Final Report: DE-FG03-98ER20298

Role of glycolytic intermediate in regulation: Improving lycopene production in *Escherichia coli by* engineering metabolic control

PI: James C. Liao

Department of Chemical Engineering, University of California Los Angeles, Los Angeles, CA 90095

Metabolic engineering1 in the postgenomic era is expected to benefit from a full understanding of the biosynthetic capability of microorganisms as a result of the progress being made in bioinformatics and functional genomics. The immediate advantage of such information is to allow the rational design of novel pathways and the elimination of native reactions that are detrimental or unnecessary for the desired purpose. However, with the ability to manipulate metabolic pathways becoming more effective2, metabolic engineering will need to face a new challenge: the reengineering of the regulatory hierarchy that controls gene expression in those pathways. In addition to constructing the genetic composition of a metabolic pathway, we propose that it will become just as important to consider the dynamics of pathway gene expression.

It has been widely observed that high-level induction of a recombinant protein or pathway leads to growth retardation and reduced metabolic activity3. These phenotypic characteristics result from the fact that the constant demands of production placed upon the cell interfere with its changing requirements for growth. We believe that this common situation in metabolic engineering can be alleviated by designing a dynamic controller that is able to sense the metabolic state of the cell and regulate the expression of the recombinant pathway accordingly. This approach, which is termed metabolic control engineering, involves redesigning the native regulatory circuits and applying them to the recombinant pathway. The general goal of such an effort will be to control the flux to the recombinant pathway adaptively according to the cell's metabolic state. The dynamically controlled recombinant pathway can potentially lead to enhanced production, minimized growth retardation, and reduced toxic by-product formation. The regulation of gene expression in response to the physiological state is also essential to the success of gene therapy. Here we illustrate an application of this approach for the enhanced production of lycopene in *Escherichia coli*.

We chose lycopene as a model compound because of its potential beneficial effects on human health. Lycopene, being an effective antioxidant4, has been proposed as a possible treatment for some cancers and other degenerative human conditions5. As a result, the in vivo synthesis of lycopene and related carotenoids has received increasing attention, and a number of reports have described their production in recombinant microorganisms6–9.

Results and discussion

Engineering a dynamic controller for the recombinant pathway.

The purpose of a dynamic controller is to regulate the flux to the engineered pathway in response to the carbon or energy availability. As shown in Figure 1, designing such an intracellular control loop requires identifying (1) a signaling molecule to reflect the relevant metabolic state, (2) a sensor to monitor this signal, (3) a controller to process the sensory input, (4) a "control valve" (i.e., a promoter) to modulate gene expression, and (5) the rate-limiting steps of thepathway that are to be modulated. Identification of a proper signaling molecule and the rate-limiting steps requires a thorough understanding of the metabolic system. The rest of the components can be derived from the two-component regulatory systems10 that contain a number of genetic regulatory modules. Our initial task was to identify a signal that could reflect the intracellular metabolic state and a sensor that would be able to monitor it. Acetyl phosphate (ACP) serves as a good candidate for such a signaling molecule, because it has been suggested to be an indicator of glucose availability11. In addition, it has already been identified as a regulator12 in various two-component regulons such as Che13, Pho14, and Ntr15. Increasing levels of ACP can potentially be a good indicator of excess glucose flux. Hence, as shown in Figure 1, our strategy was to employ ACP as a signal to regulate the expression of rate-controlling enzymes in the desired pathway, both to utilize fully the excess carbon flux and to redirect the flux away from the toxic product, acetate. To sense ACP and control gene expression accordingly, we employed the Ntr regulon in *E. coli*16 (Fig. 2A), which is responsible for adapting the cell to nitrogen-deficient conditions but plays an insignificant role under most bioreactor conditions because of the nitrogen-rich environment. The sensor in the native Ntr regulon is NRII (*glnL* product)17, which senses the nitrogen status (glutamine and 2-ketoglutarate) of the cell, and transmits the signal to the response regulator, NRI (*glnG* product) through phosphorylation. NRI-P then activates transcription from the *glnAp2* promoter18 by binding to its cognate binding sites on the DNA. NRI itself is capable of sensing the ACP level, although this effect becomes significant only in the absence of NRII15. In the presence of NRII, ACP exerts no apparent control over the activity of *glnAp2*, and NRII must be inactivated in order to make this promoter ACPinducible15. Our strategy for designing an artificial control loop focused on deleting NRII and allowing NRI to both sense ACP and control gene expression (Fig. 2A). In this manner, the *glnAp2* promoter serves as a control valve that controls gene expression according to the ACP level.

