RESEARCH ARTICLE

Optimization of integrated chromatography sequences for purification of biopharmaceuticals

Anton Löfgren [ORCID] | Mikael Yamanee-Nolin | Simon Tallvod | Joaquín G. Fons |
Niklas Andersson [ORCID] | Bernt Nilsson

Department of Chemical Engineering, Lund University, Lund, Sweden

Correspondence
Bernt Nilsson, Department of Chemical Engineering, Lund University, P.O. Box 124, SE211 00 Lund, Sweden.
Email: bernt.nilsson@chemeng.lth.se

Abstract

With continued development of integrated and continuous downstream purification processes, tuning and optimization become increasingly complicated with additional parameters and codependent variables over the sequence. This article offers a novel perspective of nonlinear optimization of integrated sequences with regard to individual column sizes, flow rates, and scheduling. The problem setup itself is a versatile tool to be used in downstream design which is demonstrated in two case studies: a four-column integrated sequence and a continuously loaded twin-capture setup with five columns.

KEYWORDS
chromatography, continuous bioprocessing, integrated column sequences, multiple column design, optimization

1 | INTRODUCTION

Continuous processes are embraced in several fields, such as petroleum, steel, and so in, but only in recent years biopharma has shown serious consideration into adopting these strategies.[1,2] Despite all advancements that have been made in the pharmaceutical field, there is still a lot of work to be done before implementing and using a complete continuous and integrated process in production. A main reason for this seems to be the added complexity in unit operations compared to conventional batch system variants.[2] However, the benefits seem to be worth it for several reasons. Firstly, there is cost savings with increased productivity, less unit operations, smaller scale production, which reduces equipment size and facility footprint, opportunity for disposable technologies, and reduced operational costs. Secondly, the potential benefit of the increased responsiveness that is drawn from low residence times and cycle times as well as the potential for standardized platforms to flexibly change production between different products. A third perspective is the increased product quality which should come from lower residence times and fewer hold tanks.[3]

Periodic cyclic steady state is achieved when unit operations are connected to each other, but still allow gradients of concentration, pH, and conductivity to elute products.[2] Studies[4-6] have been performed with several chromatography columns in integrated sequences that connected the columns only during elute and load phases, and then pumping through two columns in row. This was performed by redirecting the flow path only during the elute phase to the next column valve, creating a looped sequence.

Including the reactor in continuous bioprocesses is a great step toward achieving a complete end-to-end process.[7] However, this comes with new challenges and complexities where the capture step receives a continuous feed stream. This problem can be dealt with by different techniques, for example, periodic countercurrent chromatography,[1] twin-capture,[8] and capture-simulated moving bed.[9]

When constructing a purification process, several recovery and purification steps are often necessary to isolate the target molecule. The downstream process must be carefully designed due to several reasons: there is no single or general method to purify all kinds of proteins; it is the most expensive step; and it also has a significant effect

on therapeutic performance.[10] Thus, optimization and design of downstream sequences are complex and require a great understanding of all parameters involved. There are helpful holistic process development strategies called expert systems, which construct an optimal purification scheme with different column types and their sequential order by analyzing data about the target molecule.[11,12] Several heuristics also exist for finding the overall best downstream design with regard to the type of columns and their order,[13] for example, using orthogonal purification strategies that separate based on different physicochemical properties such as anion exchange followed by cation exchange instead of two separations of similar type.

The purpose of this paper is to formulate additional heuristics that are useful when designing integrated sequences that consist of direct column-to-column flows. This complements the previously mentioned expert systems and already established design heuristics with a novel perspective for integrated sequences. Two different cases will be studied which both build on the same theoretical background but have different objectives. The first case is the optimization of an integrated and connected batch sequence with regard to column volumes and time. The second case is about adjusting a sequence that handles a constant stream from a continuous reactor dealt with by twin capture followed by an integrated sequence of columns and the objective is to lower the column volumes and reduce the buffer consumption.

