Protein burden in *Zymomonas mobilis*: negative flux and growth control due to overproduction of glycolytic enzymes

Jacky L. Snoep,^{1,2,4} Lorraine P. Yomano,¹ Hans V. Westerhoff^{2,3,4,†} and Lonnie O. Ingram¹

Author for correspondence: H. V. Westerhoff. Tel: +31 20 4447230. Fax: +31 20 4447229.
e-mail: HW@BIO.VU.NL

¹ Department of Microbiology and Cell Science, University of Florida, Gainesville, FL 32611, USA

² Division of Molecular Biology, H5, The Netherlands Cancer Institute, The Netherlands

³ E. C. Slater Institute, BioCentrum, University of Amsterdam, Amsterdam, The Netherlands

⁴ Department of Microbial Physiology, BioCentrum, Free University, Amsterdam, The Netherlands

Increasing the expression of various glycolytic operons in *Zymomonas mobilis* caused a significant decrease rather than increase in the glycolytic flux and growth rate. Because the relative decrease depended on the amount of overexpressed protein, and was independent of which enzyme was overexpressed, we attributed it to a protein burden effect. More specifically, we examined if the decrease in glycolytic flux could be explained by a decreased concentration of other glycolytic enzymes (for which glucokinase was used as a marker enzyme). Using the summation theorem of metabolic control theory we predicted the extent of this protein burden effect. The predictions were in good agreement with the experimental observations. This suggests that the negative flux control is caused either by a simple competition of the overexpressed gene with the expression of all other genes or by simple dilution. Furthermore, we determined the implications of protein burden for the determination of the extent to which an enzyme limits a flux. We conclude that a protein burden can cause a significant underestimation of the flux control coefficient, especially if the enzyme under investigation is a highly expressed enzyme.

Keywords: protein burden, *Zymomonas mobilis*, flux control, glycolysis, metabolic control analysis

INTRODUCTION

Attempts to optimize metabolic processes, by overexpression of enzymes that are thought to be important in the determination of the overall rate of formation of the desired product, often lead to negative results (Schaaff *et al.*, 1989; Niederberger *et al.*, 1992). This can be due to: (1) an intuitive overestimation of the actual importance of the enzyme in the production process – a clear distinction should be made between essential and controlling enzymes (Jensen *et al.*, 1993a, b); (2) a shift of control upon overexpression (De Hollander, 1994; Small & Kacser, 1993); or (3) additional effects caused by the overexpression (Bailey, 1993). In the first case a quan-

titative metabolic control analysis should lead to the identification of the enzymes that do exert flux control (Kacser & Burns, 1973; Heinrich & Rapoport, 1974; Groen *et al.*, 1982; Fell, 1992). In the second case combined overexpression of a group (module, Schuster *et al.*, 1993) of enzymes should help (Small & Kacser, 1993).

In the third case, the unspecific negative effects can be divided into energetic effects (costs to produce extra protein) and competitive effects (if the protein-synthesizing machinery is limiting). In this paper we will use the term protein burden for the negative effect on any part of cell function caused by the overexpression of a protein independent of its catalytic activity. In order to evaluate the importance of an enzyme for the control of any flux under study, it is important to distinguish specific negative effects of the catalytic activity from this 'non-specific' protein burden. Thus, a quantitative analysis of the negative effects of expression of recombinant protein is needed.

The protein burden effect has been recognized for quite some time but was shown to be difficult to quantify

† Present address: Faculty of Biology, Department of Microbial Physiology, Vrije Universiteit, de Boelelaan 1087, NL-1081 HV Amsterdam, The Netherlands.

Abbreviations: ADH, alcohol dehydrogenase; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; G6PDH, glucose-6-phosphate dehydrogenase; GLK, glucokinase; PDC, pyruvate decarboxylase; PGK, phosphoglycerate kinase; PGM, phosphoglycerate mutase.

(Koch, 1983). Initial experiments by Novick & Weiner (1957) showed a negative effect of the 'gratuitous' expression of genes of the *lac* operon. Andrews & Hegeman (1976) discuss some experiments attempting to measure protein burden. They concluded that there is a measurable reduction in growth rate among strains with high levels of nonfunctional protein but at low levels of expression these effects are small. The mechanism via which this burden is translated in the cell is not known. Many studies have mentioned potentially negative energy aspects of, for instance, the maintenance of plasmids (Seo & Bailey, 1985) or the production of the recombinant protein (Koch, 1983). For a review on host-vector interactions in *Escherichia coli* see Bailey (1993). However, it was not substantiated that these processes had actual control over the aspect of cell function that was studied (usually growth rate). Koch (1983) concluded that permease insertion into the cytoplasmic membrane (for the case of the *lac* operon which has always been used as a model system) is deleterious and slows growth more than could be accounted for by the synthesis of unneeded protein.

In metabolic control analysis, the extent to which an enzyme controls a flux, the enzyme's flux control coefficient (Burns *et al.*, 1985), is defined operationally as the percentage increase in flux resulting from a 1% activation of the enzyme (Kell & Westerhoff, 1986; Kell, 1987). An important way experimentally to activate an enzyme is to enhance the expression of the corresponding gene so as to increase the concentration of the enzyme (Walsh & Koshland, 1985). However, overexpression of the enzyme concentration might lead to an underestimation of the flux control due to protein burden. In control theory the negative effect on cell function and flux caused by protein burden would be translated into a negative control coefficient.

In this study we report on effects of overexpressing several of the glycolytic enzymes of *Zymomonas mobilis*, the activities of which appeared to have no control on the glycolytic flux (Arfman *et al.*, 1992; Yomano *et al.*, 1993). We could attribute the negative effect on glycolytic flux to the protein burden resulting from the synthesis of plasmid-encoded protein.

METHODS

Bacterial strains, plasmids and growth conditions. *Z. mobilis* strains were grown at 30 °C in complex medium as previously described (100 g glucose l⁻¹) but without nalidixic acid (Arfman *et al.*, 1992). Tetracycline (10 mg l⁻¹) was added as appropriate for selection. Where indicated, media included 2 mM IPTG. Plasmid construction and recombinant strains have been described previously by Arfman *et al.* (1992) and Yomano *et al.* (1993). Minimal coding regions of the respective genes were amplified from chromosomal *Z. mobilis* DNA using the polymerase chain reaction and included terminal restriction sites for directional insertion into pLOI706EH. This vector contains an RSF1010 replicon, the *tac* promoter and the *lacI*^q repressor gene (Arfman *et al.*, 1992).

