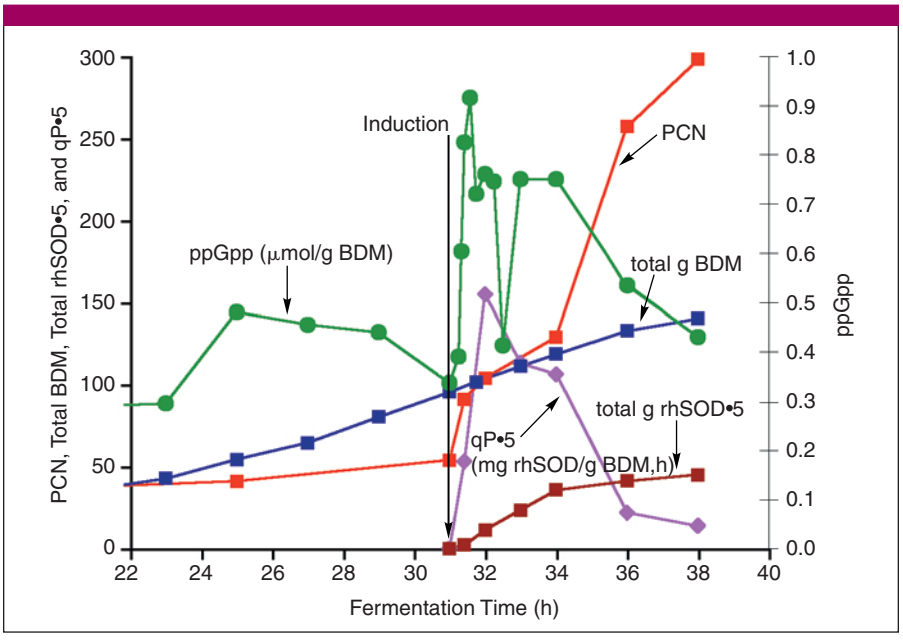

Figure 1. Standard fed-batch process for the production of recombinant human superoxide dismutase using *E. coli* (HMS174(DE3)pET11ahSOD) with a growth rate of $\mu = 0.1\ h^{-1}$ maintained by exponential feed

Fermentation Media

Fed-batch fermentation with an exponential substrate feed provides a growth rate of μ=0.1 h^{-1} during 3.5 or 4.0 doubling times. Batch volume was four liters at the start of the exponential phase (when cells enter the stationary phase); exponential feed was also started at this point. Feed control was achieved by increasing the pump speed according to the exponential growth algorithm, $x = x_0 e^{\mu t}$, with superimposed feedback control of weight loss in the substrate tank. A deep frozen (−80 °C) working cell bank (1 mL, OD_{600} = 1) was transferred aseptically to the bioreactor for inoculation.

The nutrient medium composition was semisynthetic (containing tryptone and yeast extract) for batches and entirely synthetic for fed-batch cultivation. Three and six grams per liter of potassium dihydrogenphosphate and di-potassium hydrogenphosphate•$3H_20$ were added to the medium. Other components were added in a ratio to the grams per bacterial dry matter (BDM) produced: 0.10 g tryptone (only used in batch), 0.05 g yeast extract (only used in batches), 0.25 g sodium citrate, 0.10 g magnesium sulfate•$7H_20$, 0.01 g calcium chloride•$2H_2O$, 50 μL trace element solution, 4.00 mg copper chloride•$2H_2O$, 3.20 mg zinc sulfate•$7H_2O$, 0.45 g ammonium sulfate, 0.37 g ammonium chloride, and 3.30 g glucose monohydrate. Trace element solution was prepared in 5 N HCl (gL^{-1}) and contained 40.0 $FeSO_4$•$7H_2O$, 10.0 $MnSO_4$•H_2O, 10.0 $AlCl_3$•$6H_2O$, 4.0 $CoCl_2$, 2.0 $ZnSO_4$•$7H_2O$, 2.0 Na_2MoO_2•$2H_2O$, 1.0 $CuCl_2$•$2H_2O$, and 0.50 H_3BO_3. Glucose and ammonium salts were autoclaved separately (7).

difficult to quantify. Overexpressed recombinant proteins exert a metabolic burden on host cells, triggering endogenous, nutrient-starvation response signals. Representative data can be derived for metabolic load, therefore, by using key variables of the *E. coli* cell's starvation, regulatory, and signal processing networks.

