ELSEVIER

Chemical Engineering Journal 87 (2002) 357–366

Chemical Engineering Journal

www.elsevier.com/locate/cej

Comparative evaluation of batch and fed-batch bioreactors for GAPDH production by recombinant *Escherichia coli* with distributed plasmid copy number☆

P.R. Patnaik*

Institute of Microbial Technology, Sector 39-A, Chandigarh 160 036, India

Received 6 April 2001; accepted 16 October 2001

Abstract

In large bioreactors, there is normally a distribution of plasmid copy numbers among the recombinant cells. This has been modelled here by a Gaussian distribution. For the production of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) by *Escherichia coli* containing the plasmid pBR Eco gap, the performances of batch and fed-batch fermentation have been compared for mean copy numbers between 8 and 12 per cell. In both modes of operation, GAPDH activity increased with copy number but biomass concentration decreased, suggesting that optimum performance may be possible by having a distribution of copy numbers in the starting culture.

Since GAPDH activity per unit biomass attained maximum values at the same point in time (3.12 h) for all copy numbers, the effect of a Gaussian distribution of copy numbers with different variances on the performance at this time was studied. Fed-batch required smaller optimum variances than batch fermentation, which allows easier preparation of seed cultures and greater segregational stability in the former case. Although fed-batch operation also generated larger peak activities of GAPDH, this improvement decreased as the copy number increased, again indicating an optimum copy number and its variance.

Keywords: Recombinant fermentation; GAPDH; Copy number variance; Batch versus fed-batch fermentation

1. Introduction

A conventional perception of genetically altered (recombinant) bacterial cells in a bioreactor is that they are of two kinds: those which harbour a plasmid introduced from another organism and those which do contain this plasmid. All plasmid-containing cells are considered to have the same number of copies of the plasmid. This view has, however, been challenged by recent studies which reveal that a population of cells may have a distribution of plasmid copy numbers. On a macroscopic scale, this means that different fractions of the population may differ in the average number of plasmids per cell.

While measurements and modelling of copy number distribution and its effects on bioreactor performance are relatively new, the existence of differences between cells was recognised long ago when Mitchison and Vincent [1] described a distribution in the ages of cells in a population. Later work [2,3] showed the existence of sub-populations differing in plasmid content and in their translational capacities for foreign protein in identical growth conditions. More direct evidence of a distribution of copy numbers among cells is available from the work of Byrne [4], Lamotte et al. [5], and Kuo and Keasling [6]. The former two applied different degrees of selection pressure by varying the concentration of an antibiotic, while Kuo and Keasling employed a membrane elution technique. All authors obtained cell fractions differing in plasmid content.

Given that cells in a population vary in the plasmid copy number, the classical approach of attributing the same number of plasmids to every cell may be an over-simplification. These models also grossly underestimate the statistically averaged copy number of an ensemble as in the example studied here [3]. This disparity between measured and predicted values increases with the size of the reactor because the broth in a large vessel is more heterogeneous than in a small reactor. Thus, simplified models validated in the laboratory become inadequate and even erroneous in production scale fermentations.

The importance of an accurate description of plasmid dynamics is illustrated by the effect of plasmids on cell growth and product formation. Cells containing only a few copies

☆ IMTECH communication No. 0037/2000.

* Fax: +91-172-690585/690632.

E-mail address: pratap@imtech.ernet.in (P.R. Patnaik).

1385-8947/02/$ – see front matter
PII: S1385-8947(01)00233-9

Nomenclature

Physiological phases

(Ac)	acetate phase
(Gl)	glucose phase

Components

Ac	acetate
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
Gl	glucose
S	substrate
x	biomass
YE	yeast extract

Variables

a	acetate concentration ($g\,l^{-1}$)
f	regulation function in Eq. (15)
g	glucose concentration ($g\,l^{-1}$)
g_{in}	glucose concentration in the feed stream ($g\,l^{-1}$)
h	regulation function in Eq. (16)
k_{Ac}(Ac)	inhibition constant for growth on acetate during the acetate phase ($g\,l^{-1}$)
k_{Ac}(Gl)	inhibition constant for growth on acetate during the glucose phase ($g\,l^{-1}$)
$k_{m,Ac}$	rate constant (h^{-1})
$k_{m,Gl}$	rate constant (h^{-1})
$k_{Y/Ac}$	inhibition constant for yield fulfilled by acetate ($g\,l^{-1}$)
K_{Ac}	affinity constant for acetate ($g\,l^{-1}$)
K_{Gl}	affinity constant for glucose ($g\,l^{-1}$)
K_h	saturation constant of the host component of plasmid replication (h^{-1})
m^i_{Ac}	coefficient of maintenance on acetate for cells containing i plasmids (h^{-1})
m^i_{Gl}	coefficient of maintenance on glucose for cells containing i plasmids (h^{-1})
$m_{M,Ac}$	maximum coefficient of maintenance in acetate (h^{-1})
$m_{M,Gl}$	maximum coefficient of maintenance in glucose (h^{-1})
n	exponent in equation for plasmid loss probability
p	GAPDH concentration ($g\,l^{-1}$)
r^i_{Ac}(Ac)	rate of acetate consumption during the acetate phase by cells containing i plasmids ($g\,l^{-1}\,h^{-1}$)
r^i_{Ac}(Gl)	rate of acetate production during the glucose phase by cells containing i plasmids ($g\,l^{-1}\,h^{-1}$)
r^i_{GAP}(Ac)	rate of GAPDH production during the acetate phase by cells containing i plasmids ($U\,ml^{-1}\,h^{-1}$)
r^i_{GAP}(Gl)	rate of GAPDH production during the glucose phase by cells containing i plasmids ($U\,ml^{-1}\,h^{-1}$)
r^i_{Gl}(Gl)	rate of glucose consumption during the glucose phase by cells containing i plasmids ($g\,l^{-1}\,h^{-1}$)
r^i_x(Ac)	rate of formation of biomass during the acetate phase by cells containing i plasmids ($g\,l^{-1}\,h^{-1}$)
r^i_x(Gl)	rate of formation of biomass during the glucose phase by cells containing i plasmids ($g\,l^{-1}\,h^{-1}$)
r^i_{YE}(Ac)	rate of yeast extract consumption during the acetate phase by cells containing i plasmids ($g\,l^{-1}\,h^{-1}$)
r^i_{YE}(Gl)	rate of yeast extract consumption during the glucose phase by cells containing i plasmids ($g\,l^{-1}\,h^{-1}$)
R	maximum plasmid replication rate ($mg\,(g\,biomass)^{-1}\,h^{-1}$)
x	total biomass concentration ($g\,l^{-1}$)
x_i	concentration of sub-population of cells containing i plasmids ($g\,l^{-1}$)
y	yeast extract concentration ($g\,l^{-1}$)
$Y_{GAP/x}$(Ac)	yield of GAPDH per unit biomass during the acetate phase ($U\,g^{-1}$)
$Y_{GAP/x}$(Gl)	yield of GAPDH per unit biomass during the glucose phase ($U\,g^{-1}$)
$Y_{M,x/Gl}$(Gl)	maximum biomass yield on glucose during the glucose phase ($g\,g^{-1}$)
$Y_{m,x/Gl}$(Gl)	minimum biomass yield on glucose during the glucose phase ($g\,g^{-1}$)
$Y_{x/Ac}$(Ac)	yield of biomass per unit mass of acetate when growing on acetate ($g\,g^{-1}$)
$Y_{x/Ac}$(Gl)	yield of biomass per unit mass of acetate when growing on glucose ($g\,g^{-1}$)
$Y_{x/Gl}$(Gl)	biomass yield on glucose during the glucose phase ($g\,g^{-1}$)
$Y_{YE/x}$(Ac)	yield as yeast extract consumed per unit biomass growing on acetate ($g\,g^{-1}$)
$Y_{YE/x}$(Gl)	yield as yeast extract consumed per unit biomass growing on glucose ($g\,g^{-1}$)

