Metabolic flux-based modeling of mAb production during batch and fed-batch operations

Penny Dorka · Christian Fischer · Hector Budman · Jeno M. Scharer

Received: 20 February 2008 / Accepted: 21 May 2008 / Published online: 17 June 2008
© Springer-Verlag 2008

Abstract This paper proposes mathematical models that predict the physiology, growth behavior and productivity of hybridoma cells in both batch and fed-batch systems. Murine hybridoma 130-8F producing anti-F-glycoprotein monoclonal antibody was employed as a model system. A systematic approach based on metabolic flux analysis (MFA) was utilized to yield a dynamic model for predicting the concentration of significant metabolites over time. Correlation analysis was performed to formulate a Biomass Model for predicting cell concentration and viability as a function of the extracellular metabolite concentrations. The coefficients of the model equation were estimated by employing the Metropolis–Hastings algorithm. The Metabolites Model was combined with the Biomass Model to get an Integrated Model capable of predicting concentration values for substrates, extracellular metabolites, and viable and dead cell concentration by utilizing only starting concentrations as input. The prediction accuracy of the model was tested by using experimental data.

Keywords Metabolic flux analysis · Dynamic model · mAb production · Correlation analysis · Quadratic programming · Biological models · Apoptosis · Necrosis

Introduction

The last decade has seen an unprecedented increase in the demand for biopharmaceuticals produced from animal cell culture processes, primarily due to their application in diagnostics and therapeutic treatments. The large-scale production of mAb by mammalian cells in batch and fed-batch culture systems is limited by the unwanted decline in cell viability and reduced productivity that may result from changes in culture conditions. Despite the increasing demand, manufacturers are faced with the challenge of meeting lower cost expectations of the health system and fierce market competition. All these factors point to the formulation of more efficient production strategies than the ones currently in practice for large-scale manufacture.

Fed-batch operations combine the advantages of batch and continuous-culture methods while minimizing the disadvantages that either of the two possesses. Similarly to batch operations they are relatively easy to perform and simple to scale-up. They ensure prolonged cell growth (high cell densities) and product formation due to extension of the culture time that is particularly important in the production of growth-associated products. By understanding the effects that culture conditions have on cell viability and subsequently antibody production it will become possible to impose the necessary conditions during a fed-batch operation to maximize product formation.

Several experimental studies have been conducted regarding mAb production by mammalian cells. Previous research has shown that a dual substrate (glucose and glutamine) limited feeding profile reduces overflow metabolism of the two major growth inhibitors-lactate and ammonia without a decline in growth rate [17]. In further
studies, the kinetics of the onset of apoptosis and its correlation with the exhaustion of nutrients in hybridoma culture has been reported [10, 19]. An abrupt increase in the rate of mAb production near cell death has been also observed [25].

Different modeling techniques have been proposed in the past to describe batch processes for mammalian cell cultures, the most successful being logistic and stoichiometric models. Recently Goudar et al. [14] and Acosta et al. [1] applied logistic modeling approaches to model the growth and metabolism of various hybridoma and CHO cell lines. Stoichiometric models and metabolic flux analysis (MFA) have been proposed to describe the steady-state fluxes (e.g. [8, 9]). However, steady-state flux models are applicable for exponential growth phase only. Provost et al. [22] and Provost and Bastin [21] have proposed a methodology based on MFA to derive a dynamic model for bioreaction processes of batch cultures of CHO cells. These works provided a systematic way to obtain models to explain experimental data. However, Provost et al. [22] had not dealt with fed-batch cultures and has not modeled all amino acids participating in the process. Also, the biomass was not modeled explicitly as a function of metabolite concentrations; measured values of biomass were used instead as inputs to the metabolites’ dynamic model. Sidoli et al. [24] have suggested that single-cell models (SCM) which model the different cell life cycles and population balance models (PBM) may describe more accurately the bioreaction processes under consideration. However, both SCM’s and PBM’s involve a very large number of parameters and consequently they are difficult to calibrate as they require large amount of data for calibration. Overcoming the shortcomings of the existing models may lead to substantial improvements in process efficiency and mAb production.

The focus of this work is the identification of mathematical models that can be used in the future for optimizing the fed-batch operation of a mammalian cell culture. A systematic modeling approach was used that consisted of the following steps: first, MFA in combination with quadratic programming (QP) was applied for the calculation of intracellular fluxes for metabolites from available extracellular concentration values. Flux distribution for batch and fed-batch modes were compared to determine whether the same model structure could be applied to both modes. Based on the set of identified significant fluxes (from MFA), the original metabolic network was reduced to a set of significant reactions. The reactions in the reduced metabolic network were then combined to yield a set of macro-reactions obeying Monod kinetics. Using QP for the exponential growth phase and the Metropolis–Hastings algorithm for the post-exponential phase, the proposed Dynamic Model was calibrated and model prediction was carried out individually for batch and fed-batch runs. Then correlation analysis was performed to formulate a biomass model for predicting cell concentration and viability as a function of the extracellular metabolite concentrations in batch and fed-batch experiments. The Metropolis–Hastings method was applied for estimation of growth and death coefficients in the equations for viable and dead cell predictions. The prediction accuracy of these model equations was tested by using experimental data that was not used for model calibration. The dynamic model was integrated with the biomass model to get an integrated model capable of predicting concentration values for substrates, extracellular metabolites, and viable and dead cell concentration by utilizing only starting concentrations values as input. The accuracy of the integrated model was tested through comparisons with experimental data that was not previously used for model calibration.

