
**Optimal Process Mode Selection for Clone Screening**

Jure Strnad,1,3,* Vatroslav Spudić,1 Matjaž Brinc,1 Jerica Rozman Pungerčar,2 and Zdravko Kravanja3

1 Department of Bioprocess Development, Lek Pharmaceuticals d.d., a Sandoz Company, Biopharmaceuticals, Kolodvorska 27, SI-1234 Mengeš, Slovenia

2 Analytical Department, Lek Pharmaceuticals d.d., a Sandoz Company, Biopharmaceuticals, Kolodvorska 27, SI-1234 Mengeš, Slovenia

3 University of Maribor, Faculty of Chemistry and Chemical Engineering, Smetanova 17, SI-2000 Maribor, Slovenia

* Corresponding author: E-mail: jure.strnad@sandoz.com
Phone: ++386 1 721 79 21

Received: 18-01-2011

**Abstract**

A significant amount of data is generated during clone screening procedures because several cultivation systems are used such as well-plates, shake-flasks, laboratory-scale bioreactors among others. The amount of data can often be staggering and thus requires the use of statistical and data mining procedures in order to extract hidden information. In regard to a biosimilar monoclonal antibody project, data from the shake-flask scale was observed when comparing batch and fed-batch processes for several individual clones derived from two different cell lines. The aim of the research presented in this paper was to determine if clone selection based on batch data during the initial screening phase was equivalent to the results from fed-batch selection. It was determined that fed-batch processes should be implemented early during the screening procedure, if a choice is made that the final manufacturing process should also be operated in a fed-batch mode, as clone selection based on both batch and fed-batch analytical data was divergent.

**Keywords:** Monoclonal antibody, batch, fed-batch, clone screening, glycosylation, charge variants.

**1. Introduction**

Antibodies, or immunoglobulins, are serum proteins that play a central role during humoral immune response, binding antigens and inactivating them or triggering an inflammatory response which results in their elimination.1 The use of monoclonal antibodies as therapeutics holds great promise for treating many human diseases, including autoimmune and inflammatory conditions, cancers, and viral diseases.2 They constitute more than 30% of the total biopharmaceutical production and are the largest class of proteins currently under clinical trials.3 The production costs of monoclonal antibodies are estimated to currently be at 300 USD per gram of product as against a required reduction in the region of 50 USD per gram.4 Recently cell culture titers of 5 g/L have been reported in the literature.5,6

In the course of biosimilar development, a step-wise approach is used for the selection of host cell lines, pools and clones, thus producing a complex product with the desired product quality profile. Typically, a screening process starts with a number of different host cell lines that are transfected with a vector construct containing the product’s gene sequence. Recombinant Chinese hamster ovary (CHO) host cell lines are generally used during biosimilar production as they display the efficient post-translational processing of complex proteins and because of the similarity between the glycosylation patterns of native human and CHO-derived recombinant proteins.7,8 The transfected host cell lines then undergo further selection and amplification procedures in order to produce cell pools, that are then analyzed for titers and product-quality attributes, as for example glycosylation and charge-variant patterns. Glycosylation is a critical protein-quality attribute that can modulate the efficacy of a commercial therapeutic glycoprotein.9–15 Ideally, product-quality attributes should already exhibit a close similarity to the target product during the screening phase because it is still poorly

Strnad et al.: *Optimal Process Mode Selection for Clone Screening*
understood as to which product-quality attributes can be predominately influenced by process changes and which are predominantly based on genetics. It is presumed that most of the product-quality attributes can be influenced by both, and thus a clear distinction between the two cause categories is difficult. Normally one or more pools are chosen as the sources of individual clones from the set of analyzed pools. The number of chosen pools, the number of generated clones, and the number of clones isolated from individual pools are arbitrary, tailored to clone screening capacities and timelines, significantly differing from project to project. At the end of the screening process, one production clone is selected and is used for process characterization and scale-up. It should be noted that the selection of highly productive mammalian cell lines causes a bottleneck regarding process development during the production of biopharmaceuticals, and is significantly limited by the number of clones that can be feasibly screened.

