

# Use of expert systems for the synthesis of downstream protein processes

M. Elena Lienqueo \*, Juan A. Asenjo

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

## Abstract

This paper describes recent developments in rational process design and their application in biotechnology for large-scale downstream processes. For implementing an expert system, it is necessary to divide the separation process in two parts: recovery and purification, because the information and available heuristic knowledge are different in each part. In the case of the recovery process, the implementation of an expert system, called Prot\_Ex, only use heuristic rules from literature and human experts. The sequences suggested were tested with a real example and it gave a satisfactory result. For the purification process two criteria have been defined for selecting the optimal sequences, the SSC criterion and the purity criterion. Both criteria were implemented in an expert system, Prot\_Ex\_Purification. This expert system was validated with experimental examples, in two cases and, the sequences suggested were adequate and valid. However, the sequences suggested by purity criterion have lesser steps than the sequences suggested by SSC criterion and the purity algorithm spends lesser time and computational resources than SSC algorithm, then the use of purity criterion is more recommendable for selection of purification sequences. Finally, we consider the proposed sequences found by the expert system a very good starting point for industrial process selection, a clear case of 'expert amplification'. © 2000 Elsevier Science Ltd. All rights reserved.

*Keywords:* Protein recovery; Protein purification; Downstream processing; Bioseparation; Expert system

## 1. Introduction

In the modern biotechnology industry, after successful fermentation or enzyme reactions the desired product must be separated and purified. The necessary steps to obtain these are commonly known as downstream processing (DSP), which can account for up to 60% of the total cost, excluding the cost of the purchased raw materials (Lee, 1992). DSP is usually composed of a sequence of recovery and purification operations, whose final aim is to obtain the required protein at a prespecified level of purity. The recovery process is characterised by the objective of obtaining the product in solution from the production system. Purification takes the multiprotein solution and purifies the individual protein product to a high purity level, which is generally more than 90% pure. In addition, on

a large scale, it is necessary to obtain the highest possible yield while minimizing the resources utilised and hence the cost (Asenjo, 1990). The main steps in a large-scale protein separation process are usually not more than eight, they are shown in Fig. 1. The principal unit operations used in DSP are shown in Table 1.

At present, the bio-product markets are becoming very competitive. Therefore, it is important to choose the optimal flowsheet as early as possible, since once the process has been approved by regulatory agencies, its characteristics are fixed and are expensive to change. Thus, it becomes necessary to use rational design tools for this purpose. In order to achieve this we can recognise two different tasks:

1. Selection of separation sequence or flowsheet generations.
2. Design of individual unit operations.

In this paper we will center our attention in the first task. Selection of separation sequences is a complex procedure in which a design evolves from a preliminary to a final stage in a trial and error fashion by repeatedly revising and refining the initial assumptions and restric-

\* II Pan American Workshop in Catalysis and Process Systems Engineering, September 2–3, 1999, Santa Fe, Argentina

\* Tel.: +1-405-3255811; fax: +1-405-3255813.

E-mail address: bagajewicz@mailhost.ecn.uoknor.edu (M. Bagajewicz).

tions. For protein recovery processes, the flowsheet generations always have the same type of unit operations (see Table 1). The problem that has to be solved is, e.g. choosing between alternative operations. For example, for cell disruption it is necessary to select between a homogenizer and a bed mill or for cell separation between cross-flow microfiltration and centrifugation. This selection is more or less done using heuristics, i.e. using rules of thumb to arrive at a rapid and reliable specification of equipment type (Leser & Asenjo, 1992). Another way to design is using appropriate mathematical correlations and models for optimising the task.

On the other hand, for the protein purification processes the flowsheet consists of a chromatographic sequence (see Table 1). This sequence has to be satisfied

with maximum yield and a minimum number of steps (1, 2 or 3). This part of the process is undertaken in classical chemical process engineering, finding a rigorous solution using numerical methods like mathematical optimising techniques (Jennings, Teo, Wang & Yu, 1995) or using an expert systems approach (Eriksson, Sandahl, Brewer & Osterlund, 1991; Forslund, 1995). For the selection of an optimal sequence, purely mathematical techniques have limited use in biotechnology because of a lack of useful design equations and databases. The expert systems approach is more attractive because it allows the use of empirical knowledge that is not rigorous in nature and is typical of that used by experts (Asenjo & Maugeri, 1992). In this paper, we show the implementation and validation of an expert system for recovery operations and another one for purification processes.

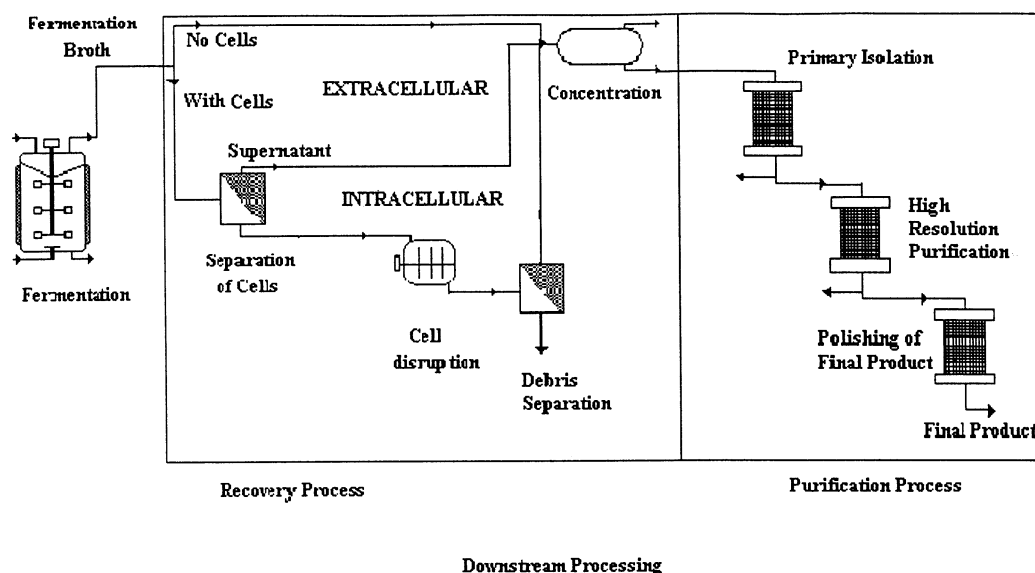


Fig. 1. General flowsheet for downstream processing of proteins.