**Materials and methods**

Murine hybridoma 130–8F producing anti-Fglycoprotein monoclonal antibody (mAb) was provided by Sanofi Pasteur Ltd. (Toronto, Canada) and was propagated in D-MEM medium (Gibco 12100) with 2% FBS (JRH 12107-78P). The medium was supplemented with proline (Sigma P-8449), lasparagine (Sigma A-4159), and l-aspartic acid (Sigma A-4534). Seed cultures were subcultured on a 3-day regime. Seed and batch cultures were grown in 250 and 500 mL spinners in a CO₂ incubator (Sanyo IR Sensor, 37 °C, 5.0% CO₂). Batch cultures were maintained for at least 7 days with frequent sampling. To avoid ammonia toxicity, fed-batch cultures were maintained under glutamine-limited conditions and the fed-batch operation was effected by injecting defined amounts of glutamine and glucose solution when its concentration inside the spinner was approaching 0. Viable and total cell concentrations were determined by the trypan blue exclusion test using a haemocytometer. Glucose, lactate, glutamine, and glutamate concentration was quantified using an YSI analyzer. Ammonia was measured using a Sigma Ammonia Kit (Sigma 171-B). Total immunoglobulin titer (mAb concentration) was determined using enzyme-linked immunosorbent assay (ELISA) with alkaline phosphatase conjugated goat antimouse IgG as the primary reagent. The amino acids in de-proteinated medium were derivatized with OPA (o-phthaldehyde and 3-mercaptopropionic acid in borate buffer) followed with FMOC (9-fluorenylmethylchloroformate in acetonitrile) and assayed using high-performance liquid chromatography (Hypersil AA-ODS column).
in each one of the dynamic mass balances to be formulated for the process. Previous work [9, 11, 21] has shown that the metabolic flux can be systematically used to find the interactions between the different metabolites occurring in this system. The starting point of the MFA consists of formulating the reaction network stoichiometric equations describing the conversion of substrates into metabolic products and biomass constituents (or macromolecular pools). The overall metabolic network for hybridoma has been taken from Gao et al. [11] and is shown in Fig. 1. It accounts for all the known major energy producing pathways in hybridoma, namely the TCA cycle, Embden–Meyerhoff fermentative pathway, glutaminolysis, and respiration that produce both energy and precursors for biosynthesis ([3]). A key component of the metabolic network is the TCA cycle or citric acid cycle that is involved in the catabolism of all three major food groups: carbohydrates, lipids, and proteins. As illustrated in Fig. 1, the complete reaction network involves \( m=30 \) metabolites and \( n=32 \) fluxes corresponding to 32 reactions. Elements of low energy gain pathways such as the pentose phosphate shunt are notably absent in Fig. 1. Our preliminary flux analysis has shown that the flux terms belonging to these primarily synthetic pathways are low in comparison to the Embden–Meyerhoff pathway or the Krebs cycle. However, an implicit account of these pathways including the pentose shunt was included in the material balance equation for biomass synthesis from glucose and the amino acids (see Table 1; Fig. 1). Although the glucose flux was second only to the glutamine flux for biosynthesis, the

![Fig. 1 Simplified metabolic network of hybridoma cells (adapted from [11])](image)

**Model development**

**Metabolic flux analysis**

The key challenge in the development of a dynamic model of metabolites and biomass is to find the correct structure of this model, i.e. to find the specific metabolites that should appear

Table 1  Macro-reactions and accompanying reaction rates

<table>
<thead>
<tr>
<th>Reaction #</th>
<th>Reaction</th>
<th>Rate expression</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1</td>
<td>Glc(G) → 2Lac</td>
<td>( r_1 = a_{1P}G_{K}X_{v,e} )</td>
<td>( a_{1}, k_{G} )</td>
</tr>
<tr>
<td>E2a</td>
<td>Glc + 2 Gln(Q) → 2 Ala + 2 Glu</td>
<td>( r_{2a} = a_{2G}G_{K}Q_{K}X_{v,e} )</td>
<td>( a_{2a}, k_{G}, k_{Q} )</td>
</tr>
<tr>
<td>E2b</td>
<td>Glc + 2Glu(E) → 2Ala + 2CO₂ + 2Lac</td>
<td>( r_{2b} = a_{2E}G_{K}E_{K}X_{v,e} )</td>
<td>( a_{2b}, k_{G}, k_{E} )</td>
</tr>
<tr>
<td>E3</td>
<td>Glc + 2Glu → 2Lac + 2NH₃ + 8CO₂</td>
<td>( r_{3} = a_{3E}X_{v,e} )</td>
<td>( a_{3}, k_{E} )</td>
</tr>
<tr>
<td>E4</td>
<td>Glu → Pro</td>
<td>( r_{4} = a_{4S_{P}}X_{v,e} )</td>
<td>( a_{4}, k_{S_{P}} )</td>
</tr>
<tr>
<td>E5</td>
<td>Asn(N) → Asp + NH₃</td>
<td>( r_{5} = a_{5S_{P}}X_{v,e} )</td>
<td>( a_{5}, k_{N} )</td>
</tr>
<tr>
<td>E6</td>
<td>Gln + Asp(D) → Asn + Glu</td>
<td>( r_{6} = a_{6Q}D_{K}E_{v,e} )</td>
<td>( a_{6}, k_{Q}, k_{D} )</td>
</tr>
<tr>
<td>E7</td>
<td>0.0508Glc + 0.0577Gln + 0.0133Ala + 0.007Arg + 0.006*Asn + 0.0201Asp + 0.004Cys + 0.0016Glu + 0.0165Gly + 0.0033His + 0.00084Ile + 0.0133His + 0.0133Leu + 0.0101Lys + 0.0033Met + 0.0055Phe + 0.0081Pro + 0.0099Ser + 0.008Thr + 0.004Tyr + 0.0096Val → biomass</td>
<td>( r_{7} = a_{7}Q_{v,e} )</td>
<td>( a_{7}, k_{Q} )</td>
</tr>
<tr>
<td>E8</td>
<td>0.0044Gln + 0.0011Ala + 0.004Arg + 0.0072*Asn + 0.0082Asp + 0.0005Cys + 0.0170Glu + 0.0145Gly + 0.0035His + 0.0054Ile + 0.0142Leu + 0.0145Lys + 0.0028Met + 0.0072Phe + 0.0148Pro + 0.0267Ser + 0.016Thr + 0.0085Tyr + 0.00189Val → mAb</td>
<td>( r_{8} = [a_{8}Q_{v,e} + a_{11}]X_{v,e} )</td>
<td>( a_{8}, a_{11}, k_{Q} )</td>
</tr>
<tr>
<td>E9</td>
<td>Gln → Glu + NH₃</td>
<td>( r_{9} = a_{9Q}X_{v,e} )</td>
<td>( a_{9}, k_{Q} )</td>
</tr>
<tr>
<td>E10</td>
<td>Glu → TCA cycle</td>
<td>( r_{10} = a_{10E}X_{v,e} )</td>
<td>( a_{10}, k_{E} )</td>
</tr>
</tbody>
</table>
bacterial flux value was found to be less than 5% of the total glucose flux.