To reconstruct this control module, we localized the region containing the NRI binding sites and the core *glnAp2* promoter region, and inserted this DNA fragment into a cloning vector. The NRI-binding sites overlap another promoter, *glnAp1*, which is regulated by cAMP-CRP16. Consequently, our DNA fragment (hereafter known as the *glnAp2* fragment) also contains *glnAp1*, but without the cAMP-CRP binding site (Fig. 2A). Without cAMP-CRP activation, *glnAP1* activity is rather low16, and data obtained with our promoter construct in a wild-type strain (JCL1595) support this observation (Fig. 2B)

Dynamic control of single-copy gene expression.

In order to examine the potential of *glnAp2* as a dynamic controller of product expression, we introduced a *glnAp2-lacZ* transcriptional fusion via □-integration into a *glnL* host strain (JCL1596), and measured the time course of the □-galactosidase (□-gal) activity. This strain contains the *glnL2001* allele15, which consists of an internal deletion between codons 23 and 182 of the *glnL* coding sequence and presumably results in a null mutation. As shown in Figure 2C, the *glnAp2*-□-gal activity increases in a time-dependent fashion similar to the excreted acetate concentration from the *glnL* host, whereas no induction of promoter activity was found for the isogenic wild-type control (JCL1595) (Fig. 2B). Thus, in the absence of NRII, *glnAp2* is capable of responding to the excess carbon flux that is indicated by acetate excretion. As the cells approached the late-exponential phase, the biosynthetic requirement decreased and the cells began to exhibit an excess carbon flux, as demonstrated by the increased generation of acetate. At this point, *glnAp2*-□-gal activity began to rise. In contrast, *glnAp2*-□-gal activity in the wild-type strain (JCL1595) was relatively low and remained constant throughout the culture, suggesting that *glnAp2-lacZ* expression is being regulated by ACP in the *glnL* strain (JCL1596), rather than by cAMP activation of *glnAp1*.

The induction profile of *glnAp2* is in stark contrast to that of the *lac* promoter (*Plac*). As shown in Figure 2D, chromosomal *Plac* activity in strain VJS632 (*lac+*) rapidly increased after induction with IPTG (isopropyl-□-D-thiogalactopyranoside) and achieved a constant level of expression in the cell, which is independent of the growth phase. As a result, this "static" induction was not able to adapt to the cell's changing needs for growth, which, at higher levels of expression, could lead to metabolic imbalance and growth retardation.

Dynamic control of multicopy gene expression.

In order to determine whether the dynamic control provided by *glnAp2* could alleviate the metabolic imbalance and growth retardation due to high-level protein expression, which is often necessary in metabolic engineering, we overexpressed two different metabolic enzymes, phosphoenolpyruvate synthase (Pps) and 3-deoxy-D-arabinoheptulosonate 7-phosphate (DAHP) synthase (AroG), under the control of the *glnAp2* promoter. As controls, these same two proteins also were overexpressed from the *tac* promoter (*Ptac*), which exhibits static control, under the same genetic background and environmental conditions. As reported previously, gratuitous overexpression of proteins3, particularly Pps19, leads to growth retardation. However, when these proteins were expressed under the dynamic control of the *glnAp2* promoter in a *glnL* background, cell growth was not inhibited, and a higher protein expression level was obtained compared to the control of *Ptac* (Fig. 3A, B). This lack of growth inhibition is attributed primarily to the dynamics of the induction profile of the *glnAp2* promoter. As shown in Figure 3C, Pps expression did not begin to increase until the stationary phase. Therefore, the *glnAp2* promoter avoided excessive metabolic burden caused by protein overexpression.

Dynamic control of the lycopene pathway.

