Process Simulation for Recombinant Protein Production: Cost Estimation and Sensitivity Analysis for Heparinase I Expressed in *Escherichia coli*

Steffen Ernst,[1] Oscar A. Garro,[1,2] Stefan Winkler,[1] Ganesh Venkataraman,[3] Robert Langer,[1,3] Charles L. Cooney,[1] Ram Sasisekharan[4]

[1]*Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*

[2]*Facultad de Agroindustrias. Universidad Nacional del Nordeste, P.R. Saenz Peña, Chaco, Argentina*

[3]*Harvard–MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*

[4]*Whitaker College of Health Sciences and Technology, Massachusetts Institute of Technology, Room E18-566, 77 Massachusetts Avenue, Cambridge, Massachusetts 01239; telephone: 617-258-9494; fax: 617-258-9409; e-mail/:ramnat@mit.edu*

Received 28 March 1996; accepted 8 August 1996

Abstract: Heparinase I from *Flavobacterium heparinum* has several potential clinical applications; the resulting high demands on protein purity and quantity can be met by recombinant expression in *Escherichia coli.* Based on laboratory scale experiments with insoluble heparinase I expression followed by renaturation, a process for production of 3 kg/year of heparinase I was designed. We present a comparative analysis of the production costs of soluble and insoluble heparinase I expression, as well as a generalized approach to sensitivity analysis, based on perturbation around a base case design scenario. This may assist focusing further development on process steps for which improvements both are feasible and result in significant cost saving. © 1997 John Wiley & Sons, Inc. *Biotechnol Bioeng* **53**: 575–582, 1997.
Keywords: recombinant protein production; *Escherichia coli;* heparinase; bioprocess simulation

INTRODUCTION

The transition from discovery to development is crucial to the commercial success of biopharmaceutical projects. Experience shows that choices made early in the development process often are difficult to change later due to regulatory constraints on process modification. The result leads to suboptimal solutions to process design problems. This may hamper the profitability and increase cost to the patient.

Correspondence to: R. Sasisekharan.
Contract grant sponsor: National Science Foundation through Biotechnology Process Engineering Center at MIT; National Institutes of Health
Contract grant numbers: GM 25810 and DK 51207
Contract, grant sponsor: Universidad Nacional del Nordeste, Argentina
Contract grant sponsor: Consejo Nacional de Investigaciones Cientificas y Technicas

Computer-aided process simulation has become a standard tool to plan, design, optimize, and evaluate chemical processes (Westerberg et al., 1979), and has also been adapted to bioprocesses. Cooney et al., 1988; Evans and Field, 1988; Petrides, 1994; Petrides et al., 1989, 1995). Process simulation involves preparing a process flowsheet, solving the corresponding mass and energy balances based on user-provided or built-in parameter estimation, estimating total cost, and analyzing sensitivity to changes in assumptions (Wells and Rose, 1986) Process simulation can provide a link between drug discovery and initial stages of process development, guiding efforts to optimize the manufacturing process expeditiously. It is especially effective when utilized early in process development.

Heparinase I (EC 4.2.2.7) from *Flavobacterium heparinum* is an interesting model for bioprocess simulation of recombinant protein production. Heparinase I degrades heparin, an acidic polysaccharide, which is widely used as an anticoagulant drug (Ernst et al., 1995). This enzyme has several potential clinical applications, which are at different stages of product development: injection of heparinase I, for elimination of heparin in patients heparinized during extracorporeal procedures, is being evaluated in clinical trails (Zimmerman et al., 1993); regional heparinization of blood in an extracorporeal device has been achieved in animal models, using a reactor with immobilized heparinase I (Langer et al., 1982); an assay for heparin based on different whole blood clotting time with heparinase I is in clinical use (Baugh et al., 1992); and heparinase I is a potent inhibitor of neovascularization (Sasisekharan et al., 1994), which may lead to its use in the treatment of angiogenesis-dependent diseases (Sasisekharan et al., 1996). These ap-

© 1997 John Wiley & Sons, Inc.
CCC 0006-3592/97/060575-08

plications put great demands on the production of high-purity heparinase I. We have recently cloned and expressed the gene for heparinase I in *Escherichia coli,* and purified the enzyme to homogeneity (Ernst et al., 1996; Sasisekhara et al., 1993). We found that, in different expression vectors, heparinase I could be expressed as a soluble enzyme, or in an insoluble form as inclusion bodies, which could be refolded for purification. Heparinase I expression as inclusion bodies followed by refolding gave higher enzyme expression levels than soluble heparinase I expression (Ernst et al., 1996), raising the question as to whether this advantage translates into a lower cost for heparinase I production when cost and efficiency of solubilization and refolding steps are considered.

Thus, recombinant heparinase I expressed in *E. coli,* is an example of a biopharmaceutical drug candidate for which a manufacturing process is in the initial stages of development. Here we use a PC-based software package, SuperPro Designer® (Intelligen, Scotch Plains, NJ), to design a manufacturing process for heparinase I based on lab-scale experimental results, and show how sensitivity analysis can be used to choose among several expression systems, and guide future work on optimizing cost efficiency of this process.