Greek letters

α,β,δ	parameters in Eq. (5)
γ	constant in equation for plasmid loss probability
κ	exponent of growth inhibition term for acetate during acetate phase
λ	exponent of yield inhibition term by acetate

φ_i	probability of plasmid loss from cells containing i plasmids each
μ_{Ac}	specific growth rate of biomass on acetate (h^{-1})
μ^i_{Ac}	specific growth rate of cells containing i plasmids on acetate (h^{-1})
μ_{Gl}	specific growth rate of biomass on glucose (h^{-1})
μ^i_{Gl}	specific growth rate of cells containing i plasmids on glucose (h^{-1})
$\mu_{M,Ac}$	maximum specific growth rate on acetate (h^{-1})
$\mu_{M,Gl}$	maximum specific growth rate on glucose (h^{-1})
ψ	exponent of growth inhibition term for acetate during the glucose phase

of a plasmid do not have to support a high metabolic load, and, therefore, there is little effect on their growth rates. However, the small number of plasmids also results in low concentrations of the cloned-gene protein. Since the recombinant protein is a key inhibitor of cell growth [2,7], a low concentration of protein is also favourable to growth. However, this negates a prime objective of a recombinant fermentation, which is to generate large amounts of protein. Thus, it is necessary to balance the conflicting objectives of good growth and high protein synthesis.

Either or both of two strategies are employed to achieve appreciable cell growth as well as protein formation. One strategy is to operate the bioreactor initially under conditions favouring cell growth and then shift to conditions promoting protein synthesis. These two phases normally require different concentrations of substrate (mainly the carbon source), which is conveniently controlled by fed-batch operation. This method has other advantages too [8], and hence many recombinant fermentations are operated in fed-batch mode [9–11].

The second strategy is to have an optimum distribution of plasmid copy numbers such that the cells are neither overburdened with plasmid and plasmid-encoded protein nor have so few plasmids that protein synthesis is not economically viable. In a sense, this strategy exploits the natural occurrence of a distribution of copy numbers in cell populations [4–6], and it was the central idea in an earlier work [12], which forms the basis of this study. Now, if the fermentation is time-dependent, the optimum distribution should also vary with time. However, it is practically difficult to ensure that the distribution of copy numbers in the population of cells is optimum at all times. So the present work and its predecessor [12] have the more modest, and feasible, objective of optimising the starting distribution of copy numbers.

In this study, the performances of batch and fed-batch fermentations have been compared for glyceraldehyde-3-phosphate dehydrogenase (GAPDH) production by a recombinant *Escherichia coli* strain when: (a) all recombinant cells have the same number of copies of a plasmid; (b) there is an optimum distribution of copy number in the bacterial population. The formation of GAPDH is growth-associated [13,14]; in view of this and the kinetics described in the next section, control of the substrate feed rate is crucial in fed-batch operation.

2. Fermentation description

This study is based on the kinetics of GAPDH production in batch fermentation by *E. coli* C600 gal K (ATCC23724) harbouring the plasmid pBR Eco gap [14]. This plasmid [15] is derived from pBR322 by insertion of the GAPDH gene into the tetracyclene gene and therefore the modified strain, *E. coli* C600 gal K (GAPDH), is resistant only to ampicillin. Kinetic studies [13] prior to the modelling effort of Nancib et al. [14] showed that threonine, leucine and cysteine promoted GAPDH synthesis up to concentrations of 0.3, 0.3 and 0.1 $g\,l^{-1}$, respectively, and then had no further effect. Thiamin was essential for growth, its maximum favourable concentration being 0.15 $g\,l^{-1}$. These aspects and the role of yeast extract, described below, were considered by Nancib et al. [14] in designing the culture media.

According to protocols described by the authors [13,14], the strain was stored in Luria broth at $-80\,^{\circ}C$, and a complex medium was used for preparing the inoculum and in the fermentation experiments. Control of pH, temperature, aeration and agitation are important; El Houtaia et al. [13] optimised these conditions at pH 7, 37 °C, 1 vvm aeration rate and 600 rpm stirring speed. Owing to the aqueous medium and the intense agitation and aeration, diffusion and mass transfer limitations were assumed to be negligible. The total cell mass concentration was determined through optical density measurements at 660 nm, and the percentage of plasmid-containing cells by replica plating in Luria broth supplemented with ampicillin. GAPDH activity was determined from absorbance data at 340 nm for NADH formed by the method described elsewhere [16]. Based on their observations, Nancib et al. [14] proposed the mechanism shown in Fig. 1 for cell growth and product formation. Although a complex medium was used, the kinetics could be described in terms of the primary carbon source, glucose, and the acetate by-product when glucose concentration fell below a threshold value.