2 | MATERIALS AND METHODS

Multiple chromatography columns are often necessary for satisfying purity in downstream processes. As stated in Section 1, integrated sequences with column-to-column pooling in several steps are a way of making the downstream process more effective as opposed to manual batch protocols. However, integration and continuous processing come with added complexity. This section will make an attempt at structuring choices when designing complete integrated sequences. Bear in mind that columns are only connected during elute and load phases.

The volume of a column might be chosen so that the eluting flow rate of the previous connected column must be reduced for the column to be able to handle it or to increase the volume so that the flow rate is accommodated; an example of which is presented in Figure 1, where it follows that the higher order columns have more variable volumes and flow rates. The implication of this is that sometimes a choice must be made of either choosing high volume and thus the possibility of maximum flow rate or the opposite which is a lower volume but then a lower flow rate. So either faster processing or lower
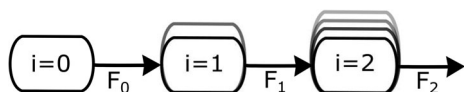

FIGURE 1 A sequence of three columns. The number of different column sizes on each i depicts the variable volume selection that increases with i. Note that the flow is only connected between columns during elute phases

column volumes can be found. This choice must be made for every column connection in a sequence, but as will be shown, in some circumstances there is only a single optimal point at a connection and thus no choice.

The length of the columns is assumed to remain constant during scale-up. This means that the maximum flow rate increases proportionally to the size of each column.

The first column in a sequence is denoted as $i = 0$, the second as $i = 1$, and so on, see Figure 1 for reference. The absolute capacity (mg/CV) of column i must be enough so that it can bind in at least the same amount as the first column, $i = 0$. This is calculated with regard to the product loss and removed impurities on every column step. The term recovery yield, u, is introduced to describe this loss. u is a parameter between 0 and 1 that describes how much of what was captured in a column is recovered and binds to the next. Thus, the absolute capacity of a column is only dependent on the capacity of the first column volume and the intermediate recovery yields.

As the maximum flow rate of a column is based on the pressure drop, and in the coupled eluate and load phases, it is reasonable that the pressure drop will increase because of the two-column train, and a robustness factor is introduced. The robustness factor, when >1, moves the optimal decision slightly away from the maximum flow rate in each step. Capacity robustness will not be discussed as it is a simple matter of increasing all column volumes equally.

x_i is defined in Equation (1) as the ratio of resin volume between two consecutive columns. Column 0 becomes the capacity baseline that the rest of the columns must match. y_i is defined in Equation (2) and the elute flow rate of the last column in a sequence is assumed to be set to maximum. The number of variables to optimize in a sequence is two times the number of columns subtracted by 1, that is, x_i and y_i for every connection. Definitions of other parameters can be found in the nomenclature list.

$$x_i = \frac{V_i}{V_{i-1}}, \tag{1}$$

$$y_i = \frac{F_{i-1,E}}{F_{i,\max}}. \tag{2}$$

2.1 | Constraints

The connection between columns must fulfill certain constraints to avoid above maximal back pressure and be designed to have enough capacity to avoid bottlenecks in the process. These requirements are expressed in mathematical form, based on x and y in Data S1.

There are two flow rate-limiting constraints where one of them sets a maximum flow rate to avoid maximum back pressure on the previous column, $i - 1$, and the other for the back pressure of the current column (including dilution), i, called *current FR limit* and *previous FR limit*, respectively. The *current FR limit* is inclined because its value depends on the current column and scales with it, while the *previous*

FR limit is a horizontal line that the flow rate cannot exceed because it is based on the previous column.

The capacity of all columns must accommodate at minimum the same amount as the first column in the sequence, which is used as baseline. The product loss of each unit is also accounted for. This third and final constraint is called the *capacity limit*.