Table 1  
Principal steps and unit operations used in downstream processing

Stages	Step	Unit operations
<i>Protein recovery</i>		
	Cell separation	Centrifugation, membrane process (filtration, microfiltration), two phase aqueous partitioning
	Cell disruption (only for intracellular proteins)	Homogenisation, bead milling, chemical and enzymatic lysis and permeabilization
	Debris separation (only for intracellular proteins)	Centrifugation, membrane processes (microfiltration, ultrafiltration), aqueous two-phase partitioning
	Concentration	Ultrafiltration, precipitation
<i>Protein purification</i>		
	Pre-treatment or primary isolation	Batch adsorption, ion exchange chromatography, affinity adsorption, hydrophobic interaction chromatography, aqueous two-phase partitioning
	High resolution purification	Hydrophobic interaction chromatography (HIC), high resolution ion-exchange chromatography, affinity chromatography, metal chelate chromatography (IMAC)
	Polishing of final product	dye-ligand chromatography HPLC Gel filtration (GF) HPLC, ion-exchange chromatography

Table 2  
Potential constraints

Constraints	Observations
Range of pH	The product is only stable in a range of pHs (e.g. 4–9)
Range of temperature	The product is stable only up to certain temperature
Proteases	Elimination of proteases, because they could degrade the protein product
Inclusion bodies	Inclusion bodies have to be solubilized and refolded

## 2. General aspects for designing a protein separation process

The design of a protein recovery and purification process shares many characteristics with other engineering design activities. To design a process or an operation requires:

1. To satisfy of a number of constraints, such as purity, quality, process temperature and pH, desired yield, etc.
2. To know properties of the materials, for instance chemical and biochemical properties, thermodynamic and fluid properties of the process material.

Using the information and knowledge of (1) and (2) to propose a sequence of equipment interconnected in a particular order.

## 3. Basic information for designing a separation process

The design of a process to economically purify a protein, maintaining a high yield, yet obtaining a virtually pure product while minimizing the cost, requires four main considerations:

1. Defining final product.
2. Characterisation of starting material.
3. Define possible separation steps and constraints.
4. Evaluation of possible process integration.

### 3.1. Defining final product

It is necessary to define the final product and to have information on its used. For instance:

1. How the product is going to be used?
  - Industrial.
  - Diagnostic.
  - Laboratory reagent.
  - Therapeutic.

Questions regarding the use and purity are vital (e.g. 80, 99 or 99.98% pure) as well as allowable ranges of specific impurity concentrations. With therapeutic proteins any impurities have to be minimized whereas

for the production bulk industrial enzymes this is not the case.

1. How big is the flow rate?

Bigger than 3 m<sup>3</sup>/h.

Smaller than 3 m<sup>3</sup>/h.

### 3.2. Characterisation of starting material

The second important point is the characterisation of the starting material. The process design will mainly depend on the physical, chemical and biochemical properties of the contaminating materials in the original broth and those of the protein that will constitute the final product. The properties of the starting material will be partially determined by:

1. Its fermentation source.
  - Bacterial.
  - Yeast.
  - Mammalian cell.
2. How big is the cell concentration?
3. The type of cultivation medium used.
  - Presence of albumin.
  - Calf serum.
  - Presence of proteases.
  - Solid bodies like whole cells or cell debris.
4. Localization of the product.
  - Intracellular.
  - Extracellular.
5. Physicochemical properties of the product and the contaminant proteins in the final protein mixture in solution.
  - Surface charge at different pHs and pI.
  - Surface hydrophobicity.
  - Molecular weight.
  - Biospecificity toward certain ligands.
6. Stability of the final product is also of extreme importance as this will affect the types of operations that can be used as well as the conditioning and processing times that can be afforded.

### 3.3. Define possible separation steps and constraints

It is necessary to have a database with all possible separation steps (Table 1) and to consider all potential constraints (Table 2).

### 3.4. Evaluation of potential process integration

Finally, the production process should try to integrate as much as possible, upstream, fermentation and downstream process. Asenjo and Leser (1996) have described possible areas and levels of process integration. Process integration can involve either the whole process, specific parts of it or particular unit operations as shown in Fig. 2. Examples of process integration between fermentation and downstream and between downstream operations are shown in Table 3.

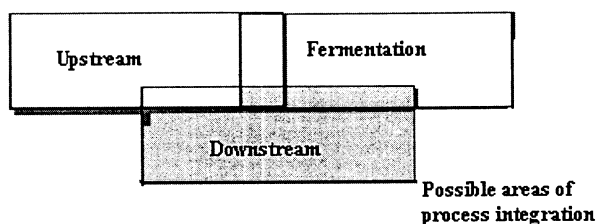


Fig. 2. The possible areas of process integration.

The knowledge showed in the previous sections can be used for defining expert system for the rational selection of downstream protein processes.

#### 4. Expert systems

Expert systems are programs that attempt to solve problems in a way similar to how a human expert would solve them. They incorporate 'rules of thumb' that experts in the field have developed through years of experience. The problems tackled are not necessarily solved in a procedural fashion, they are often vague, complex, and can contain incomplete or inexact information (Nebendahl, 1988).