**Calculation of conversion rates and fluxes**

Mass balances for the intracellular and extracellular metabolites involved in the system can be represented as follows:

\[
\frac{d\psi(t)}{dt} = RX_v(t),
\]

where \(\psi\) is the vector of intracellular and extracellular metabolite concentrations and \(t\) is the time elapsed since the starting of the batch or fed-batch culture. \(R\) is the vector of specific uptake/production rate of substrates/metabolites. \(X_v(t)\) is the viable cell concentration and is a function of culture time, \(t\). An assumption of balanced growth condition has been made. This implies that the system is at a quasi-steady state and consequently the net conversion rate of all intracellular metabolites is equal to 0. This is a common steady state and consequently the net conversion rate of all intracellular metabolites is eliminated by assuming that the system operates under balanced growth conditions and consequently, the reaction rates corresponding to intracellular metabolites (1–6 according to the numerical designation given in Fig. 1) are assumed to be equal to 0 [12, 21].

The need for concentration measurements for intracellular metabolites is eliminated by assuming that the system operates under balanced growth conditions and consequently, the reaction rates corresponding to intracellular metabolites (1–6 according to the numerical designation given in Fig. 1) are assumed to be equal to 0 [12, 21].

The biomass evolution with time for batch or fed-batch runs has been broadly divided into two phases: exponential and post-exponential phases. For both modes of operation all data points corresponding to prior the first peak viable cell density are considered to be the exponential phase. Conversely, the data points from the first peak viable cell density and beyond constitute the post-exponential phase. This is schematically shown in Fig. 2 for a fed-batch process where several peaks in biomass occur during the above defined post-exponential phase due to successive feeding of nutrients that aim to maintain the culture alive. The distinction between an exponential phase and post-exponential phase was done to reflect expected changes in metabolism between the growth period occurring before the first peak in biomass and the period after this peak.

The successive peaks in the post-exponential phase of the fed-batch operation requires special attention as addition of feed at several instances during the run caused abrupt drop in biomass concentration due to sudden dilution of the mixture followed up by an increase due to growth. For batch systems there are obviously only one exponential phase where the biomass concentration increases and one post-exponential phase where the biomass decreases since there is only one single peak in biomass cell density as a function of time.

Assuming a quasi-steady state, the conversion or production rate, \(R(m)\) of a nutrient or metabolite in a biological system can be expressed in the following manner [5]:

\[
R(m) = \sum_i \alpha_{im}j_i,
\]

where \(R\) is the uptake or production rate of metabolite \(m\) (=1,2,…,M) in reaction \(i\) (=1,2,…,N).

In Eq. 6, \(\alpha_{im}\) is the stoichiometric coefficient for metabolite \(m\) in reaction \(i\). \(\alpha_{im}\) has a negative value for substrates that get consumed and is positive for metabolites.
that are produced. The term, \( j_i \), is the intracellular flux for reaction \( i \). In matrix notation, Eq. 6 may be written as:

\[
\mathbf{R} = \mathbf{A} \mathbf{j},
\]

where \( \mathbf{R} \) is the vector of average uptake or production rates, \( \mathbf{j} \) is the flux vector of all species (intracellular and extracellular) which uptake or production rates can be measured or estimated and \( \mathbf{A} \) is the matrix of stoichiometric coefficients of the corresponding reacting species involved in the metabolic network system as shown in Fig. 1. The sparse matrix \( \mathbf{A} \) is given as:

\[
\mathbf{A} = \{ a_{i,m} \}_{i=1,2,...N; m=1,2,...M}. \tag{8}
\]

There were a total of 30 metabolites in the metabolic flux network. However, since the extracellular concentrations of cystine and ammonia were not measured there was a total of 28 metabolites which production or uptake rates can be either measured for the extracellular species or set to 0 for the intracellular species based on the quasi-steady assumption discussed earlier. Hence, in the current study, \( N = 32 \) and \( M = 28 \).

