Optimal Design of Protein Production Plants with Time and Size Factor Process Models

Jorge M. Montagna, Aldo R. Vecchietti, and Oscar A. Iribarren

INGAR−Instituto de Desarrollo y Diseño, Avellaneda 3657−3000 Santa Fe, Argentina

José M. Pinto

Department of Chemical Engineering, University of São Paulo, São Paulo SP, Brazil 05508−900

Juan A. Asenjo*

Centre for Biochemical Engineering and Biotechnology, Department of Chemical Engineering, University of Chile, Beauchef 861, Santiago, Chile

In this work we propose an optimization model for the design of a biotechnological multiproduct batch plant. A first level of detail posynomial model is constructed for each unit, as well as decisions regarding the structural optimization of the plant. A particular feature of this model is that it contains composite units in which semicontinuous items operate on the material contained by batch items. This occurs in the purification steps, in particular with the microfilters operating between retentate and permeate vessels, and with the homogenizer and ultrafilters operating on the material contained in a batch holding vessel. Also, the unit models rely on batch operating time expressions that depend on both the batch size and the size of semicontinuous items. The model takes into account all of the available options to increase the efficiency of the batch plant design: unit duplication in-phase and out-of-phase and intermediate storage tanks. The resulting mathematical model for the minimization of the plant capital cost is a mixed integer non-linear program (MINLP), which is solved to global optimality with an implementation of the outer approximation/ equality relaxation/ augmented penalty (OA/ER/AP) method. A plant that produces four recombinant proteins in eight processing stages is used to illustrate the proposed approach. An interesting feature of this example is that it represents an attempt to standardize a plant for the production of both therapeutic and nontherapeutic proteins; the model applied is generic and can thus be applied to any such modular plant. Results indicate that the best solution in terms of minimal capital cost contains no units in parallel and with intermediate storage tank allocation.

Introduction

There has been an increased interest in the development of systematic methods for the design of batch processes in specialty chemicals, food products, and pharmaceutical industries (Reklaitis, 1992). Most processes in the modern biotechnology industry correspond to batch plants, and with the rapid development of new products (i.e., both therapeutic and nontherapeutic proteins) there is an increased need for the use and establishment of modular and/or existing plants to synthesize and separate a variety of both more established and novel recombinant proteins. In addition, operation scheduling has been recently recognized as a necessary tool for the efficient design of biotechnological processes (Crougham et al., 1997).

The main host for recombinant proteins for many years has been *Escherichia coli*. However, the developments with yeast cells have grown at a very rapid pace, which has resulted in several important commercial products such as insulin, hepatitis B vaccine, and also, more recently, chymosin. The fact that many recombinant proteins made in yeast can be made to be secreted out of the cell and that yeast allows for at least partial glycosilation is an added bonus for this host. Therefore, in this paper we are exploring the use of the tools of process design optimization to investigate the behavior of a multiproduct batch plant to be used for the production of four recombinant proteins synthesized in yeast.

Robinson and Loonkar (1972) studied the problem of designing multiproduct plants operating in single product campaign mode and with a single unit in each processing stage. Sparrow et al. (1975) extended the nonlinear programming model to include both the design of discrete equipment sizes and the selection of the number of parallel units, by solving it through the use of heuristics and branch and bound. The same problem was further formulated by Grossmann and Sargent (1979) as a mixed integer nonlinear programming (MINLP) model. Knopf et al. (1981) and Yeh and Reklaitis (1987) accounted for the presence of semicontinuous units. Voudouris and

* e-mail: juasenjo@cec.uchile.cl.

10.1021/bp990145d CCC: $19.00
Published on Web 02/18/2000

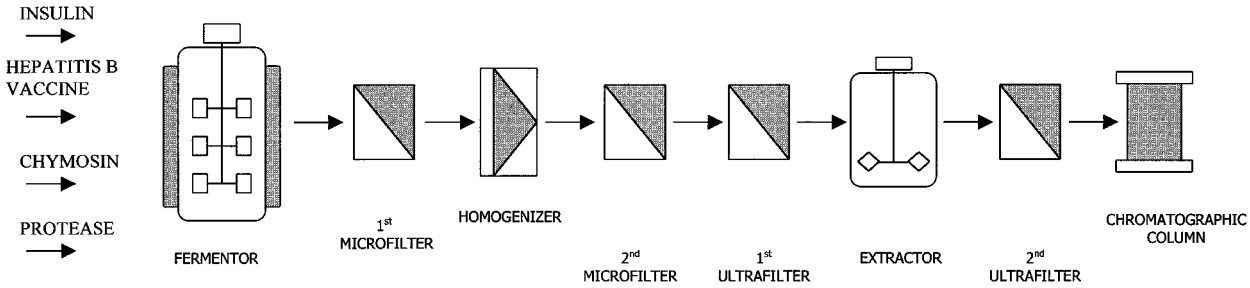

Figure 1. Flowsheet of the batch plant for the production of proteins.

Grossmann (1992) proposed reformulations of the previous design models where discrete sizes are explicitly accounted for.

Most of the unit models in the literature on batch process design are based on expressions that relate the batch sizes linearly with the equipment sizes. Also, the processing times are usually expressed as nonlinear functions of the batch size. Given certain restrictions on these mathematical expressions, the models can be referred to as posynomials, which possess a unique optimum (Grossmann and Sargent, 1979). Salomone and Iribarren (1992) proposed posynomial models in which the constants are obtained as a result of the optimization of the process decision variables with simplified models. Salomone et al. (1994) generalized the approach by allowing the process parameters to be generated from either experimental data and/or dynamic simulation.

In this work we propose an optimization model for the design of a biotechnological multiproduct batch plant. A first level detailed posynomial model is constructed for each unit of the plant. A particular feature of this model is that it contains composite units where semicontinuous items operate on the material contained by batch items. This occurs in the purification steps, in particular with the microfilters operating between retentate and permeate vessels, and with the homogenizer and ultrafilters operating on the material contained in a batch holding vessel. For these units we generalize the approach in Salomone et al. (1994), which uses batch operating time expressions that depend on both the batch size and the size of semicontinuous items, extending it to allow more than one batch unit per stage. In the posynomial models presented in the literature, time is only allowed to depend on batch sizes. The model presented is general, it takes into account all the available options to increase the efficiency of the batch plant design: unit duplication in-phase and out-of-phase and intermediate storage tanks.

The outline of this paper is as follows. First, we describe the protein production plant in terms of the products, units, and their topology. Then, unit models are described, as well as the parameters and data generated from the plant information. On the basis of the unit models we propose an optimization model for the design of the multiproduct batch plant. We briefly describe the solution method and illustrate the model with an example problem in which parallel units and intermediate storage tanks are considered.