EXPERIMENTAL RESULTS

We have previously expressed heparinase I with an N-terminal histidine tag in shake-flask *E. coli* cultures (Ernst et al., 1996). Using the expression vector, pET-28a (Novagen, Madison, WI), the enzyme appears in an inactive, insoluble form, which can be refolded and purified to homogeneity by nickel-chelate chromatography (Ernst et al., 1996). The cumulative yield was 43%, and the recovered purified heparinase I was 14.4 mg/L of culture (Ernst et al., 1996). Also, heparinase I was expressed in a soluble form using the vector pET-15b (Novagen), which gave approximately 18-fold lower enzyme expression levels, calculated on an activity basis, than insoluble heparinase I expression (Ernst et al., 1996). Based on these results, insoluble heparinase I expression was pursued, and this process was scaled up to 6-L volume using glycerol as fermentation carbon source. The yield on glycerol was 0.41 g dry cell weight (DCW)/g, the cell density at harvest was 11.2 g/L and the refolded heparinase I was 17 IU/mL of culture (unpublished observations). We use these experimental values, as well as published information (Ernst et al., 1996) as basis for process design for purification of heparinase I expressed in insoluble and soluble form, to compare the economics of the two processes.

PROCESS DESIGN

The SuperPro Designer assists in designing scale-up of a laboratory process to meet a desired plant throughput. The program contains models for a set of unit operations and reactors typically used in bioprocessing (Petrides, 1994; Petrides et al., 1989, 1995). These can be combined using a graphical interface with drag-and-drop features, to synthesize a flowsheet for the process. The user must provide stoichiometry and conversion for each reaction, as well as yields for the separation processes; the program solves the mass and energy balances, sizes equipment, estimates purchase costs, and reports stream and equipment data, as well as capital and manufacturing costs.

Insoluble Expression: Base Case

We estimate the markets for the clinical applications of heparinase I to be up to 1.8 kg for injection following open-heart surgery (3 mg/patient (Zimmerman et al., 1993) times 590,000 patients/year [American Heart Association, 1995]); 1.0 kg for immobilized heparinase I reactors for use during acute dialysis procedures (10 mg/reactor times 100,000 patients/year [unpublished estimates, G. Ameer, S.E., R.L., and R.S.]); and less than 50 g for use as an analytical reagent.

A process for production of 3 kg of heparinase I per year was designed based on the insoluble expression procedure (Fig. 1a). Fermentation medium is sterilized in a continuous heat sterilizer. The fermentation time is 12 h, including a 2-h induction. The final concentration of *E. coli* is 11.2 g/L (Au_{600} = 34), and heparinase I accumulates intracellularly as inclusion bodies at a concentration of 23 mg/g DCW. The broth is cooled and kept at 4°C for subsequent processing. Cells are harvested in a disk-stack centrifuge, washed, and disrupted in a high pressure homogenizer (three passes at a pressure drop of 800 bar). The inclusion bodies are recovered in the same disk-stack centrifuge, washed in three volumes of buffer, and recentrifuged. The inclusion body suspension is solubilized in 4 *M* guanidine hydrochloride for 45 min at a concentration of 6 mg protein/mL, and the remaining debris is removed using a polishing dead-end filter.

Heparinase I is refolded to its active form by dilution to 0.30 mg protein/mL in Tris-buffer with 0.5 *M* NaCl for 12 h, resulting in 129 mg of active heparinase I per liter of culture. After being passed through another dead-end filter, heparinase I with a histidine tag is purified by nickel-affinity chromatography (Ernst et al., 1996). The resin binding capacity is taken as 5 mg/mL (Schmitt et al., 1993) and heparinase I is eluted in four column volumes with a recovery yield of 42% (Ernst et al., 1996). The eluate is desalted and the volume reduced to half in a diafilter with a yield of 89% (Ernst et al., 1996). The histidine tag is removed from heparinase I by enzymatic cleavage with thrombin for 12 h at a thrombin:heparinase I ratio of 1:2000 (w/w) (Novagen, Madison, WI). Due to further refolding of a fraction of inactive heparinase I after cleavage of the histidine tag, the activity *increases* by 43% at this step at laboratory scale (Ernst et al., 1996). The protein is concentrated by a factor of 20 (83% yield) in a diafilter, which also removes the cleaved histidine tag peptide (2 kDa). A second affinity chromatography column binds any remaining heparinase I with the histidine tag, and the cleaved heparinase I is col-

