bioRxiv preprint doi: https://doi.org/10.1101/803023; this version posted October 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Phenotypic design choices for enhanced productivity in microbial chemical production processes

Kaushik Raj[1,*], Naveen Venayak[1,*], and Radhakrishnan Mahadevan[1,2,**]

[1]Department of Chemical Engineering and Applied Chemistry, University of Toronto, Canada
[2]Institute of Biomaterials and Bioengineering, University of Toronto, Canada
[*]These authors contributed equally
[**]Corresponding author, krishna.mahadevan@utoronto.ca

Abstract

Microbial metabolism can be harnessed to produce a broad range of industrially important chemicals. Often, three key process variables: **T**iter, **R**ate and **Y**ield (**TRY**) are the target of metabolic engineering efforts to improve microbial hosts toward industrial production. Previous research into improving the TRY metrics have examined the efficacy of having distinct growth and production stages to achieve enhanced productivity. However, these studies assumed a switch from a maximum growth to a maximum production phenotype. Hence, the choice of operating points for the growth and production stages of two-stage processes is yet to be explored. The impact of reduced growth rates on substrate uptake adds to the need for intelligent choice of operating points while designing two-stage processes. In this work, we present a computational framework that scans the phenotypic space of microbial metabolism to identify ideal growth and production phenotypic targets, to achieve optimal TRY values. Using this framework, with *Escherichia coli* as a model organism, we compare two-stage processes that use dynamic pathway regulation, with one-stage processes that use static intervention strategies. Our results indicate that two-stage processes with intermediate growth during the production stage always result in the highest productivity. By analyzing the flux distributions for the production enhancing strategies, we identify key reactions and reaction subsystems that need to be downregulated for a wide range of metabolites in *E. coli*. We also elucidate the importance of flux perturbations that increase phosphoenolpyruvate and NADPH availability among strategies to design production platforms. Furthermore, reactions in the pentose phosphate pathway emerge as key control nodes that function together to increase the availability of precursors to most products in *E. coli*. Due to the presence of these common patterns in the flux perturbations, we propose the possibility of a universal production strain that enhances the production of a large number of metabolites.

Keywords: dynamic pathway engineering, two-stage processes, industrial bioprocesses, phenotypic choices, production platforms, substrate uptake effects

bioRxiv preprint doi: https://doi.org/10.1101/803023; this version posted October 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

1 Introduction

The use of microbes for the production of chemicals through metabolic engineering has garnered significant interest in the past few decades. The naturally modular arrangement of metabolic networks makes microbial strains amenable to be used as chemical production platforms[1]. Metabolic networks have a bow-tie architecture which allows a large number of metabolites to be produced from a few universal precursors[2]. This has allowed us to successfully engineer microbes to be biocatalysts for the production of a wide range of commodity chemicals[3,4], pharmaceuticals[5,6], biofuels,[7,8] and other natural and non-natural compounds[9]. While few such processes have been successful at an industrial scale[10,11], large strain development costs and scale-up issues could deem many processes economically infeasible[12,13]. Given the cost of a target feedstock and product, the feasibility of industrial fermentation processes is typically determined by three process metrics - **T**iter: concentration of product at the end of a fermentation batch (given in $mmol/L$ of product), **R**ate/productivity: the rate of product secretion (given in $mmol/L.h$ of product), and **Y**ield: the amount of product produced per unit amount of substrate (given in $mmol$ product/$mmol$ substrate) - collectively termed the **TRY** metrics[14]. Titer and yield affect the operating expenditure of the process by impacting product separation and substrate costs respectively, while productivity affects the capital expenditure by determining the scale of the reactor required. Microbial production processes undergo several rounds of strain, pathway and process optimization to reach acceptable TRY targets[15,16].

