

Review: difficulties and advancements in the design of integrated

protein purification systems and a view

Nadine Gordimer, J. M. Coetzee

Abstract

Successful protein purification relies on selecting the best purification methods and combining them logically in order to achieve the target purity in the shortest possible time. However, rationalising the development of protein purification processes has its difficulties. Among the issues addressed in this work are those related to protein purification. Synthesis and design of protein purification processes are examined, along with the advantages

and disadvantages of the most recent approaches. Finally, the future of protein purification process development is discussed in this section. Attribution-Noncommercial-ShareAlike

Keywords : high throughput experiments, expert systems, and algorithmic method to protein purification

INTRODUCTION

In the global economy, the biopharmaceutical sector is one of the fastest-growing industries. 1.2 grammes of protein class of biopharmaceuticals that should not be underestimated foodstuffs and biotechnological goods as well applications. 3 Advancements in technology have occurred throughout the last 25 years. recombinant DNA and hybridoma technologies have revolutionised the field of genetics. large-scale manufacturing of nearly any product protein at higher concentrations by fermentation routes,4 such that the biologic bottleneck may be shifted Process improvement for protein purification feed items derived from living organisms. 5,6 This is what I mean the results of a great deal of experience and trial-and-error experiments, which typically lead to unsatisfactory outcomes feedstock and ancillary processes that are inefficient Utilization of resources. Pharmaceutical companies have been tasked with developing newer (better) products in the face of this background. Faster and cheaper production of highquality (or safer) items. 9 With this purpose in mind, it's possible to supplement the content investing in the present paradigm of process development analytical and scientific methods that are always being developed. more systematic tools that have been developed methods of design and development that are both logical and efficient. These are some examples of design tools: approaches for high-throughput experimentation and tools for the design of computeraided processes. 12 to 15 years old It's a fact. incorporating these new design tools is also crucial. from the very beginning of the procedure for creating modifications in the manufacturing process may have both positive and negative effects. Regulatory and technical hazards are both present. 8 But in spite of all that, the FDA, the US Food and Drug Administration Now known as PAT (Process Analysis Technology)

by acknowledged on a broad scale as both technically and economically advantageous difficult, making up a significant portion of the overall cost of production 7 Protein research and development for biopharmaceutical applications Product quality standards must be met by application. Regulatory authorities and the time-to-market are both tightening their criteria.8 Safe and cost-effective processes need that hence, it is imperative that you can be located fast in a location that is incredibly remote. huge area for design work. 7 As a result, protein purification has become more important. mostly, process design and development Based on heuristics, the results are tested experimentally

uses these new design tools and promotes the usage of them platform technologies are OK as long as you have a comprehensive understanding of the crucial parameters of the process. the most significant influence on the qualities' quality There is no doubt that this is a produced product. 16 As a chemical engineer, separation is a common task. tools and strategies for synthesising and designing processes developed to a point in the last few decades maturation, which has benefited industrial production. the process design of both practitioners and academics education.17 There is a plethora of information available on the subject. synthesis and design of chemical processes a thorough investigation was performed. 17 and 18 While this may be true, Purification methods for the general population are still lacking. approach for proteins obtained from biological sources. Nevertheless, progress is being made in this sector, and it is bringing with it benefits. For a thorough examination. We came up with this. Attempt to accomplish here.

**CHALLENGES IN PURIFICATION
TECHNIQUE DETAILS**

Purification of proteins is discussed in detail here, with a focus on the various phases and their aims. In addition, the difficulties of developing a protein purification method are explored.