Referring back to the system of Eq. in 7, experimental data were available to determine the values of the vector \( \mathbf{R} \), while the expression for the stoichiometric matrix \( \mathbf{A} \) was derived from the reaction scheme in Fig. 1. Thus, values for the flux vector \( \mathbf{j} \) were easily determined using regression by the QP approach.

\[
\text{Min}(\mathbf{R} - \mathbf{A}_j)^T(\mathbf{R} - \mathbf{A}_j) \tag{9}
\]

It is noteworthy that only 28 extracellular and intracellular measurements were known, whereas 32 flux values needed to be determined. Henceforth, the system of equations given by 7 was underdetermined i.e., fewer equations than unknowns. A unique solution for the system was possible by imposing additional constraints on the system. These additional constraints were based on the assumption that all reactions should proceed in the direction indicated in the metabolic flux network shown in Fig. 1. This implies that all fluxes will have values at least greater than or equal to 0. Since there were 32 unknowns with 28 equations, a unique solution is assured as long as the number of active constraints in the solution of the QP is greater than or equal to 4. All the solutions found in the current work satisfied this criterion. Also, since cystine concentration measurements were not available, its consumption rate was taken to be equal to 0 as opposed to being excluded, because otherwise, a physically unrealistic flux distribution was obtained as a solution, i.e. the flux corresponding to cystine consumption was unexpectedly high.

The MFA presented above is based on measured consumption and production rates of extracellular metabolites. Consequently, it is expected that the experimental noise of measurements will surely affect the calculated flux distribution. In order to reduce sensitivity to noise, experimental values of \( \mathbf{R} \) for a number of batches in the case of batch operation or for measured values of \( \mathbf{R} \) at different times during the run in case of a fed-batch experiment were considered together to conduct the QP regression proposed earlier. For example, if \( s \) sets of experimental batches are considered for batch operation, the augmented system of equations may be expressed as follows:

\[
\begin{bmatrix}
\mathbf{R}_{\text{set1}} \\
\vdots \\
\mathbf{R}_{\text{sets}}
\end{bmatrix}
= 
\begin{bmatrix}
\mathbf{A} \\
\vdots \\
\mathbf{A}_{\text{modj}}
\end{bmatrix}
\tag{10}
\]

\[
\Rightarrow \mathbf{R}_{\text{mod}} = \mathbf{A}_{\text{modj}}
\]

where \( s \) is the number of batches or alternatively it could be a number of different sections considered in a fed-batch experiment as shown in Fig. 2.

Correspondingly, a solution for the resulting system of equations given by 10 can be obtained from the following minimization problem:

\[
\text{Min}(\mathbf{R}_{\text{mod}} - \mathbf{A}_{\text{modj}})^T(\mathbf{R}_{\text{mod}} - \mathbf{A}_{\text{modj}}) \tag{11}
\]

s.t. \( j \geq 0 \)

The problem in 11 can be solved by a QP algorithm as performed by Gao et al. [11]. The result of the QP calculation is a metabolic flux network outlining the various biochemical reactions along with an estimate of the steady-state rate (i.e., the flux) at which each reaction in the diagram occurs. The original metabolic network is generally large and if used for formulating a dynamic model it would result in a very large model with a correspondingly large number of kinetic parameters. The identification of a large set of parameters from data will be prone to error due to measurement error. Therefore, there is an incentive to obtain a reduced metabolic network that will result in a smaller dynamic model than the one expected if the whole original network would be considered. Depending on the contribution of each reaction in the network, the smaller flux values that are a negligible portion of the net flux through the system can be eliminated to obtain a reduced metabolic network. The results of this reduction process are given later in the results section.

**Reaction kinetics and development of the dynamic model**

Based on the metabolic flux network described above a dynamic model was developed for predicting metabolite concentrations over the duration of the culture and is specifically referred to as ‘Dynamic Model’ elsewhere in the text to differentiate it from other dynamic models developed in this work.

The dynamics involved in a stirred spinner is given by the following macroscopic material balance model:
\[
\frac{d\xi(t)}{dt} = K\mathbf{r}(t) + \mathbf{u}(t) \tag{12}
\]

In the equation above, the symbol \(\xi = (\xi_1, \xi_2 \ldots \xi_d)^T\) is the vector of concentrations of extracellular species and is a subset of the vector \(\psi\) given by Eq. 1. \(\mathbf{r} = (r_1, r_2 \ldots r_d)^T\) is the macro-reaction rate’s vector. \(K\) is the matrix containing the stoichiometric coefficients of the reaction species involved in the elementary macro-reactions, i.e. the minimal set of reactions needed to describe the obtained reduced flux network. The second term on the right hand side of the equation \(\mathbf{u}(t)\) represents the vector of net exchange rate of species with the surroundings. The term is 0 for all species except for gaseous \(\text{CO}_2\). However, measurements corresponding to \(\text{CO}_2\) were not available and such measurements are difficult to make. As a result, the term \(\mathbf{u}(t)\) has been effectively dropped from Eq. 12 and the equation can be re-written in the matrix form as follows:

\[
\frac{d\xi(t)}{dt} = K\mathbf{r}(t) \tag{13}
\]

Monod kinetics has been assumed for all the reactions according to the general expression:

\[
r_i(t) = a_i \left( \prod \frac{\text{metabolite}_i}{\text{metabolite}_i + k_i} \right) X_{v,e} \tag{14}
\]

where \(r_i(t)\) is the reaction rate for \(i\)th reaction; \(\text{metabolite}_i\) and \(k_i\) are the corresponding substrate concentration and half-saturation constant, respectively. \(a_i\) is the maximum rate at which the reaction can proceed, while \(X_{v,e}\) is the experimental viable cell concentration.