Because the dynamic control provided by *glnAp2* could be used to enhance the expression level of specific genes while avoiding growth retardation, we applied this control module to a recombinant pathway for lycopene biosynthesis, which branches out from the isoprenoid pathway in *E. coli*. Isoprenoid biosynthesis in *E. coli* uses pyruvate and glyceraldehydes 3-phosphate (G3P) as precursors20 (Fig. 4A). We reconstructed a recombinant lycopene pathway in *E. coli*6 by expressing the genes *dxs* coding for 1-deoxy-D-xylulose 5-phosphate synthase) from *E. coli*, *gps* (coding for geranylgeranyl diphosphate (GGPP) synthase) from *Archaeoglobus fulgidus*, and *crtBI* (coding for phytoene synthase and desaturase, respectively) from *Erwinia uredovora*21. These genes were inserted into pCL192022, a low-copy-number plasmid, to form pCW9, and simultaneously overexpressed. In this pathway, *gps* and *idi* (encoding isopentenyl diphosphate isomerase) have been identified to control the flux to the final product6. In addition, we have shown that a gluconeogenic enzyme, phosphoenolpyruvate synthase (*pps* gene product), controls the balance between the two precursors, pyruvate and G3P, and thus controls the flux to the isoprenoid pathway. In order to control the carbon flux dynamically, we used the *glnAp2* promoter to control the expression of *idi* and *pps*. The *idi* gene was expressed under *glnAp2* in a promoterless vector derived from pJF11823 to form p2IDI. As a control, *idi* was cloned under the *tac* promoter in plasmid pJF118 to form pTacIDI. These plasmids were separately introduced into a *glnL* strain (BW18302) containing pCW9. As shown in Figure 5A, the p2IDI-containing strain (*glnAp2-idi*) produced 100 mg L-1 lycopene after 26 h in a defined medium containing glucose. The strain containing P*tac-idi* (pTacIDI), on the other hand, produced only a small amount of lycopene under this same condition. Additionally, the p2IDI strain produced almost threefold less acetate than pTacIDI (Fig. 5B), which indicates that the carbon flux to acetate was being rechanneled to lycopene as expected. The pTacIDI strain itself produced a similar fermentation pattern as the BW18302 host control, which suggests that expression of the lycopene pathway in this strain did not alter the metabolic flux distribution significantly (Fig. 5B, C). These results, which are summarized in Table 1, illustrate a directed change in the physiological regulation that allows the cell to utilize the excess carbon flux for product formation without jeopardizing cell growth (Fig. 5D). Pyruvate excretion from the p2IDI strain, however, was still high, even though it was less than that for pTacIDI through most of the time course (Fig. 5C). Since pyruvate is one of the common precursors of the isoprenoid pathway, the excretion of this metabolite from the cell indicates that carbon is still not being efficiently diverted to lycopene. As demonstrated previously, overexpression of the gluconeogenic enzyme Pps (Fig. 4B) allows the recycling of pyruvate24,25 and leads to a higher lycopene flux (data not shown). However, overexpression of Pps also causes severe growth inhibition under glycolytic conditions (Fig. 2B)19. Thus, to avoid metabolic imbalance, Pps was overexpressed from *glnAp2* from another compatible plasmid, pPSG184. The final lycopene-producing strain then contained an artificial regulon controlling *idi* and *pps* (*glnAp2-idi* + *glnAp2-pps*), and the remainder of the lycopene pathway (*dxs*, *gps*, *crtBI*) under P*lac*. Coexpression of Pps with the lycopene pathway increased the final titer of lycopene by 50% and caused the productivity to increase by threefold, from 0.05 mg mL-1 h-1 to 0.16 mg mL-1 h-1 (Fig. 5A). This is in contrast to the companion strain containing both pTacIDI and pPS184 (P*tac-idi* + P*tac-pps*), where no significant

improvement was observed with the introduction of *pps* and substantial growth inhibition occurred (Fig. 5D). Apparently, coexpression of Pps in the *glnAp2* regulon makes conditions favorable for lycopene production by increasing the utilization of pyruvate (Fig. 5C and Table 1).

These results show that the dynamic control provided by the *glnAp2*-based artificial regulon allows the expression of the ratecontrolling enzymes (Idi and Pps) for lycopene biosynthesis to be coordinated with the metabolic state. Furthermore, these results suggest a useful strategy for metabolic engineering. Rather than trying to delete the regulatory circuits as suggested in the past26–28, or trying to overwhelm them by overexpressing key enzymes, we believe that, in some cases, metabolic imbalance can be dealt with by rerouting the control in a rational and constructive manner. This approach can be used not only to optimize metabolic pathways, as demonstrated here, but also to modify other cellular activities that are important to biotechnology, such as cell cycle regulation29, protein translocation, and high-level protein production. In addition, it may become an important strategy for gene therapy in controlling gene expression according to the physiological state. In the context of metabolic engineering, however, the dual application of pathway amplification followed by rational design of control circuits could become an effective way of designing microorganisms for industrial applications.