Demonstrations of the general correlation between two connected columns through x and y are represented visually in Figure 2 by the two subplots where the grey area is outside the constraints and the white area within the constraints. Note that a chosen design value must be above $y = 0$, below the *previous FR limit*, and the *current FR limit* lines and on the right-hand side of the *capacity limit*. The most cost efficient choice is a low x_i and a high y_i which makes the entire *current FR limit* line optimal choices, or the single corner in the other situation.

When x values are close to 1 in a sequence, the volume is a product sum of variables with little variance which makes the problem almost linear, and with completely linear cases there exists strictly either one or two superior points for each connection. These are named *corner* solutions. It is possible to calculate if there are one or multiple optimal points in a connection, the equation can be found in Data S1.

The intricate and interesting aspect is that choices over the sequence interact with each other. For instance, when a choice of x and y is made at $i = 1$, the choice at $i = 2$ will be a different one, as the column volume $i = 1$ is entirely different and absolute capacities change accordingly. This implies that the number of different options at i are multiplied with the number of options $i - 1$. So the number of different purification sequence designs is exponential in relation to the number of columns, except for the single corner solutions.

As there are two variables for every column connection, x and y, a sequence of four columns has a total of six variables in the optimization. The possible so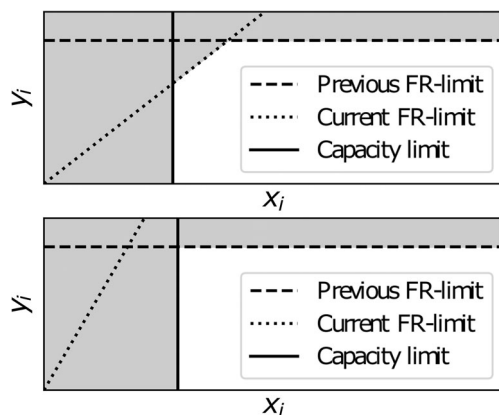lutions to the optimization will be presented in x − y plots wherein every subfigure a point is coordinates for x and y, and thus represent a choice for each column connection.

Ultrafiltration and diafiltration techniques are easily implemented features that can be integrated in the optimization. They would simplify the optimization as they require a flask that the pool from a column is loaded to or from, dividing the integrated sequence in two.

2.2 | Methods

2.2.1 | Case Study I: A general three-column downstream purification

It is very common with downstream processes based mainly on three chromatography purification steps, for capture, intermediate, and polishing. A typical case is the purification of monoclonal antibodies with Protein A-based capture step, a cation exchange intermediate step for aggregate removal, and a final polishing step, often an anion exchange. This is generalized in Case I example with three columns connected in sequence.

The first column is fixed to 24.5 mL; other details of this purification scheme can be found in Data S1. The pool is diluted by a factor of 1.5 in both connections making the *current FR limit* line more horizontal. This case study will compare the combinations of all corner solutions to a Pareto study explained further.

The results will also be compared to a disconnected column sequence, where the pool from one column is not immediately redirected to the next. Some assumptions are made in this comparison: minimal column sizes with regard to capacity are assumed, the maximum flow rate of the column is used during elution and loading, and the same dilution factors as in the connected sequence are used.

2.2.2 | Case Study II: Continuous twin-capture

The second case is an integrated purification platform with five columns that is an extension to Case Study I, demonstrating the possibility of the equation system to schedule a sequence to handle a constant feed stream. The feed stream comes from a continuous bioreactor and has a feed flow rate of 8 L a day and a concentration of 10 mg protein/L. There are two capture columns that alternate receiving the feed. During the loading time of one of the capture columns the rest of the sequence will be running, see the Gantt chart in Figure 3. There are four phases in every column: load, wash, elute, and regeneration, and the elute phase happens simultaneously as the load phase of the following column. The columns should be as small as possible to reduce resin costs while being within the time frame that allows a cycle to be fully purified while the other capture column is loaded.