The assumption of balanced growth made previously implies that the reactions taking place in the system operate at a constant maximum rate during the exponential phase. Thus, during the exponential phase \(a_i\)'s can be obtained by choosing the half-saturation constants \(k_i\)'s in Eq. 14 to be significantly smaller in comparison to substrate concentration for major part of the experimental run, i.e. \(r(t) \approx a_i X_{v,e}(t)\). However, attention was paid that the values were not too small to avoid any numerical stiffness problems [11]. The dynamic model expressed in Eqs. 13 and 14 can thus be stated as:

\[
\frac{d\xi(t)}{dt} = K\mathbf{X}_{v,e}(t) \tag{15}
\]

Additionally, following Eq. 1 the mass balances for the extracellular metabolites specified by the vector \(\xi\) can be expressed as:

\[
\frac{d\xi(t)}{dt} = \mathbf{R}_\xi X_{v,e}(t), \tag{16}
\]

where \(\mathbf{R}_\xi\) is the conversion rate vector and is a subset of vector \(\mathbf{R}\) given by Eq. 7 consisting of the reactions left after the original metabolic flux network has been properly reduced.

Next, from Eqs. 15 and 16, a system of algebraic equations can be derived as follows.

\[
\mathbf{K}a = \mathbf{R}_\xi. \tag{17}
\]

As explained before, experimental data is used to compute the values for \(\mathbf{R}_\xi\) and the reduced reaction scheme for determining expression for \(\mathbf{K}\). Again, the system in Eq. 17 is over-determined for the cases under study and correspondingly the \(a\)-values can be estimated by applying constrained QP to solve the following minimization problem:

\[
\text{Min}(\mathbf{R}_\xi - \mathbf{K}a)^T(\mathbf{R}_\xi - \mathbf{K}a) \quad \text{s.t. } a \geq 0. \tag{18}
\]

For the post-exponential phase it is not possible to assume that the reaction proceeds at the maximum rate. Thus, in this case, the maximal rates \(a_i\)'s and the half-saturation constants \(k_i\)'s were calibrated by a Markov chain Monte Carlo method using the Metropolis–Hastings algorithm [4]. The algorithm has been employed to solve difficult, non-linear parameter estimation problems arising in various disciplines. A symmetric, triangular prior distribution was employed as simple and convenient means for sampling the parameter space. First the given parameter “\(i\)” and its value (\(\theta_i\)) was assigned by random (Monte Carlo) draw from the prior triangular distribution. Using the sampled parameter value, the predicted concentrations were computed and the predictions were compared with measurements by calculating the likelihood function based on the normalized sum of squares between the observations and predictions: The maximum likelihood values of the parameters were used as “\(a\)’-s for final modeling.

**Development of the Biomass Model**

One of the major objectives of this study was to model the growth and death of the cells. The underlying motive was to develop a general model to express the viable and dead cell concentrations as a function of metabolite concentrations with respect to time [1, 16, 29]. Thus, the fundamental model structure was expected to be of the general form expressed in the equation below:

\[
\frac{1}{X_v} \frac{dX_v}{dt} = k_g f(\xi) - k_d g(\xi) \tag{19}
\]

\[
\frac{1}{X_d} \frac{dX_d}{dt} = k_d g(\xi),
\]

where \(X_v\) and \(X_d\) are the viable and dead cell concentration, respectively, and \(k_g\) and \(k_d\) are the parameters of the system.

The term biomass model shown in Eq. 19 will be referred henceforth as the dynamic model that predicts the viable and dead cell concentration over the duration of the
culture and the measured metabolite concentrations over time are used as inputs to this model.

The key challenge in this model was to identify the specific nutrients or products that affect the growth and death rates of the cells from experimental data. Although no starting assumptions were made as to whether the model would be multiplicative or not, it was expected that the $X_v$ and $X_d$ will show dependence on more than one variable/metabolite. To find these dependencies, it was decided to use correlation analysis between the experimentally measured terms in the left-hand side of Eq. 19 and different functional forms of nutrient and product concentrations. Since there was no prior information to indicate the nature of mathematical functions that would account for behavior of $f(\xi)$ and $g(\zeta)$, Monod functions for metabolite concentration were initially used for testing correlation. The software STATISTICA (StatSoft) was used to find these explicit functional forms. Results of the correlation analysis for the system under study are presented in the “Results and discussion” section. Finally, the parameters in the functional forms describing the growth and death rates given in 19 were found by a Metropolis–Hastings approach mentioned earlier where the objective function to be minimized was as follows:

$$\text{Min} \sum (X_{v,e} - X_{v,p})^2 + \sum (X_{d,e} - X_{d,p})^2$$

(20)

where $X_{v,e}, X_{v,p}, X_{d,e}, X_{d,p}$ are the experimental and predicted values of the viable and dead cell concentrations, respectively.

Results and discussion

MFA-reduced network

To obtain the reduced metabolic network for the fed-batch case, average flux distributions were calculated based on average reaction rates. These average rates are calculated by averaging the measured rates in each of the different sections of the exponential and post-exponential phases shown in Fig. 2 for the fed-batch operation or the corresponding exponential and post-exponential phases for the batch experiments. Based on the calculated average reaction rates, the corresponding fluxes were obtained from the QP solution based on 11. For comparison, the flux distributions are shown for a batch run in Fig. 3a and for a fed-batch run in Fig. 3b.

It is clear from Figs. 3 that the distribution of fluxes in batch and fed-batch cultures is strikingly similar. Thus, the fluxes that are dominant during the batch operation are equally dominant during the fed-batch operation. Based, on the flux distribution, the elimination of insignificant fluxes from the complete metabolic network was undertaken for the fed-batch case. The fluxes that were eliminated summed up to less than 1% of the sum of all fluxes. Since the same fluxes were found to be insignificant in both culture modes, a common reduced metabolic network for both batch and fed-batch was obtained as shown in Fig. 4. The fact that the reduced metabolic network is the same for both batch and fed-batch processes implies that the same dynamic model structure model can be assumed because the same macro-reactions are expected to be important for both modes of operation.