Their results showed that the fermentation could be divided into two phases. In the first phase, glucose is utilised and there is rapid (exponential) growth and GAPDH formation. Acetate is a by-product of this phase; both acetate and GAPDH formation are growth-associated. Eventually, when glucose is depleted substantially (below 0.01 $g\,l^{-1}$ in their experiments), the cells begin utilising acetate as a carbon source. Yeast extract is required to promote the metabolism of acetate [17], as illustrated in Fig. 1. Since acetate is an inhibitor of cell growth, its consumption is useful. However,

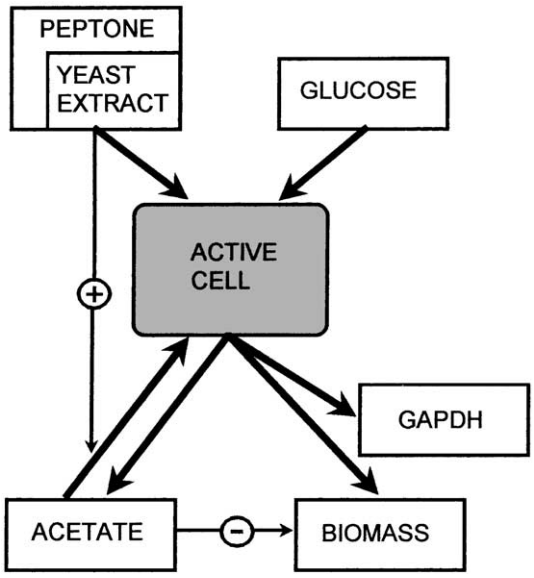

Fig. 1. Schematic diagram of the mechanism of cell growth and product (GAPDH) formation. Reproduced from Nancib et al. [14] (copyright Elsevier Science, UK, 1993).

acetate cannot be utilised as soon as it is formed because glucose is the preferred substrate. Therefore, controlling the concentration of glucose is a critical aspect of the fermentation, for which fed-batch operation should be preferred.

3. Model development

3.1. Copy number distribution

Lamotte et al.'s [5] studies employing different concentrations of ampicillin and those of Kuo and Keasling [6] using a membrane elution method showed that pBR322 and plasmids derived from it exist in different copy numbers in *E. coli* cells. These results provided experimental support to the mathematical analysis by Bentley and Quroga [3], who showed that the plasmid copy numbers in a population of bacterial cells may be represented by a Gaussian distribution of the fraction of cell mass versus the copy number. Since pBR Eco gap is derived from pBR322 [15], such a distribution is more appropriate than considering all recombinant cells to have the same number of plasmids. Mathematically, the distribution has the form:

$$p = \frac{1}{\sigma\sqrt{\pi}} \exp\left[-\frac{(c-\bar{c})^2}{2\sigma^2}\right] \tag{1}$$

where p is the probability density of a copy number c in a population with a mean copy number $\bar{c}$ and a variance σ^2.

Although the copy number may vary from cell to cell, it is practically useful to express this in macroscopic form. To do so, the population may be divided into finite sub-populations with an average copy number for each sub-group. The sizes (i.e. variations in plasmid copy number) of the sub-groups may be chosen according to the accuracy required. In addition to enabling copy number variations to be incorporated conveniently into models of cell kinetics, such a representation overcomes a limitation in applying a Gaussian distribution to discrete cells. A Gaussian distribution is a continuous function whereas a cell may have only an integral number of plasmids. By using average copy numbers of sub-populations, this number is allowed to be a non-integer, thereby satisfying the Gaussian requirement.

To apply Eq. (1) to a group of cells, we have to integrate it to obtain the standard Gaussian probability distribution:

$$P(U \le u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} \mathrm{e}^{-W^2}\,\mathrm{d}W \tag{2}$$

where $U = (c - \bar{c})/\sigma$ is a normalised variable. $P(U \le u)$ is the cumulative probability of a population having cells with plasmids fewer than $c = \sigma u + \bar{c}$ in number. For a set of sub-populations, we divide $P(U)$ symmetrically on either side of c into segments of width Δc. Then $\Delta P_1 = \{P(U \le u_2) - P(U \le u_1)\}$ is the probability of occurrence of a sub-population with copy numbers between $c_1 = \sigma u_1 + \bar{c}$ and $c_2 = \sigma u_2 + \bar{c}$. If Δc is chosen sufficiently small, ΔP_1 approximates a sub-population with a mean copy number $(c_1+c_2)/2$. Generalising this, ΔP_i represents the cumulative probability of a sub-population with an average number of $(c_i + c_{i+1})/2$ plasmids per cell.

Given a starting culture with a cell mass concentration x_0, the concentrations of cell fractions with different mean copy numbers are calculated by distributing x_0 proportionate to ΔP_i [12]. Thus, the initial concentration of cells with a mean copy number $(c_i + c_{i+1})/2$:

$$x_{i0} = \frac{x_0 \Delta P_i}{\sum_{i=1}^{M} \Delta P_i} \tag{3}$$

where M is the number of intervals into which the range of copy numbers is divided.

3.2. Cell kinetics and bioreactor model

Nancib et al. [14] proposed a kinetic model for GAPDH production in batch operation, which considered only two kinds of cells—some cells were devoid of the plasmid and others contained the same number of plasmids. Since there is evidence [3,5,6] for a distribution of plasmid copy numbers, their model may be generalised to account for this distribution. Because control of the substrate concentration in the bioreactor by varying the feed rate is a key to generating large amounts of recombinant protein [9–11], Nancib et al.'s batch reactor model was extended to fed-batch operation. Since the basis of the rate equations for cell growth, glucose consumption, GAPDH formation and related processes have been explained in detail by Nancib et al. [14], they are not repeated here and only the generalised version of their model is presented. However, it may be remembered that growth occurs on both glucose and the by-product acetate. In a batch experiment, consumption of acetate starts

after the concentration of glucose has become smaller than $0.01\,g\,l^{-1}$ [14]. Whether and, if so, how often this happens in a fed-batch experiment depends on the initial conditions and the rate of supply of glucose. This aspect has been investigated for fermentations with only one kind of cells through a cybernetic approach [18,19]; its applicability to systems with two or more types of cells, as in a recombinant fermentation, remains to be investigated and is outside the scope of this work.

3.2.1. Kinetics

3.2.1.1. Glucose phase. Cell growth in this phase follows saturation kinetics with respect to glucose and it is inhibited by acetic acid. When all recombinant cells have the same number of copies of the plasmid, Nancib et al. [14] expressed the specific growth rate as

$$\mu_{\mathrm{Gl}} = \left[\frac{\mu_{\mathrm{M,Gl}} g}{K_{\mathrm{Gl}} + g}\right]\left[\frac{1}{1 + \{a/k_{\mathrm{Ac}}(\mathrm{Gl})\}^{\psi}}\right] \tag{4}$$

Since the (mean) copy number differs from one sub-population of cells to another, so will their specific growth rates. Unlike the majority of plasmids in *E. coli*, pBR322 and its derivatives promote the specific growth rate as the copy number increases [20,21]. Among the models available for this kind of dependence, the model proposed by Satyagal and Agarwal [22] has been shown [23] to be the most satisfactory. This model has the form:

$$\mu = \frac{\alpha}{\beta + c} - \delta \tag{5}$$

where c is the copy number. Since α, β and δ are fitted parameters, the specific growth rate of a sub-population with i plasmids per cell may be expressed as the product of Eqs. (4) and (5) [3]:

$$\mu^{i}_{\mathrm{Gl}} = \left[\frac{\alpha}{\beta + c_i} - \delta\right]\left[\frac{\mu_{\mathrm{M,Gl}} g}{K_{\mathrm{Gl}} + g}\right]\left[\frac{1}{1 + \{a/k_{\mathrm{Ac}}(\mathrm{Gl})\}^{\psi}}\right] \tag{6}$$