Expert systems contain three basic elements: a knowledge base, an inference engine, and a user interface. The architecture of an expert system is shown in Fig. 3. The knowledge base contains the information, which the program uses to reach decisions. The inference engine is the program that manipulates the knowledge base to reach these decisions. Finally, the user interface allows the program and the user to communicate with each other in an effective manner (Harmon & King, 1985). There are software systems, called shells, used for the manipulation of heuristics, databases, and algebraic equations (like design equations). For instance, Nexpert Object™ and Clips. There are many examples of the use of expert systems in chemical and biochemical engineering (Siletti, 1989; Mulholland, Walker, Manis, Hinriks, Buydens, Blaffer & Schoemaker, 1991; Jakus, 1992; Forslund, 1995).

Table 3  
Examples of possible process integration

Level of integration	Action
Fermentation-downstream (the adequate design of the fermentation step can provide conditions to facilitate the downstream)	Cell immobilisation, cell recycling using membranes, use of expanded bed chromatography after fermentation
Inside downstream (recovery and purification)	Use of liquid-liquid separation, use of expanded bed for protein recovery, use of hydrophobic interaction chromatography after precipitation with ammonium sulphate

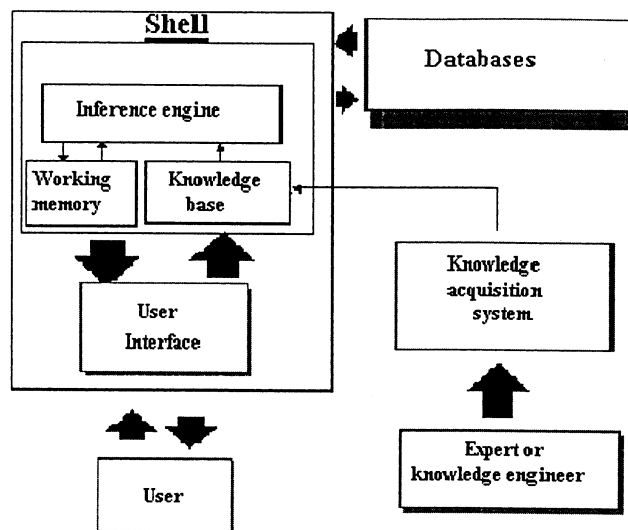


Fig. 3. The basic architecture of an expert system.

#### 5. Implementation of expert system

The implementation of an expert system for separation processes for proteins was divided in two different parts:

1. A first expert system, called Prot\_Ex, for recovery process selection.
2. A second one, Prot\_Ex\_Purification, for purification process selection.

This division was carried out because both systems need different kind of information and the available heuristic knowledge is different, too (Asenjo, Herrera & Byrne, 1989). Prot\_Ex uses only heuristic knowledge whereas Prot\_Ex\_Purification needs an important amount of quantitative data to make a good selection.

In a consultation both expert systems are integrated and controlled by the main expert system. The main expert system:

- Receives the information from the user, database or other sources.
- Organizes and gives the information that each sub-expert system needs for suggesting a sequence.
- Receives the sequences suggested by each sub-expert system
- Gives the global sequences suggested to the user or other programme.

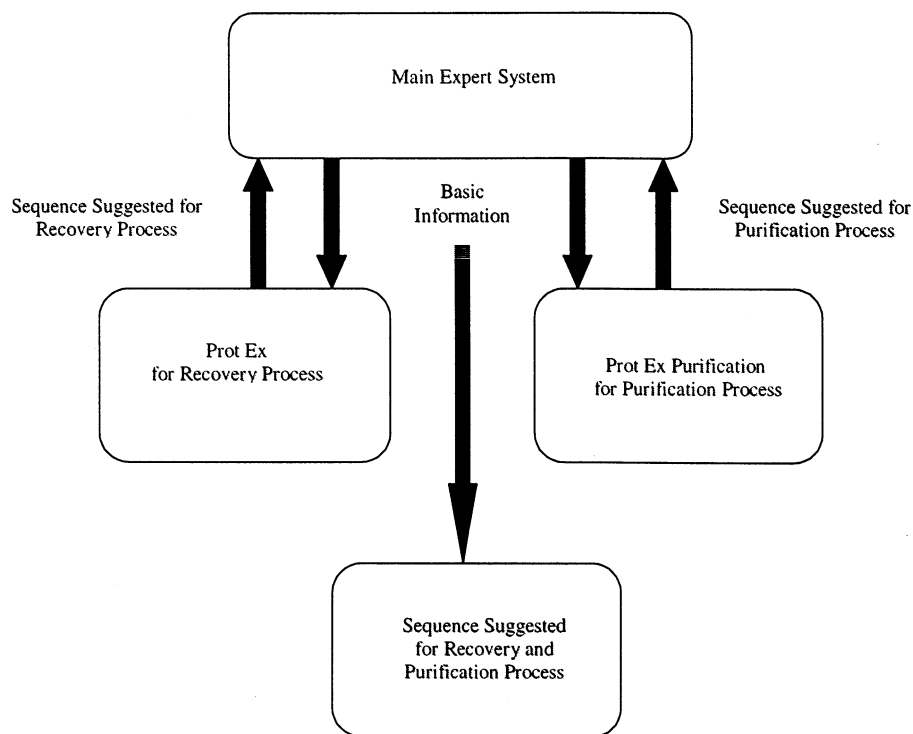


Fig. 4. Diagram of main expert system and relation with sub-expert systems Prot Ex and Pro Ex purification.

The general diagram of the main expert system and sub-expert system with its relation is shown in Fig. 4.