Wild-type microbial strains have evolved to grow at maximal rates, directing little carbon flux towards production of target compounds[17]. Metabolic engineering attempts to change the operating point (or phenotype) of a strain to enhance target chemical production by throttling growth associated fluxes and/or tuning native metabolism to balance pathway energy and cofactor requirements. Given a stoichiometric model and substrate/nutrient uptake rates, the feasible range of chemical production in a microbial strain can be visualized using its production envelope which maps the maximum product flux at all possible growth rates of the microbe (Figure 1). Strain engineering strategies to improve TRY metrics can be broadly classified into static and dynamic pathway engineering strategies. Static pathway engineering involves making gene deletions that either couple the production of a target compound with the microorganism's growth[18] or, simply redirect more carbon flux through production pathways. These strategies are typically implemented as one-stage (OS) production processes where the strain remains at a single operating point throughout the course of the process (Figure 1a). Such processes result in a higher yield by ensuring high relative pathway flux. Recently, there has been an increased interest in dynamic pathway engineering, which involves temporally controlling carbon flux through growth and production pathways. This can be achieved through the use of biological logic or sensor and actuator systems composed of cellular components[19–21]. Such strategies are implemented as two-stage (TS) production processes which start with cells in their growth stage and at some point during the fermentation, production pathway genes are expressed to switch to the production stage (Figure 1b). Such a decoupling of growth and production stages is thought to reduce batch times by reaching maximal biomass concentrations faster and thereby increase productivity[22,23].

While stoichiometric models are effective for determining relative production metrics such as yield, absolute metrics such as end-titer and productivity are also governed by variations in substrate uptake rates. However, metabolic models with constant substrate uptake rates are routinely used for simulations to monitor metabolite production rates at different phases of metabolism. The impact of reduced substrate uptake rate during stationary phase metabolism[24] is often overlooked while designing microbial production

bioRxiv preprint doi: https://doi.org/10.1101/803023; this version posted October 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

processes. Studies have shown that the rate of glucose uptake varies by significant amounts depending on the genotype of the strain and phase of metabolism [24–30] (Supplementary Figure S1a). A drop in substrate uptake rate during the production stage of a TS process would result in reduced product flux and therefore, a lower productivity, defeating the purpose of such a process (dotted lines in Figure 1b). This effect was recently shown in a theoretical study that compared the performance of TS and OS processes for D-lactic acid production in *E. coli* [31]. This study showed that reduced substrate uptake rates can limit the use advantages of a TS process and there is a very narrow range of conditions where a TS process can outperform an OS process. They call for the use of methods that increase stationary phase substrate uptake, such as engineering ATP futile cycles to expand the range of conditions in which TS processes offer enhanced productivity. This study and many others consider only TS processes that switch from wild-type growth to a non-growing production phenotype during the stationary phase of metabolism have been studied. However, given the interplay between substrate uptake and growth rates, phenotypes with intermediate growth could hold significant value. Intermediate phenotypes have been examined in the past to identify operating points that result in balanced TRY values in OS processes [32].