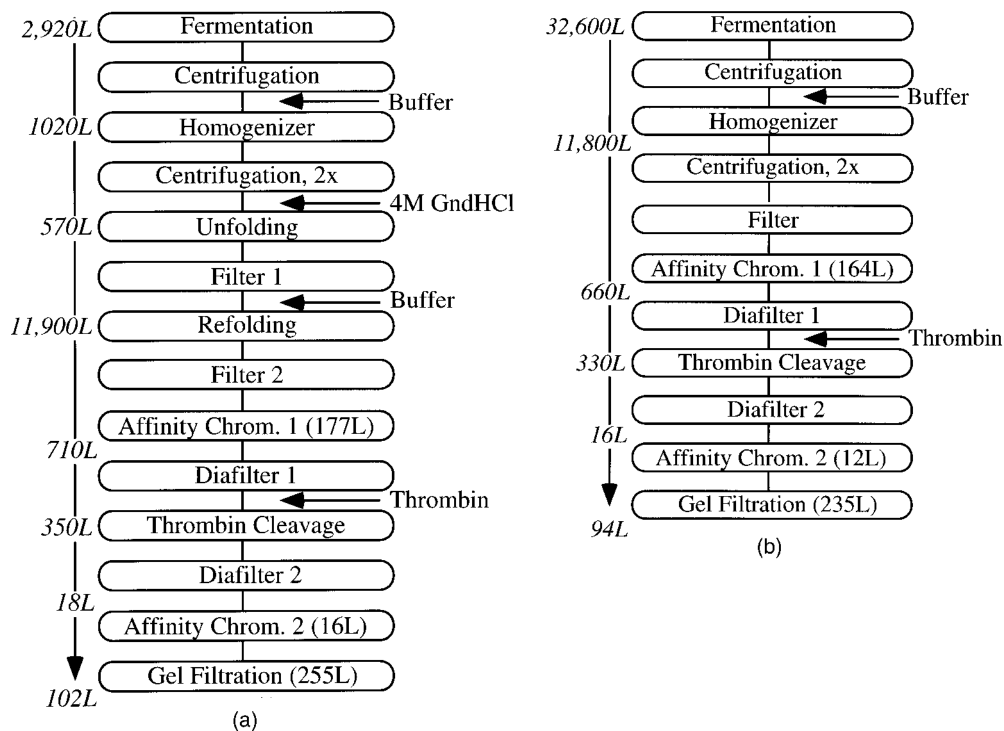

Figure 1. Block diagram for recombinant heparinase I production. (A) Production by insoluble expression of heparinase I. (B) Production by soluble expression of heparinase I.

lected in the flowthrough fraction at a yield of 78% (Ernst et al., 1996). Finally, the protein is purified on a gel-filtration column with a yield of 95%.

SuperPro Designer contains models to estimate equipment purchase prices and direct fixed capital cost (1995 dollars) (Petrides, 1994). The costs of raw materials and consumables are estimated from list prices of the largest quantities available using power-law scaling criteria. Maintenance cost is set to 25% of the purchase cost of the equipment per year (Peters and Timmerhaus, 1991), and laboratory cost allocated to capital costs (15% of other capital costs) and to labor cost (100% of other labor costs). SuperPro Designer calculates labor costs from a user-supplied number for operator hours per equipment hour. We estimated this number for each piece of equipment after careful consideration of scheduling of the entire process to assure that the total number of operators, as well as the allocation of their time for each piece of equipment, was reasonable. It is assumed that the process will operate as part of a larger plant, and that equipment-related expenses can be reduced according to utilization.

The process is designed to run in campaigns with a total of 24 batches/year. The fermentation volume is 2.9 m^3; the largest volume occurs after dilution for refolding (12 m^3), which is reduced to 710 L after the first affinity chromatography step, 18 L after the second diafiltration step, and 102 L at 1.2 mg/mL after gel filtration. Processing of each batch takes 3.5 days of one-shift operation with fermentation, refolding, and thrombin cleavage running overnight, and requires five operators. Two batches can be run per week for 12 weeks, occupying 45% of yearly equipment availability including turnover time (thus equipment purchase prices and maintenance expenses are multiplied by the utilization factor of 0.45).

Table I. Purchase cost for major pieces of equipment for insoluble and soluble expression.

	Insoluble expression		Soluble expression	
Equipment	Size	Purchase cost ($)	Size	Purchase cost ($)
Heat sterilizer	1 m^3/h	154,000	5.4 m^3/h	262,000
Fermentor	3.9 m^3	116,000	43 m^3	388,000
Disk-stack centrifuges	σ: 22,300 m^2	104,000	σ_{total}: 512,000 m^2	921,000
Refolding tank	14 m^3	114,000	—	—
Dead-end filter 2	4 × 40 m^2	124,000	4 × 38 m^2	120,000
Affinity column 1	177 L	254,000	164 L	251,000
Gel-filtration column	255 L	170,000	235 L	167,000

The most expensive pieces of equipment are listed in Table I. The capital investment is not dominated by any single piece of equipment, hence it is deemed unnecessary to scale down equipment size of any step by processing fractions of a batch sequentially at that step. The direct fixed capital cost producing heparinase I at this scale is found to be $12.5M, and the manufacturing cost $560,800/kg (including capital depreciation). The breakdown of manufacturing cost on process steps and expense categories is shown in Table II. Annualized capital cost is found from 12-year straight-line depreciation of the direct fixed capital cost, without amortization, and then distributed on a process step proportionally to the purchase costs of that step. Capital cost is the dominant expense category (55%), followed by raw materials and consumables (19%). Fermentation (26%) and primary affinity chromatography (23%) are the most expensive individual steps. Harvesting and refolding comprise 13% of the manufacturing cost, however, no single step in this part of the process accounts for more than 8% of total cost. Thus, from lab-scale data on expression and purification of recombinant heparinase I, we constructed a process flowsheet for production of 3 kg/year, and estimated the cost of production. This alone can give an indication of profitability of the process; however, the main advantage of having a flowsheet is the ability to easily modify the base case to perform a sensitivity analysis of total cost with respect to variations in the assumptions used for the base case. Thus, one can computationally test alternative process scenarios.