During the pool and clone screening processes, cultivation vessels of different scales are used, as for example well-plates, shake-flasks, laboratory-scale bioreactors, etc., sometimes under different conditions. However, these variations in clone cultivation are still poorly understood and it is speculated that they could well have a significant effect on the final clone selection. One such factor could be the process mode, for example batch, fed-batch, or perfusion. Early screening steps are normally carried out in a batch mode, however, as currently fed-batch processing at the manufacturing scale is the dominant mode of operation for mammalian cell culture processes, it is very important to decide which mode to apply at a particular screening stage with respect to the manufacturing scale mode. This paper focuses on data generated from both the batch and the fed-batch modes when analyzing a variety of cell clones, to assess if clone selection results differ in different process modes. The effect of process mode change from batch to fed-batch was evaluated by analyzing the product’s quality attributes and categorizing them into two groups: those with increasing and those with decreasing attribute values when changing the process mode. Additionally, a simple ranking procedure for all the clones was applied in order to present the influence of a different process mode on final clone selection.

2. Experimental

2.1. Cultivation Systems

An analytical characterization of 30 clones derived from a CHO dihydrofolate reductase negative (dhfr−) cell line was carried out for a biosimilar project. The product gene sequence was inserted by nucleofection, whilst selection and amplification were carried out by using a standard dhfr− protocol. Pools 17 and 26 were chosen beforehand in the primary selection process. Then, 22 clones were isolated from pool 17 and 8 clones from pool 26 for growing in both batch and fed-batch modes within sterile single use shake-flasks (Corning Incorporated Life Sciences, USA) placed in an ISF1-X orbital shaker incubator (Kühner AG, Switzerland) unit under controlled environmental conditions.

The process settings for batch cultivation were a seeding density of 0.15 · 10⁶ viable cells/mL within the pre-production stage medium, and 250 mL shake-flasks at 37 °C, 10% CO₂, 110 rpm for 10 days. The process settings for fed-batch cultivation were a seeding density of 0.40 · 10⁶ viable cells/mL within the production stage medium, and 500 mL shake-flasks at 37 °C, 10% CO₂, 150 rpm for 14 days with a complex generic feeding strategy. During the fed-batch process two different feeds of fixed volumes composed of concentrated glucose and amino acid mixtures were added starting at two different time points. Since the effects of different feeding protocols on the screening process were not the focus of this paper, the feeding regime was not changed throughout the experimental work. At the end of both the batch and the fed-batch processes, culture harvesting was carried out and individual harvests were used for analytical characterization.

Both media used during the research work were prepared in-house, animal component free, and chemically-defined. They contained essential substrates and growth factors, such as glucose, vitamins, mineral salts, lipids, insulin, glutamine, and several other amino acids needed for cell growth.

2.2. Analytical Methods

Analytical methods for product characterization were performed on the purified samples. Prior to purification all harvests were centrifuged for 5 minutes at 1000 g and filtered, in order to remove any remaining cells and cell debris. The product from the clarified harvests was captured, concentrated, and purified using Protein A chromatography.

The protein A purification was performed using the Freedom EVO pipeting station (Tecan Group Ltd., Switzerland) with 200 µL RoboColumns (Atoll GmbH, Germany) filled with MabSelect™ SuRe chromatography resin (GE Healthcare Ltd., UK). The product was eluted with a 50 mmol/L, pH 3.5 sodium acetate buffer into sample tubes with 10 mmol/L TRIS base neutralization buffer.

Monoclonal antibody content (titers) in the cell harvests and purified samples were determined using Protein A affinity liquid chromatography (ALC). An Agilent 1200 series high-performance liquid chromatography (HPLC) system was used with an UV detector and a Poros® 20 µm Protein A ID Cartridge (Applied Biosystems, USA). The product was eluted by lowering the pH from 7.5 to 2 and detected at λ = 280 nm.

Glycan structures (Gmap) were profiled using normal phase high-performance liquid chromatography (NP-HPLC). The N-glycan structures bound to the Fc region of the monoclonal antibody were first enzymatically relea-
sed, then labelled with 2-aminobenzamide (2-AB, Fluka, Switzerland), and finally separated chromatographically. An Agilent 1200 rapid resolution system was used with an ACQUITY UPLC BEH Glycan (Waters, USA) normal phase chromatographic column. The labelled glycan structures were separated on the basis of polarity.

Charge variants were separated and quantified by cation exchange chromatography (CEX). An Agilent 1200 series high-performance liquid chromatography system with a ProPac® WCX-10G analytical column (Dionex, USA) was used for the analysis. Differently-charged variants were separated based on the distinct electrostatic interactions of the variants with the stationary phase, and eluted using a gradient of increased sodium chloride concentration in the mobile phase.