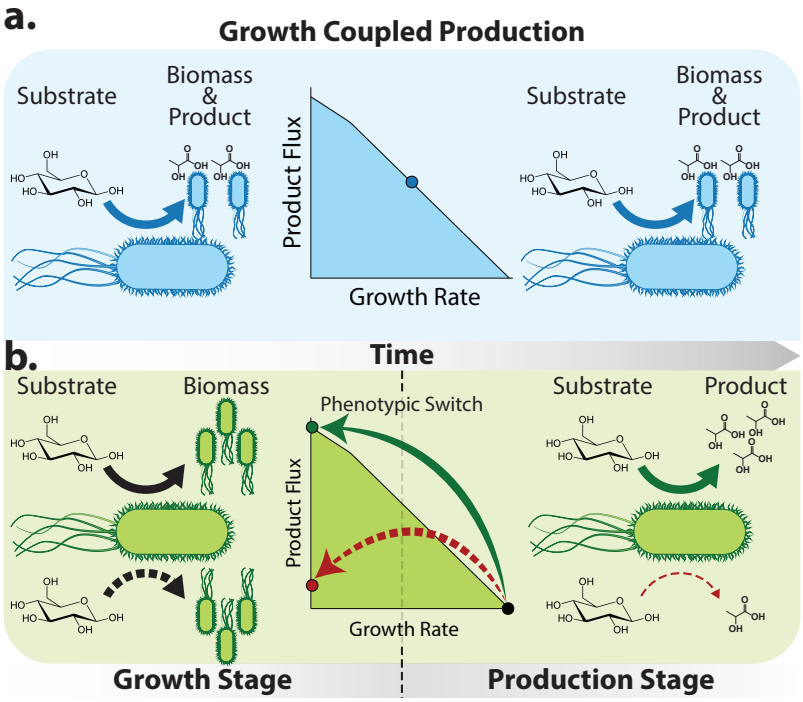

Figure 1: Static vs dynamic pathway engineering: Strain engineering strategies can be classified into **a.** Static engineering strategies where genetic perturbations that allow cells to grow and produce the target compound simultaneously (growth coupled production) are implemented. This enables the cells to produce the compound in a one-stage process. **b.** Dynamic engineering strategies where growth and production pathways are decoupled temporally. In such strategies, the process starts with a growth stage, accumulating biomass and switches over to a production stage to produce the target compound. Reduced substrate uptake during the production stage can result in lower product flux (dotted lines) than that expected assuming constant substrate uptake rates (solid lines). Hypothetical operating points for each production strategy are shown in the respective production envelopes.

In this work, we compare TS and OS production processes that make use of the entire range of feasible production operating points rather than those with maximum growth or maximum product yield. To this

bioRxiv preprint doi: https://doi.org/10.1101/803023; this version posted October 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

end, we develop Two Step Dynamic Strain Scanning and Optimization (TS-DySScO), a modular computational framework to compare microbial production processes. TS-DySScO uses TS dynamic flux balance analysis to determine the process metrics obtainable using hypothetical operating points calculated within the solution space of a microbe's metabolic model. With this information, the framework can determine the best process type and phenotypic choices that result in the maximum value of a predetermined objective. We use TS-DySScO to discover enhanced TS processes that result in high productivity while considering the substrate uptake effects of reduced growth. We also identify flux perturbations that occur consistently in production strategies for all natural products, giving rise to the possibility of a universal chassis for metabolic engineering.