**A detailed explanation of the purification
procedure for proteins**

The purification phases' needs fluctuate with time, as do the materials and volume handled throughout a specific protein purification process. Thus, the purification process may be broken down into steps depending on the key goals at each stage to assist the synthesis and design of proteins. The Petrides^{20,21} block diagram of downstream processing is adopted with slight changes in Fig. 1 for illustration purposes. Each of the four parts in this generalised block diagram represents a distinct stage in the downstream processing, ranging from the extraction of the product (intracellular or extracellular) through its purification and formulation, each of which is specified by a number of well-defined process steps. Cell harvesting, cell disruption, cell debris removal and biomass removal are shown in Fig. 1 as typical unit activities utilised in industry for each of the processing phases. These unit procedures take use of differences in the components' physical or molecular properties. The most important bioseparation unit activities have been outlined, highlighting their most important features and driving factors. ^{22,23} The numerous unit operations and their effects on downstream processing have been thoroughly detailed in multiple monographs^{24–26} and recently reviewed by Hubbuch and Kula. ²⁷ Recovering the protein and non-protein components of the mixture is the primary goal of the recovery phases. Clarification is a term used to describe the process of removing the product from cells or cell detritus, which may be performed by physical methods (e.g., filtering, centrifugation) or a combination of both, as shown in Figure 1. It is possible to accomplish partial purification of the product by the use of extraction techniques. Insoluble inclusion bodies (IBs) of the target protein must be removed before it may be refolded or renatured. ^{28–31} There are no large particles in the stream that exits the

recovery section, therefore the target protein may be retrieved in either dilute or concentrated form in a complicated combination of contaminating proteins. Concentration is needed if protein solutions are present in a weaker form. The purification portion receives it if it does not. Capture, intermediate purification, and polishing are the three primary processes of purification, each with a distinct goal, as shown in Table 1. These purification procedures normally utilise the different chromatographic operations due to their great resolving power, however the use of alternate unit operations that may give superior throughputs has been advocated. ^{32,33} Membrane filtration, precipitation, crystallisation, and membrane chromatography are only a few of the methods available. Getting the protein in the form the end user wants is the primary goal of product formulation. At this point, any necessary additives are added to the protein product in order to meet the needs of the ultimate application or to enhance the protein's stability and shelf life. The process of protein synthesis as well as related topics have been thoroughly examined. ^{34,35}

**Protein purification is plagued with
difficulties.**

Every large-scale protein purification method has the same challenge: to purify a protein effectively, inexpensively, and to a suitable purity and quantity. Depending on the protein's intended usage, the level of purification required will vary. As a result, understanding the protein's intended use is critical for the subsequent processing step. Nutraceuticals, medicines, industrial enzymes, and diagnostics are some of the commercially accessible proteins that fall under these categories. Ghosh has compiled a list of proteins that fall under several groups. Compared to other protein categories, therapeutic proteins intended for injection often have the strictest purity criteria, surpassing 99 percent.³⁷ ³ The following are the primary characteristics that make protein purification difficult: There are generally low amounts of protein products in the biological feed streams for which they are intended. Many contaminants are included in complicated feed mixes, making it difficult for the products to be identified.

Table 1. Objectives of the purification steps

Purification step	Objectives
Capture	Isolation, concentration, volume reduction
Intermediate purification	Removal of bulk impurities or main protein contaminants
Polishing	Removal of trace impurities, closely related contaminants and protein aggregates.

A few have qualities that are almost identical as the product itself. The presence of so-called critical contaminants, which must be eliminated, must be distinguished from impurities that may be tolerated.

Physicochemical parameters (such as molecular weight, charge, hydrophobicity), thermodynamic properties (such as solubility), and flow properties (such as flow rate) of the crude mixture containing the product are often poorly described (e.g. viscosity, diffusivities). When a protein product is in soluble form or in the form of inclusion bodies (IBs), it must be reinitiated.

Stability of the product is a factor.

In general, most protein products are thermolabile, which makes them sensitive to severe pH and chemical compounds like surfactants and organic solvents. Customers frequently have high standards for the quality and quantity of the things they buy.

The method of operation is as follows:

Batch or semi-continuous bioprocess unit operations are used for protein purification and need frequent cleaning.

Problem of synthesis using a combinatorial process:

Each processing step has several unit operations to pick from, and there are a plethora of ways in which these unit operations might be connected to form potential process flowsheets, creating a combinatorial challenge in process synthesis. The ideal operating conditions for a particular unit operation or process may have a number of alternatives, posing an optimization challenge.