A new constraint is introduced in this case study, which is that the entire sequence starting from washing the capture column to regenerating the last column must be shorter than the load time of the capture column. To conclude, the objective is still to minimize the total volume of all the columns, but with the added task of handling the continuous feed stream. The volume of the capture column, V_0, is

FIGURE 2 $x − y$ plots. All points in the white region are feasible choices inside the constraints. In the top figure, all points on the i maximum flow line as well as the two corners that lie inside the feasible region are optimal choices, thus multiple choices. In the bottom figure, there is only a single optimal point that is a corner where the capacity line crosses the $i − 1$ maximum flow rate line, that is, no choice
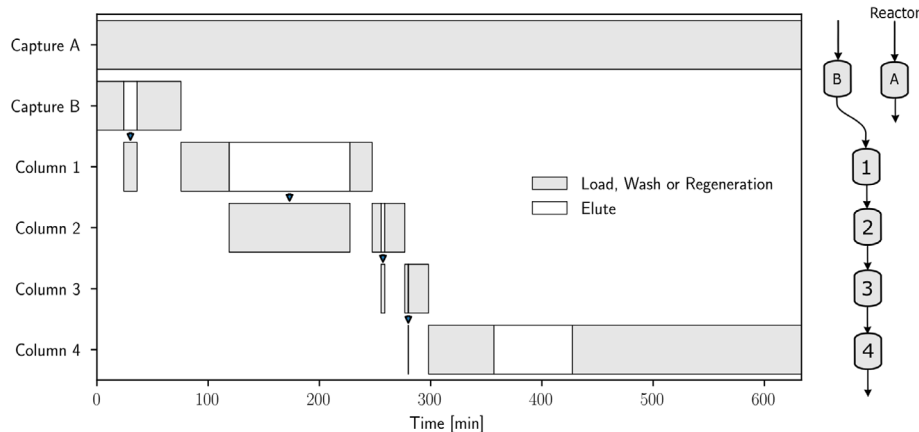
FIGURE 3 Case II. A Gantt chart representation of all steps in the continuous twin-capture sequence. Column A or B is loaded while the rest of the sequence is running. After a complete sequence, connections to A and B are switched and B is loaded instead while a new sequence starts with a Column A wash step. Note that on every column step the load phase is active simultaneously as the elute step of the previous column, apart from the capture column

therefore added as an additional variable while it was fixed in the previous case. All necessary data for the sequence are presented in Table S2 found in Data S1.

2.2.3 | Multiobjective optimization

In both case studies, there are two competing objectives to be minimized. In the first case, the objective is to reduce the total column volume and the total process time. As these two tasks are competing and are not compatible with each other, there will instead of a single optimal point be several. This is why a new objective function is introduced which tests for both objectives but at different weights to the individual objectives and finds the most efficient choice for each weight. This is called a Pareto optimization.

The overall objective for the second case is to find a solution that handles the constant feed stream. As buffer consumption also is a nonnegligible cost when producing pharmaceuticals, it is worthwhile to investigate how a set of different column volumes impacts it. Therefore, a Pareto optimization described in Data S1 is performed also in the second case study to show this relationship.

2.3 | Solver

The optimization is a nonlinear problem as the solution depends on a product sum of x values, see Equations (4) and (7) in Data S1. This generally means that the optimal point does not have to be at the constraint boundary or in any of the corners.

The optimization was carried out using the Scientific Computing Tools for Python, SciPy, version 1.1.0, https://scipy.org/. The function *optimize.minimize* from *SciPy* was used, with the method sequential least squares programming, which is an iterative method for solving nonlinear optimization problems consisting of both constraints and bounds.

3 | RESULTS AND DISCUSSION

The robustness factor moves the flow rate down slightly in all connections. The grey lines in Figures 5 and 7 are plotted to show potential

choices without any robustness factor. Throughout this section, solutions are referred to as complete sequences and choices as the individual steps within the solutions.