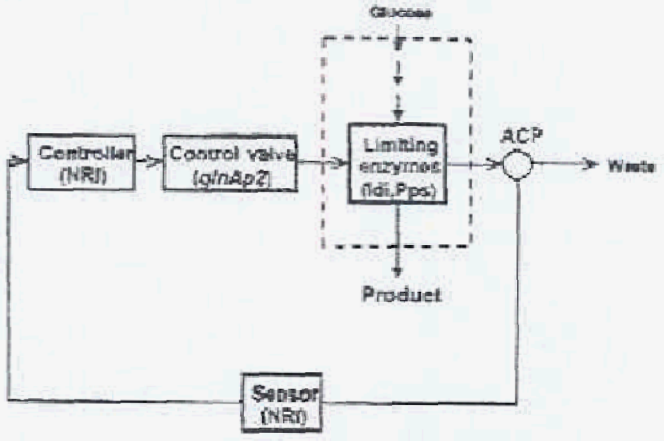

Figure 1. Design strategy of the dynamic controller represented using the block diagram of control engineering. The process to be controlled is the metabolic pathway from glucose to a desired (lycopene) and some waste products (e.g. acetate). The measured variable (or signal) is ACP, which serves as a signal for excess flux to the waste product. The sensor is the NRI protein, which also serves as a controller by binding to the DNA and modulating transcription. The control valve is the *glnAp2* promoter, which controls limiting steps (Idi and Pps) in the metabolic system. The dashed box indicates the metabolic system to be controlled.

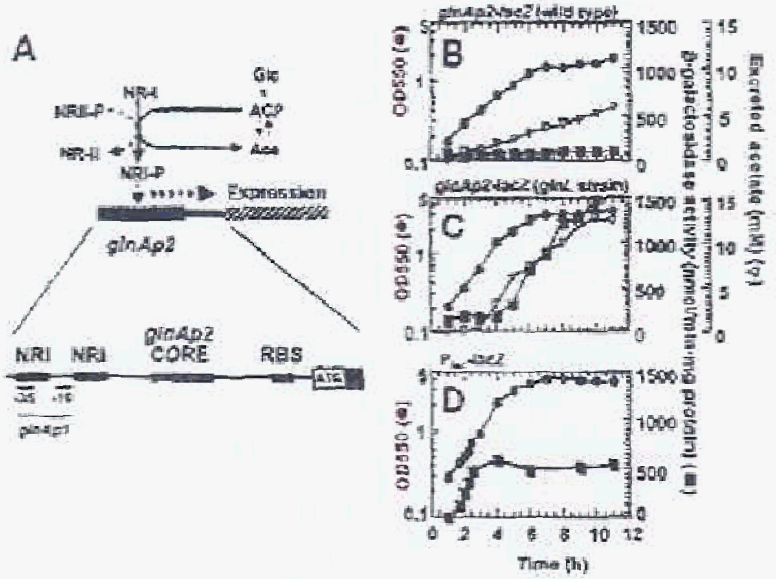

Figure 2. Characterization of the dynamic controller. (A) The *glnAp2* promoter. Inducibility of *glnAp2* by ACP only in the absence of NRII (*glnL* gene product). The architecture of the promoter construct used in this study is given at the bottom. NRI, NRI-binding site; *glnAp2* CORE, *glnAp2* core sequence; RBS, *glnA* ribosome-binding site; ATG, translational start. The regions labeled "-35" and "-10" indicate the location of the two RNA polymerase contact sites of the upstream *glnAp1* promoter. (B, C) β-Gal activity from the single-copy *glnAp2-lacZ* construct (B) cell density (in terms of optical density at 550 nm), and (C) acetate secretion in strains JCL1595 (*glnL*-) and JCL 1596 (*glnL*), respectively. (D) β-Gal activity from a wild-type single-copy P_{tet}-*lacZ* and the growth kinetics of strain VJS632 during aerobic growth in glucose. Symbols represent cell growth (●), β-gal activity (■), and extracellular acetate (▽).