Comparing the distribution of average fluxes during batch and fed-batch operation, it is evident that the ratio of flux 8 (Pyr $\rightarrow$ Lac) to flux 2 (Pyr $\rightarrow$ TCA cycle) in batch culture is 2.41, while in fed-batch culture it equals 1.29, almost twice as high as in batch culture. Moreover, there is a slight increase in flux 1 (Glc $\rightarrow$ Pyr) in fed-batch mode because there was no glucose limitation during fed-batch experiment. All these facts indicate that energy generation via the TCA cycle becomes more prevalent during the fed-batch mode as compared to simple batch operation. The reason could be a possible metabolic shift with higher emphasis on cellular maintenance rather than growth or metabolite accumulation. There is published evidence to support the previous statement [9]. Thus, although the set of significant fluxes remains the same in batch and fed-batch operations, the order of magnitude is different.

The glutamine flux remains high in both batch and fed-batch systems. This flux consists of the conversion to glutamate by either direct deamidation or transamination with pyruvate to form alanine or with aspartic acid to form asparagine. Glutamate, in turn, is also deaminated directly or undergoes transamination reaction with pyruvate to form alanine. The transamination reaction involving glutamine and pyruvate, although known to be catalyzed by a single aminotransferase enzyme, was not numbered explicitly in Fig. 4, because this reaction can also be expressed byjudicious combination of equations 25, 27 and the reversible deamination of glutamine by aminohydrolase. Some glutamate is converted to proline as a product, but most glutamate finds eventual entry into the TCA cycle. The terminal carbonaceous product of the cycle, namely CO$_2$ was not modeled explicitly. These processes continue during both batch and fed-batch operations. The net result is that the concentrations of alanine and proline in the medium increase significantly, particularly during the initial exponential phase of the culture.

Values of calibrated coefficients

Based on the obtained metabolic flux network a set of macro-reactions is obtained and shown in Table 1. The rate expressions to be used in the dynamic mass balances corresponding to these reactions are also given in Table 1.
coefficients for equation E7 and E8 (Table 1) were based on the data by Gambhir et al. [12]. The flux values were adjusted to reflect the metabolite uptake for biomass and mAb synthesis for the current application [11].

Following Table 1, the stoichiometric matrix $K$ in Eq. 13 is given in Table 2.

Model calibration and testing

For the exponential phase, the half-saturation constants, $k_{ij}$ ($i$ = a certain metabolite and $j$ = number of the elementary reaction) in Table 3 were chosen by inspection to avoid any stiffness problems [11] yet they were small enough to be negligible when the nutrient concentration was reasonably high.

The half-saturation constants for the post-exponential phase shown in Table 3 were obtained together with the $\alpha$-coefficients by using the Metropolis–Hastings algorithm and are shown in Table 3. The $\alpha$-coefficients in the reaction rate expressions in 15 and shown in Table 4, were calculated by QP for the exponential phase and by the Metropolis–Hastings technique for the post-exponential phase. The estimated $\alpha$-values were, in turn applied on additional batch and fed-batch experiments. The values for the $\alpha$-coefficients obtained through the constrained optimization given by 18 have been given in Table 4.

Unlike the three-phase model required CHO-320 cell line by Provost et al. [22], this cell line was modeled as a two-phase system consisting of an initial exponential phase.
and a second post-exponential phase. The same parameter set was employed for modeling the initial exponential phase of both the batch and the fed-batch culture. It should be noted that the parameters for a batch and a fed-batch culture were optimized separately and the values for a given parameter were averaged when the difference between the two estimates was within 20%. Similarly, the parameters for the initial exponential phase were employed for the post-exponential periods if a change in parameter values was not necessary. For most reactions the post-exponential values were essentially identical with values derived for the exponential phase. However, the parameters representing the rate of anoxic use of glucose and glutamine (E3 in Table 1), the conversion of glutamine to proline (E4) and the metabolism of asparagine (E5) were significantly lower after the initial exponential phase. The parameter, \( a \), is the maximum specific growth rate. It is evident from E7 in Table 1 that glutamine is the growth-limiting substrate under the experimental conditions. The production of mAb was partially growth associated (partially glutamine dependent) and could be represented by the Luedecking–Piret model [18]. The non-growth-associated component of the model (\( a_{11} \)) varied somewhat between the batch and fed-batch experiments. This probably reflects a more stable level of non-growing (viable apoptotic) but producing cells in fed-batch culture.

**Fig. 4 Reduced metabolic network**

**Table 2** Stoichiometric matrix \( K \) for the modeled constituents

<table>
<thead>
<tr>
<th>( r_1 )</th>
<th>( r_{2a} )</th>
<th>( r_{2b} )</th>
<th>( r_3 )</th>
<th>( r_4 )</th>
<th>( r_5 )</th>
<th>( r_6 )</th>
<th>( r_7 )</th>
<th>( r_8 )</th>
<th>( r_9 )</th>
<th>( r_{10} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glc</td>
<td>–1</td>
<td>–1</td>
<td>–1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>–0.0508</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Gln</td>
<td>0</td>
<td>–2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>–1</td>
<td>–0.0577</td>
<td>–0.0104</td>
<td>–1</td>
<td>0</td>
</tr>
<tr>
<td>Lac</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Glu</td>
<td>0</td>
<td>2</td>
<td>–2</td>
<td>–2</td>
<td>–1</td>
<td>0</td>
<td>1</td>
<td>–0.0016</td>
<td>–0.0107</td>
<td>1</td>
</tr>
<tr>
<td>Asn</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>–1</td>
<td>1</td>
<td>–0.006</td>
<td>–0.0072</td>
<td>0</td>
</tr>
<tr>
<td>Asp</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>–0.0201</td>
<td>–0.0082</td>
<td>0</td>
</tr>
<tr>
<td>Ala</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>–0.0133</td>
<td>–0.011</td>
<td>0</td>
</tr>
<tr>
<td>Pro</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>–0.0081</td>
<td>–0.0148</td>
<td>0</td>
</tr>
<tr>
<td>Biomass</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mab</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