3.1 | Case I

All optimized results overlap the corner solutions, although one of the solutions was never selected by the algorithm as an optimal solution, see Figure 4. Solution *vv*, volume-then-volume, and *tt*, time-then-time, are minimizing the volume and the process time, respectively, while solution *vt* and *tv* are mixes of the two, where *tv* was exempted in the Pareto optimization.

The choices of solution *tv* are shown in Figure 5 with the dots, first choosing a low volume and flow rate and then the second choice is higher volume and flow rate. Solution *vt* is the opposite, where the greater volume is initially selected and then the lower volume with the result that the last volume is still higher overall, since x is the relation between volumes of column i and $i-1$. So even if it would be
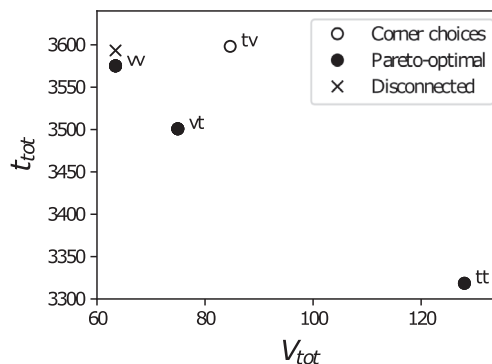

FIGURE 4 Case I Pareto front of solutions. All circles were obtained when selecting corner choices in the $x-y$ plots. The symbols represent the choice in each solution: *vv* is volume-then-volume, *vt* is volume-then-time, *tv* is time-then-volume, and *tt* is time-then-time. The full circles are choices selected by the Pareto optimization. As *tv* was the only corner choice not selected, it is not Pareto-optimal. The disconnected solution is where the columns are not coupled and the pool is not directly applied to the next column
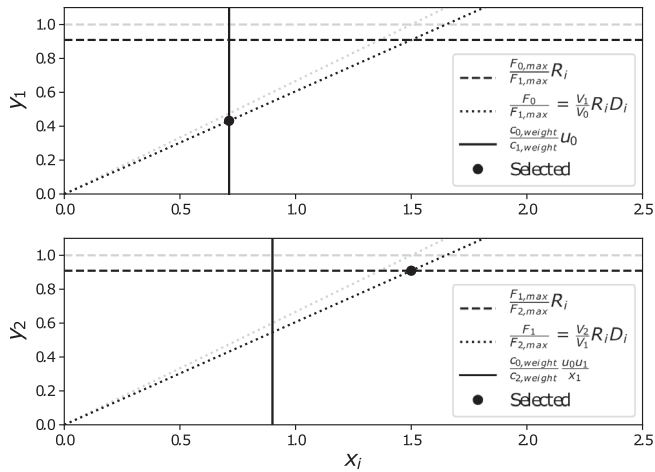
FIGURE 5 Case I. The choices in solution tv from Figure 4. The choice in the upper part is the lowered flow rate, allowing the column volume to be lower. The second choice is the higher flow rate resulting in a slightly higher column size in the last step. Note that the grey lines are the nonrobust constraints shown to visualize the effect of the robustness factor, in this case $R = 1.1$, which gives a lowered flow rate of 10%

the same x value, the absolute volume is still greater. This result is a demonstration that early choices have an impact on later choices, as discussed in the section 2.1.

A comparison of connected and disconnected column connections was made and is represented in Figure 4. The disconnected solution for this case is close to the connected solution. However, this depends greatly on the dilution factor. Small values of the dilution factor make the connected solution more efficient and vice versa. This is explained by that the incline of the *current FR limit* forces down y which increases t_{tot} and thus defeats the purpose of connecting the columns. With dilution factors equal to 1 and comparing the lowest volumes, t_{tot} for disconnected is 3,493 min and for connected 3,396 min (not shown in any figure). Our conclusion is that connected sequences are generally more beneficial if no or low dilution is required between steps, otherwise the flow rate becomes lowered so much that a disconnected sequence might be faster, assuming a time efficient dilution.