Approaches to the Synthesis of Large-Scale Protein Purification Process

Selecting the best unit operations for each processing stage and connecting them into one or more integrated processes that properly purify the product while using the least amount of resources is required to design an affordable protein purification process. Developing a (bio)separation process synthesis relies on the engineer's expertise and ingenuity. The following issues are addressed by the 21 (bio)separation process synthesis strategies:

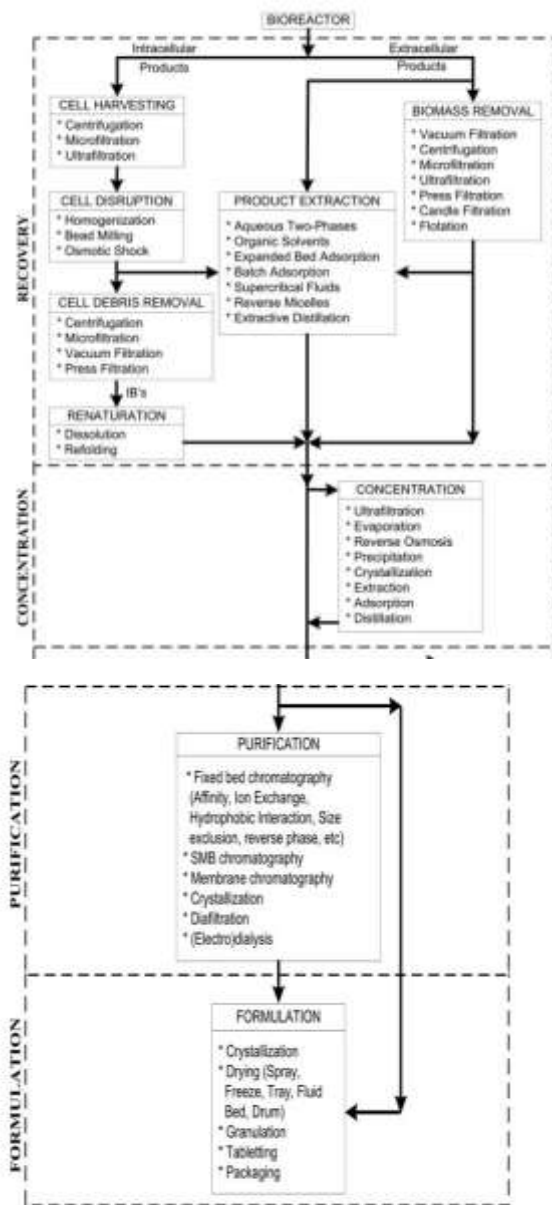


Figure 1. Generalized block diagram of downstream processing. □

Are there any possible unit operations for the

separation problem?

- Is there a way to have this all work together as one process?
- Is it possible to create a representation of the process that is both comprehensive and intelligent enough to omit unnecessary or duplicate processes?
- The various procedures must be assessed and evaluated in a manner which balances speed and accuracy, so how can this be done effectively?

Methods based on heuristics or on knowledge

Using heuristics, or knowledge-based procedures, we may make decisions based on past experience, intuition, and what is already known about the problem at hand.

18 Inexperienced engineers often depend on heuristics or rules of thumb to choose and link purification unit activities. In the purification process, certain heuristics may be used universally, while others can only be applied at particular stages. Asenjo and coworkers have outlined the most popular generic heuristics for downstream processing. 26,38 In their expert system, Asenjo and colleagues synthesised full downstream process flowsheets by combining expert knowledge with certain short-cut computations. They demonstrated that the protein recovery stages could be synthesised using just expert knowledge. 22,38 A few short-cut computations were needed to synthesise the purification stages, which required knowledge on physical properties and certain physicochemical properties. 39 In this case, the target and contaminant proteins' physicochemical properties (such as molecular weight, isoelectric point, hydrophobicity, and titration curve) are first obtained⁴⁰, and chromatography units are chosen based on their ability to exploit differences in physicochemical properties in terms of the deviation factor between the target and contaminant proteins. Unit operations are then sequenced by means of a systematic procedure including semi-quantitative assessments of the chromatographic procedures that are under consideration. 39.5 to 44.9 Assuming triangular Gaussian elution peaks, the theory is simplified. 42,44 The chromatography models were oversimplified and unit activities were studied sequentially, ignoring interactions between the purification units, a shortcoming of this technique. Expert systems, which aim to encapsulate heuristics in computer code, have been developed. 20,22,38,39,41 The purification

processes (capture, intermediate purification, and polishing) might benefit from some extra practical guidelines: 37,45