Soluble Expression

A second design case was assembled based on expression of soluble, active heparinase I (Ernst et al., 1996). The expression level is 1.48 mg/g dry cell weight (Ernst et al., 1996), and the cell density at harvest is 11.2 g/L. The cell lysate is clarified by centrifugation and dead-end filtration, and then loaded directly on the first nickel affinity column (Fig. 1b). For purification, the design and assumptions are identical to those of the base case for insoluble expression, except for the thrombin cleavage step, which does not result in increased activity (Ernst et al., 1996). For production of 3 kg/year in 24 batches, the fermentation volume is 32.6 m^3, which is reduced to 12 m^3 for homogenization, and results in 94 L of pure product at 1.3 mg/mL.

Table I lists purchase prices for major pieces of equipment, the most expensive of which are the centrifuges for processing cell lysate ($225,000 to 350,000 each), and the fermentor ($388,000). Heparinase I production by soluble expression is found to require $24.1M in direct fixed capital, and $1,024,400/kg in manufacturing cost including depreciation. When this cost is broken down similarly to the base case for insoluble expression, it is found that fermentation (42%) and harvesting (36%) are the most expensive process steps, and that capital cost (59%) is the dominant cost category (Table III). This process is 80% more expensive than the process for insoluble expression, mainly due to the low expression level of heparinase I, which results in large process volumes during fermentation and harvesting, leading to higher capital and raw materials costs.

SENSITIVITY ANALYSIS

Soluble vs. Insoluble Expression

From the initial designs of processes for soluble and insoluble heparinase I expression, which were based on reported data (Ernst et al., 1996), it appeared that insoluble expression was more cost effective. This preference for insoluble expression, however, may change if key parameters in the two cases change. We investigated the effect of increasing the expression level of soluble heparinase I, and of changing protein concentration during refolding, which in a process for production of tissue plasminogen activator (t-PA) in *E. coli* turned out to be the most important parameter to determine cost (Datar et al., 1993). Figure 2 shows the specific cost as a function of these two parameters. If the expression level for soluble heparinase I expression can be increased to around 100 mg/L (compared to 129 mg/L for insoluble heparinase I expression), the production costs of the two processes are equal. Inversely, if refolding could be carried out only at very dilute conditions (<50 mg/L), the process based on soluble heparinase I expression will be more economical even at the present expression level. In the case of t-PA, protein refolding had to

Table II. Breakdown on process step and expense category of the cost of production for the base case of insoluble expression of heparinase I ($/kg product).

Step	Capital cost	Raw materials, consumables, and utilities	Labor	Maintenance	%
Fermentation	$92,000	$13,700	$26,900	$13,600	26
Harvest and refolding	94,800	48,900	16,400	15,500	31
Affinity chromatography 1	58,800	32,800	26,000	9600	23
Filtration and thrombin cleavage	23,100	7600	7100	3800	7
Affinity chromatography 2 and GF	39,200	4000	20,700	6400	13
Percent	55%	19%	17%	9%	100

Groups of steps have been pooled for simplicity. The total cost $560,800/kg.

Table III. Breakdown on process step and expense category of the cost of production by soluble expression of heparinase I ($/kg product).

Step	Capital cost	Raw materials, consumables, and utilities	Labor	Maintenance	%
Fermentation	$215,200	$152,400	$28,100	$34,100	42
Harvest	266,700	10,800	45,600	50,100	36
Affinity chromatography 1	57,900	30,200	26,000	9,600	12
Filtration and thrombin cleavage	22,500	6700	7100	3800	4
Affinity chromatography 2 and GF	38,300	3100	20,600	6400	7
Percent	59%	20%	12%	10%	100

Groups of steps have been pooled for simplicity. The total cost is $1,024,400/kg.

be carried out at 2.5 mg/L (Datar et al., 1993), indicating that a soluble production route would be less expensive even at modest expression levels; however, this was not investigated (Datar et al., 1993). For heparinase I, we lowered the temperature of the fermentor to 15°C during induction to reduce inclusion body formation, and achieved soluble heparinase I at a level of 4.1 mg/g DCW (unpublished observations). This would lead to a manufacturing cost of $653,200/kg, which is only 16% more than for insoluble heparinase I expression.