Process Description

Products. Figure 1 shows the flowsheet of a multiproduct batch plant for the production of recombinant proteins. Although there are differences in most process flowsheets for recombinant proteins depending mainly on the product characteristics, we have made an attempt to "standardize" such a process. We have even included a liquid−liquid extraction step as an initial separation/purification stage. This separation has been very successful in the initial purification of many proteins and enzymes, including hydrophobic ones (chymosin and α-amylase) (Hayenga et al., 1991; Schmidt et al., 1994), recombinant ones such as IGF-1 (Hart et al., 1994, Andrews et al., 1991) ,and recombinant protein particles synthesized in yeast (Andrews et al., 1995). The products involved in the plant are insulin, hepatitis B vaccine, chymosin, and a cryophilic protease.

Human insulin is a therapeutic protein used in the treatment of diabetes. Hepatitis B vaccine is an engineered recombinant protein that has gained tremendous acceptance and is produced by yeast fermentation (in small plants) worldwide. Chymosin is the bovine rennet protease used for clotting casein in the cheese industry and has successfully been cloned in several different microorganisms including yeasts. A cryophilic protease is presently under development. More than one such protease has been extracted from marine sources, and they are presently being cloned in different microorganisms including yeasts (Andrews et al., 1999). The aim is to produce proteases with high activity at low temperatures to be used in detergents and as wound debriding agents.

Insulin and hepatitis B vaccine are therapeutic proteins that will require further purification steps. This common multiproduct plant will produce technical grade insulin and vaccine. Chymosin is a food product, which will require much less purification. The same is true for the cryophilic protease to be used in detergents and as a wound debriding agent. An estimation of production targets and product prices for the products is shown in Table 1, based on published data (Datar and Rosen, 1990; Petrides et al., 1996; Andrews et al., 1999).

Process. All proteins are produced as the cells grow in the fermenting stage. Because the vaccine and the protease have been considered to be intracellular, the first microfiltration step is used to concentrate the cell suspension, which is afterwards sent to the homogenization stage for cell wall disruption to liberate the intracellular proteins. The second microfiltration stage is used to remove the cell debris from the solution of proteins.

The ultrafiltration stage prior to the extraction is used for concentrating the solutions in order to minimize the extractor volume. In the liquid−liquid extractor, salt concentration is manipulated to first drive the product to a poly(ethylene glycol) (PEG) phase and back again into aqueous saline solution. In this process, many of the proteins other than the product are removed (Huenupi et al., 1999).

Ultrafiltration is used again for concentrating the solution before the chromatographic step in case this is required, and finally the last stage is chromatography in which selective binding is used to further separate the product of interest from other proteins.

Table 1. Product Prices and Demands

product	name	production (kg/year)	price (dollars/kg)
1	insulin	1500	8000
2	vaccine	1000	7500
3	chymosin	3000	1000
4	protease	6000	500

Insulin and chymosin are extracellular. They are present in the permeate from the first microporous filtration membrane unit where the cells are removed. To reduce the amount of valuable product lost in the retentate, extra water is added to the cell suspension. The filtration operation with make up of water is also called diafiltration and dilutes the solution of proteins.

These extracellular products skip the homogenization and microfiltration (for cell debris removal) stages, but then ultrafiltration is necessary to concentrate the dilute solution prior to extraction. The final steps of extraction, ultrafiltration, and chromatography are similar to those of the intracellular products.

Units. The following are the units used in each processing stage.

Fermentors. All products are proteins produced by genetically engineered *Saccharomices cerevisiae* (baker's yeast). Maintenance of the appropriate recombinant strains, cell propagation, and inoculum preparation are similar in all four cases; hence, this has not been included in the present evaluation. Such a simplification can be well justified. Therefore, the first stage is fermentation for the production of yeast and the recombinant protein.

Microfilters. These are usually tangential flow filtration units that use relatively large cross sectional areas. They are used here to separate cells or cell debris from the liquid that is able to cross the membrane orifices and is called permeate. The cells or debris remains in a concentrated suspension, which is known as retentate (Asenjo, 1990a; Asenjo and Patrick, 1990; Zeman and Zydney, 1996). This unit is used in place of conventional filters that produce a cake of solids. The disadvantage of the conventional filters is that the solids involved in the process are quite compressible, which makes pressure drop through the cake ineffective to accelerate the filtration. Another alternative is the use of a centrifuge; however, there are may reasons to choose a microfiltration unit for this purpose (Asenjo and Patrick, 1990).

Homogenizer. These units break cell walls by shear stress produced by pressure drop. The breakage is produced by forcing the flow of a concentrated cell suspension through a valve with a very large pressure drop over a very narrow gap (Asenjo, 1990a; Asenjo and Patrick, 1990). This is used to liberate the intracellular proteins or protein particles. This would be the case for the vaccine and the protease.

Ultrafilters. These units are similar to microfilters, but filtration membrane pores are much smaller. They are used to separate the proteins (in the retentate) from smaller size molecules and the liquid in which they are dissolved. The objective of these dewatering stages is to reduce the size requirement of the expensive purification stages such as extraction and chromatography.

Extractor. These are stirred tanks in which two nonmiscible liquids are mixed so that the products originally in one phase distribute between both (and preferentially partition to the other phase while the main contaminants do not). Afterwards the stirrer is stopped to allow droplets of dispersed phase to settle out. In this particular case, we contact two aqueous phases: a saline phase and a PEG-rich phase. This aims to separate the protein of interest from others in solution.

Table 2. Size Factors S_{ij} (r, retentate; p, permeate)

stage j	unit	S_{ij} (m^3/kg) insulin	vaccine	chymosin	protease
1	fermentor	1.250	0.625	0.415	0.3125
2	microfilter I	r: 1.25 p: 2.5	r: 0.625 p: no	r: 0.415 p: 0.830	r: 0.3125 p: no
3	homogenizer	no	0.155	no	0.08
4	microfilter II	no	r: 0.155 p: 0.31	no	r: 0.08 p: 0.16
5	ultrafilter I	2.50	0.31	0.830	0.16
6	extractor	0.40	0.20	0.135	0.10
7	ultrafilter II	0.40	0.20	0.135	0.10
8	chromatographer	0.05	0.05	0.05	0.05

The following consecutive steps obtain the purification. To the aqueous mixture of proteins are added a phase-forming salt-rich water phase (sulfate or phosphate) and a polymer-rich water phase . In many cases a salt that will induce partition of the product protein (e.g., NaCl) into the other (PEG) phase is also included in the salt-rich water phase. Both phases are vigorously mixed for a few seconds, and as surface tension between them is very small a good dispersion is quickly obtained (Mistry et al., 1996). Settling separates the phases and the saline solution (sulfate or phosphate phase) with most of the contaminating proteins and a large fraction of the NaCl is discarded. The PEG solution with the protein of interest is mixed with a new saline solution (sulfate or phosphate), with no NaCl, that will allow the back-extraction of the protein into the salt phase (Huenupi et al., 1999). The phases are separated as a saline phase that contains the purified product and a PEG phase that is recycled.

Chromatography Column. This purification stage is based on the selective binding capacity of column packing that depends on each particular protein. The column is first fed with the solution containing the protein of interest, which is bound to the packing. Then the column is fed with an eluent stream with a different salt concentration or pH to recover the product. Finally, a washing regeneration stream is passed through the column to return the packing to its initial state.