**Table 3** Values of the half-saturation constants for the post-exponential phase

<table>
<thead>
<tr>
<th>Constant</th>
<th>Batch mode (in mmol/L)</th>
<th>Fed-batch mode (in mmol/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( k_G )</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>( k_Q )</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>( k_E )</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>( k_N )</td>
<td>0.02</td>
<td>0.12</td>
</tr>
<tr>
<td>( k_D )</td>
<td>0.05</td>
<td>0.05</td>
</tr>
</tbody>
</table>

and derived for the exponential phase. However, the parameters representing the rate of anoxic use of glucose and glutamine (E3 in Table 1), the conversion of glutamine to proline (E4) and the metabolism of asparagine (E5) were significantly lower after the initial exponential phase. The parameter, \( a_7 \), is the maximum specific growth rate. It is evident from E7 in Table 1 that glutamine is the growth-limiting substrate under the experimental conditions. The production of mAb was partially growth associated (partially glutamine dependent) and could be represented by the Luedecking–Piret model [18]. The non-growth-associated component of the model (\( a_{11} \)) varied somewhat between the batch and fed-batch experiments. This probably reflects a more stable level of non-growing (viable apoptotic) but producing cells in fed-batch culture.

**Bioprocess Biosyst Eng (2009) 32:183–196**
\[
\frac{dx}{dt} = k_g \left( \frac{Q}{Q + k_{Q0}} \right) - k_d \left( \frac{Lac}{Lac + k_{dl}} \right)
\]

Values for the estimated growth and death rate coefficients and saturation constants obtained by solving the minimization given by Eq. 20 are shown in Table 6. The rate of growth shown in Eq. 21 is dependent on the glutamine concentration. Lactate has been identified as a growth inhibitor. In fact, the rate of death was expressed as a function of the lactate concentration. The nature of the lactate inhibition appears to be un-competitive. The large half-saturation constant for lactate indicates that inhibition is effective only at high lactate concentrations. It may be that lactate was a surrogate inhibitor and inhibition was caused by adverse pH depression or osmomolarity as a consequence of high lactate concentrations. The pH was uncontrolled in these studies and it usually dropped from 7.1 to 6.5 or below by the end of the culture period. Alternatively, high lactate level in the medium may reflect high cytoplasmic concentrations. Internal lactate has been reported to inhibit lactate oxidoreductase activity, hence inhibit glycolysis [15, 23]. In either case, lactate accumulation needs to be curtailed for optimal performance in fed-batch culture.

It should be emphasized that the parameter values in Table 6 for batch and fed-batch modes of operation are the same and the model built on these values is able to capture the process behavior for both. To test the validity of the model, a separate data set was used for calibration and prediction of the biomass model for the batch as well as the fed-batch system. The predicted concentrations for biomass and the key metabolites are shown in Fig. 6 for batch, and Fig. 7 for fed-batch experiments. These predictions are model verification, since the observed data were not employed for model calibration.

The viable biomass concentration and the major metabolites (glucose, glutamine and lactate) were predicted well in both the batch and fed-batch systems. As reported by Legazpi et al. [16] and Acosta et al. [1], a Monod type specific growth rate expression was suitable to describe the dependence of the specific growth rate on glutamine concentration. The maximum specific growth rate of 0.051 h\(^{-1}\) and the half-saturation constant of 0.05 mM glutamine (see Table 6) are comparable with the corresponding values (0.055 h\(^{-1}\) and 0.08 mM, respectively) reported by Acosta et al. [1]. Exponential growth began after a short lag period and continued until the growth limiting substrate, glutamine was exhausted. In fed-batch culture, fresh medium enriched with either 4 or 8 mM glutamine was added to the culture at three (as shown in Fig. 2) or more occasions. The addition of medium, of course, resulted in a dilution of the viable cell concentration but the viable cell population density was maintained for considerably longer time than in batch culture. Fed-batch operation extended the maintenance of a viable population density of at least 5 \( \times \) 10\(^5\) cells/ml by 20–48 h in comparison to batch culture.

However, the net growth rate was reduced during fed-batch operation after the initial exponential phase even when non-limiting levels of glutamine and glucose were present. The slower growth was attributed to lactate accumulation reaching inhibitory levels.

Most glucose (approximately 70%) was converted to lactate by the Embden–Meyerhoff glycolytic pathway. It seems that there was little cellular dependence on energy produced by the oxidation of glucose via the TCA cycle, particularly during the initial exponential phase of growth. Acosta et al. [1] reported similar lactate yields approaching stoichiometric conversion of glucose to lactate (2 mol lactate/mol glucose) at lower glucose concentrations in the medium. In our hands lactate accumulated continually; the reported lactate utilization during the later stages of CHO

<table>
<thead>
<tr>
<th>Expression</th>
<th>Influencing factor</th>
<th>Correlation coefficient highest value (average value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth: ( \frac{dx}{dt} )</td>
<td>( \frac{Q}{Q + k_{Q0}} )</td>
<td>0.93 (0.84)</td>
</tr>
<tr>
<td>Growth: ( \frac{dx}{dt} )</td>
<td>( \frac{Lac}{Lac + k_{dl}} )</td>
<td>-0.90 (0.81)</td>
</tr>
<tr>
<td>Death: ( \frac{dx}{dt} )</td>
<td>( \frac{Lac}{Lac + k_{dl}} )</td>
<td>0.92 (0.83)</td>
</tr>
</tbody>
</table>