Using this equation, the rate of growth of this sub-population is simply

$$r^{i}_{x}(\mathrm{Gl}) = \mu^{i}_{\mathrm{Gl}} x_i \tag{7}$$

The rate of glucose consumption is the sum of the requirements for growth and maintenance.

$$r^{i}_{\mathrm{Gl}}(\mathrm{Gl}) = \frac{r^{i}_{x}(\mathrm{Gl})}{Y_{x/\mathrm{Gl}}(\mathrm{Gl})} + m^{i}_{\mathrm{Gl}} x_i \tag{8}$$

Since acetate inhibits growth in the glucose phase, this is accommodated in a yield coefficient dependent on acetate concentration.

$$Y_{x/\mathrm{Gl}}(\mathrm{Gl}) = Y_{\mathrm{m},x/\mathrm{Gl}}(\mathrm{Gl}) + \frac{Y_{\mathrm{M},x/\mathrm{Gl}}(\mathrm{Gl}) - Y_{\mathrm{m},x/\mathrm{Gl}}(\mathrm{Gl})}{1 + \{a/k_{\mathrm{Y/Ac}}\}^{\lambda}} \tag{9}$$

Nancib et al.'s [14] data showed that the maintenance coefficient decreased hyperbolically as the specific growth rate increased. They expressed this variation in the simple form:

$$m^{i}_{\mathrm{Gl}} = \frac{m_{\mathrm{M,Gl}}}{1 + \mu^{i}_{\mathrm{Gl}}/k_{\mathrm{m,Gl}}} \tag{10}$$

The rates of acetate formation, consumption of yeast extract and formation of GAPDH may be written in a manner similar to Eq. (8), but without the maintenance term.

$$r^{i}_{\mathrm{Ac}} = \frac{r^{i}_{x}(\mathrm{Gl})}{Y_{x/\mathrm{Ac}}(\mathrm{Gl})} \tag{11}$$

$$r^{i}_{\mathrm{YE}}(\mathrm{Gl}) = r^{i}_{x}(\mathrm{Gl}) Y_{\mathrm{YE}/x}(\mathrm{Gl}) \tag{12}$$

$$r^{i}_{\mathrm{GAP}}(\mathrm{Gl}) = 10^{-3} r^{i}_{x}(\mathrm{Gl}) Y_{\mathrm{GAP}/x}(\mathrm{Gl}) \tag{13}$$

Eqs. (6)–(13) have generalised the model of Nancib et al. [14] by expressing individual rates for each fraction of the recombinant cell population. Thus, $i = 1, 2, \ldots, M$ here, and for the acetate phase, with M being the number of fractions (based on the copy number distribution).

3.2.1.2. Acetate phase. The rate of growth is similar to that in the glucose phase.

$$r^{i}_{x}(\mathrm{Ac}) = \mu^{i}_{\mathrm{Ac}} x_i = \left[\frac{\alpha}{\beta + c_i} - \delta\right]\left[\frac{\mu_{\mathrm{M,Ac}} a}{K_{\mathrm{Ac}} + a}\right]\left[\frac{f h x_i}{1 + \{a/k_{\mathrm{Ac}}(\mathrm{Ac})\}^{\kappa}}\right] \tag{14}$$

The main difference between this equation and Eq. (6) is by the introduction of two more parameters f and h. They express the dependence of growth on yeast extract and glucose, respectively. Since growth on acetate occurs only if yeast extract is present

$$f = \begin{cases} 1 & \text{if } y > 0\,g\,l^{-1} \\ 0 & \text{if } y = 0\,g\,l^{-1} \end{cases} \tag{15}$$

h is a switching parameter to express the observation that cells begin to use acetate for their growth only after glucose concentration has fallen below a threshold value, which was $0.01\,g\,l^{-1}$ in Nancib et al.'s [14] studies. Thus,

$$h = \begin{cases} 1 & \text{if } g \le 0.01\,g\,l^{-1} \\ 0 & \text{if } g > 0.01\,g\,l^{-1} \end{cases} \tag{16}$$

Since the utilisation of acetate is a consequence of growth, f and h also enter this rate equation, which is otherwise similar to that for glucose.

$$r^{i}_{\mathrm{AC}}(\mathrm{Ac}) = \left[\frac{r^{i}_{x}(\mathrm{Ac})}{Y_{x/\mathrm{Ac}}(\mathrm{Ac})} + m^{i}_{\mathrm{Ac}} x_i\right] f h \tag{17}$$

The maintenance coefficient has the same hyperbolic form as Eq. (10), and the rates for yeast extract consumption and GAPDH formation are similar to the glucose phase.

$$m^{i}_{\mathrm{Ac}} = \frac{m_{\mathrm{M,Ac}}}{1 + \mu^{i}_{\mathrm{Ac}}/k_{\mathrm{m,Ac}}} \tag{18}$$

$$r^i_{YE}(Ac) = r^i_x(Ac) Y_{YE/x}(Ac) \quad (19)$$

$$r^i_{GAP}(Ac) = 10^{-3} r^i_x(Ac) Y_{GAP/x}(Ac) \quad (20)$$

3.2.2. Reactor model

3.2.2.1. Batch operation. For each sub-population, a mass balance may be written identical to that of Nancib et al. [14].

$$\frac{dx_i}{dt} = (1-\varphi_i)\mu_i x_i, \quad i = 1, 2, \ldots, M \quad (21)$$

Here φ_i is the plasmid loss probability, which may be expressed as [24].

$$\varphi_i = \gamma \left[\mu_i - R \frac{(\mu_i/K_h)^n}{1 + (\mu_i/K_h)^n} \right] \quad (22)$$

and $\mu_i = \mu^i_{Gl} + \mu^i_{Ac}$ is the overall specific growth rate of the sub-population.

Plasmid-free cells arise from plasmid-harbouring cells through segregational loss of plasmids (Eq. (22)) and also grow on their own. Therefore, their net rate of formation is

$$\frac{dx_0}{dt} = \sum_{i=1}^{M} \varphi_i \mu_i x_i + \mu_0 x_0 \quad (23)$$

where the subscript 0 signifies the absence of plasmid. In the absence of any independent study of μ_0, this was expressed similar to Eqs. (6) and (14) by setting $c_i = c_0 = 0$.