Stage Models

In this section we describe the unit models from a conceptual standpoint and also the procedure to derive the data needed for solving the mathematical model. These data are summarized in Tables 2 and 3. Most of the separation processes information is taken from Asenjo (1990b) and Asenjo and Patrick (1990), and the posynomial modeling approach is taken from Salomone and Iribarren (1992).

The general batch process literature (Biegler et al., 1997) describes batch stages j through a sizing equation and a cycle time that are applied for a product i as follows:

$$V_j \geq S_{ij} B_i \tag{1a}$$

$$T_{ij} = \text{constant} \tag{1b}$$

where V_j is the size of stage j, e.g., m^3 of the vessel; B_i is the batch size for product i, e.g., kg of product exiting from the last stage; S_{ij} is the size factor of stage j product i, i.e., the size needed at stage j to produce 1 kg of final product i; and T_{ij} is the time required to process a batch of product i in stage j.

Table 3. Time Factors T_{ij} (B_i: kg)

stage j	unit	T_{ij} (h) insulin	vaccine	chymosin	protease
1	fermentor	24	24	24	24
2	microfilter I	12.5 A^{-1} B_i	2.5 A^{-1} B_i	4.15 A^{-1} B_i	1.25 A^{-1} B_i
3	homogeneizer	no	0.465 cap^{-1} B_i	no	0.24 cap^{-1} B_i
4	microfilter II	no	3.1 A^{-1} B_i	no	1.6 A^{-1} B_i
5	ultrafilter I	105 A^{-1} B_i	5.5 A^{-1} B_i	35 A^{-1} B_i	3 A^{-1} B_i
6	extractor	1.5	1.5	1.5	1.5
7	ultrafilter II	18 A^{-1} B_i	8 A^{-1} B_i	4.75 A^{-1} B_i	3 A^{-1} B_i
8	chromatographer	0.5	0.5	0.5	0.5

Consider the fermentor and the insulin product as an example. If we estimate a final concentration of 50 kg dry biomass/m^3, that 0.4 of this biomass is proteins and 0.05 of these proteins is insulin, and an overall yield estimate of the process of 0.8 (0.8 of the insulin produced in the fermentor exits the chromatographic column), then the size factor for the fermentor for producing insulin can be estimated as

$$S_{ij} = \frac{\mathrm{m}^3}{50\ \mathrm{kg} \times 0.4 \times 0.05 \times 0.8} = 1.25\ \mathrm{m}^3/\mathrm{kg} \quad (2)$$

Similarly, vaccine, chymosin, and cryophilic protease were estimated to be 0.1, 0.15, and 0.2 of total proteins of the biomass, respectively.

The batch stage description is completed by estimating a processing time T_{ij} for stage j when producing product i. For the fermentor, we estimate $T_{ij} = 24$ h for all products, which includes time for charging, cell growth, and discharging.

This model of batch stages given by constraint 1 is the simplest one. Its level of detail suffices for the fermentor and the extractor. These units are truly batch items that hold the load to be processed and whose operations are governed by kinetics, and hence, the operating time does not depend on the batch size.

As a first approximation for the extractor, we take a phase ratio of 1 for all products. Therefore, the required extractor volume is twice the inlet batch volume, while the inlet and outlet aqueous saline batches are of the same volume. It is also assumed, as a result of preliminary balances, that this operation reduces the total amount of proteins to about twice the amount of the target protein. With respect to the kinetic effects we take as first estimates (Mistry et al., 1996) the following times: 15 min stirring to approach phase equilibrium, 30 min settling to get almost complete disengaging of the phases, and 20 min for charging and discharging.

A special consideration must be done in the case of the microfiltration, homogenization, and ultrafiltration stages. Although the mathematical model considers them batch stages, their corresponding equipment consists of holding vessels and semicontinuous units that operate on the material that is recirculated into the holding vessel. The batch items are sized as described before. For example, for the homogenizer processing cryophilic protease, we estimated that the fermentor broth is concentrated 4 times up to 200 kg/m^3 at microfilter 1 and considered a yield of 1 because the intracellular protease is fully retained at the microfilter. Then the size factor of the homogenizer vessel is 4 times smaller than the fermentors, i.e., $S_{ij} = 0.08$ m^3/kg protease.

The sizing equation for semicontinuous items can also be found in the general batch processes literature (Reklaitis et al., 1983):

$$R_j = D_{ij}\frac{B_i}{\theta_{ij}} \quad (3)$$

where R_j is the size of the semicontinuous item k, usually a rate of processing. For example, in the case of the homogenizer, it is the capacity in cubic meters of suspension per hour, but in the case of the filters R_j is their area of filtration A_j (m^2). B_i is again the batch size, θ_{ij} is the operating time that the semicontinuous item j needs to process a batch of product i, and D_{ij} is the duty factor (a size factor for semicontinuous items), i.e., the size needed at stage j to process 1 kg of product i in 1 h. For example, if we adopt three passes through the homogenizer, its duty factor is the vessel size factor 0.08 m^3/kg $\times$ 3, i.e., $D_{ij} = 0.24$ m^3/kg. The meaning of a capacity of 0.24 m^3/h is that it allows 1 kg of final product cryophilic protease to be processed in 1 h.

The general batch processes literature considers semicontinuous units to work in series with batch units so that their operating time are the times for filling or emptying the batch units. However, in the process considered, pumps are the only semicontinuous units, which transfer batches between the units. As the pumps' cost does not have a relevant impact on the plant design, they were not explicitly modeled. The times for filling and emptying batch items were estimated and included in the batch cycle times.

On the other hand, the process does have special semicontinuous units with an important economic impact on the cost. They are the homogenizer and ultrafilters, but their operating time is the batch processing time of the respective stage. These types of aggregated units are shown in Figure 2. Their mathematical model has been introduced by Salomone et al. (1994). A size factor for the batch item and a time expression for the stage that depends on both the batch size and the size of the semicontinuous item are as follows:

$$V_j \geq S_{ij}\, B_i \quad (4a)$$

$$T_{ij} = T^0_{ij} + T^1_{ij}\frac{\mathrm{B}_i}{\mathrm{R}_j} \quad (4b)$$

where R_j refers to the size of the semicontinuous item that operates on the batch size at stage j. T^0_{ij} and T^1_{ij} are appropriate time factors that take into account contributions to the total cycle time of the stage that are either fixed amounts of time or proportional to the batch size and inversely proportional to the size of the semicontinuous item.

For the homogenizer, R_j is its capacity, T^1_{ij} the duty factor of the homogenizer itself, and T^0_{ij} includes the estimated times for filling and emptying the homogenizer holding vessel.