Measurement of glycolytic flux. Glycolytic flux was measured as CO₂ evolution from 1 l cultures as previously described by Arfman *et al.* (1992) but using an acoustic flow meter (model

ADM 2000, J&W Scientific). Nalidixic acid was omitted resulting in a 15% higher rate of flux than previously reported by Arfman *et al.* (1992). Results are expressed as µmol CO₂ evolved min⁻¹ (mg cell protein)⁻¹ (q_{CO_2}). Reported values for q_{CO_2} were determined at culture densities of OD₅₅₀ 1.0–2.0.

Measurement of enzyme activities. Cultures were harvested by centrifugation at approximately OD₅₅₀ 2.0. For glucokinase (GLK), glucose-6-phosphate dehydrogenase (G6PDH), phosphoglycerate mutase (PGM), pyruvate decarboxylase (PDC), phosphoglycerate kinase (PGK), and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) assays, pellets were washed and resuspended in 20 mM MES/KOH (pH 6.5) containing 2 mM MgCl₂, 50 mM NaCl and 10 mM β-mercaptoethanol. For alcohol dehydrogenase (ADH) assays, 10 mM sodium ascorbate and 0.5 mM ferrous ammonium sulfate were added to this buffer. After disruption in a French pressure cell, specific activities were determined for GLK (Scopes *et al.*, 1985), NAD⁺-dependent G6PDH (Scopes *et al.*, 1985), PGM (Pawluk *et al.*, 1986) with the addition of 0.1 mM 2,3-bisphosphoglyceric acid, PDC (Neale *et al.*, 1987), GAPDH (Pawluk *et al.*, 1986), PGK (Pawluk *et al.*, 1986) and for 'total ADH' (Neale *et al.*, 1986). Protein was measured using the dye-binding assay (Bradford, 1976).

RESULTS

Overproduction of glycolytic enzymes leads to a protein burden effect on CO₂ production

The catabolic flux from glucose to ethanol has been studied quite extensively in *Z. mobilis* (Osman *et al.*, 1987; Arfman *et al.*, 1992; Algar & Scopes, 1985). Our current study focused on effects of protein burden and we limited ourselves to six of the glycolytic enzymes. The selected enzymes are relatively simple to assay and have been well characterized. Initially they were found to be present at concentrations that were not controlling glycolytic flux (Arfman *et al.*, 1992; Yomano *et al.*, 1993). We overexpressed these enzymes using a controllable expression vector and studied effects on the glycolytic flux (as measured as CO₂ evolution rate) and specific growth rate.

As was shown before, *Zymomonas* strain CP4 containing the empty vector [CP4(pLOI706EH)] exhibited no difference in growth rate, glycolytic flux, or glycolytic enzyme activities as compared to the same strain without that vector (Arfman *et al.*, 1992; Yomano *et al.*, 1993). However, overexpression of the glycolytic enzymes from that same vector caused a decrease in growth rate and glycolytic flux, an effect which was stronger in the presence of IPTG (Table 1).

In order to distinguish between a negative effect due to the catalytic activity of the enzymes and a negative effect of the burden placed on the cell to make additional protein, we calculated for each case how much protein was overexpressed. The calculation was based on the specific activity of the purified enzyme and the increase in specific activity in the recombinant strain as compared to the strain with the empty plasmid, CP4(pLOI706EH). Dividing these two activities yielded the change in amount of recombinant protein per total cell protein: Δ (activity per unit total protein)/(activity per unit pure enzyme) = Δ (unit pure enzyme/unit total protein). The specific activities of the purified enzymes were taken from the

Table 1. Growth rate (μ), glycolytic flux (J) and enzyme activities in *Z. mobilis* recombinants

Values are means $\pm$ estimated SEM from two or more experiments. The estimated SEM was calculated as $(SEM)^2 = \Sigma(x_i - \bar{x})^2 / n(n-1)$, where x_i refers to the observed value, $\bar{x}$ to the mean value and n to the number of observations.

Strain, plasmid, gene	μ (h ⁻¹)	J [$\mu\text{mol CO}_2 \text{ min}^{-1}$ (mg protein) ⁻¹]	Recombinant enzyme activity [IU (mg protein) ⁻¹]*	GLK activity [IU (mg protein) ⁻¹]
CP4	0.54 $\pm$ 0.01	2.03 $\pm$ 0.03		1.5 $\pm$ 0.11
CP4(pLOI706EH)	0.53 $\pm$ 0.01	2.06 $\pm$ 0.02		1.5 $\pm$ 0.04
CP4(pLOI794) (<i>gap</i>)	0.48 $\pm$ 0.01	2.03 $\pm$ 0.02	5.2 $\pm$ 0.66 (2.1 $\pm$ 0.14)	1.5 $\pm$ 0.04
+ 2 mM IPTG	0.34 $\pm$ 0.01	1.54 $\pm$ 0.02	35.6 $\pm$ 1.40	1.2 $\pm$ 0.10
CP4(pLOI795) (<i>pgk</i>)	0.50 $\pm$ 0.01	1.97 $\pm$ 0.02	4.8 $\pm$ 0.79 (2.8 $\pm$ 0.22)	1.4 $\pm$ 0.05
+ 2 mM IPTG	0.30 $\pm$ 0.01	1.35 $\pm$ 0.02	30.9 $\pm$ 3.33	0.9 $\pm$ 0.10
CP4(pLOI697) (<i>pgm</i>)	0.50 $\pm$ 0.01	2.07 $\pm$ 0.02	17.9 $\pm$ 1.18 (14.3 $\pm$ 0.47)	1.6 $\pm$ 0.04
+ 2 mM IPTG	0.51 $\pm$ 0.01	2.05 $\pm$ 0.01	26.9 $\pm$ 0.32	1.6 $\pm$ 0.03
CP4(pLOI714) (<i>pdh</i>)	0.48 $\pm$ 0.01	1.86 $\pm$ 0.02	10.3 $\pm$ 1.04 (3.2 $\pm$ 0.12)	1.4 $\pm$ 0.07
CP4(pLOI715) (<i>adhB</i>)	0.50 $\pm$ 0.01	2.02 $\pm$ 0.02	13.0 $\pm$ 0.94 (2.9 $\pm$ 0.38)	1.5 $\pm$ 0.13
+ 2 mM IPTG	0.47 $\pm$ 0.01	1.82 $\pm$ 0.02	105 $\pm$ 6.60	1.3 $\pm$ 0.04
CP4(pLOI716) (<i>adhA</i>)	0.50 $\pm$ 0.01	2.03 $\pm$ 0.02	9.8 $\pm$ 2.79 (2.9 $\pm$ 0.38)	1.5 $\pm$ 0.06
+ 2 mM IPTG	0.42 $\pm$ 0.01	1.58 $\pm$ 0.03	51.0 $\pm$ 4.35	1.2 $\pm$ 0.01