The results from this case study can help to design a variety of integrated sequences as the case is quite general with its three connected bind-and-elute columns.

The results from this case study can be directly applied to complete downstream purification units as was discussed in Section 1.[4-6] For research purposes, integrated and automated processes that can be sped up can increase the research output when testing a variety of candidates. At large manufacturing facilities it might be worthwhile to keep volumes down if steps are connected as resin material is a great cost at full-scale.

3.2 | Case II

The solution with the highest total volume and lowest buffer consumption ($w = 0$) had a capture column size of 304 mL which

decreased to 213 mL ($w = 1$) while traversing the Pareto in Figure 6. The cycle time for Case Study II changes with the capture column size from 820 min ($w = 0$) to 575 min ($w = 1$). It is reasonable that lower volumes also had lower cycle times, and also that the buffer consumption increased, per mg product, as the buffer exchangers were larger to be able to handle the flow rate making it possible to stay within the cycle time constraint.

Comparing y_2 in Figure 7, the optimal solution is on the flow rate limit constraint for lower w values and not in a corner. This special balance might be because the connection between a buffer exchanger and a bind-and-elute column, that is, large to small column, gives a very small *capacity limit*. With $w = 0.2$, the corner with the greater x value corrects the total volume while a low flow rate limits the cycle time. Apparently, these corners are worse with these special circumstances, which influence the algorithm to select a point on the *current FR limit* line.

Lower capture column size corresponds to lower total volume of the sequence. Even though x_2 is higher at $w = 0.8$, the initial capture column size, and x_3 are low enough so that the total column size is lower than at $w = 0.2$, see Figures 7 and 6.

As there were big size differences of columns in this case, all connections were not simple choices and instead there were gliding optimal points over the Pareto front. The implication is that choices are more complex in a connected step from a large to a small column.

This case study was more specialized but allowed the reader to see how a downstream section, with a constantly received volumetric flow rate as a constraint, might be scheduled with the help of the nonlinear optimization described in this paper. The practical significance of this is that integrated sequences that have a large variety of inherent choices that must be made can be optimized to function within a constraint by tweaking volumes, flow rates, and cycle times.
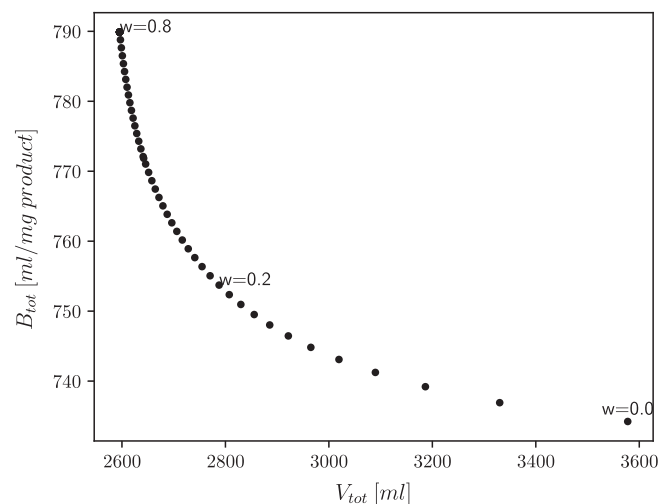