2.3. Data Evaluation

Batch and fed-batch data for all clones was evaluated using correlation plots. Additionally, 95% confidence bands around the regression line, calculated by using data from all clones were used to evaluate the difference or similarity of batch and fed-batch data in all the correlation plots. If the individual plot diagonal which designated ideal similarity between the batch and fed-batch modes, lied in the regression line’s 95% confidence region, the batch and fed-batch data was assessed as being similar.

Decreased or increased structure values when the process mode was changed from batch to fed-batch, were evaluated by calculating relative differences expressed as percentages. For example, the pool 17 average fed-batch value for the 2K species was 2.17% and the average batch value was 3.59%. The difference was, therefore, –1.42% and was divided by their average value of 2.88% and then multiplied by 100 to calculate the end relative value of –49%. The 2K species, therefore, was decreased when the process mode was changed from bath to fed-batch processing.

To obtain the individual clone scores, the structural values ($x_i$, batch and $x_i$, fed-batch) for fifteen glycan species were evaluated using correlation plots.
re subtracted, and then divided by the reference value (k) in order to calculate relative proportions. All proportions were subsequently re-calculated as absolute values, and summarized (equation 1).

\[
\text{Score}^{\text{batch}} = \sum_{i=1}^{15} \frac{x_i - k}{k} \quad \text{and} \quad \text{Score}^{\text{fed-batch}} = \sum_{i=1}^{15} \frac{x_i - k}{k}
\]

The simple scoring procedure was used only to generate a better graphical representation of the data and was not used as a selection tool.

3. Results and Discussion

The product-quality attribute values for each clone from both pools were separately plotted as correlation plots for each attribute. The batch values for product-quality attributes were plotted onto the abscissa, and the fed-batch attribute values onto the ordinate. If the process mode did not have an influence on the values of the product-quality attributes, the points representing individual clones would be clustered around the plot’s diagonal and the diagonal would also fall into the 95% confidence region. However, if the process mode did influence the values, the points would deviate toward one of the axes and the diagonal would be outside of the 95% confidence region. Clones that exhibited a significant deviation from the majority of data points are individually marked in all the Figures.

The glycosylation pattern for the researched monoclonal antibody is quite complex. The major complex biantennary glycan structures are schematically presented in Figure 1.

Individual glycan structural values are graphically presented in Figure 2. As mentioned, the diagonal lines on the plots represent the positions where the batch and fed-batch data are identical. Protein-linked glycans modulate the biochemical attributes of proteins, such as bioactivity, folding, and immunogenicity.\(^{22,23}\)

In general, the value distributions for the glycan map results were inconclusive, but indicated differences between batch and fed-batch results. The most significant observation was that the clones from pool 26 were generally more closely clustered together than those from pool 17 and, therefore, exhibited a lower performance variability for individual clones. In other words, whichever clone from pool 26 had been chosen, the likelihood that it would exhibit similar characteristics as the other clones from pool 26 would be very high. On the other hand clones from pool 17 exhibited in some cases two distinct populations, as for example for the bG0 (-N), bG0 (-F), bG0, Mannose 5 and bG2 species and, therefore, a higher performance variability. In practice this would mean that a chosen clone would only represent the characteristics of the population from which it had been chosen and could not be representative for the behaviour of all clones. Based on the 95% confidence bands, only for the bG0 (-F), bG0, Mannose 5, bG1 (1-6) and bG2 species a similarity between batch and fed-batch data could be confirmed. For the majority of the presented species the corresponding data was not equal. Choosing the most perspective clones from a number of pools is extremely complex as it is not entirely understood as to which product-quality attributes are fixed by the genetic set-up of the clones and, therefore, cannot be significantly influenced by process settings, as against which clone characteristics are dependent on process conditions and, therefore, can be adjusted during the process development phase. Another significant observation was also that some low-amount structures were detected during one process mode but were unobserved during the other process modes, as for example Mannose 3 or Mannose 8 data points found on the abscissa. Such cases occurred for those structures that were considered to be insignificant for biological activity and product safety. It should also be noted that the glycan distributions for all 30 clones during this phase of the selection process were not comparable to the reference product (data not shown). However, this was not surprising as the clones were cultivated in shake-flasks, which are not as controllable as bioreactors.