Lengeler's model organizes metabolic control hierarchically on three levels — operons, modulons, and stimulons (11). *Operons* are involved in the control of single enzymes and enzyme clusters, and *modulons* control the coordination of functional metabolic units. On both levels, the available information is very specific, which means operons and modulons are of only limited use in process optimization. However, *stimulons* are more complex regulatory entities: They are involved in cellular signal processing. Stimulons act on the highest level of metabolic regulation by coordinating widely distributed genes and global regulatory networks to redirect energy consumption and to confer resistance against adverse conditions.

The stringent response network, with its signal and effector molecule ppGpp, is involved in starvation and stress responses. Therefore ppGpp is a promising target molecule for monitoring metabolic load. To quantify ppGpp and relevant energy metabolism nucleotides (such as adenine and guanine), we combined an HPLC method with a rapid sampling procedure (to cope with the short half-life of the analytes) (7).

Monitoring aptitude. Figure 1 shows ppGpp aptitude for monitoring metabolic loads derived from recombinant gene expression. Fermentation of our model protein, rhSOD, in fed-batch fermentation of our *E.coli* strain and plasmid pET11ahSOD, shows the typical behavior of an expression system that is too strong. The expression rate (qP) exceeds the host cell's capacity, so the system collapses (see the path of the total biomass in Figure 1). As mentioned, maximum yields are unattainable because the biosynthesis cannot be maintained for sufficient time.

The concentration of ppGpp tightly follows the increase of recombinant protein qP after induction. Analogous experiments (data not shown) using a vector expression system with a less strong host (*E.coli* JM105 pKK hSOD) also resulted in increased ppGpp levels after induction, but to a significantly lower extent, corresponding to much lower protein production rates (7).

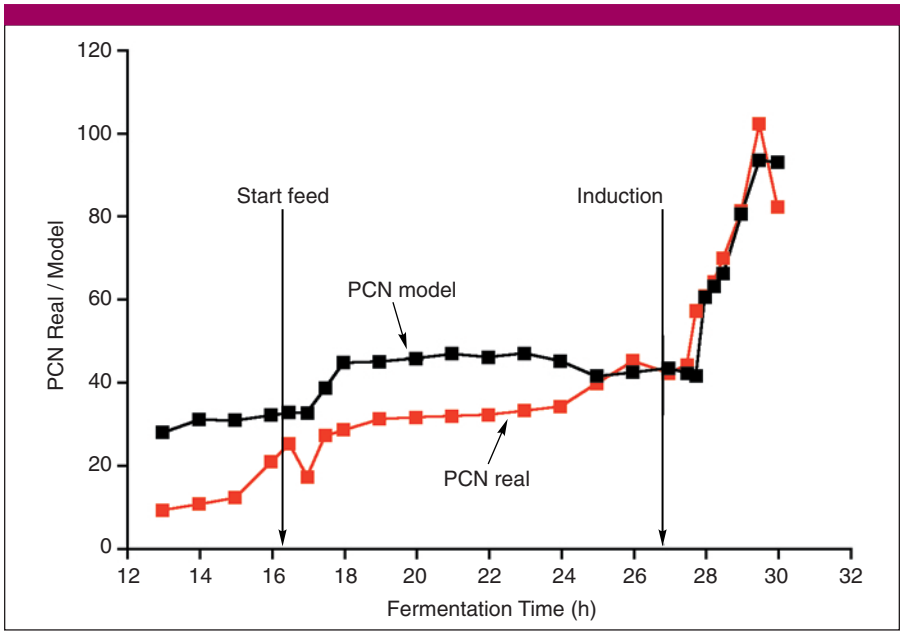

Figure 2. Modeling of plasmid copy number (PCN) from online data uses neural network simulations. Fed-batch fermentation of *E. coli* (HMS174(DE3)pET11ahSOD) has an exponential feed of μ=0.2 h^{-2}.