Perturbation Around the Base Case

In the bioprocess industry, process development is often a trade-off between optimization and time to market launch. It is therefore imperative to focus one's research on improving parts of the process that are most likely to have the largest

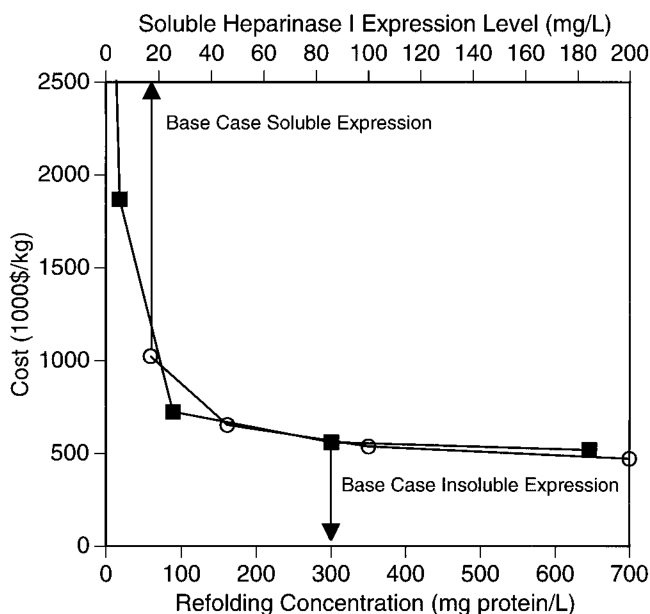

Figure 2. The sensitivity of cost of soluble heparinase I expression to the expression level (circles) and sensitivity of cost of insoluble heparinase I expression to the concentration during refolding (squares). Arrows indicate the location of the base cases for soluble and insoluble heparinase I expression, respectively.

impact in the shortest time. Thus, two parameters should be evaluated when deciding upon which step to focus on. First, the effect on cost of improving the performance by a given fraction should be investigated. This can be expressed as a partial derivative of the cost function with respect to the input parameter that will be improved. Second, the probability of achieving such an improvement in performance of that parameter should be assessed. This assessment is often subjective; however, we attempt to quantitatively include this aspect in the sensitivity analysis.

For sensitivity analysis, we operate mainly with process improvements that improve the product recovery after a given step (R_j), and therefore affect the design of other steps in the process. We define the sensitivity parameters $\alpha_{R,j}$ as the partial derivative of the cost function:

$$\alpha_{R,j} = \frac{\partial C'_n}{\partial R_{j,n}} \text{ (evaluated at } R_{j,n} = 1)$$

C' is the specific cost ($/kg product), and the index n refers to values normalized with respect to the base case. In general, costs of a process step have been found to depend on the scale as a power-law function, $C_i = C_0(S_i/S_0)^x$, where S_i is the scale of process step i and C_0 is the cost of that process step at a reference sale (S_0) (Peters and Timmerhaus, 1991). The power x depends on the category of expense, for expample, capital, labor, maintenance, etc. Thus, for variations in the process from the base case:

$$C = \sum_{i=1}^{N} \sum_{k=1}^{n} C_{0,i,k} \left(\frac{S}{S_0}\right)^{x_k}$$

where the summations cover N process steps and n cost categories with different scaling powers x_k. A change in recovery of a step j, followed by adjustment of the projected batch size to meet the specified production goal, will generally lead to scaling of *previous* steps ($S_i/S_{0,ij} = R_{0,j}/R_j$, for $i = [1 \ldots j]$), but no change in *subsequent* steps $S_i/S_{0,i} = 1$, for $i = [j+1 \ldots N]$. However, if the change in recovery leads to a change in process volume in subsequent steps, these may have to be scaled also. For instance, the eluate volume from the affinity chromatography step depends on the column volume, and will therefore decrease if the column has been scaled down due to increased recovery. This

would lead to a scaling down of the subsequent diafiltration step, which in turn reduces the process volume until a specified concentration has been reached, eliminating scaling of any step further downstream. Thus, for sensitivity analysis of cost with respect to the recovery from affinity chromatography (step j), the summation should be extended to cover the first diafiltration (step $J = j + 1$). Taking the partial derivative of the cost function with respect to R_j we get:

$$\frac{\partial C}{\partial R_j} = \sum_{i=1}^{J} \sum_{k=1}^{n} -x_k C_{0,i,k} \left(\frac{R_j}{R_{0,j}}\right)^{-x_k-1} \frac{1}{R_{0,j}}$$

where $J \geq j$, and depends on the scaling criteria for the steps immediately following j. For a constant production rate:

$$\alpha_{R,j} = \frac{\sum_{i=1}^{J} \sum_{k=1}^{n} x_k C'_{0,i,k}}{C'_0}$$

where $C'_{0,i,k}$ is the specific cost of step i and category k of the base case; x_k depends on the category of expense. We use 0.6 for capital cost and maintenance; 0.25 for labor; and 1 for raw materials, consumables, and utilities (Peters and Timmerhaus, 1991).

The effect of recovery in different steps is investigated by calculating values of α_R using the breakdown of cost on steps and expense categories from the base case. (For fermentation, recovery can be improved by growing to higher cell density at harvest or by increasing the specific expression level of the target protein; in purification steps, the recovery is directly proportional to the yield of that step.) $-\alpha_{R,j}$ and the cumulative specific cost ($\Sigma_{1\,..\,j}\, C'_{0,i}$) are plotted in Figure 3. These two parameters follow each other quite closely, such that the highest values of $-\alpha_{R,j}$ are found late in the process, where a change in recovery will cause all previous steps to be rescaled. This means that, for a given relative change in recovery of a step, one finds the highest saving in cost if that step is late in the process, downstream of the most expensive steps. At the last step, the sensitivity parameter $-\alpha_{R,N}$ is equal to the sensitivity of cost to scaling of the entire process, a value of 0.62.