### 5.1. Prot\_Ex for recovery process

The expert system for the recovery process was implemented using only heuristic knowledge. The heuristic rules were obtained from the literature and renowned industrial experts were consulted to clarify specific points for which knowledge is missing or there is ambiguous information. To do this, it was necessary to provide the experts with a questionnaire (Table 4).

For recovery processes, it is only necessary to choose between two or three options. For example, for separation of cells it is necessary to select between centrifugation and cross-flow microfiltration. This selection is based on the following variables:

- Thermal sensitivity of the product.
- Shear liability of cells.
- Size of the cells.
- Process throughput.
- Density difference between solid and liquid phase.
- Cell concentration in the broth.

Depending on the value of each variable it is possible to choose one or the other option.

Considering the previous points, the expert system Prot\_Ex was implemented. It has more than six hundred logical rules to emulate human reasoning. Its implementation was carried out in a commercial programme, Nexpert Object™ from Neuron Data. Consul-

tations were carried out to validate the expert system Prot\_Ex. The basic information used for the expert system is shown in Table 5. Table 6 displays a comparison between a sequence suggested by Prot\_Ex and a good process for recovery of somatotropin produced in *Escherichia coli* (Wheelwright, 1991).

For this example, the recovery sequences suggested by Prot\_Ex and the industrial process were very similar. The difference with the published process is in step 1, where centrifugation was used instead of microfiltration. For small cells such as *E. coli* microfiltration usually has clear economic advantage (Asenjo &

Table 4

Typical questions presented to biotechnology experts for implementation of the expert system for recovery processes (Leser, 1996)

Aspect	Question
Therapeutic product	What would your recommendation be if the product utilisation is therapeutic?
Cell separation	Are the common 'broth conditioning processes', like heating, freezing, coagulation, flocculation, enzymatic or other used to facilitate cell separation in recombinant DSP or not?
Cell disruption	Would you agree that the best suggestion of a mechanical method for breaking cells is 'high pressure homogenizer' for most real situations?
Precipitation of nucleic acid	What are the preferred methods for precipitation of nucleic acids?

Table 5  
Basic information for designing a purification process of Somatotropin (Lienqueo et al., 1996)

Defining final product					
Protein's name	Somatotropin				
pI	7.86 (Estimated)				
Molecular weight	22.0 KDa (Estimated)				
Surface hydrophobicity	0.93 M (Estimated)				
Concentration	25.0 mg/ml				
Titration curve					
pH	4.0	5.0	6.0	7.0	8.0
Charge (coulomb/molecule) $10^{-25}$	4.77	2.42	1.03	0.12	-0.03
Localization of the product	Intracellular, inclusion bodies				
Utilization	Industrial				
pH stability range	4–9				
Final purity level	94%				
Specific affinity	Not known				
Characterisation of starting material					
Fermentation source	<i>E. coli</i>				

Patrick, 1990). Therefore, we considered that the sequences suggested by Prot\_Ex are adequate and it can be used as a starting point for the selection of a good industrial process.

### 5.2. Prot\_Ex\_Purification for selection of purification processes

The selection of an efficient purification process is a critical step in downstream processes. These steps are not usually chosen in a rational manner, the common method being 'trial and error'. Therefore, the use of some basic heuristic rules for separation processes has been proposed. These rules are: (Asenjo & Patrick, 1990; Prokopakis & Asenjo, 1990):

1. Choose the separation process based on the different physicochemical properties, such as:
  - 1.1. Surface charge as a function of pH (titration curve) and pIs.
  - 1.2. Surface hydrophobicity.
  - 1.3. Molecular weight.
  - 1.4. Biochemical properties such as biospecificity with different ligands.
  - 1.5. Stability at different temperatures and pHs.
2. Eliminate those contaminant proteins and compounds that are found in greater percentage first.
3. Use a high-resolution step, as soon as possible. The chromatographic techniques ranked according to their efficiency are:
  - 3.1. Affinity.
  - 3.2. Ion exchange.
  - 3.3. Hydrophobic interaction.
  - 3.4. Gel filtration.
4. Do the most arduous purification step at the end of the process (final polishing).

Considering those rules two selection criteria have been defined: Separation Selection Coefficient (SSC) criterion and Purity criterion.

The first one, SSC criterion, was developed by Asenjo and collaborators (Asenjo, 1990; Asenjo & Maugeri, 1992; Leser & Asenjo, 1992). The SSC criterion considers that the key parameter is SSC, which characterizes the ability of the purification operation to separate two proteins. SSC value is defined as:

$$SSC = DF\eta\Theta_l \quad (1)$$

Deviation factor (DF), this variable relates the difference between a property of the target protein (for example, the dimensionless retention time in a specific chromatographic technique) and the same property of the contaminant.

$$DF = |K_{D\text{product}} - K_{D\text{contaminant}}| \quad (2)$$

Dimensionless retention time ( $K_D$ ), this variable represents the behaviour of the proteins in a separation carried out by gel filtration, ion exchange or hydrophobic interaction chromatography. Mathematical relationships for predicting  $K_D$  have been derived using the physicochemical properties of proteins (Lienqueo, Leser & Asenjo, 1996).