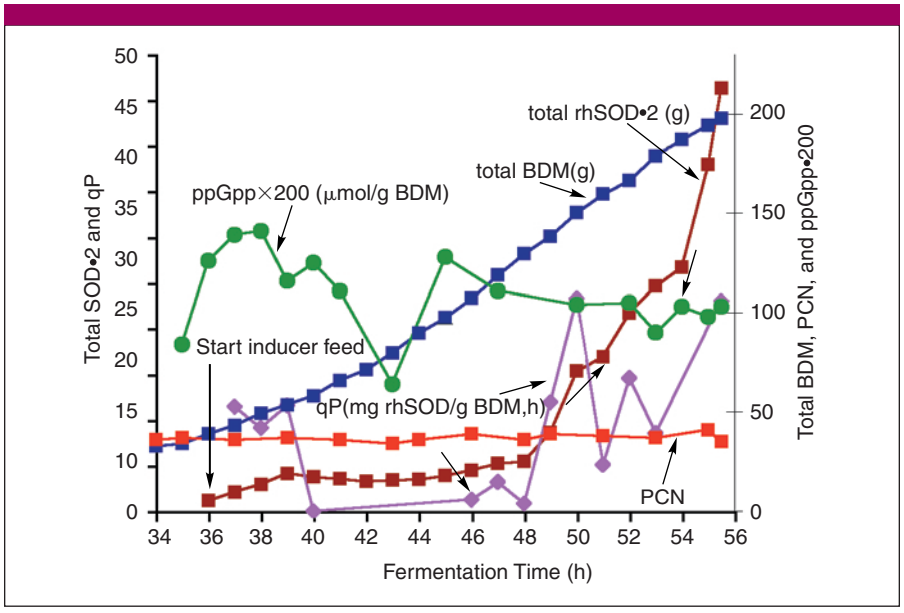

Figure 3. Fed-batch process for the production of recombinant human superoxide dismutase using modified *E. coli* (HMS174(DE3)pET11ahSODclone3); shows stabilized plasmid replication; exponential substrate feed with a growth rate of 0.1 h^{-1}; inducer feed (μmol IPTG/g BDM) was increased in three steps to assess stability of the plasmid copy number: 0.6–1.17 at 35–43 h, 1.17–3.26 at 43–48 h, and 3.26–19.6 at 48–55 h.

Inducer Improvements

Continuous addition of inducer significantly improves bioprocess operations by:

- fine tuning the recombinant gene expression rate,
- maintaining the productivity of the host vector system over a prolonged period of time, and
- extending the production of recombinant proteins.

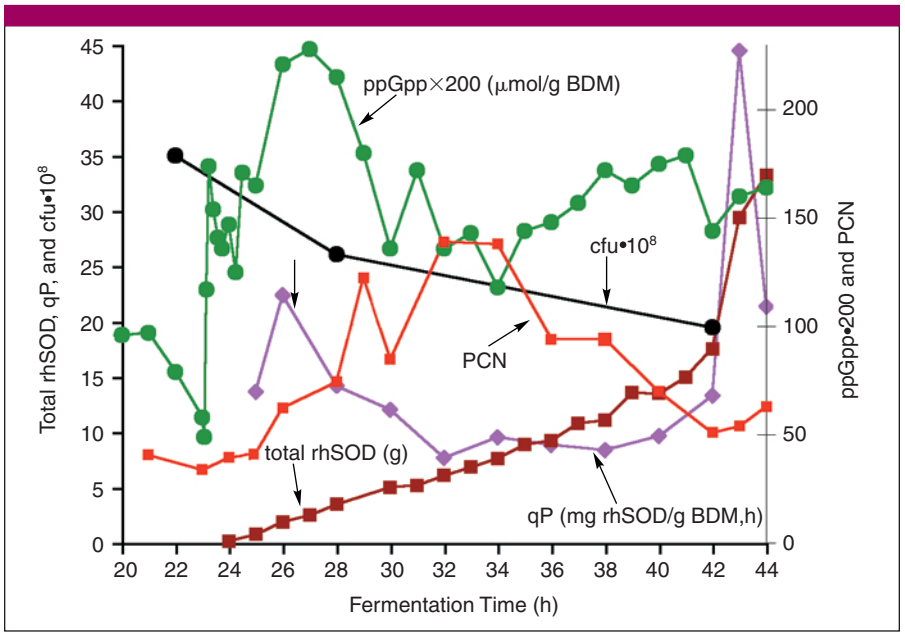

Figure 4. Fed-batch process for the production of recombinant human superoxide dismutase using modified *E. coli* (HMS174(DE3)pET11ahSOD) with a continuous feed of inducer (IPTG) to control transcription rate