In the case of ultrafilters, a fixed permeate flux model was considered with a rate of 20 L/m^2 of membrane

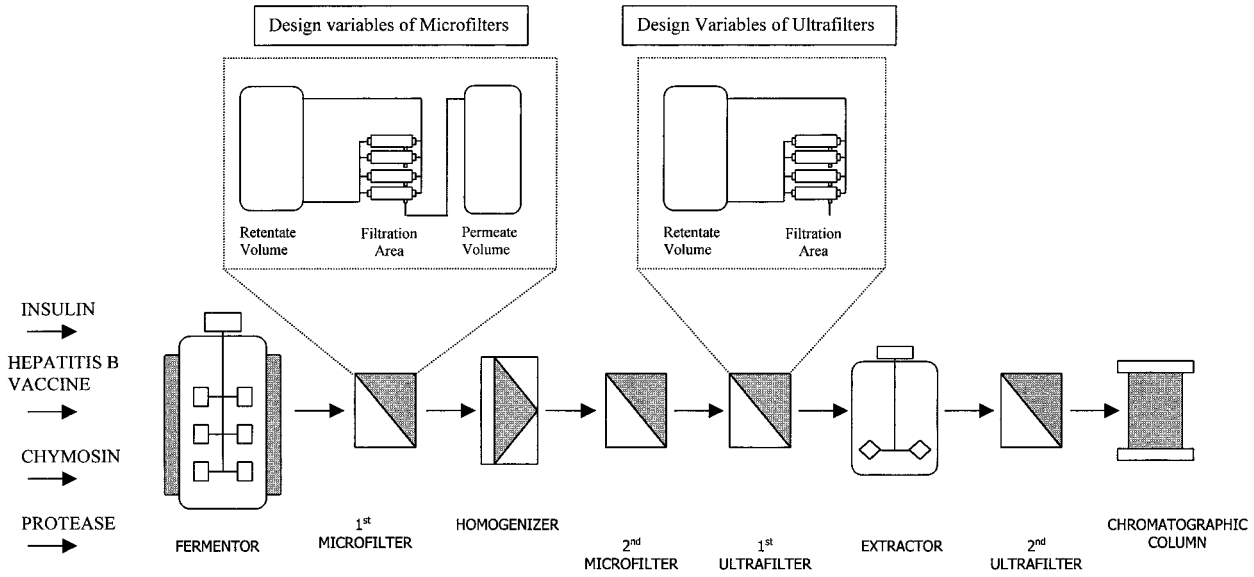

Figure 2. Flowsheet of the proteins batch plant showing the aggregate units.

Table 4. Cost of Equipment (U.S. dollars)

unit	size	cost
fermentor	V_j (m^3)	63400 $V^{0.6}$
micro- and ultrafilters	$V_{retentate}$ (m^3)	5750 $V_r^{0.6}$
	$V_{permeate}$ (m^3)	5750 $V_r^{0.6}$
	A_{filter} (m^2)	2900 $A^{0.85}$
homogenizer	$V_{holding}$ (m^3)	5750 $V^{0.6}$
	Cap (m^3/h)	12100 $cap^{0.75}$
extractor	V_{extr} (m^3)	23100 $V^{0.65}$
chromatography	V_{chrom} (m^3)	360000 $V^{0.995}$

area/h (Asenjo, 1990a). In this case, the size of the semicontinuous item R_j is the filtration area. T_{ij}^0 is again the time for filling and emptying the retentate holding vessel, and T_{ij}^1 is the inverse of the permeate flux times the ratio (m^3 permeate/kg of product). This ratio is estimated from a mass balance taking into account that the ultrafilters are used for water removal from solutions up to 50 g/L of total proteins. Ultrafilters are used to reduce the volume required at the liquid extractor and the chromatographic column. The upper bound on concentration is a constraint that avoids protein precipitation.

The microfilter model is quite similar to that of the ultrafilter, but there are two batch items associated to them instead of one, the retentate and the permeate vessels, plus the semicontinuous item area of filtration.

For microfilter 1 a fixed permeate flux of 200 L/m^2 h is adopted. For extracellular insulin and chymosin, we estimate a total permeate (feedwater plus make up water) twice the feed, while for intracellular protease and vaccine we estimate it in 75% of the feed (the retentate is concentrated four times).

For microfilter 2 a fixed permeate flux model is also used. In this case, the flux is smaller than the one in microfilter 1 because the pore size to retain cell debris is smaller than the one for whole cells. As a first estimation we take 100 L/m^2 h and a total permeate (feed plus make up water) twice the feed.

With respect to the chromatographic column, an adsorptive type chromatography is considered, with a binding capacity of 20 kg/m^3 of column packing. The size factor of this unit is the inverse of that binding capacity. As a first approximation, a fixed total operating time of 0.5 h was estimated for loading, eluting, and washing regeneration.

Finally, the stage model is completed with a cost model that expresses the cost of each unit as a function of its size, in the form of a power law. These expressions are summarized in Table 4, with most of the cost data taken from Petrides et al. (1995).

Optimization Model for the Design of the Multiproduct Batch Plant

In this section, we describe the mathematical optimization model for designing the multiproduct batch plant. The model includes the stage models described in the previous section plus additional constraints that are explained in this section. The plant consists of M batch stages (in our case 8 batch stages). Each stage j has a size V_j (m^3), and more than one unit can be installed in parallel. They can work either in-phase (starting operation simultaneously) or out-of-phase (starting times are distributed equally spaced between them).

The duplication in-phase is adopted in case the required stage size exceeds the specific upper bound. In this case G_j units are selected, splitting the incoming batch into G_j smaller batches, which are processed simultaneously by the G_j units. After processing, the batches are added again into a unique outgoing batch. Otherwise, duplication out-of-phase is used for time-limiting stages; if a stage has the largest processing time, then it is a bottleneck for the production rate. Assigning M_j units at this stage, working in out-of-phase mode, reduces the limiting processing time and thus increases the production rate of the train. For this case, the batches coming from the upstream stages are not split. Instead, successive batches produced by the upstream stage are received by different units of stage j, which in turn pass them at equally spaced times onto the downstream batch stage.

The allocation and sizing of intermediate storage has been included in the model to get a more efficient plant design. The goal is to increase unit utilization. The insertion of a storage tank decouples the process into two subprocesses: one upstream from the tank, and the other downstream. This allows the adoption of independent batch sizes and limiting cycle times for each subprocess. Therefore, the previously unique B_i is changed to batch sizes B_{ij} defined for product i in stage j. Appropriate constraints adjust the batch sizes among different units.

The objective is to minimize the capital cost of the plant. The decision variables in the model are as fol-

lows: at each batch stage the number of parallel units in-phase and out-of-phase and their size, and the installation or absence of intermediate storage between the batch stages and their size. The plant is designed to satisfy a demand of Q_i (kg) of each product i, for the P products considered, within a time horizon H (h).

In summary, the objective function to be optimized is

$$\text{Min Cost} = \sum_{j=1}^{M} M_j G_j a_j V_j^{\alpha_j} + \sum_{j=1}^{M} VT_j^{\eta_j} \tag{5}$$

where a_j and α_j, c_j and η_j are appropriate cost coefficients that depend on the type of equipment being considered. VT_j is the size of the storage tank allocated after stage j.