Table 6  
Comparison between sequences suggested by Prot\_Ex and industrial process for recovery of somatotropin (Lienqueo et al., 1996)

Steps	Industrial process	Sequence suggested by Prot_Ex
1	Centrifugation	Crossflow microfiltration
2	High-pressure homogenization	High-pressure homogenization
3	Centrifugation-pellet wash	Disk centrifugation
4	Solubilization	Solubilization
5	Renaturation	Renaturation
6	Microfiltration	Ultrafiltration
7	Concentration and diafiltration	

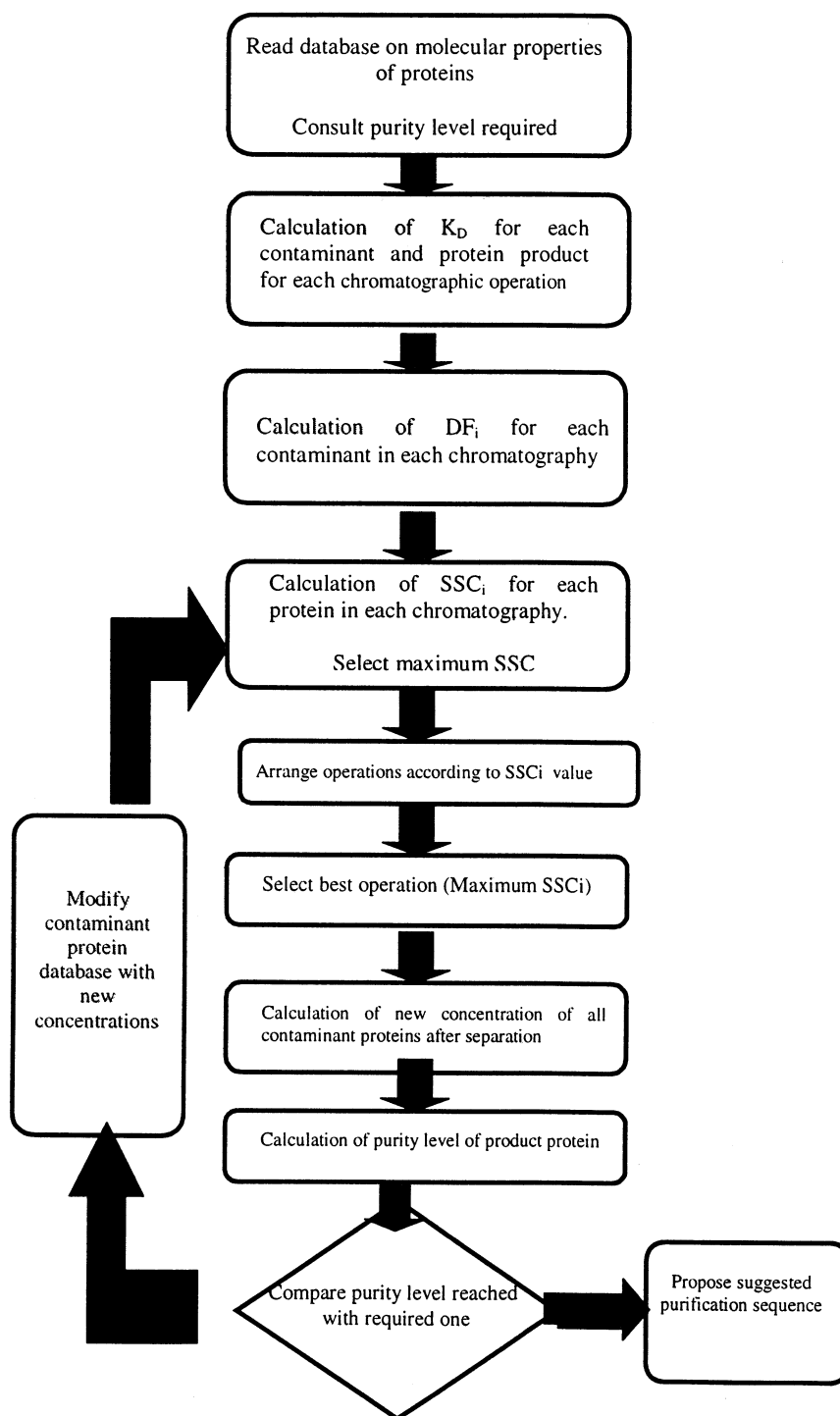


Fig. 5. Algorithm of SSC criterion.

Efficiency factor ( $\eta$ ) This parameter gives account of the unequal capability of the different separation processes to separate different proteins. Its value is constant for each type of separation and the chromatographic material used and have been measured experimentally (Lienqueo et al, 1996).

Concentration factor ( $\Theta_i$ ) Its value represents the relative concentration of each contaminant. It will af-

fect the selection criteria. In this way contaminants which are present in a high concentration have to be eliminated first.

$$\Theta_i = \frac{\text{concentration of contaminant } i}{\text{concentration of all contaminants}} \quad (3)$$

The amount of contaminant eliminated after a chromatographic step has been mathematically calculated

for different situations (Lienqueo, Salgado & Asenjo, 1999)

Then the SSC criterion chooses as best chromatographic operation the one that has the highest SSC value (Leser, Lienqueo & Asenjo, 1996). After determining which chromatographic technique gives the maximum SSC value, it is necessary to calculate the new concentration of all contaminant after the chromatographic technique selected has been applied and

construct a new database of contaminants concentration. With this new database the system calculates the purity level and this value is compared with the level of purity required. The optimal sequence of steps is chosen until the required level of purity is reached. The algorithm used is shown in Fig. 5.