- Avoid interstage conditioning by sequencing chromatographic units.
- Use a different purification method for each stage.
- Only as a last polishing step, or to switch buffers, may size exclusion chromatography (SEC) be employed for capture.
- Therapeutic proteins need at least one viral clearance step.
- It is recommended to utilise the purification sequence IEX-HIC-SEC in cases when there is no prior knowledge of the protein characteristics of the target and the contaminant proteins.

Methods based on heuristics are always going to be of a higher quality. As a result, relying only on heuristics does not ensure that all viable process alternatives are included in the design search space, nor does it ensure that the process that is ultimately chosen is the best of the ones that were considered. Heuristic-based approaches or expert systems, on the other hand, are very useful in situations when little or no information is available. They're also a breeze to put on. It is possible to quickly reduce the combinatorial issue to something that can be dealt with using heuristics, but this should be followed by a more quantitative examination and assessment of the remaining "more promising" process options in order to choose and build the optimal one.

Optimizing or Algorithmic Methods

The mathematical programming approaches used to optimise the process flowsheet are the foundation of the algorithmic methodologies. Advances in mathematical programming tools and major advancements in computer power allowed for this indepth mathematical study of the process synthesis issue. 17,19 All viable process alternatives must be identified and their mathematical superstructures built up in order to begin the process. The genuine optimum process may not be identified if the formulation of the superstructure is not as extensive as feasible since it is not included in the search space. Mathematical programming approaches, such as those based on the mixed integer non-linear (MINLP) formulation^{46,47}, or those based on generalised disjunctive programming (GDP),⁴⁸, or a mixture of both, are required for these methods to work well. 49 Research in the field of integrated biochemical processes is lacking in the area

of superstructure formulations for optimum synergy. Unit activities in batch and semicontinuous modes are often integrated into biochemical processes under various conditions in both continuous and batch modes. This complicates the present optimization-based algorithms' ability to simultaneously describe process conditions, process alternatives, and plant scheduling. 50,51 A two-stage optimization-based strategy was presented by Samsatli and Shah⁵⁰ to address this issue. Dynamic optimization is used to identify the unit operation and equipment capacity processing rates and circumstances in the initial step. 50,52 In the second stage, comprehensive scheduling and design changes are made utilising information from the previous stage in order to accurately determine the sequence and time of unit activities. 50,53 An optimization strategy based on a graphical depiction of the viable operating windows for integrated protein recovery bioprocesses and a pareto technique to locate the optimum operating points within the feasible areas has been developed by Titchener Hooker and coworkers^{54 – 56}. For the development of methods for the simultaneous optimization of structural and process variables in an integrated protein manufacturing plant, Asenjo and coworkers⁵⁹ have used a MINLP formulation^{57,58}. Researchers from Bogle and coworkers^{23,60} presented an integrated biochemical process design methodology in which downstream processing unit operations were first screened using physical and chemical property data before a systematic evaluation of generated superstructure was carried out using an implicit enumeration tool that converts the problem of MINLP into an algorithmic graph generation and search task. The optimum synthesis of chromatographic protein purification methods has been described using MILP formulations. 61 The excessively huge number of potential linkages between purification procedures also restricts the adoption of superstructure-based computational approaches to protein purification process creation. There are only four or five chromatographic methods commonly employed to purify proteins, as indicated in Fig. 1. An attractive superstructure representation is tough and complicated to achieve when there are as many as 320 possible connections between chromatographic units (without employing the same unit action more than one time in a flowsheet) for a particular separation. Even more importantly, many of the interconnections between the various components of the superstructure are simply not viable, making it more difficult to establish an optimal procedure. For protein purification process synthesis, optimising just the superstructure is inefficient. As a result, algorithmic approaches need a thorough grasp of the whole process, as well as

trustworthy unit operating models, solute characteristics and the auxiliary material properties, without which this approach is ineffective. With a methodical approach, however, it might be quite effective in narrowing down the design search area while still taking practical issues into account. When it comes to protein purification, advanced computer simulation tools have made it easier to employ algorithmic or optimization-based design methods. Between the ages of 12 and 14.