* Recombinant enzyme activity is the specific *in vitro* activity of the plasmid-encoded enzyme; in parentheses the activity of the same enzyme is given as measured in CP4(pLOI706EH).

Fig. 1. Overproduction of plasmid-encoded protein affected glycolytic flux (J) and growth rate (μ). The solid line illustrates the predicted relationship between the amount of overproduced protein and glycolytic flux (■) (or growth rate, □) assuming that the enzyme had no control on flux (or growth rate) as derived in Appendix 1, equation 7; a similar relationship would be expected between $\Delta(e_i/e)$ and $\Delta(\mu/\mu)$. Six glycolytic enzymes (cf. Table 1) were overexpressed using a controllable expression vector.

enzymes was overexpressed, glycolytic flux decreased rather than increased.

By definition protein burden is independent of the enzyme that is overexpressed; it is not the specific activity of the enzyme that is causing the negative effect but the amount of protein that is incorporated in the overexpressed enzyme. This should be reflected in a correlation between the amount of recombinant protein that is overexpressed Δ (recombinant protein/total protein) and the percent decrease in flux and growth rate. Fig. 1 suggests that this correlation exists: its data points refer to various enzymes with different specific activities and different expression levels. Not the specific activities *per se* but the extra production of protein caused the decrease in flux. A similar correlation was observed between the relative amount of protein that is overproduced and the relative decrease in growth rate (Fig. 1). The observed correlation in Fig. 1 cannot now be interpreted mechanistically, i.e. the protein burden can be due to an energetic effect (extra costs to synthesize additional protein) or to an effect on the capacity of the biosynthetic apparatus (the cell not being able kinetically to cope with synthesizing both its 'own' and the plasmid-encoded proteins).

Estimation of the extent of the protein burden effect

When a protein without direct flux control is overexpressed, this may lead to a reduction in the relative concentration (a 'dilution') of other proteins. These other proteins may well include the proteins that exert control on the flux under study. In fact, because total flux control by all enzymes equals 1 (for exceptions see below), one should expect a reduction in flux when one overexpresses an enzyme that does not control that flux directly. To test

literature: GAPDH, 205 IU (mg protein)⁻¹ (Pawluk *et al.*, 1986); PGK, 800 IU (mg protein)⁻¹ (Pawluk *et al.*, 1986); PGM, 2000 IU (mg protein)⁻¹ (Pawluk *et al.*, 1986); PDC, 130 IU (mg protein)⁻¹ (Neale *et al.*, 1987); ADH II, 950 IU (mg protein)⁻¹ (Neale *et al.*, 1986); ADH I, 240 IU (mg protein)⁻¹ (Neale *et al.*, 1986). For all cases of overexpression, the relative change in glycolytic flux was then plotted versus the relative amount of overexpressed protein (Fig. 1). Independently of which of the six

Fig. 2. Overproduction of plasmid-encoded protein affected GLK activity. The solid line illustrates the predicted relationship between the overproduction of protein and the decrease of a chromosomally encoded enzyme activity as derived in Appendix 1, equation 4 and Appendix 2, equation 15. Data represent six glycolytic enzymes which were overexpressed using a controllable expression vector and GLK activity was measured in cell-free extracts.

the first part of this hypothesis, i.e. that other enzymes are diluted, we measured the 'dilution' of a marker glycolytic enzyme, GLK. Indeed a decrease in GLK activity in a dose-dependent manner was observed upon overexpression of plasmid-encoded protein (Fig. 2).

When elaborating the concept of negative flux control through the protein burden effect, the precise result might be expected to depend on how expression of the neutral protein affects the concentration of the flux-controlling proteins. We therefore analysed two models in detail.

(1) The competitive model. There is very strict competition in the sense that the increase in synthesis of e_1 is compensated for by an equal decrease in the total synthesis of all other proteins: $\delta e = 0$. Here e represents total protein in the cell and e_1 is the protein in the overexpressed enzyme. In a special case of the latter model, the 'equal competition model', all proteins decrease in the same proportion.

(2) The noncompetitive model. There is no competition for protein synthesis, and also otherwise no effect of the increased synthesis of enzyme 1 on the synthesis of the other proteins.

The effect of overexpression of recombinant protein e_1 on activities of other proteins (e_j ; $j \neq 1$) can be predicted according to either of these models. In the equal competition model all enzyme concentrations other than e_1 will be decreased to the same relative extent and this decrease will be equal to the change in e_1 relative to all other proteins. Since in the equal competition model the total protein content remains constant, changes in enzyme concentrations are directly related to changes in enzyme activities (cf. Appendix 1, equation 4)

$$\frac{\Delta \alpha_j}{\alpha_j} = \Delta \left(\frac{e_1}{e} \right) * -\frac{e}{e - e_1} \quad (j \neq 1)$$

Here α_j is the specific activity of any enzyme, e_j , in the pathway (except enzyme 1).

In Appendix 2 we show that in the noncompetitive model exactly the same relation (equation 15) between the relative change in α_j and the overexpression of e_1 (total protein)⁻¹ is expected as in the equal competitive model, provided that there is no effect of any energetic burden on protein synthesis.