The size of each unit has to be large enough to be able to process every product:

$$V_j \geq \frac{S_{ij} B_{ij}}{G_j} \quad \forall i = 1, ..., P; \forall j = 1, ..., M \tag{6}$$

where S_{ij} is the size factor for product i in stage j. In case of parallel units working in-phase, the division of B_{ij} by the number of units G_j takes into account the reduction in the batch size to be processed by these units.

The operating time T_{ij} to process product i at stage j has the general following form:

$$T_{ij} = T_{ij}^0 + T_{ij}^1 \frac{B_{ij}}{R_j} \quad \forall i = 1, ..., \text{P}; \forall j = 1, ..., \text{M} \tag{7}$$

where T_{ij}^0 and T_{ij}^1 are appropriate constants that depend on both the product and the stage. Expression 7 accounts for a fixed and variable contribution to the total operating time. The last term in eq 7 depends on both the batch size and the size of the semicontinuous item associated to this batch stage, as was already discussed in the previous section.

The limiting cycle time for product i in the subprocess h, $TL^h{}_i$, is the largest processing time in this production train:

$$TL_i^h \geq \frac{T_{ij}}{M_j} \quad \forall i = 1, ..., \text{P}; \forall j \in J_j; \forall h \tag{8}$$

where J_h is the set of units which conform the subprocess h.

The division by the number of units in parallel working out-of-phase, M_j, takes into account the reduction in the cycle time of this stage, due to the operation of M_j units that alternatively process the consecutive batches.

To avoid accumulation of material, the processing rate of both subprocesses downstream and upstream of the storage tank must be the same:

$$\left(\frac{B_i^d}{TL_i^d}\right) = \left(\frac{B_i^u}{TL_i^u}\right) \quad \forall i = 1, ...,P \tag{9}$$

Constraint 9 equalizes the production rate upstream and downstream of the storage tank. To express 9 in a simpler form, the inverse of the production rate of product i (E_i), is defined as

$$E_i = \frac{TL_i^h}{B_{ij}} \quad \forall i = 1, ..., P; \forall j \in J_h; \forall h \tag{10}$$

Expression 10 is used to replace TL_i^h in constraint 8, dropping constraint 9.

The production constraint is posed as follows: during the time horizon H the plant must produce the target production quantities Q_i of each product i. The number of batches of each product i to be produced during time H is Q_i/B_i, and the production of each batch demands a time TL_i. The following constraint holds:

$$\sum_{i=1}^{P} Q_i E_i \leq H \tag{11}$$

The size of the storage tank VT_j, allocated after batch stage j, is given by the following expression (Modi and Karimi, 1989):

$$VT_j \geq ST_{ij}(B_{ij} + B_{i,j+1}) \quad \forall i = 1, ..., P; \forall j = 1, ..., M - 1 \tag{12}$$

where ST_{ij} is the size factor corresponding to the intermediate storage tank, with identical definition to the batch stages.

As no a priori tank allocation is given, binary variables y_j are used to select their allocation. The value of variables y_j is 1 if a tank is placed in position j, or zero otherwise. Constraint 12 is generalized to size the tank only if it exists:

$$VT_j \geq ST_{ij}(B_{ij} + B_{i,j+1}) - F_j(1 - y_j) \quad \forall i = 1, ..., P; \forall j = 1, ..., M - 1 \tag{13}$$

where F_j is a constant value sufficiently large such that when y_j is 0 (the tank does not exist), the constraint is trivially satisfied for any value of VT_j. In particular, the cost minimization will drive $VT_j = 0$. When the tank exists ($y_j = 1$), the term with F_j vanishes, and the original constraint (12) holds.

If the storage tank does not exist between two consecutive stages, then their batch sizes are constrained to be equal. Otherwise, this constraint is relaxed. This effect is imposed by the following constraints (Ravemark, 1995):

$$1 + \left(\frac{1}{\Phi} - 1\right) y_j \leq \frac{B_{ij}}{B_{i,j+1}} \leq 1 + (\Phi - 1) y_j \quad \forall i = 1, ..., P; \forall j = 1, ..., M - 1 \tag{14}$$

where Φ is a constant value corresponding to the maximum ratio allowed between two consecutive batch sizes.

In summary, the multiproduct plant design model that includes the options of parallel units in-phase and/or out-of-phase and provision of intermediate storage, consists of the objective function 5 subject to constraints 6, 8, 11, 13, and 14, plus the upper and lower bounds that may apply.

An important feature of the model is that both the objective function and the constraints are posynomial expressions that possess a unique local (and thus, global) solution (Grossmann and Sargent, 1979).

This basic model has been adapted to handle the particular feature of the composite stages (homogenizer, ultrafilters, and microfilters). In this case, constraint 6 is applied not to a general batch stage size but to each of the items that compose it. So in the case of microfilters, constraint 6 applies to both the retentate and the permeate vessels. A new parameter SR_{ij} was introduced to represent the size factor of the retentate vessel, while

Table 5. Intermediate Storage Cost Coefficients and Size Factors

unit[a]	$ST_{i,j}$ size factor for product i in stage j			
	insulin	vaccine	chymosin	protease
fermentor	1.25	0.625	0.415	0.3125
microfilter I	2.50	0.155	0.83	0.08
homogenizer	2.50	0.155	0.83	0.08
microfilter II	2.50	0.31	0.83	0.16
ultrafilter I	0.40	0.20	0.135	0.10
extractor	0.40	0.20	0.135	0.10
ultrafilter II	0.05	0.05	0.05	0.05
chromatography	0	0	0	0

[a] Tank cost coefficients: $c_j = 5750$; $\eta_j = 0.6$.

S_{ij} was left for the permeate vessel. Also in this case, the objective function must account for all the stage components. The notation a_j and α_j was left for the cost coefficients of the permeate vessel, b_j and β_j for the retentate vessel, and d_j and γ_j for the filtration area. A similar approach was implemented for the ultrafilters (retentate vessel and ultrafiltration area) and homogenizer (holding vessel and the homogenizer itself).

Model Solution Algorithm

The problem has been executed with different alternatives to evaluate the options included in the model. Given that the problem modeled has nonlinear objective function and constraints as well as 0−1 binary variables, the resulting mathematical model is a mixed integer nonlinear program (MINLP). This model includes 104 binary variables and has been convexified using the transformations proposed by Kocis and Grossmann (1988). The MINLP model has been solved using DICOPT++, which is included in the GAMS optimization modeling software (Brooke et al., 1992). The algorithm implemented in DICOPT++ relies on the outer approximation/equality relaxation/augmented penalty (OA/ER/AP) method that was proposed by Viswanathan and Grossmann (1989).

The OA/ER/AP solution method consists of the decomposition of the original MINLP problem into a sequence of two subproblems: a nonlinear programming (NLP) subproblem and a mixed integer linear programming (MILP) subproblem also known as the Master problem. At each iteration of the algorithm one NLP subproblem and one Master subproblem are solved. For the NLP solution the binary variables are fixed, and the value of the objective function obtained is an upper bound for an MINLP minimization problem. The Master subproblem is generated by the cumulative linearization (outer approximation) of the nonlinear equations at the NLP solution point. The Master subproblem solution provides the binary variable values for the following iteration. The objective function value is a lower bound for an MINLP (convex) minimization problem. The method stops when the upper and the lower bound difference is less than or equal to the specified tolerance.