The second level of sensitivity analysis involves a subjective assessment of the feasibility of achieving improvements in given input parameters. Here we generalize this assessment by relating the sensitivity of cost to improvement of a parameter to the maximum value attainable for that parameter. We define the loss of step j (L_j) as:

$$L_j = \frac{R_{max,j} - R_j}{T_{max,j}}$$

where $R_{\max,j}$ is the maximum recovery possible. For purification steps the maximum recovery is generally 100% (except when purification results in refolding or other forms of activation of the target protein); however, for synthetic processes, one must use benchmark standards for optimal performance. We use 50 g/L for cell density and 40 mg/g dry cell weight for protein expression level (determined as active soluble protein after refolding) (Bech Jensen and Carlsen, 1990; Fieschko and Ritch, 1986; Jung et al., 1987; Stader, 1995; Yee and Blanch, 1992; Zabriskie et al., 1987), leading to $R_{\max,j} = 2000$ mg/L. For synthetic process steps, the maximum recovery possible is largely dependent on the properties of the target protein, and generalized values should be used with caution. We can now define a sensi-

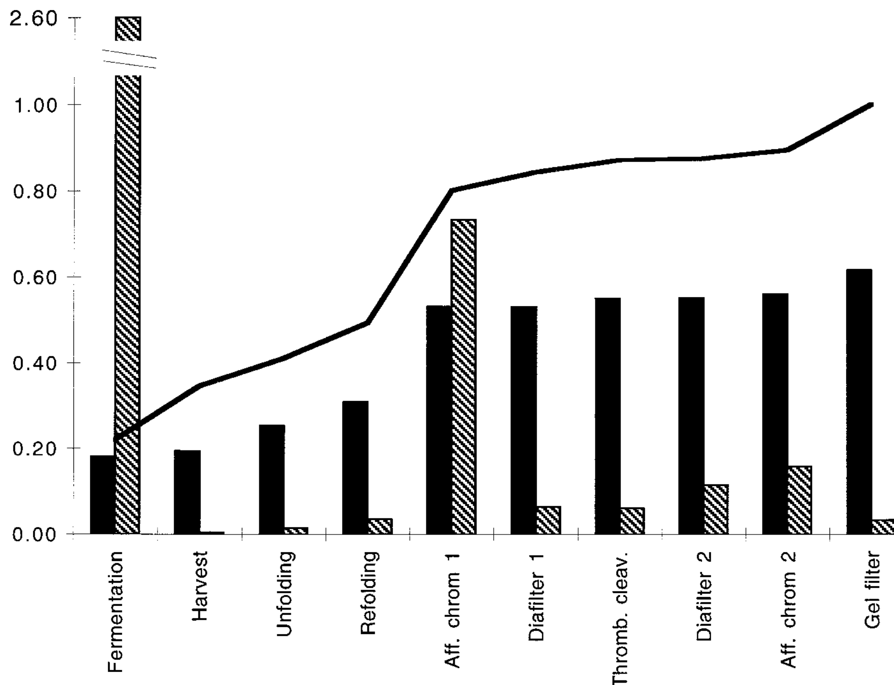

Figure 3. Plot of sensitivity parameters $\alpha_{R,j}$ (solid bars), $\alpha_{L,j}$ (hatched bars), and cumulative cost (line) as functions of process step.

tivity parameter α_L, that relates the relative change in cost to the relative change in loss:

$$\alpha_{L,j} = \frac{\partial C_n'}{\partial L_j}, \text{ which means that: } \alpha_{L,j} = \left(1 - \frac{R_{max}}{R_j}\right) \alpha_{R,j}$$

In other words, instead of expressing how cost is affected by *increasing recovery* by a given factor, we express how cost is affected by *reducing losses* by a given factor. When the recovery of a step is 50% of maximum, the two numbers are equal; when recovery is above 50%, $\alpha_{L,j}$ is a better indicator of the step that further development should focus upon. $\alpha_{L,j}$ for each step of the process is plotted in Figure 3. For two steps, fermentation and primary affinity chromatography, $\alpha_{L,j}$ attains values significantly higher than any other step. This indicates that, for these two steps, there are both high potential cost reduction and room for improvement in the recoveries obtained before maximum recoveries are reached. For fermentation, $\alpha_{L,j}$ is much larger than $\alpha_{R,j}$, indicating that the recovery from this step is far from its potential maximum, whereas, for affinity chromatography, the two sensitivity parameters are almost identical, indicating that R is close to 50% of R_{max}.

From this analysis we conclude that one should focus on improving fermentation and affinity chromatography recoveries. We investigated design cases in which a range of values of the heparinase I expression level and the affinity chromatography yield were assumed, and calculated the manufacturing cost (Fig. 4). Improving the yield to 90% would result in a 25% reduction in cost; increasing the expression level of active, refolded heparinase I to 2000 mg/L would lead only to a 12% reduction in cost. In subsequent optimization of the fermentation, we achieved a cell density at harvest of 20 g/L with an effective expression level of active, refolded heparinase I of 445 mg/L (data not shown). The corresponding direct fixed capital is \$11.0M, and the manufacturing cost \$506,600/kg, a saving of 10%. Additional increases in heparinase I expression level will not have a significant impact on cost (Fig. 4).