The rates of change of glucose, acetate, GAPDH and yeast extract are the sums of the contributions from the M recombinant sub-populations and the plasmid-free sub-population.

$$\frac{dg}{dt} = -\sum_{i=0}^{M} r^i_{Gl}(Gl) \quad (24)$$

$$\frac{da}{dt} = \sum_{i=0}^{M} r^i_{Ac}(Gl) - \sum_{i=0}^{M} r^i_{Ac}(Ac) \quad (25)$$

$$\frac{dp}{dt} = \sum_{i=1}^{M} r^i_{GAP}(Gl) + \sum_{i=1}^{M} r^i_{GAP}(Ac) \quad (26)$$

$$\frac{dy}{dt} = -\sum_{i=0}^{M} r^i_{YE}(Gl) - \sum_{i=0}^{M} r^i_{YE}(Ac) \quad (27)$$

The total biomass concentration x at any time is evidently

$$x = \sum_{i=0}^{M} x_i \quad (28)$$

Note that the index $i = 0, 1, \ldots, M$ in Eqs. (24), (25) and (27), but $i = 1, 2, \ldots, M$ in Eq. (26). This difference arises because cells devoid of the plasmid ($i = 0$) do not synthesise GAPDH but they still grow, consume glucose and produce acetate (see Fig. 1).

3.2.2.2. Fed-batch operation. The mass balances follow directly from Eqs. (21) and (24)–(27) by adding the flow terms.

$$\frac{dx_i}{dt} = (1-\varphi_i)\mu_i x_i - \frac{Q x_i}{V}, \quad i = 1, 2, \ldots, M \quad (29)$$

$$\frac{dx_0}{dt} = \sum_{i=1}^{M} \varphi_i \mu_i x_i + \mu_0 x_0 - \frac{Q x_0}{V} \quad (30)$$

$$\frac{dg}{dt} = \frac{Q(g_{in} - g)}{V} - \sum_{i=0}^{M} r^i_{Gl}(Gl) \quad (31)$$

$$\frac{da}{dt} = \sum_{i=0}^{M} r^i_{Ac}(Gl) - \sum_{i=0}^{M} r^i_{Ac}(Ac) - \frac{Qa}{V} \quad (32)$$

$$\frac{dp}{dt} = \sum_{i=1}^{M} r^i_{GAP}(Gl) + \sum_{i=1}^{M} r^i_{GAP}(Ac) - \frac{Qp}{V} \quad (33)$$

$$\frac{dy}{dt} = -\sum_{i=0}^{M} r^i_{YE}(Gl) - \sum_{i=0}^{M} r^i_{YE}(Ac) - \frac{Qy}{V} \quad (34)$$

$$\frac{dV}{dt} = Q \quad (35)$$

4. Results and discussion

Using data from Nancib et al. [24] and the method described by Bentley and Quroga [3], plasmid copy numbers were determined to be between 8.49 and 12.04. These were regressed to Eq. (5) and the values of α, β and δ, together with those of other parameters reported by Nancib et al. [14], are listed in Table 1. For convenience and to enable comparison with earlier work [12], the population of recombinant cells was divided into five groups with mean copy numbers of 8–12.

There are two approaches to determine the best time-dependent feed rate. One approach relies on the nature of the bioreactor model (including cell kinetics) and of the profiles of the specific rates (for growth, production formation and substrate consumption) as functions of the substrate concentration. Hamiltonian methods [25] and dynamic programming [26], which have been commonly used, are of this kind. However, these methods are computationally expensive when there are many variables, as in the model presented in this paper, and do not always lead to stable solutions or a global minimum [27]. Genetic algorithms [27,28] overcome the convergence problems, but they have not yet been adequately tested for their computational superiority and ease of on-line implementation. This limitation is illustrated by the hybridoma reactor studied by Roubos et al. [27], where dynamic programming was faster than a genetic algorithm in some situations but slower in others.

Ease of implementation is an important consideration for production scale bioreactors, where hardware and economic

Table 1
Values of the parameters[a]

Parameter	Units	Value
f	–	0 or 1
g_{in}	$g\,l^{-1}$	10
h	–	0 or 1
$k_{Ac}(Ac)$	$g\,l^{-1}$	4.0
$k_{Ac}(Gl)$	$g\,l^{-1}$	4.0
$k_{m,Ac}$	h^{-1}	0.07
$k_{m,Gl}$	h^{-1}	0.12
$k_{Y/Ac}$	$g\,l^{-1}$	1.986
K_{Ac}	$g\,l^{-1}$	20.0
K_{Gl}	$g\,l^{-1}$	0.7
K_h	h^{-1}	0.118
$m_{M,Ac}$	h^{-1}	0.07
$m_{M,Gl}$	h^{-1}	0.34
n	–	1.8
V^0	$mg\,(g\,cell)^{-1}\,h^{-1}$	0.215
$Y_{GAP/x}(Ac)$	$U\,g^{-1}$	13.460
$Y_{GAP/x}(Gl)$	$U\,g^{-1}$	13.460
$Y_{M,x/Gl}(Gl)$	$g\,g^{-1}$	3.0
$Y_{m,x/Gl}(Gl)$	$g\,g^{-1}$	0.4
$Y_{x/Ac}(Ac)$	$g\,g^{-1}$	2.0
$Y_{x/Ac}(Gl)$	$g\,g^{-1}$	2.3
$Y_{YE/x}(Ac)$	$g\,g^{-1}$	4.5
$Y_{YE/x}(Gl)$	$g\,g^{-1}$	0.31
α	h^{-1}	102.5
β	–	20.5
δ	h^{-1}	2.92
γ	–	0.0015
λ	–	2.242
κ	–	2.5
$\mu_{M,Ac}$	h^{-1}	0.7
$\mu_{M,Gl}$	h^{-1}	1.0
ψ	–	2.5

[a] All values except those of α, β and δ were taken from Nancib et al. [14]. The values of α, β and δ were obtained by regressing equation (5).

limitations may require compromising the theoretically highest possible productivity of the recombinant protein so as to gain in stability and sensitivity [29]. For the parameter values in Table 1, the specific growth rate and the specific rate of formation of GAPDH have strongly non-linear profiles (not shown), and the former has an uncommon bimodal distribution. Any optimisation method must account for such features over a range of copy numbers with Gaussian distributions. For such complex problems, Montague and Ward [30] recommended a sub-optimal approach using the chemotaxis algorithm [31]. Here the feed rate is expressed empirically as a polynomial function of time:

$$Q = \sum_{j=0}^{N} a_j \left(\frac{t}{t_f}\right)^j \tag{36}$$

where t_f is the time at which the fermentation is stopped. The number of terms N and the values of the coefficients a_j are determined through random number generations, similar to but in a more simple manner than in a genetic algorithm [27]. Eq. (36) also makes the feed rate a direct function of time, thereby making on-line implementation easier than in model-dependent methods [25,26].