2 Materials & Methods

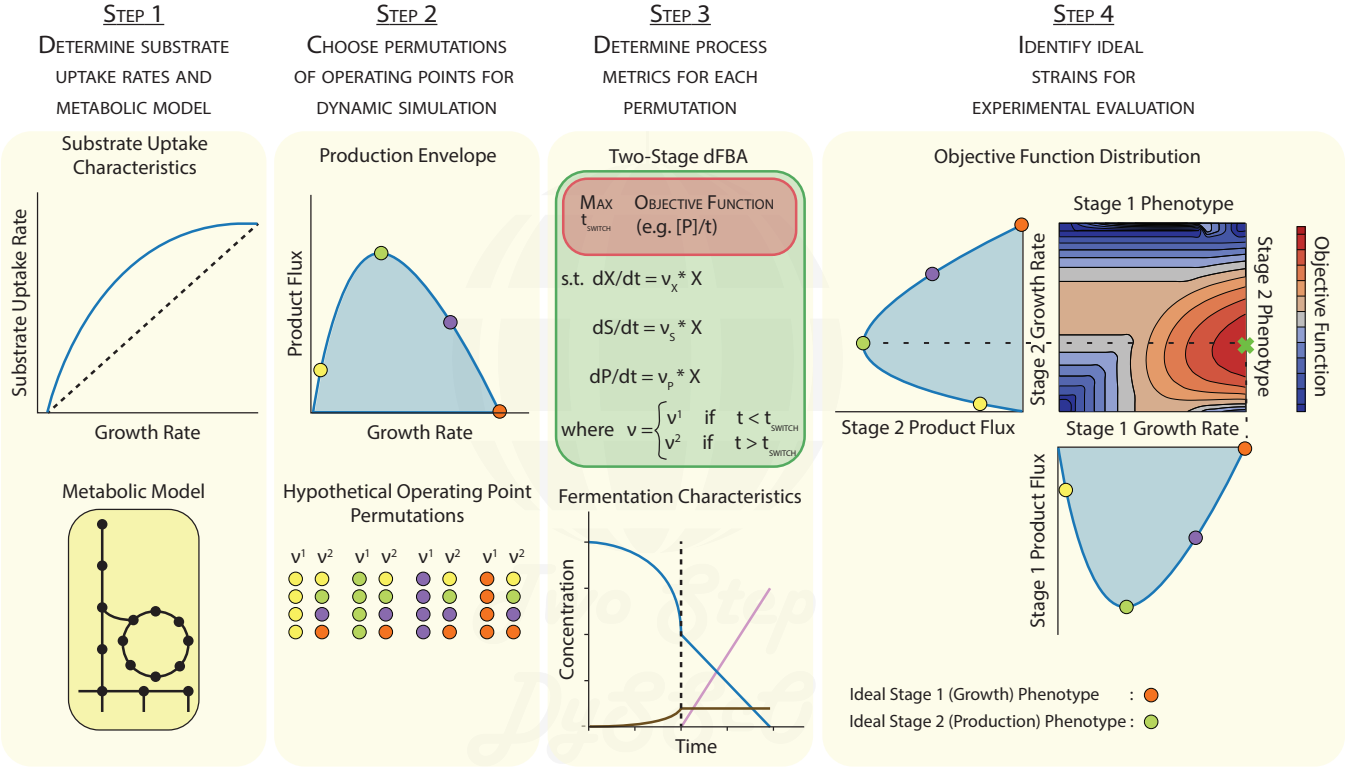

Figure 2: TS-DySScO workflow: **Step 1** - A metabolic reconstruction of the microorganism and an approximation of the substrate uptake characteristics at different growth rates are given as inputs to the formulation. Currently, TS-DySScO accepts all COBRA compatible metabolic models and any mathematical function to model substrate uptake variations. **Step 2** - A realistic production envelope is generated for the product of interest. This entails maximizing the product flux at all possible growth rates of the organism. A user defined number of operating points are chosen on the production envelope and all possible permutations of these operating points are determined for two-stage analysis. **Step 3** - A two-stage dynamic flux balance analysis that maximizes a user-defined objective function is conducted to determine the process metrics and fermentation characteristics for each permutation of operating points. **Step 4** - A distribution of process metrics for all two-stage processes is plotted and the optimal production strain is identified. The optimal phenotype for each stage has been projected onto the respective production envelope.

bioRxiv preprint doi: https://doi.org/10.1101/803023; this version posted October 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

2.1 TS-DySScO formulation

The TS-DySScO workflow formulated in this study is briefly summarized in Figure 2 and described below.

2.1.1 Establishing substrate uptake characteristics

The first step towards thoroughly analyzing the performance of the production strategies using a metabolic model is establishing a relationship between the growth rate and the rate of substrate uptake in the organism being studied. There are several studies that have attempted to elucidate the relationship between substrate uptake and growth rates in *E. coli*. Many of these have examined this relationship using a chemostat under glucose limiting conditions[33–36]. Under these conditions, the rate of glucose uptake is limited by the dilution rate prevailing in the reactor and not by the effects of genetic perturbations in the cells. Therefore, for our analysis we only consider studies with batch fermentations under glucose excess conditions[24–30].

$$q_s = q_{s,min} + (q_{s,max} - q_{s,min}) * \left(-1 + \frac{2}{(1 + e^{-K_{uptake} * \mu})}\right) \quad \text{(Eq. 1)}$$

where

$q_{s,min}$: substrate uptake rate at zero-growth rate

$q_{s,max}$: substrate uptake rate at maximum growth rate

K_{uptake} : substrate uptake parameter

We chose to use a logistic curve that spans between the minimum recorded glucose uptake rate during stationary phase and the maximum glucose uptake rate used in the metabolic model, to estimate the substrate uptake rates at various growth rates (Eq. 1). We have presented a more detailed analysis of the relationship between substrate uptake rate and growth rate in the supplementary information (Supplementary text 1.1, Figure S1b).