Considering that the most important value is the final purity level and that now we had developed an algorithm to calculate the purity after a purification step,

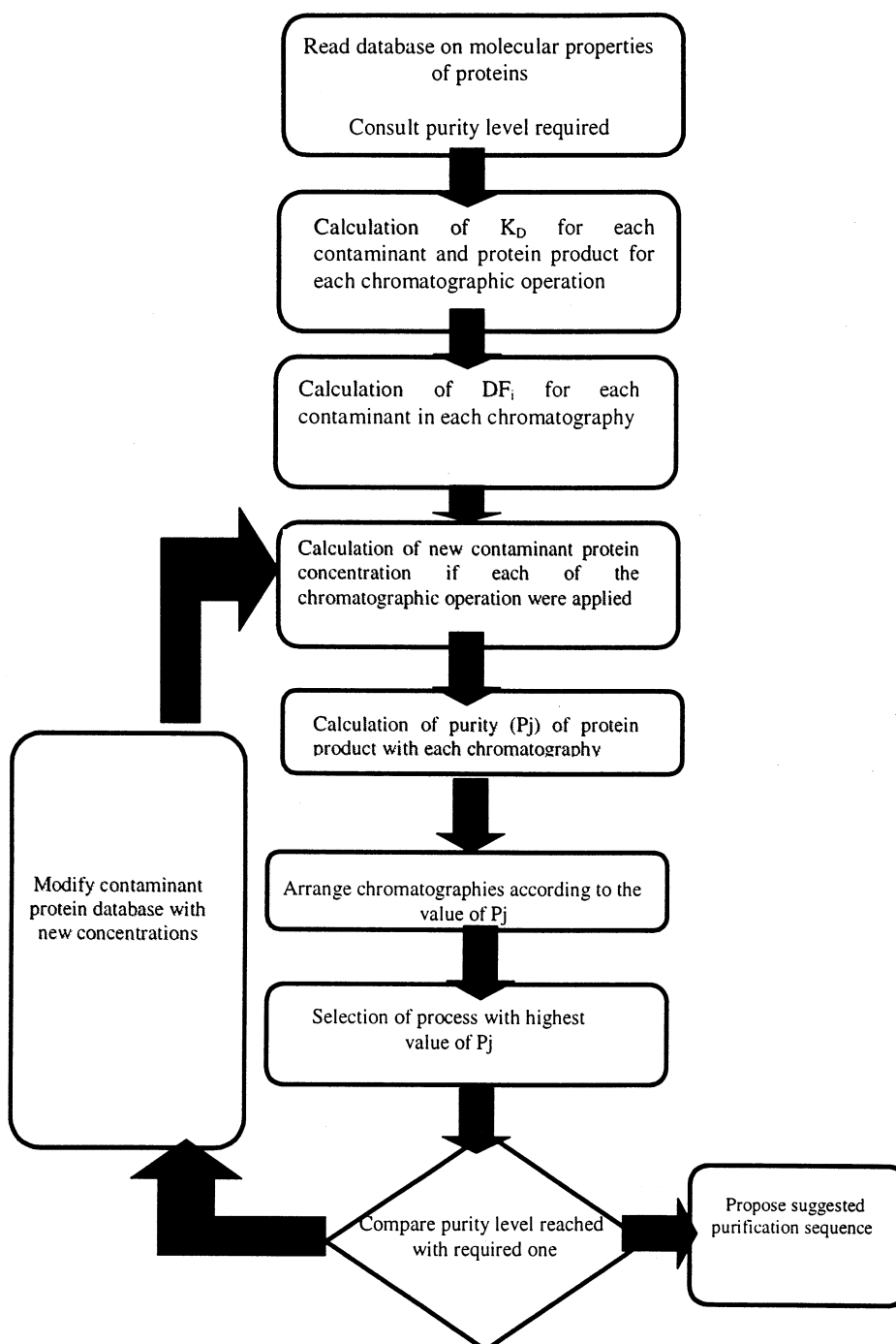


Fig. 6. Algorithm of purity criterion.



Table 7  
Main variables used for SSC and purity criteria (Lienqueo et al., 1999)

Variable	Symbol	Definition
SSC criterion		
Separation selection coefficient	SSC	$SSC = DF\eta\Theta_i$
Deviation factor	DF	$DF =  K_D\text{product} - K_D\text{contaminant} $
Dimensionless retention time	$K_D$	Relationships with physicochemical properties of the proteins. (Lienqueo et al., 1999)
Efficiency factor	$\eta$	Values for each chromatographic techniques shown in Table 8
Peak width	$\Sigma$	Values for each chromatographic techniques shown in Table 8
Concentration factor	$\Theta_i$	$\Theta_i = \frac{\text{concentration of contaminant } i}{\text{concentration of all contaminants}}$
Purity criterion		
Purity	$P_j$	$P_j = \frac{\text{concentration of the target protein}}{\Sigma\text{concentration of the proteins present}}$

Table 8  
Expressions and parameters used for SSC criteria and purity criteria (Lienqueo et al., 1999)

Chromatographic techniques	Efficiency factor $\eta$	Peak width $\Sigma$
Anion exchange	1.00	0.15
Cation exchange	1.00	0.15
Hydrophobic interaction	0.86	0.22
Gel filtration	0.66	0.46

we implemented the purity criterion as a possible selection criterion. This criterion compares the final purity level obtained after a particular chromatographic techniques has been applied.

The purity concept is defined as:

$$\text{Purity} = \frac{\text{concentration of the target protein}}{\Sigma\text{concentration of the proteins present}} \quad (4)$$

After determining which chromatographic technique gives the highest purity level, the system chooses this as the technique to use at this step. A sequence of steps is chosen until the required level of purity is reached. Finally, the system creates a list with the defined sequence of operations. The algorithm is shown in Fig. 6

The main parameters, variables and correlations used for SSC and purity criteria are summarized in Tables 7 and 8. Both criteria consider that the cost of each chromatographic option is equal, except for affinity chromatography. Therefore, the optimal system is the

one that has the least number of total steps. Considering these criteria an expert system was implemented, Prot\_Ex\_Purification, and its implementation was carried out in the same shell than Prot\_Ex.

For testing these criteria, consultations were carried out to validate the expert system Prot\_Ex\_Purification using both criteria. Table 9 and Table 10 show the basic information used for purification of bovine serum albumin (BSA) and  $\beta$ -glucanase, respectively and Table 11 and Table 12 show a comparison between sequences suggested by Prot\_Ex\_Purification and experimental sequence for both examples.