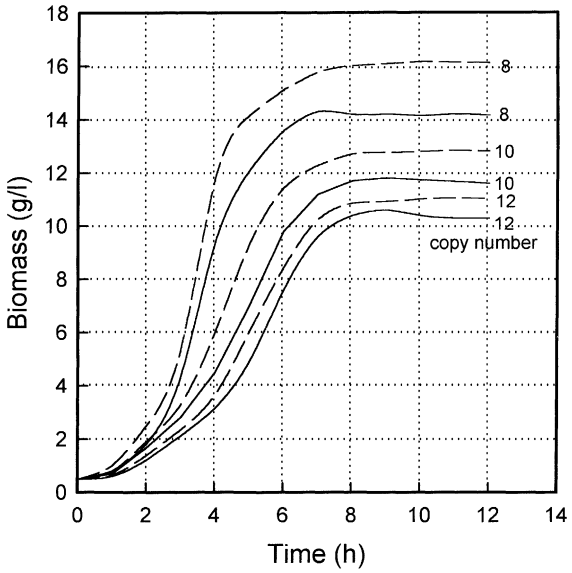

Fig. 2. Variation of biomass concentration with time when all cells have the same initial number of plasmids (indicated against the plots). Continuous lines are for batch operation and broken lines are for fed-batch. Initial concentrations: $a = 1.2\,g\,l^{-1}$; $g = 10\,g\,l^{-1}$; $p = 10\,U\,ml^{-1}$; $x = 0.5\,g\,l^{-1}$; $y = 5\,g\,l^{-1}$.

Eqs. (4)–(36) do not permit analytical solutions; they were therefore solved numerically by a fourth-order Runge–Kutta–Gill method with variable step-size. Fig. 2 shows the variation in biomass concentration for a population of cells with a uniform copy number. They could be either a homogeneous population, as assumed by Nancib et al. [14,24], or a slice of a heterogeneous population. This interpretation also applies to Figs. 3 and 4, which are for GAPDH. The plots in Figs. 2 and 3 reach saturation values for both batch and fed-batch operations. The onset of this constant region signifies the depletion of glucose below the critical concentration ($0.01\ g\,l^{-1}$ in Nancib et al. [14]) and utilisation of acetate, the latter being promoted by the presence of yeast extract [17].

There are two important differences between Figs. 2 and 3. First, with increasing copy number the biomass concentration decreases (Fig. 2), but GAPDH activity increases (Fig. 3). Cells with a few copies of a plasmid support a lower metabolic load than those with a high copy number and therefore have higher growth rates. This difference in metabolic demand is largely due to less protein synthesis when the copy number is small [21]. However, since protein synthesis is a key requirement, sufficiently large copy numbers are required. The second difference is that improvement of cell growth though fed-batch fermentation varies inversely with the copy number, while improvement of GAPDH production increases with the copy number. This difference too may be explained in a similar manner. What these two differences imply is that a high protein synthesis rate may have

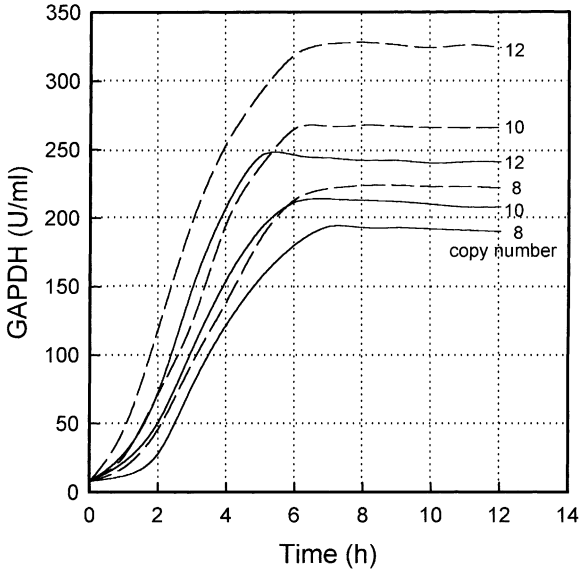

Fig. 3. Variation of volumetric GAPDH activities with time when all cells have the same initial number of plasmids. Other features are the same as explained in Fig. 2.

to be offset against a high growth rate; since they require different regions of a spectrum of copy numbers, the fermentation is optimised by a distribution in the copy number.

The opposing effects of high protein synthesis and low growth rate of biomass result in the interesting profiles of GAPDH activity per unit biomass shown in Fig. 4. Initially the rate of GAPDH formation exceeds the growth rate. However, eventually the metabolic load exerted by the recombinant protein becomes unsustainable [21]. This results in reduced synthesis capability; although the growth of biomass also slows down, this adds to the existing biomass, thereby reducing the GAPDH concentration. Hence, GAPDH activity rises initially but declines later, thereby passing through peak values during the course of fermentation. The coincidence of the occurrence times of the maxima implies that at any instant of time the *relative* rates of GAPDH formation and biomass growth are the same for all copy numbers even though the *absolute* rates are different. Since there are no reports yet of such studies for other recombinant fermentations, it is not known whether the latter feature is true of all plasmids in *E. coli* or depends on the nature of the plasmid vis-a-vis the host cell.

A common time for the peaks is, however, convenient for bioreactor operation and to study the effect of copy number variance. This was done earlier [12] for batch fermentation, and it was shown that for each copy number there is an optimum variance that maximises β-galactosidase activity at the time of occurrence (3.12 h from the start) of the peaks in Fig. 4. A similar analysis was carried out here for fed-batch fermentation. To each copy number in the set [8–12], a Gaussian distribution around this mean value and a variance between 0 and half of the mean was applied. The upper end of the variance was chosen from experience such that it was large enough to ensure that the optimum variance was within this range. This is substantiated by the results in Fig. 5, where the optimum variance for batch and fed-batch operations are compared.

Regardless of whether GAPDH activity is maximised per unit volume of the broth or per unit mass of cells, the optimum variance for fed-batch operation is smaller than for batch operation at all copy numbers. This has an important practical implication. As suggested before [12], the optimum distribution of copy numbers may be applied to a bioreactor by discretising it into finite segments along the copy number axis. Then inocula with copy numbers equal to the mean value of each segment may be prepared and these sub-cultures may be combined to obtain the final seed culture. The smaller is the required variance of the copy number, the fewer are the segments into which it needs to

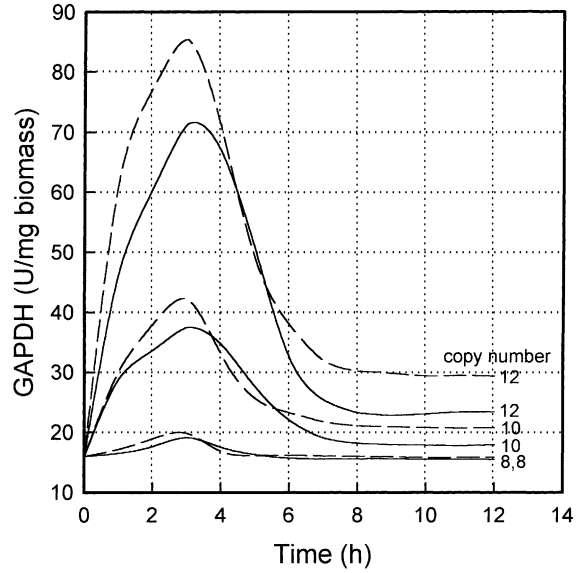

Fig. 4. Variation of GAPDH activities per unit biomass as fermentation progresses and all cells have the same number of plasmids at the start. Other features are as explained in Fig. 2.