However, even significant increases of ppGpp, assumed to trigger stringent response mechanisms, could not protect the host cell metabolism from being overstrained and finally collapsing. That could be a result of very high transcription rates for the target mRNA by the T7 polymerase, which curtails the synthesis of mRNAs for the translation of cellular proteins (7). The behavior suggests that recombinant gene expression rates must be controlled during transcription.

Plasmid Replication

Figure 1 shows the dramatic PCN increase at high expression rates. Two antisense RNAs, RNA II (which acts as a primer, initiating duplication of DNA strands) and RNA I (which binds to RNA II before priming, inhibiting duplication), effectively limit plasmid replication (5,12).

Uncharged transfer RNAs (tRNAs) diminish the inhibitory effect of RNA I because of their extensive homology sequence to RNA I and RNA II, which interferes with plasmid replication control. At high recombinant gene expression rates, uncharged tRNAs are elevated through increased amino acid and by shifted demand for individual amino acids. High tRNA levels increase plasmid replication.

The resulting increased gene dosage strongly stimulates recombinant gene expression because T7 polymerase binds to its own promoters, enhancing transcription rates. Higher transcription rates trigger a chain reaction that leads to metabolic overburden for the host cell.

Model for PCN estimates. Process optimization always approaches the limits of the individual host vector system by operating fermentation processes at high qP, and the plasmid replication rate always increases under these parameters. Because plasmid replication is intrinsically coupled to the level of host cell metabolites, our process operation focus was on gaining access to plasmid replication regulators. The mathematical model we developed for monitoring and estimating PCNs is based on a neural network. We used fermentation data sets obtained online (CO_2 production rate, alkaline–base consumption, and OD) and offline (PCN data sets of five identical fermentations) to train the neural network simulations. To create the self-organizing maps (SOMs), we used Viscovery SOMine (Eudaptics Software, www.eudaptics.com) (13).

Figure 2 shows the aptitude of the model by comparing real and modeled PCN data for a standard fed-batch fermentation process. The modeled PCN data even show a significant increase of PCN after induction. Nevertheless, we used a published analytical method for determining PCNs for the experiments shown in Figures 3, 4, and 5 (8). Although modeling accelerated and improved data acquisition, we were unable to abolish the undesirable PCN increase at high recombinant gene expression rates using process control measures. Therefore, to prevent the harnessing effect from excessive plasmid replication, we targeted our modifications on the origin of the replication.

Strategies for PCN stabilization were derived from the plasmid replication scheme of ColE1. Because overreplication is mainly triggered by the binding of uncharged tRNAs to either RNA I or RNA II, our strategies focused on restoring the replication control function of replication control. Such strategies must be designed so that the homology is maintained between the two antisense RNAs but destroyed between the antisense RNAs and the tRNAs. So we exchanged the sequence of seven

Table 1. Summary of key results of individual steps of process optimization

Type of Process Optimization Step	Total Yield of Recombinant Protein in Fed-Batch Process[a]	Improvement Compared with Standard Process[b]
Standard process[b]	8.99	NA
Fed-batch process 1[c]	23.10	2.5
Fed-batch process 2[d]	33.20	3.6
Fed-batch process 3[e]	18.97	2.1

[a] Grams per 10 liters of broth volume
[b] Fed-batch with pulsed induction for recombinant gene expression
[c] Using gene construct with stabilized plasmid (pET11ahSODclone3) replication and continuous addition of increasing amounts of inducer (IPTG)
[d] With continuous feed of optimized amounts of inducer (0.9 μmol/g BDM IPTG)
[e] With continuous feed using lactose as the inducer

nucleotides in loop two with its complement without reversing the sequence (5,6).