### 5.3. Comparison between SSC and Purity criteria

In both examples the sequences suggested are valid and adequate. Nevertheless, in the first case, the sequence suggested by the purity criterion has one fewer step than that suggested by the SSC criterion, that mean less capital cost. This situation takes place because the SSC criterion considers the contaminant that gives the highest SSC value for one protein only. The purity criterion chooses the optimum chromatographic step considering all the contaminant proteins present. Hence, this situation can occur when there are several contaminants present in similar quantities as was the case in this example.

On other hand, the purity criterion spends less time and computational resources than the SSC criterion, since the purity algorithm gives the optimal sequences in less time than the SSC algorithm.

### 5.4. Comparison between sequences suggested and experimental sequences

In both examples the sequences suggested result in experimentally valid ones and are a good starting point for an industrial process. There are differences between the estimated and experimental final purity level, which are less than 20%. These differences could originate in different causes:

- Hydrophobic interaction between the protein, this effect was not considered in the development of correlations for estimated dimensionless retention time for hydrophobic interaction chromatography (HIC) (Table 8).
- Asymmetric chromatographic peaks, the peak shape could be different to a triangle shape.

The differences between estimated and experimental purity level could be minimised:

- Improving the prediction of dimensionless retention time for HIC.
- Using a Gaussian approximation in the representation of the chromatographic peak instead of triangle simplifications.

## 6. Conclusions

It is possible to use artificial intelligence tools, as expert systems, for selection of optimal protein separation processes. For implementing an expert system, it is necessary to divide the separation process in two parts, because the information and available heuristic knowledge are different in each case. The first part for the recovery of proteins uses only heuristic rules from human experts and literature, the expert system was called Prot\_Ex. The second part of a purification process is an hybrid expert system, which uses heuristic rules and mathematical correlations and models for selecting the optimal protein purification sequences,

expert system called Prot\_Ex\_Purification. Both expert systems, implemented in the shell Nexpert Object™, were integrated by a main expert system.

Considering the cases studied, like the recovery of somatotropin, the purification of BSA and the purification of  $\beta$ -glucanase, the sequences proposed by the expert system are a good starting point for a practical industrial process. Nevertheless, there are differences between estimated and experimental purity level, these could be minimized by improving correlations for predicting retention time and approximation of the peak shape.

In the specific case of Prot\_Ex\_Purification, the sequences suggested by the purity criterion have less

Table 9  
Basic information for desining a purification process of BSA from artificial mixture (Lienqueo et al., 1999)

### Defining final product

Protein's name	Bovine serum albumin (BSA)							
pI	4.9							
Molecular weight	67.0 KDa							
Surface hydrophobicity	0.86 M							
Concentration	2.00 mg/ml							
Titration curve								
pH	4.0	5.0	6.0	7.0	8.0			
Charge (Coulomb/molecule) $10^{-25}$	1.03	-0.14	-1.16	-1.68	-2.5			
Localization of the product	Extracellular							
Utilization	Industrial							
pH stability range	4–9							
Final purity level	94%							
Specific affinity	Not known							
<i>Characterisation of starting material</i>								
Fermentation source	Artificial mixture (BSA, ovalbumin, thaumatin, soybean trypsin inhibitor)							
Contaminant	Initial concentration (mg/ml)	mw (Kda)	Hydrophobicity (M)	Titration curve (coulomb/molecule) $10^{-25}$				
				pH 4.0	pH 5.0	pH 6.0	pH 7.0	pH 8.0
Ovalbumin	2.0	43.8	0.54	1.40	-0.76	-1.65	-2.20	-2.36
SBTI	2.0	24.5	0.90	1.22	-0.76	-1.54	-2.17	-2.13
Thaumatins	2.0	22.2	0.89	1.94	1.90	1.98	1.87	0.91

Table 10  
Comparison between sequences suggested by Prot\_Ex\_Purification and experimental process for purification of BSA from a mixture of four proteins (Lienqueo et al., 1999)

Step	Sequence suggested by Prot_Ex_Purification				Experimental validation process	
	SSC criterion		Purity criterion			
	Chromatography step	Purity	Chromatography step	Purity	Chromatography step	Purity
1	Cation exchange pH 6.0	33%	Anion exchange pH 7.0	64%	Anion exchange pH 7.0	64%
2	Hydrophobic interaction	49%	Hydrophobic interaction	95%	Hydrophobic interaction	80%
3	Anion exchange pH 7.0	97%				

Table 11  
Basic information for designing a purification process of  $\beta$ -1,3-glucanase (Lienqueo et al., 1999)

Defining final product								
Enzyme name	$\beta$ -1,3-glucanase ( $\beta$ gl IIa)							
pI	4.1							
Molecular weight	31 Kda							
Surface hydrophobicity	0.00 M							
Concentration	0.62 mg/ml							
Titration curve								
pH	4.0	5.0	6.0	7.0	8.0			
Charge (Coulomb/molecule) $10^{-25}$	1.46	-0.62	-1.02	-2.33	-2.52			
Localization of the product	Extracellular							
Utilization	Industrial							
pH stability range	4–9							
Final purity level	70%							
Specific affinity	Not known							
Characterisation of starting material								
Fermentation source			<i>B.subtilus</i> ToC46					
Contaminant	Initial concentration (mg/ml)	mw (Kda)	Hydrophobicity (M)	pH 4.0	Titration curve (coulomb/molecule) $10^{-25}$			
					pH 5.0	pH 6.0	pH 7.0	pH 8.0
1	2.74	41.0	1.5	1.46	0.26	-0.87	-1.65	-2.04
2	2.74	32.9	1.5		0.00	-2.70	-3.51	-3.51
3	0.25	35.5	0.2		-0.55	-0.22	-0.73	-1.82
4	0.42	62.5	0.0		-1.06	-1.17	-2.79	-3.32
5	0.09	40.6	0.0		-0.55	-0.22	-0.73	-1.82
6	0.09	69.6	0.0		-0.55	-0.22	-0.73	-1.82
7	0.25	40.6	0.0		1.46	-0.47	-1.06	-1.04
8	0.25	69.6	0.0		1.46	-0.47	-1.06	-1.04