The predicted relationship based on the two models between the relative change in GLK activity and the change in e_1 relative to all other proteins is given by the solid line in Fig. 2. It appears that the observed 'dilution' of GLK is similar to that expected on the basis of the models. That the dilution is actually somewhat greater than expected may be partly due to the energy cost of the additional protein synthesis if the noncompetitive model holds. At least semiquantitatively, Fig. 2 thus confirms the first aspect of the protein burden concept: overexpression of recombinant protein resulted in a decrease of the activity of other proteins.

The second aspect of the protein burden concept is that this decrease could be responsible for the decrease in glycolytic flux observed in Fig. 1. In terms of metabolic control theory (Kacser & Burns, 1973; Westerhoff & Van Dam, 1987) flux should decrease proportionately when all enzymes that exert control are inactivated to the same relative extent; the enzymes that were overexpressed are thought to have no control on flux and the combined remaining enzymes should thus have a combined control close to 1 (assuming an ideal pathway in which the sum of all control coefficients is 1; for different cases see Van Dam *et al.*, 1993; Kholodenko & Westerhoff, 1994, 1995). This prediction is independent of any specific protein burden model. Taking the relative changes in GLK activity as representative for the relative changes in activity of the enzymes that control the glycolytic flux, the prediction is that the relative change in flux equals the relative change in GLK activity (equation 6, Appendix 1). In Fig. 3 the observed relative changes in GLK activity and glycolytic flux are plotted. Indeed the fractional decrease in GLK activity showed the expected correlation (with a slope close to 1) with the fractional change in glycolytic flux (Fig. 3). Not only the glycolytic enzymes but also the biosynthetic enzymes should be diluted and indeed a similar decrease in growth rate as in glycolytic flux was observed (Fig. 3).

Interestingly, one should be able to estimate how the flux will be reduced upon overexpression of protein as one can predict how the nonoverproduced proteins will be diluted. In Appendices 1 and 2 we derived the expected relationship between the relative amount of overproduced protein and the decrease in activity of other proteins (equations 4 and 15 for the competitive and non-competitive model, respectively). Again we assumed that the proteins that are overexpressed have no direct control on the flux of interest and thus that the sum of control coefficients of the other proteins will be 1. The line in Fig. 1 is drawn on the basis of the expected relationship

Fig. 3. Correlation between the relative decrease in GLK activity and the relative changes in μ (□) and J' (■) upon overexpression of glycolytic enzymes. The solid line represents the predicted relationship assuming no control on flux (or growth rate) of the overexpressed protein (equation 6, Appendix 1). GLK is used as a marker enzyme to measure the effect on all other enzymes.

(equation 7, Appendix 1). Clearly, it comes close to fitting the experimental flux data.

Protein burden and the analysis of metabolic control

The extent to which an enzyme (e_1) controls the flux through a pathway has been defined by Kacser & Burns (1973), Burns *et al.* (1985) and Westerhoff & Van Dam (1987) as the percentage increase in steady state flux resulting from a 1% activation of the enzyme at constant activities of all other enzymes. For an alternative definition see Schuster & Heinrich (1992) and Kholodenko *et al.* (1995). More precisely

$$C_{e_1}^J = \frac{e_1}{J} * \left(\frac{dJ}{de_1} \right) = \left(\frac{d \ln J}{d \ln e_1} \right)_{(de_j=0)} = \left(\frac{\partial \ln J}{\partial \ln e_1} \right)_{e_j}$$

One way in which such a flux control coefficient has been measured has been by increasing the concentration of the enzyme by manipulation of gene expression. In practice, both the flux J and the enzyme concentration are taken per unit biomass, and the control coefficient is estimated by dividing the relative change in specific flux by the relative change in specific enzyme activity

$$C_{e_1}^J(\text{app}) = \frac{\delta \ln J'}{\delta \ln e_1}$$

with: $J' = J/e$ and $e'_1 = e_1/e$. Importantly, in this second operational definition, it is not ascertained that the other enzyme activities do not change. This difference becomes relevant when the gene expression is changed (we shall not deal with this case here, see Westerhoff *et al.*, 1990).

In Appendix 1 (equation 10) it is shown how the fractional change in specific flux ($J' = J/e$) is related to the relative change in the specific concentration of the overexpressed protein (e_1/e)

$$C_{e_1}^J(\text{app}) = \frac{e}{(e - e_1)} * C_{e_1}^J - \frac{e_1}{e - e_1}$$

Fig. 4. Dependence of the difference between apparent and true flux control coefficient [Z-axis, $C_{e_1}^J(\text{app}) - C_{e_1}^J$] on enzyme concentration (X-axis, e_1/e) and control coefficient of that enzyme (Y-axis, $C_{e_1}^J$).

This equation shows that whenever $C_{e_1}^J < 1$, $C_{e_1}^J(\text{app}) < C_{e_1}^J$. Consequently, if one determined the control coefficient of e_1 with respect to J by increasing the concentration of e_1 , without correcting for the protein burden effect, one would obtain $C_{e_1}^J(\text{app})$ rather than $C_{e_1}^J$: the latter would be underestimated. The underestimation would become more pronounced at higher concentrations of the recombinant protein relative to total protein.

Surprisingly, in the noncompetitive model exactly the same underestimation of the flux control coefficient would be made (Appendix 2). Although according to this model there is no change in absolute flux and enzyme concentration, there is still a decrease in flux and enzyme concentration (total cell protein) $^{-1}$ due to the increase in total protein content of the cell. Consequently, due to the way we express glycolytic flux and enzyme concentration [i.e. (mg total cell protein) $^{-1}$], there is no apparent difference between the two models. However, although the decrease in glycolytic flux that is predicted according to the two models might be the same, the actual cause is fundamentally different; in the noncompetitive model there is a decrease in specific activity of the other enzymes due to the increase in total protein content while in the competitive model this is due to a decrease in the synthesis of these enzymes.

The equations containing control coefficients as derived in Appendices 1 and 2 refer to (per definition) small changes in e_1 . However, the assumption that e_1 has no control on flux and that all control is thus located in the rest of the enzymes of the pathway allows derivation of equations that are valid also with larger changes in e_1 (equations 1–7 of Appendix 1; equations 13–15 of Appendix 2). The expected relationships between the relative change in J and the change in e_1 (total protein) $^{-1}$ indicate that the values as predicted with the two models fit quite well with the experimental data.