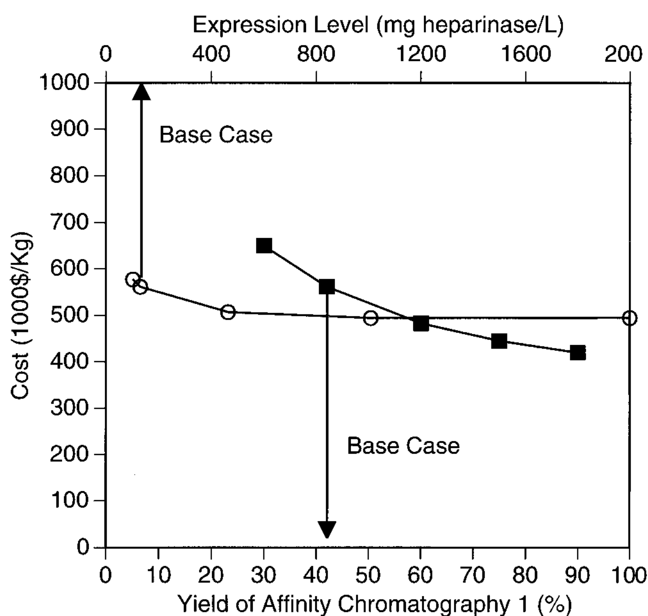

Figure 4. Sensitivity of cost with respect to expression level of heparinase I (circles) and yield of the first affinity chromatography column (squares) for insoluble heparinase I expression. Arrows indicate the location of the base case for insoluble expression of heparinase I.

DISCUSSION

Development of recombinant expression in *E. coli* often results in a choice between soluble or insoluble protein expression, depending on the choice of expression vector and fermentation conditions. For production of potential biopharmaceuticals, this choice may be irreversible due to regulatory constraints; e.g., completed toxicology and clinical trials may be invalidated by a subsequent process change. Thus, it is important to consider the economic consequences for the entire process at an early process development stage. Typically, soluble protein expression results in lower product accumulation during fermentation, which leads to large volumes of broth for a given production rate, and consequently higher processing cost; however, insoluble protein expression necessitates additional processing steps for solubilization and refolding.

For recombinant heparinase I, we evaluated scaling-up two laboratory processes, for soluble and insoluble protein expression, respectively. We found that heparinase I expressed in insoluble form can be refolded at sufficiently high concentrations to be less expensive than heparinase I expressed in soluble form, even if the soluble enzyme expression levels could be increased to around 80% of the insoluble enzyme expression level. This contrasts a case study of a process for production of recombinant t-PA from *E. coli* inclusion bodies (Datar et al., 1993). This process involves refolding at dilute conditions, such that the cost of refolding tanks constitutes 75% of total equipment purchase cost (Datar et al., 1993), suggesting that *soluble* t-PA expression might be a cost-effective alternative. These studies illustrate that the choice between soluble and insoluble protein expression processes is highly dependent on the concentration at which protein refolding can be carried out for the insoluble expression process.

We have presented a general approach to sensitivity analysis of the cost function, which is easily implementable for any integrated process simulation base case, provided a reliable breakdown of cost on process steps and expense categories can be performed. This sensitivity analysis is based on a simplified cost function to evaluate perturbations around a base design case, which incorporates a more accurate cost estimation, and it readily gives an overview of the effects of changing various parameters. As process data accumulates during a development project, more accurate models and parameter values will be available, and more sophisticated simulation and sensitivity analyses will be possible.

When investigating large deviations from the base case

(Fig. 2 and 4), the simplified cost function may no longer be valid. For example, when heparinase I expression level is increased beyond a certain point, little additional cost reduction can be achieved without changing other parameters. This reflects both that fermentation-related expenses constitute a minor part of total cost at high protein expression levels, and that the cost functions built into SuperPro Designer approach a minimum purchase price for fermentors smaller than a few hundred liters. In general, process steps that are downstream of the most expensive steps (high $\alpha_{R,j}$) or for which the recovery is far from maximum ($R_j << R_{\max}$) will be most advantageous to optimize. We suggest combining cost analysis with estimates of the maximum yields/recoveries possible for the various steps to identify the process parameter(s) that will most likely give significant reductions in cost as a result of further optimization, as opposed to immediately focusing on improving the most expensive step in a process.

The initial construction of a flowsheet is facilitated greatly by the Windows-based user interface of SuperPro Designer, as compared with other process design software packages (such as BPS/Aspen Plus, Aspen Technologies, Cambridge, MA). We utilized this to construct a base-case scenario for recombinant production of heparinase I, which helped us choose the most cost-effective expression system. The base-case process design served as a starting point for sensitivity analysis, which guided further optimization of the process. Thus, we have demonstrated how the use of process simulation can be integrated in the development of a protein production process.

We are grateful to Intelligen, Inc. for making available the software which was used in this work.