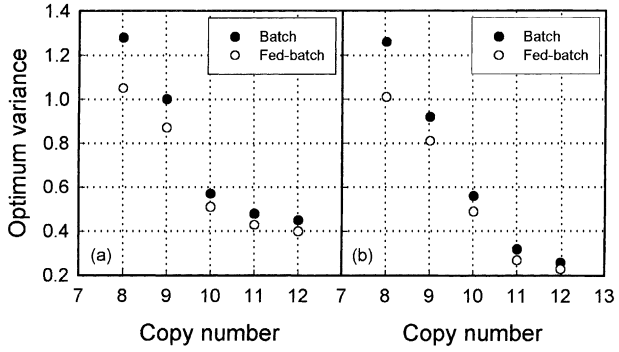

Fig. 5. Optimum variances for different plasmid copy numbers for batch and fed-batch fermentations: (a) GAPDH in U/(ml broth); (b) GAPDH in U/(mg biomass).

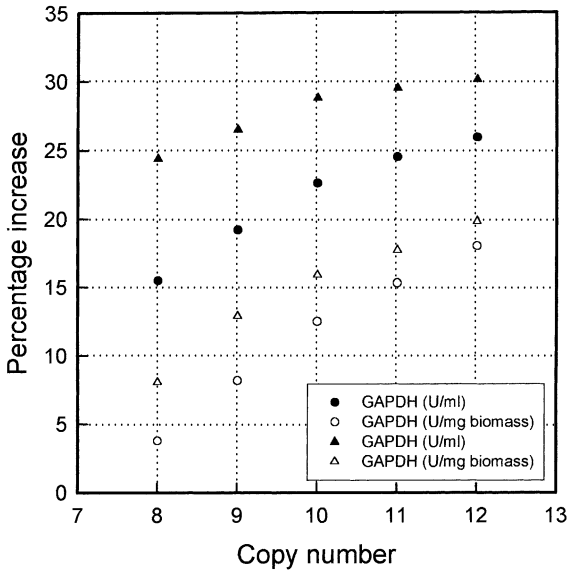

Fig. 6. Percentage increase in the peak GAPDH activity achieved by fed-batch fermentation over batch fermentation at different plasmid copy numbers. The circles are for fermentations with the same initial copy number for all the cells and the triangles are for optimum copy number variance among the cells.

be divided so as to obtain a good approximation of the continuous Gaussian distribution. From this perspective, inocula with distributed copy numbers are easier to prepare for fed-batch fermentation and are closer to the optimum distribution than is the case for batch fermentation. A small variance also favours retention of plasmid stability during cell division [32].

In addition to this starting advantage, fed-batch also generates more GAPDH activity than batch fermentation does. Since the GAPDH profiles for batch fermentation have already been published [12], Fig. 6 shows the extents of improvement possible through fed-batch. The activity is enhanced significantly throughout the range of copy numbers, regardless of whether all recombinant cells in the seed culture have the number of plasmids or there is an optimum variance of copy numbers. Although the extent of improvement increases with the copy number in both situations, this is moderated by other observations. The superiority of fed-batch over batch operation is lower for GAPDH activity per unit cell mass than it is per unit volume of broth, and it decreases as the copy number increases (because corresponding pairs of points approach each other). Thus, while results for uniform copy numbers (Fig. 4) suggest that fed-batch operation becomes more favourable as the copy number increases, the additional benefit of an optimum starting distribution reduces with increasing copy number. Similar to the trade-off between cell growth and GAPDH synthesis, the optimum copy number (and its distribution) for which fed-batch is the best choice thus involves balancing the improvements shown in Figs. 4 and 6. This is an economic consideration.

5. Concluding remarks

Although demarcation of cells into plasmid-free ones and plasmid-bearing cells with the same number plasmids in each cell is a simple approach, in recombinant fermentations carried out on a realistic scale there is actually a distribution of plasmid copy number among the cells. This study analysed the role of such a distribution on the performance of a bioreactor operated in batch and fed-batch modes. Based on prior work [5,6], a Gaussian distribution was employed for ensemble averaged copy numbers between 8 and 12.

Without a distribution in the starting copy number, fed-batch fermentation generated more biomass and GAPDH activity than batch fermentation did. In both the cases, larger copy numbers resulted in less growth and more GAPDH synthesis. This observation suggests that it may be necessary to counter-balance a high growth rate and a high protein synthesis rate in choosing an optimum plasmid copy number. The optimum copy number is subject to metabolic limitation because too large a plasmid load may eventually become unsustainable [2,21].

Of particular interest was the result that when GAPDH activity was expressed on the basis of cell mass, it attained maxima at about the same time (3.12 h) for all copy numbers. These maxima were further increased by applying Gaussian distributions to the copy numbers. For batch as well as fed-batch fermentations there was an optimum variance for each value of the average copy number of the initial population of cells. This optimum decreased as the copy number increased and the variances for fed-batch fermentation were always smaller.

The smaller variances required for large copy numbers and for fed-batch operation have two advantages. They allow easier preparation of seed cultures, as explained in Section 4, and they provide greater segregational stability [32]. The benefit of fed-batch fermentation with a distributed copy number is, however, moderated by the observation that the incremental benefit over that possible through batch fermentation decreases with increasing copy number.

It will be of interest to investigate whether optimum variances and the relative performances of fed-batch and batch fermentations observed for pBR Eco gap in *E. coli* are specific to high copy number plasmids or are also true for low copy numbers. Apart from differences in their effects on cell growth and protein synthesis [2,21,23], these two classes of plasmids also differ in their replication patterns. The replication of low copy number plasmids is tuned to the cell cycle while plasmids which are present in many copies replicate exponentially throughout the division cycle [6]. This difference is thought to be one reason for the high segregational stability of low copy plasmids. This study shows that, for pBR Eco gap, fed-batch fermentation requires a lower variance of the copy number, which also contributes to plasmid stability. Now, this plasmid is derived from pBR322 [15], whose replication is not synchronised with the cell cycle [20]. Therefore, it is relevant to ask whether a distribution

in copy number will also favour fermentations employing plasmids with cell-cycle specific replication, which are inherently more stable. Since a distribution in plasmid copy number is a natural feature of large scale fermentations, the questions raised here are important in determining how a bioreactor should be operated best under realistic conditions.