Fig. 4 shows how the difference between the apparent and true flux control coefficient of enzyme 1 is predicted to vary with the specific concentration of enzyme 1 (e_1/e) for

various magnitudes of its control coefficient. Only if enzyme 1 has full control, the apparent control coefficient equals the true control coefficient of 1. When enzyme 1 contributes less than 5% to the total protein concentration of the cell, the apparent control coefficient underestimates the true control coefficient only marginally. However, if enzyme 1 contributes 15% and its true control coefficient is zero, its apparent control coefficient is approximately -0.2 . Fig. 4 shows that the higher the enzyme concentration and/or the more negative its flux control coefficient the greater the difference between the true and apparent control coefficient.

DISCUSSION

Negative effects on cell function (cell growth in particular) have been widely observed upon overexpression of protein from plasmids. Explanations that have been proposed for this protein burden have eluded experimental verification. This was partly because the precise mechanism of the protein burden effect was uncertain and proposed mechanisms were unlikely, and partly because it was difficult to predict how extensive the protein burden effect should be for any proposed mechanism. Because protein synthesis constitutes one of the major free-energy sinks of the growing prokaryote (Stouthamer, 1979), some explanations of the protein burden effect focused on the possibility that the extra protein synthesis compromises the free-energy state of the cells. The living cell might be optimal in terms of its protein composition and any extra protein synthesis would waste free energy already destined for useful processes (Marr, 1991). However, in many cases the free-energy availability does not seem to be the main concern for an organism (Stucki, 1980; Westerhoff *et al.*, 1983; Tempest & Neijssel, 1984; Jensen *et al.*, 1993a, b; Linton, 1991). Estimation of the extent of this energy-related protein burden effect again suffered from difficulties in estimating the free-energy cost of the synthesis of the extra proteins, and the free-energy dependence of growth rate (Dykhuizen & Hartl, 1983; Koch, 1983; Diamond, 1986). Perhaps the most important aspect of the present paper is that it quantifies a protein burden effect that does not depend on where precisely the limitation lies. From the summation principles embodied in metabolic control analysis, one can derive from the extent to which the protein that is overexpressed controls (limits) a flux or growth rate, the extent to which all other enzymes limit the flux. This then allows one to predict the protein burden effect of overexpressing a protein. The present paper also showed that this prediction was in line with the experimental results. Importantly, our results are independent of assumptions as to what limits flux or growth rate. Furthermore, they do not depend on postulating an optimum state of the cells, such as was done in other work that predicted the relative magnitudes of control coefficients (Brown, 1991; Schuster & Heinrich, 1992; Wilhelm *et al.*, 1994).

More specifically, we showed that the overexpression of plasmid-encoded protein led to the dilution of other enzymes and this had a negative effect on both glycolytic flux and growth rate. We derived equations describing the

effect of protein burden on the determination of flux control coefficients. We based these equations on the assumption of ideal metabolism, such that the sum of the enzymes' flux control coefficients equals 1. Exceptions to this rule occur in cases of metabolic channelling (Kell & Westerhoff, 1985; Kholodenko & Westerhoff, 1993), in organized multienzyme systems, and group transfer (Van Dam *et al.*, 1993). The equations can be readily adjusted to these cases, but we have opted for simplicity here. If these complications were inserted, then the protein burden effect might be increased (up to a doubling). This could be in line with Fig. 1.

Differences between the two models describing protein burden (the noncompetitive and the competitive model) are insignificant because of the way in which the metabolic flux and enzyme concentrations are expressed. If flux and enzyme concentrations were expressed as an absolute value, say per cell rather than per total protein, then no protein burden would be observed according to the noncompetitive model. In the competitive model, the resulting effect on protein burden is independent of how changes in enzyme concentrations or metabolic fluxes are expressed, as the total protein concentration does not change.

Most likely, the assumption that all enzymes are diluted to the same extent, in the equal competition model, is an oversimplification: for instance, if the ribosomal-binding sites of the different proteins would not be saturated (Jensen & Pedersen, 1990) and/or have different affinities for ribosomes, not all proteins concentrations would decrease in the same proportion. Recently it was observed in *E. coli* that upon gratuitous overexpression of β -galactosidase or truncated *tufB*, the expression of six other proteins proportionally decreased, which is in agreement with our data. In contrast, synthesis of the heat-shock proteins DnaK and GroEL increased, suggesting a different regulation of the synthesis of these stress proteins. In the study a dramatic decrease in the protein-synthesizing capacity was observed at high levels of overexpression and the authors suggested that the cessation of both growth and translation is not simply a result of a proportionate dilution of normal proteins (Dong *et al.*, 1995).

Usually effects of protein burden in optimization of catabolic fluxes will be of minor importance due to relatively low concentrations of the enzymes in catabolic routes. In some organisms, however, this should be different; indeed protein burden was evident upon overexpression of glycolytic enzymes in *Z. mobilis*. This organism reserves about 50% of its cytoplasmic protein for glycolytic enzymes (An *et al.*, 1991). Consequently, these enzymes are present at high concentrations and the organism may well be more sensitive to protein burden than, for instance, *E. coli*.

In studies considering control of flux or growth rate in *E. coli* (e.g. Ruijter *et al.*, 1991; Chao *et al.*, 1993; Van der Vlag *et al.*, 1994), protein burden was not considered. The reported control coefficients may therefore have been underestimations of the true control coefficients. How-

ever, the enzymes that were studied in these cases were initially present in low concentrations and the manipulation of concentrations was subtle. Consequently, the underestimations of the control coefficients should have been small, and insufficient to account for the missing control (Van der Vlag *et al.*, 1994). The negative effects on growth rate as observed by Jensen *et al.* (1993b) did occur at high concentrations of the overexpressed protein. Upon increasing the H⁺-ATPase concentration from 1.7 % (wild-type level) to 6.0 % of total protein, a decrease in growth rate by 34 % was observed. This negative effect is much greater than could be accounted for by the protein burden effect. In experiments in which large changes in enzyme concentrations are used in order to determine control coefficients, protein burden may interfere with a correct interpretation of the results. For instance in Schaaff *et al.* (1989) high copy number vectors were used for overexpression of glycolytic enzymes in yeast and rather large increases in enzyme activity were concluded to have no significant increase in flux. For instance, the 13.9-fold increase in GLK activity resulted in only a 1.07-fold stimulation in the ethanol-production rate. Here a protein burden effect may have caused an underestimation of the importance of the enzyme.