High-throughput experimentation-based methods

The development of a protein purification technology is aided by high-throughput experimental approaches. The characteristics of the target and contaminant proteins in the complicated crude mixture from main biological sources such as bacteria, yeast, mammalian, and Chinese Hamster Ovary (CHO) cells are usually unknown at the beginning of protein purification process development. Since protein purification process development is mostly done empirically, it necessitates numerous experiments, such as screening different solubility conditions for protein crystallisation or testing various classes of stationary phase materials and mobile phase conditions (such as pH, salt type, salt concentration), to identify chromatographic media and conditions that provide sonication with a high degree of sonication efficiency. Traditional laboratory-scale tests have lately been supplanted by high-throughput screening (HTS) approaches. Different chromatographic procedures have been selected and systematically sequenced using 10,11,62–65 HTS of chromatographic medium, without needing any information on the characteristics of the components being separated. 11,66 Smaller sample sizes and quicker processing times (due to parallelization of experimental runs and shorter processing periods) are two major benefits of HTE platforms over conventional laboratory sized experiments for process development. Some of the main downsides are the necessity for powerful experimental design procedures and efficient sample analysis tools. 10 In spite of this, robotics-based integration of chromatographic purification and sample analysis for speeding the development of recombinant protein purification processes has shown success in this approach. 67

Inventions that combine different approaches.

Hybrid techniques combine two or more of the several protein purification processes outlined above, taking

use of each one's strengths in the process's purification of proteins. Ahamed et al technique 's to protein purification is a nice example of a hybrid approach. 68 According to Fig. 2, a protein purification method is rationally synthesised and designed using a mix of HTE and modelling tools. Rapid collection of solute characteristics from crude protein mixtures is essential for process modelling using HTE.. As an analytical method for fractionating and characterising crude protein mixtures and for pH adjustment of an ion exchange step, pH gradient ion-exchange chromatography has recently been shown to be useful in this regard. 69 Proposals for reducing the number of model parameters and the amount of testing necessary to simulate the purifying units involved in thermodynamic generalisation were made. 68 Fast generation of different process choices and screening of them to pick the best one need a logical synthesis approach. HTE based on HTS approaches is used to optimise individual unit activities within the flowsheet after the optimal process has been determined and prior to its thorough design. Design of experiments (DoE) must be used to plan the experiments and handle the data obtained as a result.

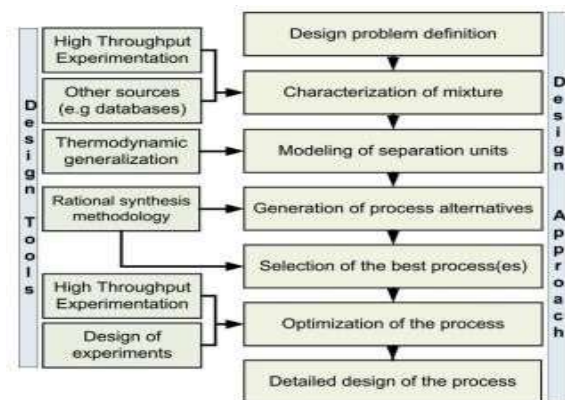


Figure 2. A hybrid approach to bioseparation process design.

There are two main advantages to this hybrid approach. The first is that it involves detailed process modeling and optimization that enables the in-silico study and comparison of various process alternatives in a way that mimics the actual industrial process. In this way, scale-up issues are simultaneously addressed. It also involves parallel development of high throughput experimentation tools for rapid data acquisition from crude protein mixtures as well as for optimization of the selected best process.