Individual charge variants are presented in Figure 3, where the pattern of pool 26’s higher consistency compared to pool 17 was observed repeatedly. Some batch data from the CEX analysis were unavailable, therefore, fewer points are presented in Figure 3 than in Figure 2.

Similarly as for the glycan structural distributions, it was observed that charge variant distributions shifted when the process mode was altered from batch to fed-batch cultivation. Charge variants are distinguishable based on the occurrence of C terminal bound lysine or proline amide molecules on the monoclonal antibody’s two heavy chain ends. Variants with no C terminal bound lysine or proline amide molecules are designated as 0K, variants with single lysine or proline amide molecules are designated as 1K, and variants with dual lysine and/or proline amide molecules are termed as 2K. Variants that elute earlier than the 0K peak are referred to as acidic variants because they are less positively-charged and, therefore, bind less tightly to the column. Variants that elute later than the 2K peak are referred to as basic peaks. Lysine variants need to be monitored to ensure manufacturing consistency, but are not expected to affect product efficacy or safety.\(^{24}\) The charge variant distributions of the clones were somewhat more similar to the target product in comparison with the glycan structural distributions. However, the discrepancy was still significant during this step of the screening process and an additional adjustment had to be carried out by optimizing process settings. Two popula-
Figure 2. Correlation plots with 95% confidence bands (green) for individual glycan structures. The regression lines are represented in violet, data points of clones generated from pool 17 are represented as red circles, clones generated from pool 26 are represented as blue triangles.
tions for pool 17 were also observed for the 1K and 2K structures. The only case in which the 95% confidence bands included the diagonal were the Acidic Variants.

Individual clone titers are presented in Figure 3, where a significant shift was observed towards fed-batch processing. The titer results were as expected, as it is well-known that titers increase when the process mode is changed from batch to fed-batch. In fed-batch cultivation the cells receive an additional influx of substrates by feeding the culture with a fresh medium or only certain individual medium components that allow for higher viable cell densities and, therefore, higher product concentrations. As a result of additional medium feeding, fed-batch processes can endure several days more than batch processes. Generally, batch cultivation is terminated because essential substrate concentrations are depleted. However, fed-batch processing can also be terminated because of high inhibitory concentrations of certain metabolites, as for example lactate, glutamate, and ammonium. Clones from pool 17 were spread over both broad fed-batch titer intervals and very narrow batch titer intervals. In practice this would mean that distinguishing clones by batch-derived titers would be difficult. On the other hand, differentiating clones by fed-batch titers would be easier. It was interesting to observe that clone C1 from pool 26 had an extremely increased titer during the fed-batch processing in comparison with the other seven clones from the same pool.

A deviation in product-quality attribute values towards the abscissa or ordinate was evident in several cases, as for example for the Mannose 8 species and titers, respectively. A mathematical evaluation of the data from both process modes was performed in order to support the graphical results and is presented in Figures 4 and 5. The pools were analyzed individually, as their characteristics differed significantly. Negative values in the plots signified that the structural value was decreased, whereas a positive structure value signified that it had increased when the process mode was changed from batch to fed-batch.

Attributes bG0 (-F), bG0, Mannose 5, bG1 (-F), bG1 (1-6), bG1 (1-3) and bG2 displayed an opposing trend regarding pool 17 and pool 26, respectively. However, both pools showed a structural value decrease for Mannose 3, Mannose 7, Mannose 8, 1K, 2K and Basic Variants, whereas the structural values for bG0 (-N), bG1 (-N), Mannose 6, bG2S1, Acidic Variants, 0K and titers increased. It was evident that, if the analysis had been made using the average values for each product-quality at-

![Figure 3. Correlation plots with 95% confidence bands (green) for individually-charged structures and titers. The regression lines are presented in violet, the data points of the clones generated from pool 17 are represented as red circles, and the clones generated from pool 26 are represented as blue triangles.](image-url)
tribute calculated from both pools, the attribute value increases and decreases would not have been representative for both pools, as some of the attributes exhibited opposing trends.

Additionally, a two-tailed $t$-test with unequal variance was applied to assess how the process mode influenced the differentiation of the pools. A graphical representation of the results is presented in Figure 6. The red line represents the limit $p$-value of 0.05. All values above this limit signify that the pools are statistically the same when referring to individual product-quality attributes, whereas all values below this limit signify that the pools can be differentiated.