The modified *E. coli* gene construct (HMS174(DE3)pET11ahSODclone3) was grown in fed-batch culture at a growth rate of $0.1h^{-1}$ (Figure 3). By modifying the origin of replication, PCN was kept constant at a low level during the entire fermentation process despite varying recombinant gene expression rates. With the lower and stable gene dosage, the production phase could be significantly prolonged to about four doubling times. Four doublings increases the yield about 2.5 times compared with standard processes (Table 1). The significant metabolic load reduction (shown in the ppGpp data) highlights the efficiency of this approach. Stabilizing plasmid replication at a low level improves the researcher's control over recombinant gene expression rate, particularly in tuning the transcription rate.

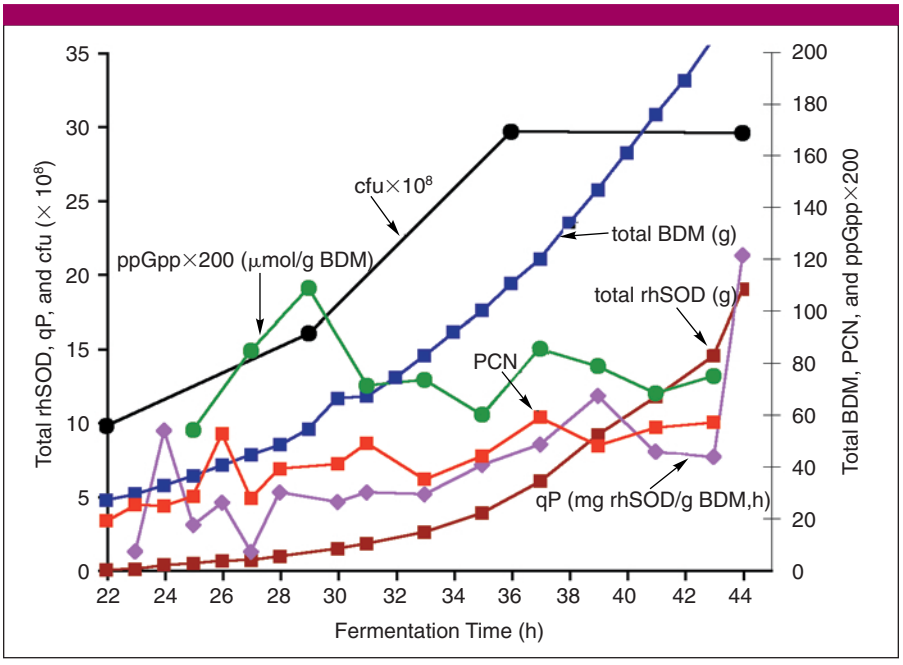

Figure 5. Fed-batch process for the production of recombinant human superoxide dismutase using modified *E. coli* (HMS174(DE3)pET11ahSOD) with a continuous feed of inducer (lactose) to control transcription rate

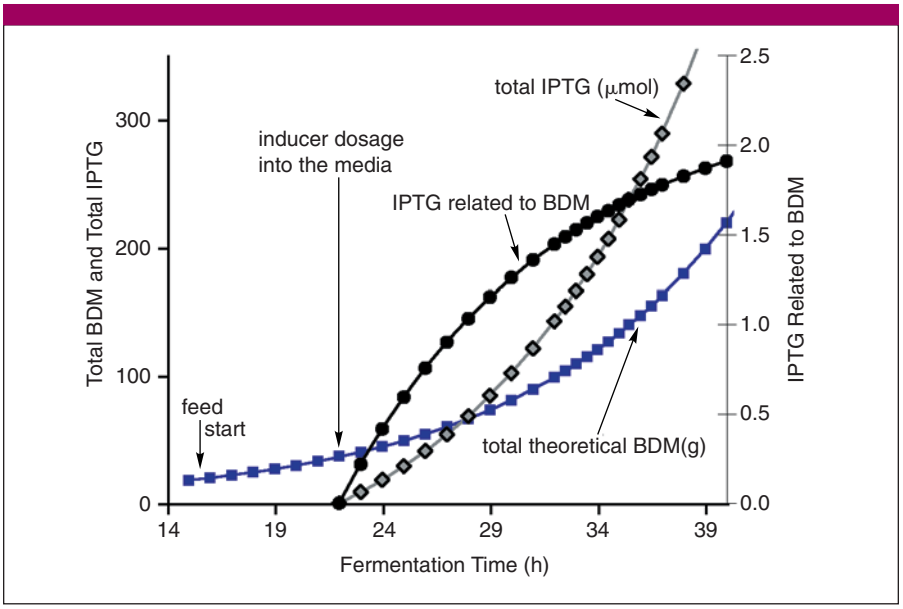

Figure 6. Model-based determination of appropriate amounts of inducer (IPTG) in ratio with the biomass in a fed-batch process for maximum exploitation of the host cell metabolism