OUTLOOK

This review would be incomplete without a look at future protein purification process development directions. The authors consciously limit their discussion here to process development of biopharmaceutical proteins that show the greatest potential for growth and which have the greatest regulatory concerns among the different classes of proteins. The changing market dynamics of protein therapeutics calls for continuous improvements in the development of processes for their manufacture and especially their purification, to meet not only today's challenges but also those of the future. Protein purification process development methods that rely on heuristics-based expert systems and those that employ HTE techniques for screening and optimization purposes, or hybrid methods based on a combination of both expert knowledge and HTE will probably be around for as long as the time-to-market pressure on therapeutic proteins continues. However, as stricter product quality requirements amidst other regulatory concerns gain more focus, we will probably see a growing need for more rational process development approaches. Eventually, hybrid approaches based on clever combinations of expert knowledge, a variety of automated HTE platforms and the algorithmic or optimization based methods supported by dedicated softwares will thrive. Bioprocess modeling and simulation tools developed over the past decades will gain wider applicability. This is especially true as greater understanding of the underlying physical mechanisms in separation techniques is acquired and their mathematical description improved. As the demand and throughput requirements on therapeutic proteins increase, a rising need for better scalable separation techniques will occur. This subject is already coming into focus.^{32,33} While chromatography will continue to be the workhorse for therapeutic protein purification for decades to come, innovations in process chromatography will involve the development of better quality resins that show greater selectivity and less susceptibility to fouling and re-use. Protein engineering strategies for selective protein purification through the use of protein tags or through the development of ligands for selective affinity purification of native target proteins⁷⁰ will also gain ground. There is an increased need for bioprocess synthesis and design methods and tools to improve to levels comparable to their chemical process counterparts. When comparing these two types of processes, two shortcomings of current bioprocess design are immediately apparent: (1) the lack of reliable thermodynamic models for molecular properties estimation; and (2) the general lack of

reliable databases of physicochemical and thermodynamic properties of macro-biomolecules, including proteins. It is in this light that emphasis has been and continues to be on the development of reliable databases as well as mechanistic, as oppose to correlative, thermodynamic models for biomolecules.71,72

SUMMARY

The key to successful and efficient protein purification is selection of the most appropriate purification techniques and their combination in a logical way to obtain the desired purification in the minimum number of steps. However, the rationalization of protein purification process development is faced with a number of challenges. These challenges include, among others, the complex nature of biological feed materials, product stability issues and the often stringent product quality requirements and short timeto-market, especially for biopharmaceutical proteins. This is further complicated by the large number of alternative unit operations and process options as well as the large number of process parameters to be explored. Current protein purification process synthesis methods can be classified as heuristics or knowledgebased, algorithmic, HTEbased, or any combination of the above. Knowledgebased methods are the fastest but rely heavily on heuristics, often resulting in suboptimal processes. The success of methods based on HTE is highly dependent on the realization of experimental techniques expressing a high degree of parallelization and automation suitable for HTS, the development of tools for sample analysis and data handling and powerful strategies for experimental design. Algorithmic or optimization-based methods enable greater process understanding but rely on accurate knowledge of the properties of components in the often complex feed mixtures. Hence, the parallel development of experimental tools for the rapid acquisition of such input data is paramount to the success of optimization-based methods. The hybrid methods are simply combinations of any of the above methods. In that sense, they are the most promising. The strengths and weaknesses of each of the reviewed strategies in addressing the process synthesis problem are summarized in Table 2. This overview shows that there remain challenges in the quest to improve bioprocess development, but also that significant progress has already been made.