Computational Results

The model developed has been solved with DICOPT++ using the data shown in Tables 1−5. A horizon time of 6000 h has been considered.

Figure 3 shows the optimal solution when no intermediate storage tanks are considered in the model. Table 6 summarizes the optimal sizes and number of parallel units obtained. For this case, five parallel units out-of-phase have been selected for the fermentor. This unit has the limiting cycle time for all products. For the chromatographic column two parallel units in-phase have been allocated. The reason is that the batch sizes obtained for the batch units in the previous stages are greater than the upper bound capacity of the chromatographic column. The selection of parallel units allows the reduction of the idle time for the stage. Table 7 shows the idle times obtained by solving the model without allowing units in parallel. Table 8 summarizes the idle times corresponding to the optimal solution of Figure 3.

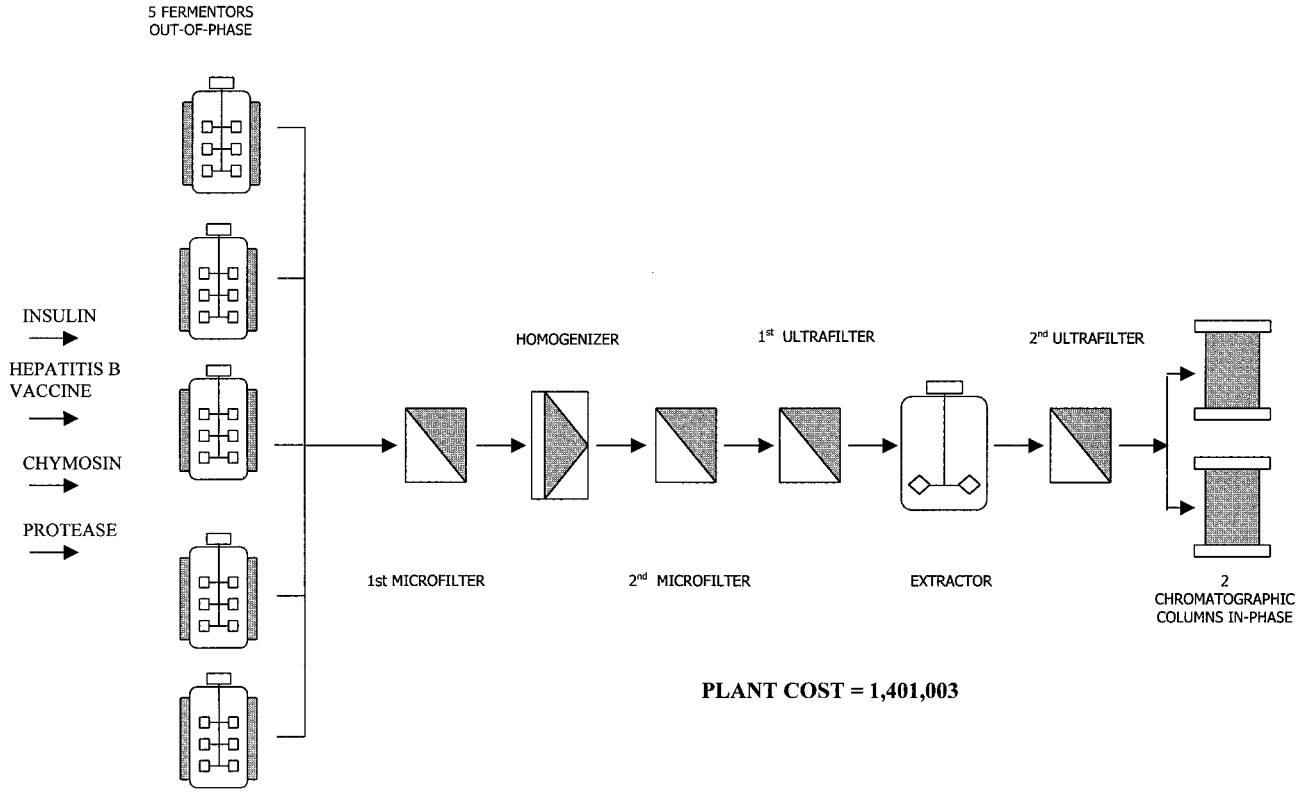

Figure 3. Protein production plant design without intermediate storage.

Table 6. Optimal Solution Problem without Intermediate Storage Tank (r, retentate; p, permeate); Total Cost: $1,401,003

stage	1	2	3	4	5	6	7	8
V_j	4.496	p: 8.992 r: 4.496	1.151	p: 2.302 r: 1.151	8.992	2.167	1.463	0.360
R_j		*A*: 14.741	cap: 0.973	*A*: 7.547	*A*: 99.784		*A*: 14.387	
M_j	5	1	1	1	1	1	1	1
G_j	1	1	1	1	1	1	1	2

Table 7. Idle Times in Plant with No Duplication of Units and No Intermediate Storage Tanks

	unit							
product	1	2	3	4	5	6	7	8
insulin	0	0			0	22.2	0	23.5
vaccine	0	13.85	0	0	20.59	22.2	2.63	23.5
chymosin	0	0.53			0.45	22.2	5.31	23.5
protease	0	16.21	5.48	5.36	21.07	22.2	12.08	23.5

Table 8. Idle Times in Plant with Parallel Units and without Intermediate Storage Tanks

	unit							
product	1	2	3	4	5	6	7	8
insulin	0	0			0.02	3.00	0	4.30
vaccine	0	2.33	0.11	0.10	3.40	3.00	0.50	4.30
chymosin	0	0			0	3.00	0.92	4.30
protease	0	2.33	0	0	3.37	3.00	1.50	4.30

This table shows a considerable reduction in the idle times when compared to the solution without units in parallel.

Figure 4 shows the optimal solution considering both units in parallel and the allocation of intermediate storage tanks. Table 9 presents the sizes for the units. Three tanks have been selected: the first after the fermentor, the second after the first ultrafilter, and the third after the second ultrafilter. The plant investment cost has been considerably reduced (41%). No units in parallel have been allocated in this case. The major reductions in cost are for the fermentor and the chromatographic column. Table 10 shows the idle times for this case. In comparison with Table 8 significant reductions can be observed. The key for this solution is the reduction in cost of the chromatographic column. The provision of intermediate storage previous to this unit decouples the plant operation. This decoupling allows the chromatographic column to work almost full time for all products. It is also obvious that the installation of storage tank, much cheaper than the column, allows a reduction in size and thus in the column capital cost.