NOMENCLATURE

- $\alpha_{R,j}$ sensitivity parameter for changes in recovery of step *j*
- $\alpha_{L,j}$ sensitivity parameter for changes in loss in step *j*
- B_n the normalized value of a quantity *B* relative to the base case: $B_n = B/B_0$
- C absolute cost ($)
- C' specific cost ($/kg product)
- L_i loss of step *i*
- R_i recovery of step *i*
- S_i scale of step *i*

References

Baugh, R. F., Deemar, K. A., Zimmerman, J. J. 1992. Heparinase in the activated clotting time assay: monitoring heparin-independent alterations in coagulation function. Anesth. Analg. **74**: 201–205.

Bech Jensen, E., Carlsen, S. 1990. Production of human growth hormone in *Escherichia coli:* expression of different precursors and physiological effects of glucose, acetate, and salts. Biotechnol. Bioeng. **36**: 1–11.

Cooney, C. L., Petrides, D., Barrera, M., Evans, L. 1988. Computer-aided design of a biochemical process, pp. 39–61. In M. P. Phillips, S. P. Shoemaker, R. D. Middlekauff, and R. M. Ottenbrite (eds.), The impact of chemistry on biotechnology: multidisciplinary discussions. American Chemical Society, Washington, DC.

Datar, R.*V., Cartwright, T., Rosén, C.-G. 1993. Process economics of animal cell and bacterial fermentations: a case study analysis of tissue plasminogen activator. Bio/technology* **11**: 349–357.

Ernst, S., Langer, R., Cooney, C. L., Sasisekharan, R. 1995. Enzymatic degradation of glycosaminoglycans. CRC Crit. Rev. Biochem. Mol. Biol. **30**: 387–444.

Ernst, S., Venkataraman, G., Winkler, S., Godavarti, R., Langer, R., Cooney, C. L., Sasisekharan, R. 1996. Expression in *Escherichia coli,* purification, and characterization of heparinase I from *Flavobacterium heparinum.* Biochem. J. **315**: 589–597.

Evans, L. B., Field, R. P. 1988. Bioprocess simulation: a new tool for process development. Bio/technology **6**: 200–203.

Fieschko, J., Ritch, T. 1986. Production of human alpha consensus interferon in recombinant *Escherichia coli.* Chem. Eng. Commun. **45**: 229–240.

Jung, G., Denèfle, P., Becquart, J., Mayaux, J.-F. 1987. High-cell density fermentation studies of recombinant *Escherichia coli* strains expressing human interleukin-1β. Ann. Inst. Pasteur/Microbiol. **139**: 129–146.

Langer, R., Linhardt, R. J., Hoffberg, S., Larsen, A. K., Cooney, C. L., Tapper, D., Klein, M. 1982. An enzymatic system for removing heparin in extracorporeal therapy. Science **217**: 261–263.

Peters, M., Timmerhaus, K. 1991. Plant design and economics for chemical engineers. 4th edition. McGraw-Hill, New York.

Petrides, D. P., Cooney, C. L., Evans, L. B. 1989. An introduction to biochemical process design. In: M. Shuler (ed.), Chemical engineering problems in biotechnology. AICHE, New York.

Petrides, D. M. 1994. BioPro Designer: an advanced computing environment for modeling and design of integrated biochemical processes. Comput. Chem. Eng. **18**: S621–S625.

Petrides, D., Sapidou, E., Calandranis, J. 1995. Computer-aided process analysis and economic evaluation for biosynthetic human insulin production—a case study. Biotechnol. Bioeng. **48**: 529–541.

Sasisekharan, R., Bulmer, M., Moremen, K. W., Cooney, C. L., Langer, R. 1993. Cloning and expression of heparinase I gene from *Flavobacterium heparinum.* Proc. Natl. Acad. Sci. USA **90**: 3660–3664.

Sasisekharan, R., Moses, M. A., Nugent, M., Cooney, C. L., Langer, R. 1994. Heparinase inhibits neovascularization. Proc. Natl. Acad. Sci. USA **91**: 1524–1528.

Sasisekharan, R., Moses, M. A., Nugent, M., Cooney, C. L., Langer, R. S. 1996. Method for inhibiting angiogenesis using heparinase, United States patent 5,567,417.

Schmitt, J., Hess, H., Stunnenberg, H. G. 1993. Affinity purification of histidine-tagged proteins. Mol. Biol. Rep. **18**: 223–230.

Stader, J. 1995. Gene expression in recombinant *Escherichia coli,* pp. 1–51. In: A. Smith (ed.), Gene expression in recombinant microorganisms. Marcel Dekker, New York.

Wells, G. L., Rose, L. M. 1986. The art of chemical process design. Elsevier, Amsterdam.

Westerberg, A. W., Motard, R., Winter, P. 1979. Process flowsheeting. Cambridge University Press, Cambridge, UK.

Yee, L., Blanch, H. W. 1992. Recombinant protein expression in high cell density fed-batch cultures of *Escherichia coli.* Bio/technology **10**: 1550–1556.

Zabriskie, D. W., Wareheim, D. A., Polansky, M. J. 1987. Effects of fermentation feeding strategies prior to induction of expression of a recombinant malaria antigen in *Escherichia coli.* J. Ind. Microbiol. **2**: 87–95.

Zimmerman, J. J., Lewis, N. T., Heft, R. A. 1993. Method for the enzymatic neutralization of heparin. United States patent 5,262,325.