Attributes Mannose 6, Mannose 8, bG0 (-N), Mannose 5, 1K and 0K differentiate pool 17 from pool 26 in both batch and fed-batch modes. Other product-quality attributes differentiate the pools only in one of the two process modes, except for the bG2, Mannose 3 and bG0 (-F) species. It is evident from Figure 6 that differentiation of both pools is harder to achieve in the batch mode. From among all the 21 product-quality attributes only 9 attributes differentiate both pools in the fed-batch mode. The conclusion of this statistical analysis was that fed-batch processing allows for a better differentiation of both pools and, therefore, should be applied early during the screening process. Altogether, it was evident that any change in process mode can influence product-quality attributes.

All 29 clones were additionally scored according to their differences from the target glycan distribution, for both process modes. One clone was omitted from the ranking as it displayed incomplete data. The final scores for clones cultivated in fed-batch mode and batch mode are presented in Figure 7, with the reference scored as zero value.

As can be seen from Figure 7, all clones exhibited somewhat different glycan profiles from the target profile during this phase of the screening process. It was observed that clone C79 from pool 26 was distinct in comparison with other clones from pool 26. In addition, the group of clones C22, C31, C38, C42 and C48 featured distinct values in comparison with the other clones from pool 17. From this simple ranking procedure, clone C25 from pool 17 would be the most promising candidate for further process development. However, it has to be stressed that production clone selection is based on several more process and product-quality attributes than shown in this simple il-
It is interesting to note that for the group of five outlying clones from pool 17 the average ranking scores derived from the fed-batch data (the average value was 22.57) were slightly lower and hence closer to the target profile than the ranking scores derived from the batch data (the average value was 33.36), which can be seen as a shifting of the five points towards the right-side of the diagonal plot in Figure 7. The diagonal was in this case not included into the 95% confidence region.

4. Conclusion

A statistical analysis of batch and fed-batch data on the shake-flask scale was made, in order to identify the influence of any process mode change on product-quality attributes. It was evident that the process mode is very influential and, therefore, any mode that is planned for the manufacturing process should be applied at an early stage during the screening procedure. Product-quality attributes were also divided into those that increased and those that decreased, if the process mode was changed from batch to fed-batch cultivation. It is evident from the analysis that the clone screening data exhibited a significant variance, therefore, calculations using averages can often be misleading and should be used carefully. It was also evident that
the results derived from fed-batch processing enabled a more efficient differentiation of clones. The presented data could also be evaluated using multivariate data analysis (MVDA) techniques instead of using individual score techniques. The application of these complex techniques is being investigated and will be a goal of our future research in the optimization of clone selection procedures.

5. Acknowledgements

We gratefully acknowledge the partial financial support from the grant No. 3211-06-000501 of the Ministry of Higher Education, Science and Technology of the Republic of Slovenia, and the program P2-0032 from the Slovenian Research Agency. We would also like to thank Lek Pharmaceuticals d.d., a Sandoz Company, for its support, especially the Biopharmaceuticals departments of Cellular and Microbial Biology, Bioprocess Development, Down-stream Processing Development and Analytics.

6. References


Povzetek

Pri izboru produkcijskega klona je potrebno zbrati in ovrednotiti veliko količino podatkov, saj pri izboru uporabljamo različne gojustvene sisteme, kot so na primer ploščice z gojustvimi jamicami, erlenmajerice, bioreaktorji itd. Obdelava vseh podatkov je ponavadi zelo dolgotrajna in pogosto je potrebno uporabiti različne matematične in statistične metode za ovrednotenje skritih informacij. V pričujočem članku smo primerjali podatke pridobljene pri gojenju skupine klonov, ki proizvajajo monoklonsko protitelo, in sicer tako šaržne rezultate kot tudi analitske rezultate pridobljene pri semi-šaržnem gojenju kulture, torej z dohranjevanjem. Cilj raziskave je bil ugotoviti ali so šaržne pridobljeni podatki ekvivalenti podatom pridobljenim s semi-šaržnim procesom in kako te vrednosti vplivajo na izbor produkcijskega klona. Ugotovili smo, da podatki, pridobljeni z različnimi procesnimi ustroji, niso primerljivi, kar kaže na potrebo, da se v izboru produkcijskega klona uporabi tisti procesni ustroj, ki je načrtovan za samo industrijsko produkcijo.