FIGURE 6 Case II Pareto front over the total volume of columns and the total buffer consumption. The location in the Pareto weights are shown at 0 and the two locations shown in Figure 7 at 0.2 and 0.8
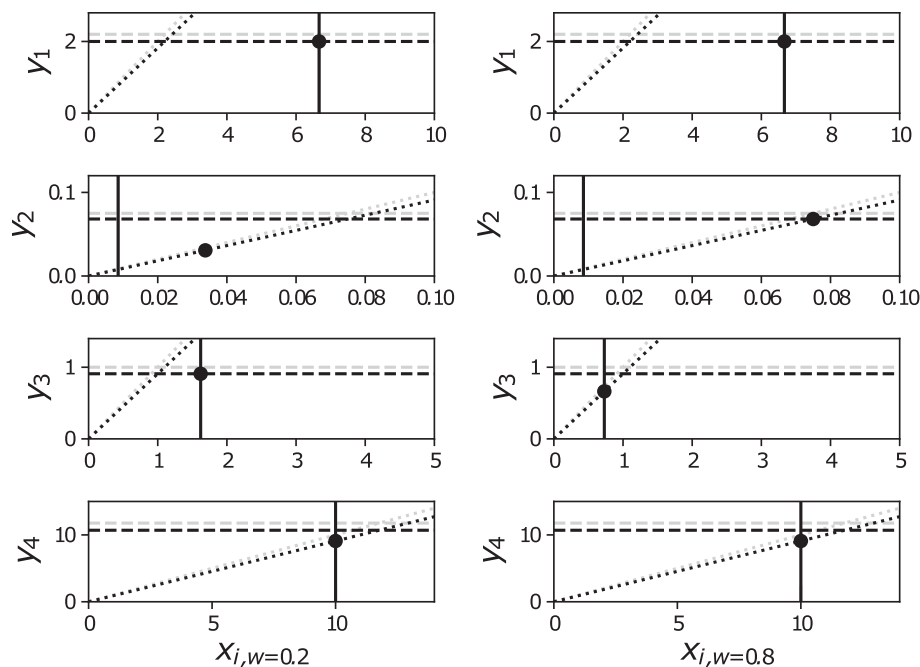
FIGURE 7 Case II. $x - y$ plots over the four-column connections with $w = 0.2$ in the left subfigures and $w = 0.8$ in the right subfigures. For $w = 0.2$, the capture column size is 235 mL and the total volume is 2,795 mL. For $w = 0.8$, the capture column size is 213 mL and the total volume is 2,593 mL

3.3 | General remarks

Interestingly, bind-and-elute columns act similarly to buffer exchange columns. The only principal difference is the magnitude of the x values and therefore the volumetric flow rate in the elute phase must be set carefully to not exceed the maximum flow rate limit of following column.

When the capacity is high enough in column $i - 1$ compared to column i, there will only be a single optimal choice. Newer versions of affinity columns usually have a very high capacity, compared to maximum flow rate, which reduces the overall problem size as the following column must be relatively large to handle the load which makes it likely that it is able to handle the maximum eluting flow rate.

4 | CONCLUSIONS

For designers of downstream purification processes, this study can be seen as a tool to use for integrated and continuous sequences and a couple of different heuristics can be drawn from the exemplified case studies.

Utilizing the demonstrated algorithm, a sequence of choices can tune a purification process toward either speed or column size. Both are important parameters during optimization and scheduling of a downstream process to keep resin material costs down, increase the speed of a system, and calibrate a complex system to handle a constant feed stream from a perfusion reactor.

Assumed optimal choices, corner choices in the article, reflected accurately the optimization of Case Study I when there were relatively equal column capacities. Even though the problem was nonlinear, assumed optimal choices were the same as those chosen by the optimization algorithm, except for one solution. Our advice from that

result is that if speed is necessary, keep volumes down early in the sequence and increase only in the final steps.

Connected sequences are faster if there is no or low dilution factors between column steps as the elute and load phases are combined. With dilution, however, the flow rate must be lowered to avoid transcending the maximum back pressure.

With the given reactor feed, the optimization, combined with constraints and column data, was shown to be able to produce an optimal operating sequence. The amount of buffer used can be traded with some reduction of total resin volume.

Manufactured columns come in a selection of column sizes and radii. Instead of looking at a continuous range of column radius, it could be more realistic to study discrete values of column sizes. Another expansion to this study could be to include packing material costs as an additional parameter for a more detailed cost analysis.