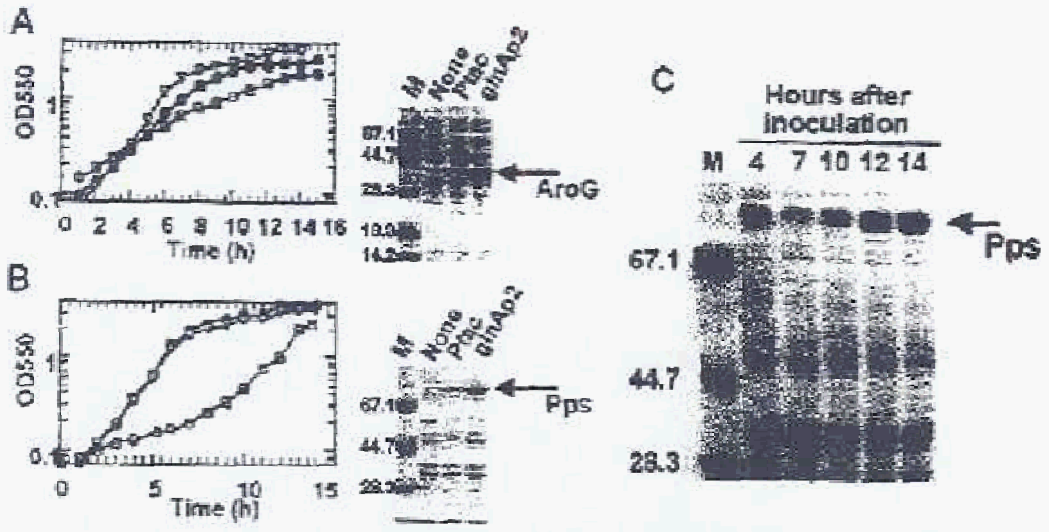

Figure 3. Protein overexpression from multicopy plasmids with and without the dynamic controller. The dynamic *glnAp2* promoter achieved a higher expression level and reduced growth retardation compared to the static IPTG-inducible tac promoter in M9 salts-glucose medium. (A) Overexpression of Pps from plasmids pS706 (P_{tac}-*pps*) (○) and p2AROG3 (*glnAp2-aroG*) (●). The host strain is BW18302, which was separately transformed with pAROG, p2AROG3, or no plasmid (▽). At 15 h after inoculation, the protein content from each strain was analyzed by SDS–PAGE using a 10% gel (right panel). Lane M, molecular weight markers with protein band sizes indicated on the gel. Lane labeled "None" contains BW18302 with no plasmid. Lane labeled "P_{tac}" contains BW18302/pAROG cell extract, and lane labeled "*glnAp2*" contains BW18302/p2AROG3 cell extract. (B) Overexpression of Pps from plasmids pPS706 (P_{tac}-*pps*) (○) and pPSG706 (*glnAp2-pps*) (●) in BW18302, compared with no plasmid control (▽). At 15 h after inoculation, protein content was analyzed by SDS–PAGE using an 8% gel (right panel). Lane M contains molecular weight marker. Lane labeled "None", contains BW18302 with no plasmid. Lane labeled "P_{tac}" BW18302/pPS706 contains cell extract. Lane labeled "*glnAp2*" contains BW18302/pPSG706 cell extract. (C) Time course of Pps expression from BW18302/pPSG706 (*glnAp2-pps*).