References

[1] J.M. Mitchison, W.S. Vincent, in: I.L. Cameron, G.M. Padilla (Eds.), Cell Synchronicity, Academic Press, New York, 1966.

[2] W.E. Bentley, N. Mirjalili, D.C. Anderson, R.H. Davis, D.S. Kompala, Plasmid-encoded protein: the principal factor in the metabolic burden associated with recombinant bacteria, Biotechnol. Bioeng. 35 (1990) 668–687.

[3] W.E. Bentley, O.E. Quroga, Investigation of subpopulation heterogeneity and plasmid stability in recombinant *Escherichia coli* via a simple segregated model, Biotechnol. Bioeng. 42 (1993) 222–234.

[4] D. Byrne, An investigation of the dynamics of plasmid loss in recombinant culture, Ph.D. Thesis, National University of Ireland, 1990, cited by: A.J. McLoughlin, Plasmid stability and ecological competence in recombinant cultures, Biotechnol. Adv. 12 (1994) 279–324.

[5] D. Lamotte, A. Gschaedler, J. Boudrant, Transitory plasmid instability during the exponential phase of growth of a recombinant *Escherichia coli*, Biotechnol. Lett. 16 (1994) 1113–1118.

[6] H. Kuo, J.D. Keasling, A Monte Carlo simulation of plasmid replication during the bacterial division cycle, Biotechnol. Bioeng. 52 (1996) 633–647.

[7] N.A. DaSilva, J.E. Bailey, Theoretical growth yield estimates for recombinant cells, Biotechnol. Bioeng. 28 (1986) 741–746.

[8] S. Shioya, Optimization and control in fed-batch bioreactors, Adv. Biochem. Eng. Biotechnol. 46 (1992) 111–142.

[9] K. Ye, S. Shin, K. Shimizu, Fuzzy neural network for control of high cell density cultivation of recombinant *Escherichia coli*, J. Ferment. Bioeng. 77 (1994) 663–673.

[10] K. Fujioka, K. Shimizu, Efficient production of gene products by repeated fed-batch fermentation using multiple fermentors, J. Ferment. Bioeng. 77 (1994) 194–198.

[11] C. Cheng, Y.L. Huang, S.-T. Yang, A novel feeding strategy for enhanced plasmid stability and protein production in recombinant yeast fed-batch fermentation, Biotechnol. Bioeng. 56 (1997) 23–31.

[12] P.R. Patnaik, Optimizing initial plasmid copy number distribution for improved protein activity in a recombinant fermentation, Biochem. Eng. J. 5 (2000) 101–107.

[13] N. El Houtaia, N. Nancib, G. Branalant, C. Branalant, J. Boudrant, Production of glyceraldehyde-3-phosphate dehydrogenase using genetically engineered *Escherichia coli*, Biotechnol. Lett. 11 (1989) 775–778.

[14] N. Nancib, R. Mosrati, J. Boudrant, Modelling of batch fermentation of a recombinant *Escherichia coli* producing glyceraldehyde-3-phosphate dehydrogenase on a complex selective medium, Chem. Eng. J. 52 (1993) B35–B48.

[15] G. Branalant, G. Flesch, C. Branalant, Molecular cloning of glyceraldehyde-3-phosphate genes of *Bacillus stearothermophilus* and *Escherichia coli*, Gene 25 (1983) 1–7.

[16] W. Ferdinand, Isolation and specific activity of rabbit muscle GAPDH, Biochem. J. 92 (1964) 1978–1985.

[17] N. Nancib, C. Branalant, J. Boudrant, Metabolic roles of peptone and yeast extract for the culture of a recombinant strain of *Escherichia coli*, J. Ind. Microbiol. 8 (1991) 165–170.

[18] D.S. Kompala, D. Ramkrishna, N.B. Jansen, G.T. Tsao, Investigation of bacterial growth on mixed substrates: experimental evaluation of cybernetic models, Biotechnol. Bioeng. 28 (1986) 1044–1055.

[19] P. Doshi, R. Rengaswamy, K.V. Venkatesh, Modelling of microbial growth for sequential utilization in a multi-substrate environment, Process Biochem. 32 (1997) 643–650.

[20] S. Lin-Chao, H. Bremer, Effect of the bacterial growth rate on replication control of plasmid pBR322 in *Escherichia coli*, Mol. Gen. Genet. 203 (1986) 143–149.

[21] J.C. Diaz Ricci, M.A. Hernandez, Plasmid effects on *Escherichia coli* metabolism, Crit. Revs. Biotechnol. 20 (2000) 79–108.

[22] V.N. Satyagal, P. Agarwal, A generalized model of plasmid replication, Biotechnol. Bioeng. 33 (1989) 1135–1144.

[23] P.R. Patnaik, An evaluation of models for the effect of plasmid copy number on bacterial growth rate, Biotechnol. Lett. 22 (2000) 1719–1725.

[24] N. Nancib, R. Mosrati, J. Boudrant, Study of population dynamic for a recombinant bacterium during continuous cultures application of data filtering and smoothing, Biotechnol. Bioeng. 39 (1992) 389–407.

[25] J. Lee, W.F. Ramirez, Optimal fed-batch control of induced foreign protein production by recombinant bacteria, AIChE J. 40 (1994) 899–907.

[26] M. de Trembay, M. Perrier, C. Chavarie, J. Archambault, Optimization of fed-batch culture of hybridoma cells using dynamic programming: single and multi-feed cases, Bioprocess Eng. 7 (1992) 229–234.

[27] J.A. Roubos, G. van Straten, A.J.B. van Boxtel, An evolutionary strategy for fed-batch bioreactor optimization concepts and performance, J. Biotechnol. 67 (1999) 173–187.

[28] P. Angelov, R. Guthke, A genetic-algorithm-based approach to optimisation of bioprocesses described by fuzzy rules, Bioprocess Eng. 16 (1997) 299–303.

[29] P.R. Patnaik, Start-up sensitivity to the initial state of a batch bioreactor for a recombinant *Escherichia coli* in a complex medium, Bioprocess Eng. 15 (1996) 215–219.

[30] G.A. Montague, A.C. Ward, A suboptimal solution to the optimisation of bioreactors using the chemotaxis algorithm, Process Biochem. 29 (1994) 489–496.

[31] H.J. Bremmerman, R.W. Anderson, An alternative to back propagation: a simple rule for synaptic modification of neural network training and memory, Internal Report, Department of Mathematics, University of California, Berkeley, CA, 1989.

[32] D.K. Summers, The kinetics of plasmid loss, Trends Biotechnol. 9 (1991) 273–278.