REFERENCES

[1] *Pavlou AK and Reichert JM, Recombinant protein therapeutics success rates, market trends and values to 2010. Nat Biotechnol 22:1513– 1519 (2004).*

[2] *BioPortfolio Limited, Biopharmaceuticals Current Market Dynamics & Future Outlook. BioPortfolio Limited, Frampton, UK (2005).*

[3] *Headon DR and Walsh G, The industrial production of enzymes. Biotechnol Adv 12:635–646 (1994).*

[4] *Walsh G, Biopharmaceutical benchmarks. Nat Biotech 18:831–833 (2000).*

[5] *Smith C, Striving for purity: advances in protein purification. Nat Meth 2:71–77 (2005).*

[6] *Thiel KA, Biomanufacturing, from bust to boom... to bubble? Nat Biotech 22:1365–1372 (2004).*

[7] *Lightfoot EN and Moscariello JS, Bioseparations. Biotechnol Bioeng 87:259–273 (2004).*

[8] *Hodge G, Investing in Process Development. wwwbioprocessingjournalcom (2004).*

[9] *FDA, Innovation or Stagnation: Challenge and Opportunity on the Critical Path to New Medical Products. US Food and Drug Administration (2004).*

[10] *Bensch M, Wierling PS, Lieres Ev and Hubbuch J, High throughput screening of chromatographic phases for rapid process development. Chem Eng Technol 28:1274–1284 (2005).*

[11] *Rege K, Pepsin M, Falcon , Steele L and Heng M, Highthroughput process development for recombinant protein purification. Biotechnol Bioeng 93:618–630 (2006).*

[12] *Gosling I, Process simulation and modeling for industrial bioprocessing: Tools and techniques. Ind Biotechnol 1:106–109 (2005).*

[13] *Petrides DP, Koulouris A and Lagonikos PT, The role of process simulation in pharmaceutical process development and product commercialization. Pharmaceutical Eng 22:1–8 (2002).*

[14] *Lu Y, Clarkson A, Titchener-Hooker N, Pantelides C and Bogle D, Simulation as a tool in process design and management for production of intracellular enzymes. IChemE Symposium Series 133:183–189 (1994).*

[15] *Petrides DP, BioPro Designer: An advanced computing environment for modeling and design of integrated biochemical processes. Comput Chem Eng 18:S621–S625 (1994).*

[16] *FDA, Guidance for Industry PAT-A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance. US Food and Drug Administration (2004).*

[17] *Barnicki SD and Stirola JJ, Process synthesis prospective. Comput Chem Eng 28:441–446 (2004). 18 Naonori Nishida GS and Westerberg AW, A review of process synthesis. AIChE J 27:(1981).*

[19] *Westerberg AW, A retrospective on design and process synthesis. Comput Chem Eng 28:447–458 (2004).*

[20] *Petrides DP, Cooney CL and Evans LB, Biochemical process design, in Chemical Engineering Problems in Biotechnology. American Institute of Chemical Engineers (1989).*

[21] Petrides D, *Bioprocess design and economics, in Bioseparations Science and Engineering, ed. by Harrison RG. Oxford University Press, New York (2003).*

[22] Leser EW and Asenjo JA, *Rational design of purification processes for recombinant proteins. J Chromatogr 584:43–57 (1992).*

[23] Wai PPC, Bogle IDL, Bagherpour K and Gani R, *Process synthesis and simulation strategies for integrated biochemical process design. Comput Chem Eng 20:S357–S362 (1996).*

[24] Harrison RG, *Bioseparations Science and Engineering. Oxford University Press, New York (2003).*

[25] Asenjo JA, *Separation Processes in Biotechnology. Marcel Dekker, New York (1990).*

[26] Asenjo JA and Patrick I, *Large-scale protein purification, in Protein Purification Applications a Practical Approach, ed. by Harris ELV and Angal S. Oxford University Press, Oxford, pp. 1–28 (1990).*

[27] Hubbuch J and Kula M-R, *Isolation and purification of biotechnological products. J Non-Equilib Thermodyn 32:99–127 (2007).*

[28] Freydel EJ, Ottens M, Eppink M, Dedem Gv and Wielen Lvd, *Efficient solubilization of inclusion bodies. Biotechnol J 2:678–684 (2007).*

[29] Choe W-S, Nian R and Lai W-B, *Recent advances in biomolecular process intensification. Chem Eng Sci 61:886–906 (2006).*

[30] Cabrera LD and Bottomley SP, *Protein expression and refolding – a practical guide to getting the most out of inclusion bodies. Biotechnol Annu Rev 10:31–50 (2004).*