In this paper we showed that the effect of protein burden on cell physiology can be significant. This is in contrast to what has been proposed before by Koch (1983). Furthermore, we were able to predict the extent to which the protein burden effect influences glycolytic flux and cell growth. This prediction was based on the feature that the sum of the control coefficients of the enzymes that are not overexpressed is 1 minus the control of the overexpressed enzyme. The reasonable fit between the experimental data and this simple model indicates that energy costs do not always have to be taken explicitly into account when dealing with the protein burden effect.

ACKNOWLEDGEMENTS

The authors wish to thank Douwe Molenaar, Guy Brown, Peter Ruhdal Jensen, Jannie Hofmeyr and Mark Garner for valuable discussions during the preparation of the manuscript. This research was supported in part by the US Department of Energy (DE-FG05-86ER13575) and the Netherlands Organization for Scientific Research.

APPENDIX 1: The equal competition model

In the competitive model the total enzyme concentration is constant

$$\Delta\left(\sum_{i=1}^n e_i\right) = \Delta e = 0 \quad (1)$$

Here Δe_i refers to the change in concentration of e_i ; this change is not necessarily small. Consequently, the increase in synthesis of e_i is compensated for by an equal decrease in the total synthesis of all the other proteins

$$\sum_{j=2}^n \Delta e_j = -\Delta e_1 \quad (2)$$

In the equal competitive model, with equal fractional changes in all e_j other than e_1 , this leads to the equation

$$\frac{\Delta e_j}{e_j} = -\frac{\Delta e_1}{e - e_1} \quad (3)$$

Since $\Delta e = 0$, changes in enzyme concentration are directly related to changes in specific enzyme activities ($\alpha_j = e_j/e$)

$$\frac{\Delta \alpha_j}{\alpha_j} = -\frac{\Delta e_1}{e - e_1} = \Delta\left(\frac{e_1}{e}\right) * \frac{e}{e - e_1} \quad (j \neq 1) \quad (4)$$

Equation 4 was used to calculate the solid line in Fig. 2.

Assuming no control of e_1 on the flux, and thus full control of all the other enzymes, i.e.

$$\sum_{j=2}^n C_{e_j}^J = 1 \quad (5)$$

$$\frac{\Delta \alpha_j}{\alpha_j} = \frac{\Delta J}{J} \quad (6)$$

equation 6 was used to calculate the line in Fig. 3.

By substituting equation 6 in equation 4 this can be rewritten as

$$\frac{\frac{\Delta J}{J}}{\Delta\left(\frac{e_1}{e}\right)} = -\frac{e}{e - e_1} \quad (7)$$

Equation 7 was used to calculate the solid line in Fig. 1.

We note that equations 1–7 do not require the changes in e_1 to be infinitesimally small. In the following we shall assume that the change in e_1 is infinitesimal. The expression of e_1 has an effect on the synthesis (equation 1) and specific activity (equation 4) of all the other enzymes. This has consequences for the determination of control coefficients, as the flux J can be dependent on all enzymes in the pathway

$$\ln J = C_{e_1}^J * \ln e_1 + \sum_{j=2}^n C_{e_j}^J * \ln e_j \quad (8)$$

Applying equations 3–8 (for infinitesimal changes in e_1) and using $de_j/e_j = d\ln e_j$ yields

$$d\ln J = \left(C_{e_1}^J - \frac{e_1}{e - e_1} * \sum_{j=2}^n C_{e_j}^J\right) * d\ln e_1 \quad (9)$$

Using the summation theorem, $\sum_{j=1}^n C_{e_j}^J = 1$, this yields

$$C_{e_1}^J(\text{app}) = C_{e_1}^J * \frac{e}{e - e_1} - \frac{e_1}{e - e_1} \quad (10)$$

Equation 10 served as the basis for Fig. 4, where the apparent control of e_1 on the flux is defined by

$$C_{e_1}^J(\text{app}) = \frac{d\ln J}{d\ln e_1} \quad (11)$$

This control coefficient differs from the flux control coefficient of e_1 which is defined as

$$C_{e_1}^J = \left(\frac{\partial \ln J}{\partial \ln e_1} \right)_{e_j} \quad (12)$$

APPENDIX 2: The noncompetitive model

In the noncompetitive model there is no effect of the synthesis of e_1 on the synthesis of the other enzymes

$$\Delta e_j = 0 \quad (13)$$

Consequently,

$$\Delta e = \Delta e_1 \quad (14)$$

Although synthesis of e_1 will not lead to a difference in concentration of e_j , due to the increase in the total amount of protein there will be an effect on α_j , which represents the specific activity of e_j

$$\begin{aligned} \frac{\Delta \alpha_j}{\alpha_j} &= -\frac{\Delta e}{e + \Delta e} = \frac{\Delta e_1 (e - e_1)}{e(e + \Delta e)} * -\frac{e}{e - e_1} \\ &= \Delta \left(\frac{e_1}{e} \right) * -\frac{e}{e - e_1} (j \neq 1) \end{aligned} \quad (15)$$

Equation 15 is identical to equation 4 and is used in Fig. 2. In the case of zero control of e_1 , equation 15 can again be rewritten as equation 7.