2.1.2 Defining hypothetical operating points

The next step is to determine all hypothetical operating points of interest using a metabolic model. In this study, we refer to a metabolic mode of an organism, represented by a unique combination of the possible growth and product secretion rates within the solution space of its metabolic model, as an operating point or 'phenotype'. For a user-defined number of points between the minimum and maximum growth rates of the organism as determined from the metabolic model, we derive the corresponding substrate uptake rates using the relationship established in the previous step. At each of these points, by constraining the growth and substrate uptake reactions to the required values and maximizing the secretion rate of the metabolite of interest, we obtain the hypothetical operating points of interest for examination.

2.1.3 Two-stage dynamic flux balance analysis

Dynamic flux balance analysis (dFBA)[37] can be used to obtain process metrics for a fermentation process by keeping track of substrate, biomass and product concentrations over the course of a fermentation batch, provided that initial concentrations of these species are known. This is done by using ordinary differential equations to simulate changes in the concentration of relevant species using their fluxes obtained from a metabolic model. Here, we modify the dFBA formulation to allow for phenotype switching between the

bioRxiv preprint doi: https://doi.org/10.1101/803023; this version posted October 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

two stages of a TS fermentation process (Eq. 2b-d). We do this by using distinct of biomass, substrate and product fluxes in the dFBA equations during the growth and production stages. By repeating this process with all possible permutations of the hypothetical operating points as growth and production phenotypes, we obtain process metrics for all possible TS processes.

$$\max_{t_{switch}} \; f(\mathbf{TRY}) \tag{Eq. 2a}$$

such that

$$\frac{d[X]}{dt} = \nu_{n,X} * [X] \tag{Eq. 2b}$$

$$\frac{d[S]}{dt} = \nu_{n,S} * [X] \tag{Eq. 2c}$$

$$\frac{d[P]}{dt} = \nu_{n,P} * [X] \tag{Eq. 2d}$$

where

$$n = \begin{cases} 1 & \text{if } t \leq t_{switch} \\ 2 & \text{if } t > t_{switch} \end{cases}$$

2.1.4 Optimizing metabolite production strategy

The time of switching between the two stages is an additional variable that affects the concentration of the species in the dFBA formulation and therefore, the process metrics that determine the performance of each TS process. In order to obtain the best fermentation metric for each combination of operating points, we formulate an optimization problem that varies the switching time (t_{switch}) to maximize a desired objective function (Eq. 2a). This can be the titer, rate, yield or a function of these metrics. We can then plot a distribution of the process metrics for each permutation to identify the optimal combination of operating points. Following this, the desirable operating point combinations can be chosen for experimental validation. For this study, we used the productivity of a fermentation process as the objective function.

2.1.5 Packaging and availability

The TS-DySScO framework is written as a python package that accepts COBRApy [38] compatible metabolic models. The modular nature of this package allows users to select the metabolic model, fermentation start parameters and, substrate uptake characteristics with ease. In order to reduce run times, we allowed for the optimization and dFBA calculations shown in Eq. 2 to be run parallelly on multi-core and multi-processor systems. The switch time optimization problem was implemented using the COBYLA method in the optimization package of scipy. The TS-DySScO framework can be installed and run on any system with a working Python 3 distribution. The framework, along with installation instructions are available on GitHub (`https://github.com/lmse/tsdyssco`). The code used to perform the various analyses described in this article are available at `https://github.com/lmse/phenotypic_design`.

bioRxiv preprint doi: https://doi.org/10.1101/803023; this version posted October 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

2.2 Implementation

All analyses were conducted using the COBRApy[38] and Cameo[39] packages on a Python 3.7 distribution. *E. coli*'s genome scale metabolic reconstruction - iJO1366[40] was used to perform all simulations to compare the two fermentation strategies. Unless otherwise specified, fermentation batches were started with 500 mM ($\approx 90\ g/L$) of D-glucose as the substrate and 0.05 g/L of biomass. These values are in the range of required substrate and biomass concentrations to achieve acceptable TRY targets[15]. Flux distributions for best performing phenotypes in each fermentation strategy were obtained by constraining growth, substrate uptake and product secreting reactions to required values and performing a parsimonious flux balance analysis (pFBA)[41] on the metabolic model using the IBM ILOG CPLEX (v12.9) solver. Data visualization was performed using the plotly package.