Table 5 Correlation analysis of growth influencing factors and functionalities

Table 6 Parameters values for the Biomass Model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value for batch mode</th>
<th>Value for fed-batch mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>( k_g )</td>
<td>0.051</td>
<td>0.051</td>
</tr>
<tr>
<td>( k_d )</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>( k_{Q0} )</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>( k_{dl} )</td>
<td>14.1</td>
<td>14.1</td>
</tr>
</tbody>
</table>
cell culture [2] was not detected with this hybridoma line. In contrast to glucose, glutamine was utilized mainly via the TCA cycle. According to our flux model, glutamine metabolism was initiated by either direct deamination with aminohydrolase (78%) or by trans-amination with aminotransferase (22%) yielding glutamate as the major product. Glutamate was modeled as substrate for several deamination, and trans-amination reactions. The conversion of glutamate to α-ketoglutarate provided an entry into the TCA cycle. This conversion amounted to nearly 75% of the total glutamate utilized. As shown in Figs. 6 and 7, the predicted concentrations of glutamate during the mid-simulation period were relatively poor. In contrast, the model predictions of glutamate in the calibration runs were very good. The less than perfect fit is probably due to the low concentration, but high turnover rate of glutamate, consequently even a small deviation in one of the reaction rates from the calibrated set could give rise to poor prediction.

Enzymatic transamination is an important process for both glutamine and glutamate metabolism. Both two-substrate reactions involve the keto acid, pyruvate as the second substrate. Amino transferases are known to posses broad substrate specificity; thus, the same enzyme may have catalyzed both reactions. Alanine, the product of transamination involving both glutamine and glutamate accumulated continually reaching over 1.5 mM. Glutamine was the amino nitrogen donor for approximately 2/3 of the alanine formed, the rest coming from glutamate. There was no evidence that alanine was utilized to any significant extent as an energy source once glutamine was exhausted. An interesting feature of this cell line is the absolute requirement for aspartic acid (Asp). The mathematical analysis of the amino acid fluxes confirmed this requirement during the growth phase. Aspartic acid was supplied in the medium in addition to being formed from asparagine by transamination.

Monoclonal antibody production was partially growth (glutamine) dependent. Although the cell-specific productivity is the highest during the exponential growth phase, this hybridoma cell line continued to synthesize mAb, albeit at a lower rate, after net growth ceased. A continued mAb production during the death phase was observed in both batch and fed-batch culture. Barford et al. [3] argued that the concept of growth and non-growth-associated protein production [18] is not valid and it is not likely to assist in understanding mAb productivity. Although some mAb production is strongly related to cell growth only [13], the continued production of mAb by growth-arrested cells of many cell lines is well established [1, 7, 29].

Fig. 5 Experimental data for viable and dead cell concentrations (*) versus simulated data (—) for biomass model for batch mode (a, c) and fed-batch mode (b, d)
were unable to model mAb production by any means other than considering both growth related (glutamine dependent) and non-growth-related productivity.

The current model did not differentiate between non-apoptotic and viable apoptotic cells. This could lead to variation in antibody production rates and might explain inaccurate predictions for cell death and mAb in the current work. It is probable that these two different classes of viable cells have different rates of reaction or to be more specific, different mAb production rate. Therefore, there is a possibility that the present model predicts an average rate of generation based on an averaging of apoptotic and non-apoptotic cells as viable population. Thus, there is a clear motivation to monitor the evolution of apoptotic and non-apoptotic cells during operation.

Although attempts were made to maintain constant levels of key nutrients, the fed-batch operation presented in this paper has not been optimized. A common objective function for optimizing the feed rate profile in fed-batch operations is the time averaged productivity (concentration × volume/time) while the decision variables are the switching time from batch to fed-batch operation, the
initial feed rate, and a time constant (usually exponential constant) in the feed rate Eq. [26]. Optimization is subject to physical constraints such as final volume, pump speed, etc. The decision variables are embedded in the differential equations and due to the non-linearity in the material balance models the optimization of feed rate profiles can be a computationally difficult control problem [26]. Stochastic algorithms such as Markov chain Monte Carlo (MCMC) procedures [26] or genetic algorithms [6] may be applied to determine the optimal feed rate trajectories and then verify the predictions with experimentation.

**Conclusions**

Dynamic models for batch and fed-batch operations were derived based on initial MFA. The model was calibrated with batch and fed-batch data. Parameters values were determined using QP for the initial exponential phase and employing Metropolis–Hastings algorithm for the post-exponential periods. The simplified elementary fluxes and consequent model structure were similar for batch and fed-batch processes. However, some differences in reaction rate constants were noted indicating a shift in energy.
metabolism during the post-exponential phase of batch and fed-batch operations. The model parameters were then applied to predict a new set of batch and fed-batch experimental observations. In general, the model based on the combined parameter set predicted the major nutrient and product concentrations well.

The viable and dead cell concentrations have been explicitly modeled. Correlation analysis has been used to investigate the dependencies of growth and death rates on nutrient and product concentrations. Using experimental initial values, the cell concentration model was coupled to the dynamic metabolite model to generate an integrated model that predicts all significant system variables, including viable cell, various metabolite, and mAb concentrations, independent of subsequent experimental values.

The model is an improvement over pre-existing models of similar nature since MFA has been applied to a fed-batch situation in order to obtain a dynamic model for this mode of operation. Also, all amino acids are described by the model in contrast with most previous studies where only carbonaceous metabolites are modeled. Based on MFA the same structure of the dynamic model has been used for both the batch and fed-batch systems. The calibrated and verified model may be employed for optimizing fed-batch processes for recombinant protein production.

Acknowledgments The authors would like to thank NSERC (The National Science and Engineering Research Council of Canada) for their financial support of this study.

References

27. Statistica, StatSoft, Inc., Tulsa, OK, USA