Transcription Regulation

Stable plasmid replication significantly improves recombinant gene expression control. That response can be used to optimize development processes. We developed strategies to adjust qP to the rate necessary to prevent damage to the host cell synthesis machinery. Gene expression rates can be continually regulated by adding limited amounts of inducer. Using this strategy, the transcription rate is directly controlled by titration of the repressor. The induction regime, however, requires that accurate amounts of inducer are fed to keep it in a constant ratio with the biomass. This is a different process from that typically used, which adds inducer in a single pulse. The benefits of continuous additions of inducer are shown in the "Inducer Improvement" sidebar.

Inducer feed regime. To determine appropriate (and tolerable) amounts of inducer for maximum recombinant gene expression, we exposed recombinant bacteria to exponentially increasing amounts of inducer, added to the substrate feed of a fed-batch culture. By using the experimental set-up shown in Figure 6, optimal amounts of inducer can be calculated for single fermentation runs (14). Appropriate amounts of inducer for succeeding runs can be estimated from deviations to the theoretical growth curve triggered by overexpression of recombinant proteins (14).

In a series of experiments (data not shown), we determined that host cell metabolism remains uncompromised by overexpression of a recombinant gene at an optimum (tolerable) amount of 0.9 µmol IPTG/g of BDM at a growth rate of $0.1h^{-1}$. Figure 4 shows the aptitude of that feed regime in combination with the appropriate amount of inducer in fed-batch fermentation. Using optimum exploitation of the host vector system and extending the fermentation process, recombinant protein yield is increased 3.6 times when compared with standard methods of pulsed inducer supply (Table 1). This process also substantially improves downstream processing by increasing batch-to-batch reproducibility and improving product consistency (homogeneity) and quality.

In further experiments, we investigated growth rates to determine the dependence of critical amounts of inducer (IPTG). To achieve equal recombinant gene expression

Table 2. Dependence of amount of inducer per gram of biomass dry matter (μmol/g BDM) at different growth rates (μ) to gain equivalent recombinant gene expression

Growth Rate μ (h^{-1})	IPTG μmol/g BDM
0.1	1.1
0.2	2.3
0.3	4.8

rate at increasing growth rates, the amount of inducer must be increased (Table 2). That phenomenon is probably the result of less efficient inducer transport systems and substrates at higher growth rates (15).

Lactose Induction

The novel induction strategy we used allows IPTG to be replaced by lactose, which is an important asset for quality assurance in industrial bioprocessing. Critical points in using lactose for induction are its transport into the cell (unlike IPTG) and its simultaneous consumption. Therefore the objective is to limit carbon during cultivation to prevent inducer exclusion and to feed appropriate amounts of lactose to provide high level recombinant gene expression.

Using our model (Figure 6), we show that approximately 120 μmol lactose/g BDM (two orders of magnitude higher, data not shown) are necessary to achieve induction rates equivalent that of the IPTG. In carbon-limited fed-batch fermentation of *E. coli* (HMS174(DE3)pET11ahSOD) with exponential substrate feeds and continuous lactose supply, yield was increased 2.1 times compared with standard processes (Figure 5). Although this proves the aptitude of lactose for recombinant gene expression, the synthesis potential of the host vector system was not yet fully exploited (which can be seen in the ppGpp data): Room for optimization still exists. The goal is to design a sophisticated inducer feed regime at the optimal balance between lactose induction and consumption. Although transcription rate control is more complex with lactose than with IPTG, lactose offers the advantage of transcription rates that are restored to tolerable levels by lactose consumption.

Targeted Process Optimization

The integrated approach we describe is useful for producing recombinant proteins on an industrial scale and for designing experiments for targeted process optimization. Our approach improves unit costs, increases efficiency, and accelerates process design and operation for producing recombinant proteins. Control of recombinant gene expression guarantees optimal exploitation of the host vector system and leads to foreseeable host cell behavior. Variation of cellular components is minimized, facilitating downstream processing.