3 Results & Discussion

3.1 Case Study: Production of D-Lactic Acid in *E. coli*

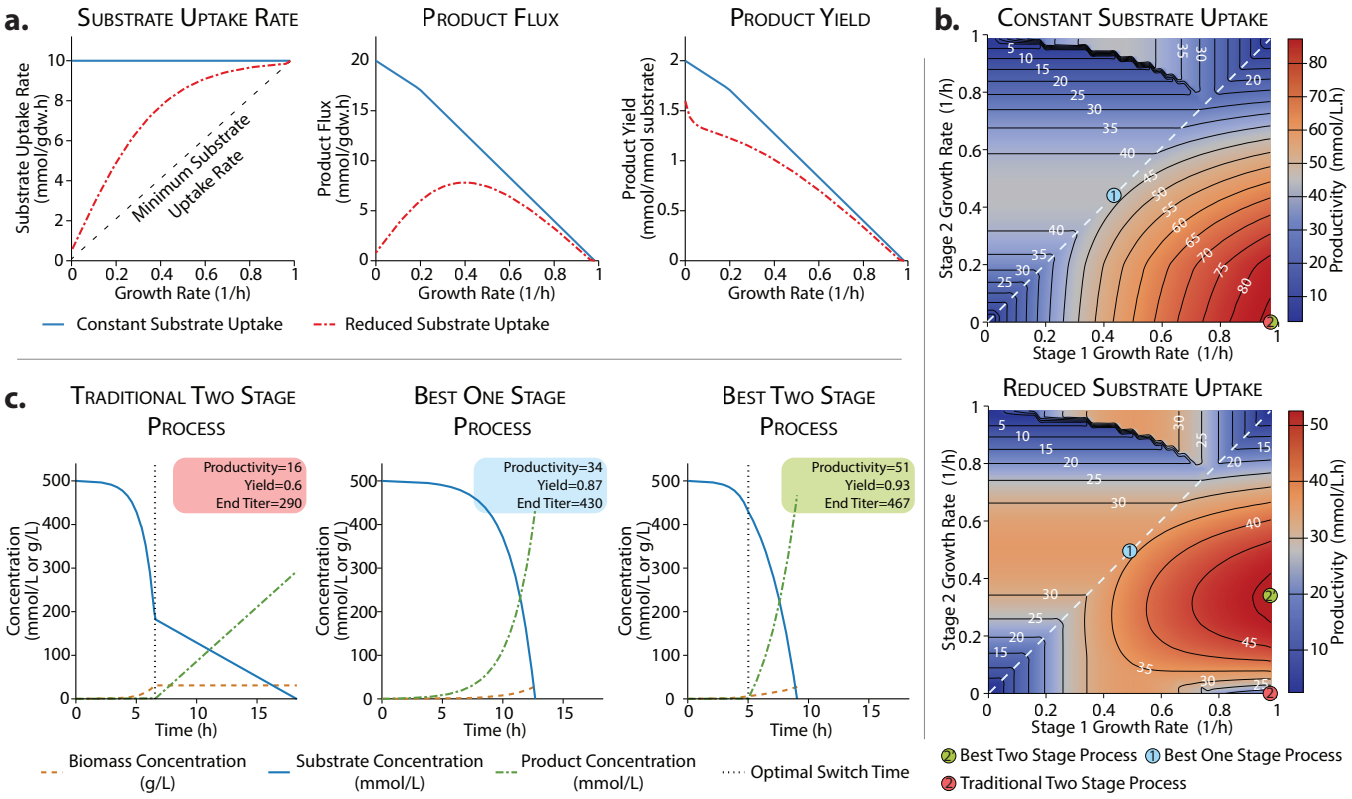

Figure 3: TS-Dyssco implemented for D-lactic acid production in *E. coli*: **a.** substrate uptake rates, product flux and product yields obtained using the iJO1366 reconstruction of *E. coli* assuming constant and reduced substrate uptake rates. **b.** Productivity distribution for TS and OS processes in *E. coli* assuming constant and reduced substrate uptake rates. Isoclines on the distributions show phenotypes with the same productivity levels. **c.** Fermentation profile for various production strategies assuming reduced substrate uptake rates.

We applied the newly formulated TS-DySScO framework to predict strategies that maximize the pro-

bioRxiv preprint doi: https://doi.org/10.1101/803023; this version posted October 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

ductivity of D-lactic acid production in *Escherichia coli*, starting with 500 mM of glucose as substrate and 0.05 g/L biomass. First, we examined the process under constant substrate uptake conditions i.e. assuming that the substrate uptake is unaffected by growth and other metabolic perturbations (Figure 3a). In these conditions, since the product flux and yield are highest when there is no growth, the best TS process for productivity is the traditional TS process, where the strain is allowed to switch from a maximum growth (wild-type) to a zero-growth phenotype (Figure 3b). Industrial fed-batch processes, where production pathway genes are induced upon reaching stationary phase and substrate is added slowly, fall under this category. As expected, the traditional TS process has a much higher productivity than the best OS process.

However, an impediment in growth rate either due to reaching stationary phase or rewiring of metabolism has been shown to alter substrate uptake rates[24–30]. If these effects are considered, the product flux in a non-growing strain is heavily impacted (shown as red dashed lines in Figure 3a). This makes a non-growing phenotype during the second stage of a TS process ineffective. We can observe this in Figures 3b,c where the traditional TS process has very low productivity - among the lowest of any possible process. As observed in a previous study[31], many OS processes (shown as dashed lines in Figure 3b) have a higher productivity than the traditional TS process. However, a fair evaluation of two-stage