In this paper we do not take into account operating costs, as this corresponds to a different hierarchical level. Here we have implemented a constant size and time factor model. To obtain constant factors we have found reasonable values for every process variable so as to cover the degrees of freedom of the process. This model is posynomial and easy to solve, its solution is guaranteed to be a global optimum, etc. The next step in the hierarchy would be to implement process performance models, i.e., the size and time factors will not be constants but functions of the process variables. This model is much harder to solve, and there is yet no global optimization algorithm developed for it. The process performance models would permit the inclusion of the economic impact of process variables. Operating cost is one of such impacts. For example, both the capital and operating costs of the ultrafiltration steps depend on the extent of concentration, but with the constant factors model this is not taken into account. With this model we can optimize batch sizes (which also means distribution of the horizon time among products) and capital costs due to idle times and volumetric underutilization of units (duplication and storage are the decision variables). Usually the economic impact of this structural optimization is much larger (in the order of a 40−50% cost reduction with respect to a reasonable but not optimal structure) than the optimization of the process variables. The optimization of process variables has an impact on the order of a 10−20% cost reduction with respect to reasonable but not optimal values for the process variables. This is the reason it is left as a posterior optimization in the hierarchy.

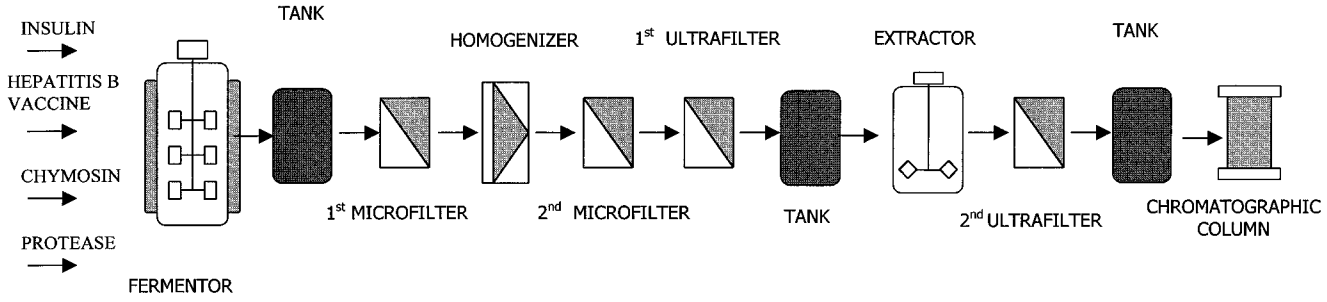

Figure 4. Protein production plant design considering units in parallel and intermediate storage tanks.

Table 9. Optimal Solution Problem with Intermediate Storage Tank (r: retentate; p: permeate); Total Cost: $828,073

stage	1	2	3	4	5	6	7	8
V_j	22.496	p: 9.020 r: 4.510	1.074	p: 2.148 r: 1.074	9.020	0.811	0.547	0.075
R_j		*A*: 14.731	cap: 0.999	*A*: 7.880	*A*: 99.391		*A*: 16.194	
VT_j	27.006				2.014		0.345	
M_j	1	1	1	1	1	1	1	1
G_j	1	1	1	1	1	1	1	1

Table 10. Idle Times in Plant with Parallel Units and Intermediate Storage Tanks

	unit							
product	1	2	3	4	5	6	7	8
insulin	0	0			0	0.01	0	0
vaccine	0	1.93	0.04	0	2.91	0	0.17	0
chymosin	0	0.01			0	0	0.31	0.17
protease	0	2.09	0	0	3.07	0	0.50	0

Conclusions

An optimizaton model for the design of a biotechnological multiproduct plant has been developed, which combines a first level of detail posynomial model for the unit stages, as well as decisions concerning the structural optimization of the plant.

A particular feature of the posynomial unit model is that it takes into account composite units where semicontinuous items operate on the material contained by batch items. This occurs with the microfilters operating between retentate and permeate vessels and with the homogenizer and ultrafilters operating on the material contained in a batch holding vessel. For these units we generalized the approach in Salomone et al. (1994), which uses batch operating time expressions that depend both on the batch size and on the size of semicontinuous items, extending it to allow more than one batch unit per stage. In the posynomial models presented in the literature, time is only allowed to depend on batch sizes.

The model presented is general, and it considers all the available options to increase the efficiency of the batch plant design, such as unit duplication in-phase and out-of-phase and intermediate storage tanks. An attempt has been made to "standardize" a potential multiproduct batch plant for the production of recombinant proteins (therapeutic and nontherapeutic). Although it is difficult to select such a standard process, the model developed is generic and can be applied to any such modular plant, i.e., including a precipitation stage instead of a liquid−liquid extraction.

The resulting mathematical model for the minimization of the plant capital cost is a mixed integer non-linear program that was solved to global optimality. Results were obtained for a plant that produces four recombinant proteins in eight processing stages. Solutions when no intermediate storage is provided indicated that units in parallel in- and out-of-phase are required to fulfill the product demands. Moreover, if allocation of intermediate storage tanks is allowed, it results in a significant reduction in the plant cost (41%) when compared to the previous case, simply by eliminating the need for parallel units.

The level of detail of posynomial models allows optimization of the plant structure and the batch sizes, but does not include the influence of process variables such as the concentration of biomass in the fermentor, the concentration in the ultrafiltration units. or the number of passes in the homogenizer. The next level of hierarchical approach would be to include process performance models to predict the size and time factors as well as operating costs as a function of the process variables. Successive levels in the hierarchy are of increasing detail but of decreasing economic impact: the optimization of process variables will typically render additional savings in the order of 10−20% as compared with the reasonable but not optimal values adopted here to compute the constant factors. On the other hand, the optimization problem addressed in this paper has a larger economic impact, on the order of 40−50% cost reduction with respect to a reasonable but not optimal structure.

Acknowledgment

The authors would like to acknowledge financial support received from VITAE within the Cooperation Program Argentina-Brazil-Chile under grant B-11487/6B004. J. A. Asenjo would also like to thank Fundacion Andes for the donation of all the Advanced Equipment to the Centre of Biochemical Engineering and Biotechnology of the University of Chile.

Notation

a_j	Cost coefficient for a batch unit in stage j.
b_j	Cost coefficient for the retentate vessel of a batch unit at stage j.
B_{ij}	Batch size of product i in stage j.
c_j	Cost coefficient for an intermediate storage tank allocated in position j.
d_j	Cost coefficient for the semicontinuous unit associated to a batch unit at stage j.
E_i	Inverse of production rate of product i.
G_j	Number of batch units in parallel in-phase in stage j.
H	Net available production time for all products.
J_h	Set of units that compose subprocess h.
M	Number of stages in the plant.
M_j	Number of batch units in parallel out-of-phase in stage j.
P	Number of products.
Q_i	Production requirement of product i.
R_j	Size of semicontinuous item j; A_j area in case of filters, Cap_j capacity in the homogenizer
S_{ij}	Size factor for product i for a batch unit at stage j.
SR_{ij}	Size factor for product i for a batch unit at stage j in the retentate vessel.
ST_{ij}	Size factor for product i for an intermediate storage tank in the location j.
T_{ij}	Processing time of product i at batch stage j.
T_{ij}^0	Coefficient in the expression for T_{ij}.
T_{ij}^1	Coefficient in the expression for T_{ij}.
TL_i	Limiting cycle time of product i.
V_j	Size of a batch unit at stage j.
VT_j	Size of the intermediate storage tank allocated in position j.
y_j	Binary variable that indicates if a tank is allocated in position j.