In addition to industrial applications, regulating recombinant gene expression offers a unique tool for investigating cell physiology for experiments such as determining the different metabolic loads resulting from varying recombinant gene expression.

Although regulating recombinant gene expression by controlling transcription rates is effective, oscillations still exist in plasmid replication rates. Combining the inducer feed strategy described and modifying plasmid replication together with modern genomic and proteomic tools further improves control and consistency in recombinant protein production systems, thus accelerating bioprocess development.

References

(1) R. Karia, Y. Lis, S.R. Walker, "The Erosion of Effective Patent Life: An International Comparison," *Medicines: Regulation and Risk*, J.P. Griffin, Ed. (Queens University Press, Belfast, 1992), pp. 287–302.

(2) F.W. Studier and B.A. Moffat, "Use of the Bacteriophage T7 RNA Polymerase to Direct Selective High-Level Expression of Cloned Genes," *J. Mol. Biol.* 189, 113–130 (1986).

(3) S.-T. Liang et al., "mRNA Composition and Control of Bacterial Gene Expression," *J. Bacteriol.* 182(11), 3037–3044, (2000).

(4) W. Kramer et al., "Kinetic Studies for the Optimization of Recombinant Protein Formation," *Ann. NY Acad. Sci.* 782, 323–333 (1996).

(5) R. Grabherr et al., "Stabilizing Plasmid Copy Number to Improve Recombinant Protein Production," *Biotechnol. Bioeng.* 77, 142–147 (2002).

(6) R. Grabherr, E. Nilsson, and K. Bayer, Expression Vectors with Modified ColE1 Origin of Replication for Control of Plasmid Copy Number (EP 00 121709.0) (2000).

(7) M. Cserjan-Puschmann et al., "Metabolic Approaches for the Optimisation of Recombinant Fermentation Processes," *Appl. Microbiol. Biotechnol.* 53(1), 43–50 (1999).

(8) S. Breuer et al., "Off-Line Quantitative Monitoring of Plasmid Copy Number in Bacterial Fermentation by Capillary Electrophoresis," *Electrophoresis* 19, 2474–2478 (1998).

(9) K. Bayer et al., "Humane Rekombinante Superoxiddismutase," *BioEngineering* 6(6), 24–30 (1990).

(10) T. Porstmann et al., "Rapid and Sensitive Enzyme Immuno Assay for Cu/Zn Superoxide Dismutase with Polyclonal and Monoclonal Antibodies," *Clinica Chimica Acta* 171, 1–10 (1988).

(11) *Biology of the Prokaryotes*, J. Lengeler, G. Drews, and H. Schlegel, Eds. (Georg Thieme Verlag, Stuttgart, Germany, 1999).

(12) B. Wróbel and G. Wegrzyn, "Replication Regulation of ColE1-Like Plasmids in Amino Acid-Starved *Escherichia Coli*," *Plasmid* 39, 48–62 (1998).

(13) E. Dürrschmid et al., "Optimized Data Exploration of Recombinant Fermentations Using Neural Network Simulations," *Proc. ACoFoP IV Intern.* (Conference on Automatic Control of Food and Biological Processes, Göteborg, Sweden), 21–23 September 1998.

(14) G. Striedner et al., "Metabolic Approaches for the Optimisation of Recombinant Fermentation Processes," *Recombinant Protein Production with Prokaryotic and Eukaryotic Cells: A Comparative View on Host Physiology* (Kluwer Academic Publishers, New York, 2001).

(15) A. Death and T. Ferenci, "Between Feast and Famine: Endogenous Inducer Synthesis in the Adaptation of *Escherichia Coli* to Growth with Limiting Carbohydrates," *J. Bacteriol.* 176(16), 5101–5107 (1994). ***BP***

Correction

To the June *BioPharm*

Liz Howard, PhD, is an Intellectual Property Partner with Orrick, Herrington and Sutcliffe in their Silicon Valley office, not in San Francisco as reported on page 21 of "Stem Cells and Xenotransplantation — Ethics, Patents, and Politics: An Industry Roundtable" (Volume 15, Number 6). We apologize for the error.