Table 12  
Comparison between sequences suggested by Prot\_Ex\_Purification and experimental process for purification of  $\beta$ -glucanase (Lienqueo et al., 1999)

Step	Sequence suggested by Prot_Ex_Purification				Experimental validation process	
	SSC criterion		Purity criterion			
	Chromatography step	Purity	Chromatography step	Purity	Chromatography step	Purity
1	Hydrophobic interaction	32%	Hydrophobic interaction	32%	Hydrophobic interaction	35%
2	Anion exchange pH 6.5	70%	Anion exchange pH 6.5	70%	Anion exchange pH 6.5	60–70%

steps than the sequences suggested by SSC criterion, then the Purity criterion is more recommendable.

On other hand, the use of expert systems for the synthesis of downstream protein processes is a clear case of 'expert amplification'. Heuristic knowledge from experts was complemented with databases and design equations.

To implement such a solution more globally there is a lack of databases of protein properties. With advances in the area of proteomics and protein engineering in a few years there could be sufficient information

for such an approach. An effort should be made to generate such data.

### Acknowledgements

Financial assistance from Fundación Andes and Vicerrectoria Académica, Universidad de Chile (Beca PG/080/97). This project was also supported by Proyecto Catedra Presidencial en Ciencias.

## References

- Asenjo, J. A., Herrera, L., & Byrne, B. (1989). Development of an expert system for selection and synthesis of protein purification processes. *Journal of Biotechnology*, *11*, 275–298.
- Asenjo, J. A. (1990). Selection of operations in separation processes. In J. A. Asenjo, *Separation processes in biotechnology* (pp. 3–16). New York: Marcel Dekker.
- Asenjo, J. A., & Patrick, I. (1990). Large-scale protein purification. In E. L. V. Harris, & S. Angal, *Protein purification applications: a practical approach* (pp. 1–28). Oxford: IRL Press.
- Asenjo, J. A., & Maugeri, F. (1992). An expert system for selection and synthesis of protein purification processes. In P. Todd, S. K. Sikdar, & M. Bier, *Frontiers in bioprocessing II* (pp. 358–379). Washington: American Chemical Society.
- Asenjo, J. A., & Leser, E. W. (1996). Process integration in biotechnology. In M. Verrall, *Downstream processing of natural products* (pp. 123–137). Chichester: Wiley.
- Eriksson, H., Sandahl, K., Brewer, J., & Osterlund, B. (1991). Reactive planning for chromatography: chemometrics and intelligent laboratory systems. *Laboratory Information Management*, *13*, 185–194.
- Forslund, G. (1995). Designing for flexibility: a case study. *Expert Systems*, *12*(1), 27–37.
- Harmon, P., & King, D. (1985). *Expert systems: artificial intelligence in business* (p. 283pp). New York: Wiley.
- Jakus, W. (1992). Artificial intelligence in chemistry. *Collect Czech Chemistry Communication*, *57*, 2414–2442.
- Jennings, L. S., Teo, K. L., Wang, F. Y., & Yu, Q. (1995). Optimal protein separation. *Computers & Chemical Engineering*, *19*(5), 567–573.
- Lee, J. M. (1992). *Biochemical engineering* (p. 320). Englewood Cliffs, NJ: Prentice-Hall.
- Leser, E. W., & Asenjo, J. A. (1992). Rational design of purification processes for recombinant protein. *Journal of Chromatography*, *584*, 43–57.
- Leser, E. W. (1996) *Prot\_Ex: an expert system for selecting the sequences of processes for downstream purification of proteins*, PhD thesis, University of Reading, England.
- Leser, E. W., Lienqueo, M. E., & Asenjo, J. A. (1996). Implementation in an expert system of selection rationale for purification processes for recombinant proteins. *Annals of the New York Academy of Sciences*, *782*, 441–455.
- Lienqueo, M. E., Leser, E. W., & Asenjo, J. A. (1996). An expert system for selection and synthesis of multistep protein separation processes. *Computers & Chemical Engineering*, *20*, 189–194.
- Lienqueo, M. E., Salgado, J. C., & Asenjo, J. A. (1999). An expert system for selection of protein purification processes: experimental validation. *Journal of Chemical Technology & Biotechnology*, *74*, 293–299.
- Mulholland, M., Walker, N., Manis, F., Hinriks, H., Buydens, L., Blaffer, T., & Schoenmakers, P. (1991). Expert system for repeatability testing of high-performance liquid chromatographic methods. *Journal of Chromatography*, *550*, 257–266.
- Nebendahl, D. (1988). *Expert systems: introduction to the technology and applications* (p. 211). New York: Wiley.
- Prokopakis, G. J., & Asenjo, J. A. (1990). Synthesis of downstream processes. In J. A. Asenjo, *Separation processes in biotechnology* (pp. 571–601). New York: Marcel Dekker.
- Siletti, C. A. (1989). *Computer aided design of protein recovery processes*, PhD thesis, (pp. 379). Massachusetts Institute of Technology, Cambridge, USA.
- Wheelwright, S. M. (1991). *Protein purification: design and scale up of downstream processing* (p. 228). München: Carl Hanser.