When we denote the flux by J' , the flux and enzyme concentrations are expressed per total protein for infinitesimal changes in e_1

$$d \ln J' = \frac{d \left(\frac{J'}{e} \right)}{\frac{J'}{e}} = \frac{e}{J'} * \left(\frac{e d J' - J' d e}{e^2} \right) = d \ln J' - d \ln e \quad (16)$$

$$\text{Using } d \ln J = C_{e_1}^J * d \ln e_1$$

$$d \ln J' = C_{e_1}^J * d \ln e_1 - d \ln e = C_{e_1}^J * \frac{d e_1}{e_1} - \frac{d e_1}{e} \quad (17)$$

$$C_{e_1}^{J'}(\text{app}) = \frac{d \ln J'}{d \ln e_1} = \frac{d \ln J'}{d \left(\frac{e_1}{e} \right)} * \frac{e}{e_1} = \frac{C_{e_1}^J * \frac{d e_1}{e_1} - \frac{d e_1}{e}}{\frac{e d e_1 - e_1 d e}{e^2}} * \frac{e}{e_1} \quad (18)$$

$$C_{e_1}^{J'}(\text{app}) = C_{e_1}^{J'} * \frac{e}{e - e_1} - \frac{e_1}{e - e_1} \quad (19)$$

Consequently, provided the flux is expressed per unit total protein the same relationship between the apparent and the true flux control coefficient is expected for the noncompetitive and the equal competition model.

REFERENCES

- Algar, E. M. & Scopes, R. K. (1985). Studies on cell-free metabolism: ethanol production by extracts of *Zymomonas mobilis*. *J Biotechnol* **2**, 275–287.
- An, H., Scopes, R. K., Rodriguez, M., Keshav, K. F. & Ingram, L. O. (1991). Gel electrophoretic analysis of *Zymomonas mobilis* glycolytic

and fermentative enzymes: identification of alcohol dehydrogenase II as a stress protein. *J Bacteriol* **173**, 7227–7240.

Andrews, K. J. & Hegeman, G. D. (1976). Selective disadvantage of non-functional protein synthesis in *Escherichia coli*. *J Mol Evol* **8**, 317–328.

Arfman, N. A., Worrell, V. & Ingram, L. O. (1992). Use of the *tac* promoter and *lacIq* for the controlled expression of *Zymomonas mobilis* fermentative genes in *Escherichia coli* and *Zymomonas mobilis*. *J Bacteriol* **174**, 7370–7378.

Bailey, J. E. (1993). Host-vector interactions in *Escherichia coli*. *Adv Biochem Eng* **48**, 29–52.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principles of protein-dye binding. *Anal Biochem* **72**, 248–254.

Brown, G. C. (1991). Total cell protein concentration as an evolutionary constraint on the metabolic control distribution in cells. *J Theor Biol* **153**, 195–203.

Burns, J. A., Cornish-Bowden, A., Groen, A. K., Heinrich, R., Kacser, H., Porteous, J. W., Rapoport, S. M., Rapoport, T., Stucki, J. W., Tager, J. M., Wanders, R. J. A. & Westerhoff, H. V. (1985). Control analysis of metabolic systems. *Trends Biochem Sci* **10**, 16.

Chao, Y.-P., Patnaik, R. R., Roof, W. D., Young, R. F. & Liao, J. C. (1993). Control of gluconeogenic growth by *pps* and *pck* in *Escherichia coli*. *J Bacteriol* **175**, 6939–6944.

De Hollander, J. A. (1994). Optimization of product formation in a microbial fermentation process. In *Biothermokinetics*, pp. 405–413. Edited by H. V. Westerhoff. Andover: Intercept.

Diamond, J. M. (1986). Why do disused proteins become genetically lost or repressed? *Nature* **321**, 565–566.

Dong, H., Nilsson, L. & Kurland, C. G. (1995). Gratuitous over-expression of genes in *Escherichia coli* leads to growth inhibition and ribosome destruction. *J Bacteriol* **177**, 1497–1504.

Dykhuisen, D. E. & Hartl, D. L. (1983). Selection in chemostats. *Microbiol Rev* **47**, 150–168.

Fell, D. A. (1992). Metabolic control analysis: a survey of its theoretical and experimental development. *Biochem J* **286**, 313–330.

Groen, A. K., Wanders, R. J. A., Westerhoff, H. V., van der Meer, R. & Tager, J. M. (1982). Quantification of the contribution of various steps to the control of mitochondrial respiration. *J Biol Chem* **257**, 2754–2757.

Heinrich, R. & Rapoport, T. A. (1974). A linear steady-state treatment of enzymatic chains. General properties, control and effector strength. *Eur J Biochem* **42**, 89–95.

Jensen, K. F. & Pedersen, S. (1990). Metabolic growth rate control in *Escherichia coli* may be a consequence of subsaturation of the macromolecular biosynthetic apparatus with substrates and catalytic components. *Microbiol Rev* **54**, 89–100.

Jensen, P. R., Michelsen, O. & Westerhoff, H. V. (1993a). Control analysis of the dependence of *E. coli* physiology on the H^+ -ATPase. *Proc Natl Acad Sci USA* **90**, 8068–8072.

Jensen, P. R., Westerhoff, H. V. & Michelsen, O. (1993b). Excess capacity of H^+ -ATPase and inverse respiratory control in *Escherichia coli*. *EMBO J* **12**, 1277–1282.

Kacser, H. & Burns, J. A. (1973). The control of flux. In *Rate Control of Biological Processes*, pp. 65–104. Edited by D. D. Davies. London: Cambridge University Press.

Kell, D. B. (1987). Forces, fluxes and the control of microbial growth and metabolism. *J Gen Microbiol* **133**, 1651–1665.

Kell, D. B. & Westerhoff, H. V. (1985). Catalytic facilitation and membrane bioenergetics. In *Organized Multienzyme Systems*, pp. 63–138. Edited by G. R. Welch. New York: Academic Press.