Greek letters

α_j	Cost exponent for a batch unit at stage j.
β_j	Cost exponent for the retentate vessel of a batch unit at stage j.
γ_j	Cost coefficient for the semi-continuous unit associated to a batch unit at stage j.
η_j	Cost exponent for an intermediate storage tank allocated in position j.
Φ	Maximum allowed ratio between consecutive batch sizes.

References and Notes

Andrews, B. A.; Nielsen, S.; Huang, R.; Hodgson, C.; Asenjo, J. A. Partitioning, Separation and Purification of Three Recombinant Proteins in Aqueous Two-Phase Systems. Presented

at the 7th International Conference on Partitioning in Aqueous Two Phase Systems, New Orleans, LA, 1991.

Andrews, B. A.; Huang, R. B.; Asenjo, J. A. Purification of virus like particles from yeast cells using aqueous two phase systems. *Bioseparation* **1995**, *5*, 105−112.

Andrews, B. A.; Salamanca, M.; Barria, C.; Achurra, P.; Thaysen, M.; Mancilla, M.; Asenjo, J. A. Purification, Characterization and Process Considerations of Cryophilic Proteases of Marine Origin. Presented at the Biochemical Engineering XI Conference (United Engineering Foundation), Salt Lake City, UT, 1999.

Asenjo, J. A. Cell Disruption and Removal of Insolubles. In *Separations for Biotechnology II*; Pyle, D. L., Ed.; Elsevier: New York, 1990a; p 11−20.

Asenjo, J. A., Ed. *Separation Processes in Biotechnology*; Marcel Dekker: New York, 1990b.

Asenjo, J. A.; I. Patrick; Large Scale Protein Purification. In *Protein Purification Applications: A Practical Approach.* E. Harris, E., Angal, S., Eds.; IRL Press: Oxford, 1990; Chapter I, p 1−28.

Biegler, L. T.; Grossmann, I. E.; Westerberg, A. W. *Systematic Methods of Chemical Process Design*; Prentice Hall: New Jersey, 1997.

Brooke A.; Kendrick, D.; Meeraus, A. GAMS. Release 2.25. User's Guide. The Scientific Press: San Francisco, 1992.

Crougham, M.; Caldwell, V.; Randlev, B.; Billeci, K.; Nieder, M. Prediction of Culture Performance through Cell Cycle Analysis: A Potential Tool in Operations Scheduling. Presented at the Biochemical Engineering X Conference, Kananaskis, Canada, 1997.

Datar, R.; Rosen, C. G. Downstream Process Economics in Separation Processes in Biotechnology. Asenjo, J. A., Ed.; Marcel Dekker: New York, 1990; pp 741−793.

Grossmann, I. E.; Sargent, R. W. H. Optimal Design of Multipurpose Chemical Plants. *Ind. Eng. Chem. Process Des. Dev.* **1979**, *18*, 343−348.

Hart, R. A.; Lester, P.; Reyfsneyder, D. H.; Ogez, J. R.; Builder, S. E. Large Scale in situ Isolation of Periplasmic IGF-1 from *E. coli*. *Biotechnology* **1994**, *12*, 1113−1119.

Hayenga, K.; Murphy, M.; Arnold, R.; Lorch, J.; Heinsohn, H. Application of Two-Phase Liquid −Liquid Extraction to the purification of Calf Chymosin from *A. awamori*. Presented at the 7th International Conference on Partitioning in Aqueous Two Phase Systems, New Orleans, LA, 1991.

Huenupi, E.; Gomez, A.; Andrews, B. A.; Asenjo, J. A. Optimization and Design Considerations of Two-Phase Continuous Protein Separation *J. Chem. Technol. Biotechnol.* **1999**, *74*, 256−263.

Knopf, O. C.; Okos, M. R.; Reklaitis, G. V. Optimal Design of Batch/Semicontinuous Processes, *Ind. Eng. Chem. Process Des. Dev.* **1982**, *21*, 79−86.

Kocis, G. R.; Grossmann, I. E. Global Optimization of Nonconvex MINLP Problems in Process Synthesis. *Ind. Eng. Chem. Res.* **1988**, *27*, 1407.

Mistry, S. L.; Kaul, A.; Merchuk, J. C.; Asenjo, J. A. Mathematical Modeling and Computer Simulation of Aqueous Two Phase Continuous Protein Extraction. *J. Chromatogr. A* **1996**, *741*, 151−163.

Modi, A. K.; Karimi, I. A. Design of Multiproduct Batch Processes with Finite Intermediate Storage. *Comput. Chem. Eng.* **1989**, *13*, 127−139.

Petrides, D.; Sapidou, E.; Calandranis, J. Computer Aided Process Analysis and Economic Evaluation. *Biotech. Bioeng.* **1995**, *48*, 529−541.

Ravemark, D. Optimization Models for Design and Operation of Chemical Batch Processes. Ph.D. Thesis, Swiss Federal Institute of Technology, Zurich, 1995.

Reklaitis, G. V.; Ravindran, A.; Ragsdell, K. M. *Engineering Optimization Methods and Applications*; John Wiley: New York, 1983.

Reklaitis, G. V. Overview of Scheduling and Planning of Batch Process Operations. NATO Advanced Study Institute−Batch Process Systems Engineering, Antalya, Turkey, 1992.

Robinson, J. D.; Loonkar, Y. R. Minimizing Capital Investment for Multiproduct Batch Plants. *Process Technol. Int.* **1972**, *17*, 861−863.

Salomone H. E.; Iribarren, O. A. Posynomial Modeling of Batch Plants. *Comput. Chem. Eng.* **1992**, *16*, 173−184.

Salomone, H. E.; Montagna, J. M.; Iribarren, O. A. Dynamic Simulations in the Design of Batch Processes. *Comput. Chem. Eng.* **1994**, *18*, 191−204.

Schmidt, A. S.; Ventom, A. M.; Asenjo, J. A. Partitioning and Purification of α-amylase in aqueous two-phase systems. *Enzyme Microb. Technol.* **1994**, *16*, 131−142.

Viswanathan, J. V.; Grossmann, I. E. A Combined Penalty Function and Outer Approximation Method for MINLP Optimization. *Comput. Chem. Eng.* **1990**, *14*, 769−782.

Voudouris, V. T.; Grossmann, I. E. Mixed Integer Linear Programming Reformulations for Batch Process Design with Discrete Equipment Sizes. *Ind. Eng. Chem. Res.* **1992**, *31*, 1315−1325.

Yeh, N. C.; Reklaitis, G. V. Synthesis and Sizing of Batch/Semicontinuous Processes. *Comput. Chem. Eng.* **1987**, *11*, 639−654.

Zeman, L. J.; Zydney, A. L. *Microfiltration and Ultrafiltration: Principles and Applications*; Marcel Dekker: New York, 1996.

Accepted December 10, 1999.

BP990145D