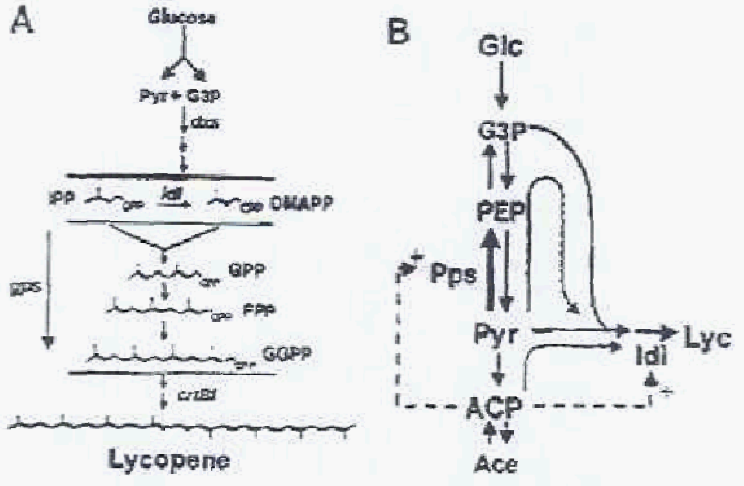

Figure 4. Metabolic control engineering of lycopene production. (A) Reconstructed lycopene pathway with the genes *dxs*, *idi*, *gps*, and *crtBI*. Pyr, pyruvate; G3P, glyceraldehyde 3-phosphate; IPP, isopentenyl diphosphate; DMAPP, dimethylallyl diphosphate; GPP, geranyl diphosphate; FPP, farnesyl diphosphate; GGPP, geranylgeranyl diphosphate. (B) Strategy to divert metabolic flux to the lycopene pathway by the dynamic control of IDI and PPS under *glnAp2*. The two types of dotted line represent the control circuits applied to *idi* and *pps* (_ _ _) and the rerouting of carbon flux to the lycopene pathway (.....). Glc, glucose; PEP, phosphoenolpyruvate; ACP, acetyl phosphate; Ace, acetate; Lyc, lycopene.

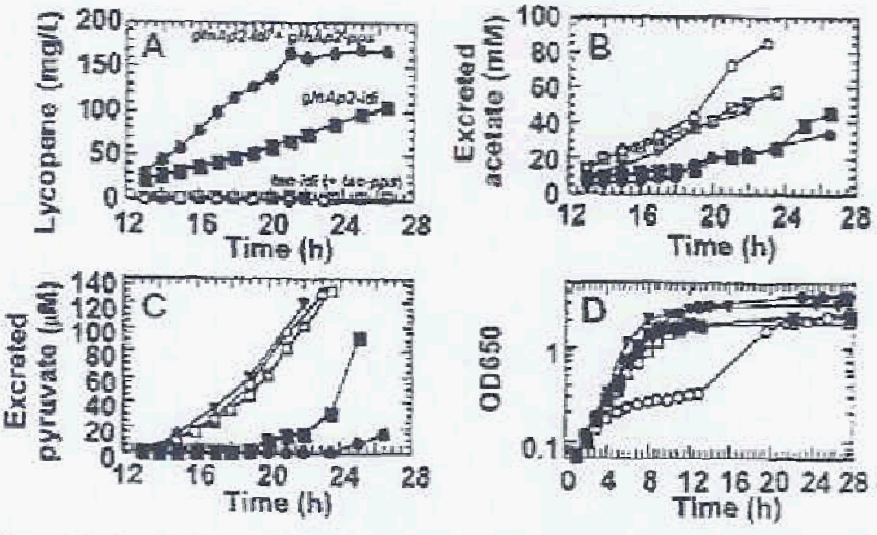

Figure 5. Dynamic control by ACP and the *glnAp2* loop enhances lycopene production from *E. coli*. (A) Production of lycopene in YE salts defined medium containing 1.5% (wt/vol) glucose. (B) Excretion of acetate. (C) Excretion of pyruvate. (D) Growth kinetics. Symbols are as follows: BW18302 host control (▼), and strains containing the lycopene pathway genes *dxs*, *gps*, and *crtBI* (expressed from pCW9) along with pTacIDI (P_{tac}-*idi*) (□), p2IDI (*glnAp2-idi*) (■), pTacIDI (P_{tac}-*idi*) + pPS184 (P_{tac}-*pps*) (○), or p2IDI + pPSG184 (*glnAp2-idi* + *glnAp2-pps*) (●).

Table 1. Carbon yield of end-product formation in batch cultures of lycopene-producing *Escherichia coli*[a]

	Carbon yield on glucose (mol C/mol C)				
Strain	Lycopene	Acetate	Pyruvate	Formate	Succinate
Host only (BW16302)	0	0.26	0.0004	0.0077	0
+pTacIDI (P_{tac}-*idi*)	0.0003	0.23	0.0005	0.018	0
+pTacIDI/pPS1S4 (P_{tac}-*idi* + P_{bc}-*DXS*)	0.0012	0.36	0.0011	0.019	0
+p2IDI (*glnAp2-idi*)	0.014	0.16	0.0012	0.040	0.011
+p2IDI/pPSG184 (*glnAp2-idi* +*glnAp2-pcs*)	0.022	0.14	0.0006	0.015	0

[a]This table is a summary of end-product formation at the point of glucose exhaustion during growth in YE defined salts medium containing 1.5% (w/v) glucose. Corresponding time courses for this experiment are found in Figure 5.