- Kell, D. B. & Westerhoff, H. V. (1986). Metabolic control theory: its role in microbiology and biotechnology. *FEMS Microbiol Rev* **39**, 305–320.
- Kholodenko, B. N. & Westerhoff, H. V. (1993). Metabolic channelling and control of the flux. *FEBS Lett* **320**, 75–78.
- Kholodenko, B. N. & Westerhoff, H. V. (1994). Sum of the flux control coefficients: what is it equal to in different systems? In *Modern Trends in Biothermokinetics*, pp. 205–210. Edited by S. Schuster, M. Rigoulet, R. Ouhabi & J.-P. Mazat. New York: Plenum Press.
- Kholodenko, B. N. & Westerhoff, H. V. (1995). The macroworld versus the microworld of biochemical regulation and control. *Trends Biochem Sci* **20**, 52–54.
- Kholodenko, B. N., Molenaar, D., Schuster, S., Heinrich, R. & Westerhoff, H. V. (1995). Defining control coefficients in non-ideal metabolic pathways. *Biophys Chem* (in press).
- Koch, A. L. (1983). The protein burden of *lac* operon products. *J Mol Evol* **19**, 455–462.
- Linton, J. D. (1991). Metabolite production and growth efficiency. *Antonie Leeuwenhoek* **60**, 239–311.
- Marr, A. G. (1991). Growth rate of *Escherichia coli*. *Microbiol Rev* **55**, 316–333.
- Neale, A. D., Scopes, R. K., Kelly, J. M. & Wettenhall, R. E. H. (1986). The two alcohol dehydrogenases of *Zymomonas mobilis*: purification by different dye ligand chromatography, molecular characterization and physiological role. *Eur J Biochem* **154**, 119–124.
- Neale, A. D., Scopes, R. K., Wettenhall, R. E. H. & Hoogenraad, N. J. (1987). Pyruvate decarboxylase of *Zymomonas mobilis*: isolation, properties and genetic expression in *Escherichia coli*. *J Bacteriol* **169**, 1024–1028.
- Niederberger, P., Prasad, R., Miozzari, G. & Kacser, H. (1992). A strategy for increasing an in vivo flux by genetic manipulations. *Biochem J* **287**, 473–479.
- Novick, A. & Weiner, M. (1957). Enzyme induction as an all-or-none phenomenon. *Proc Natl Acad Sci USA* **43**, 553–566.
- Osman, Y. A., Conway, T., Bonetti, S. J. & Ingram, L. O. (1987). Glycolytic flux in *Zymomonas mobilis*: enzyme and metabolite levels during batch fermentation. *J Bacteriol* **169**, 3726–3736.
- Pawluk, A., Scopes, R. K. & Griffiths-Smith, K. (1986). Isolation and properties of the glycolytic enzymes from *Zymomonas mobilis*. *Biochem J* **238**, 275–281.
- Ruijter, G. J. G., Postma, P. W. & Van Dam, K. (1991). Control of glucose metabolism by enzyme II_{Glc} of the phosphoenolpyruvate-dependent phosphotransferase system in *Escherichia coli*. *J Bacteriol* **173**, 6184–6191.
- Schaaff, I., Heinisch, J. & Zimmermann, F. K. (1989). Overproduction of glycolytic enzymes in yeast. *Yeast* **5**, 285–290.
- Schuster, S. & Heinrich, R. (1992). The definition of metabolic control analysis revisited. *Biosystems* **27**, 1–15.
- Schuster, S., Kahn, D. & Westerhoff, H. V. (1993). Modular analysis of the control of complex metabolic pathways. *Biophys Chem* **48**, 1–17.
- Scopes, R. K., Testolin, V., Stoter, A., Griffiths-Smith, K. & Algar, E. M. (1985). Simultaneous purification and characterization of glucokinase, fructokinase and glucose-6-phosphate dehydrogenase from *Zymomonas mobilis*. *Biochem J* **228**, 627–634.
- Seo, J.-H. & Bailey, J. E. (1985). Effects of recombinant plasmid content on growth properties and cloned gene product formation in *Escherichia coli*. *Biotechnol Bioeng* **27**, 1668–1674.
- Small, J. R. & Kacser, H. (1993). Responses of metabolic systems to large changes in enzyme activities and effectors. I. The linear treatment of unbranched chains. *Eur J Biochem* **213**, 613–624.
- Stouthamer, A. H. (1979). A search for correlation between theoretical and experimental growth yields. *Int Rev Biochem* **21**, 1–47.
- Stucki, J. W. (1980). The optimal efficiency and the economic degrees of coupling of oxidative phosphorylation. *Eur J Biochem* **109**, 269–283.
- Tempest, D. W. & Neijssel, O. M. (1984). The status of Y_{ATP} and maintenance energy as biologically interpretable phenomena. *Annu Rev Microbiol* **38**, 459–486.
- Van Dam, K., Van der Vlag, J., Kholodenko, B. N. & Westerhoff, H. V. (1993). The sum of the control coefficients of all enzymes on the flux through a group-transfer pathway can be as high as two. *Eur J Biochem* **212**, 791–797.
- Van der Vlag, J., Van Dam, K. & Postma, P. W. (1994). Quantification of the regulation of glycerol and maltose metabolism by II_{Glc} of the phosphoenolpyruvate-dependent glucose phosphotransferase system in *Salmonella typhimurium*. *J Bacteriol* **176**, 3518–3526.
- Walsh, K. & Koshland, D. E., Jr (1985). Characterization of rate-controlling steps in vivo by use of an adjustable expression vector. *Proc Natl Acad Sci USA* **82**, 3577–3581.
- Westerhoff, H. V. & Van Dam, K. (1987). *Thermodynamics and Control of Biological Free Energy Transduction*. Amsterdam: Elsevier.
- Westerhoff, H. V., Hellingwerf, K. J. & Van Dam, K. (1983). Thermodynamic efficiency of microbial growth is low but optimal for maximal growth rate. *Proc Natl Acad Sci USA* **80**, 305–309.
- Westerhoff, H. V., Koster, J. G., Van Workum, M. & Rudd, K. E. (1990). On the control of gene expression. In *Control of Metabolic Processes*, pp. 399–412. Edited by A. Cornish-Bowden & M. L. Cárdenas. New York: Plenum Press.
- Wilhelm, T., Hoffmann-Klipp, E. & Heinrich, R. (1994). An evolutionary approach to enzyme kinetics; optimization of ordered mechanisms. *Bull Math Biol* **56**, 65–106.
- Yomano, L. P., Scopes, R. K. & Ingram, L. O. (1993). Cloning, sequencing, and expression of the *Zymomonas mobilis* phosphoglycerate mutase gene (*pgm*) in *Escherichia coli*. *J Bacteriol* **175**, 3926–3933.

Received 23 February 1995; revised 5 May 1995; accepted 17 May 1995.