NOMENCLATURE

B	Buffer consumption (mL/mg product)
BE	Buffer exchanger (−)
c	Capacity by weight or volume (mg product/mL packing) or (mL load/mL packing)
CV	Column volumes (−)
D	Dilution factor (−)
F	Volumetric flow rate (CV/min)
n	Number of connections (−)
R	Robustness factor (−)
t	Time (min)
u	Recovery yield (−)
V	Column volume (mL/column)
w	Weight factor (−)
x	Column volume quotient (−)

y Flow rate factor (–)

Greek Symbols

φ Phase volumes CV

ORCID

Anton Löfgren https://orcid.org/0000-0002-8182-8816
Niklas Andersson https://orcid.org/0000-0002-6032-5494

REFERENCES

1. Warikoo V, Godawat R, Brower K, et al. Integrated continuous production of recombinant therapeutic proteins. *Biotechnol Bioeng*. 2012; 109(12):3018-3029.
2. Rathore AS, Kateja N, Agarwal H. Continuous downstream processing for production of biotech therapeutics. In: Subramanian G, ed. *Continuous Biomanufacturing: Innovative Technologies and Methods*. Weinheim, Germany: John Wiley & Sons; 2017:261-288.
3. Patil R, Walther J. Continuous manufacturing of recombinant therapeutic proteins: upstream and downstream technologies. *New Bioprocessing Strategies: Development and Manufacturing of Recombinant Antibodies and Proteins*. Framingham, MA: Springer; 2017:277-322.
4. Andersson N, Lfgren A, Olofsson M, Sellberg A, Nilsson B, Tiainen P. Design and control of integrated chromatography column sequences. *Biotechnol Prog*. 2017;33(4):923-930.
5. Lofgren A, Andersson N, Sellberg A, Nilsson B, Lofgren M, Wood S. Designing an autonomous integrated downstream sequence from a batch separation process—an industrial case study. *Biotechnol J*. 2018; 13(4):e1700691.
6. Steinebach F, Ulmer N, Wolf M, et al. Design and operation of a continuous integrated monoclonal antibody production process. *Biotechnol Prog*. 2017;33(5):1303-1313.
7. Godawat R, Konstantinov K, Rohani M, Warikoo V. End-to-end integrated fully continuous production of recombinant monoclonal antibodies. *J Biotechnol*. 2015;213:13-19.
8. Angarita M, Müller-Spth T, Baur D, Lievrouw R, Lissens G, Morbidelli M. Twin-column CaptureSMB: a novel cyclic process for protein A affinity chromatography. *J Chromatogr A*. 2015;1389:85-95.
9. Baur D, Angarita M, MüllerSpth T, Steinebach F, Morbidelli M. Comparison of batch and continuous multicolumn protein A capture processes by optimal design. *Biotechnol J*. 2016;11(7):920-931.
10. Costa S, Almeida A, Castro A, Domingues L. Fusion tags for protein solubility, purification and immunogenicity in *Escherichia coli*: the novel Fh8 system. *Front Microbiol*. 2014;5:63.
11. Nfor BK, Verhaert PD, van der Wielen LA, Hubbuch J, Ottens M. Rational and systematic protein purification process development: the next generation. *Trends Biotechnol*. 2009;27(12):673-679.
12. Vasquez-Alvarez E, Pinto J. Efficient MILP formulations for the optimal synthesis of chromatographic protein purification processes. *J Biotechnol*. 2004;110(3):295-311.
13. Baumann P, Hubbuch J. Downstream process development strategies for effective bioprocesses: trends, progress, and combinatorial approaches. *Eng Life Sci*. 2017;17(11):1142-1158.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Löfgren A, Yamanee-Nolin M, Tallvod S, Fons JG, Andersson N, Nilsson B. Optimization of integrated chromatography sequences for purification of biopharmaceuticals. *Biotechnol Progress*. 2019;35:e2871. https://doi.org/10.1002/btpr.2871