strategies should include the entire available phenotypic space. Even under reduced substrate uptake conditions, there are several TS processes that have a higher productivity than the best performing OS process. These processes can be achieved by allowing the strain to grow at a reduced rate during the production stage, rather than completely eliminating growth. The switch time optimization formulation results in earlier switching between the phenotypes when the strain is allowed to grow during the production stage (Figure 3c and Supplementary Figure S2). The TRY metrics attained in the best TS process assuming reduced substrate uptake is higher than that of the OS process with the highest productivity (Figure 3b and Supplementary Figure S3). The best TS process predicted requires the strain to be able to dynamically switch from wild-type growth to a phenotype with intermediate growth and production (growth coupled production). However a process with the highest productivity may not be the most economically optimal choice due to variations in substrate and product prices. The process metric distributions can be collectively used to choose phenotypes for TS processes depending on substrate and product prices prevailing at a particular time.

3.2 TS processes have highest productivity for all natural metabolites in *E. coli*

Having established that TS processes result in a higher D-lactate productivity, we wished to examine if this trend held true for other native metabolites in *E. coli*. We anticipated that the different production flux profiles for each product would result in variations in the process metrics. Hence, we used TS-DySScO with the fermentation start parameters previously described, to predict process optimality for 70 native exchange metabolites in the iJO1366 reconstruction of *E. coli*'s metabolism. The best TS process has the highest productivity for all products analyzed, with the OS process and traditional TS process trailing behind (Figure 4). Surprisingly, the best TS process outperforms OS processes even if substrate uptake rates are assumed to be constant for the OS process.

In general, products with more carbon atoms have a lower molar productivity. However, two products, namely 5-methylthioribose and spermidine have unusually low productivities, which will be examined in

bioRxiv preprint doi: https://doi.org/10.1101/803023; this version posted October 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

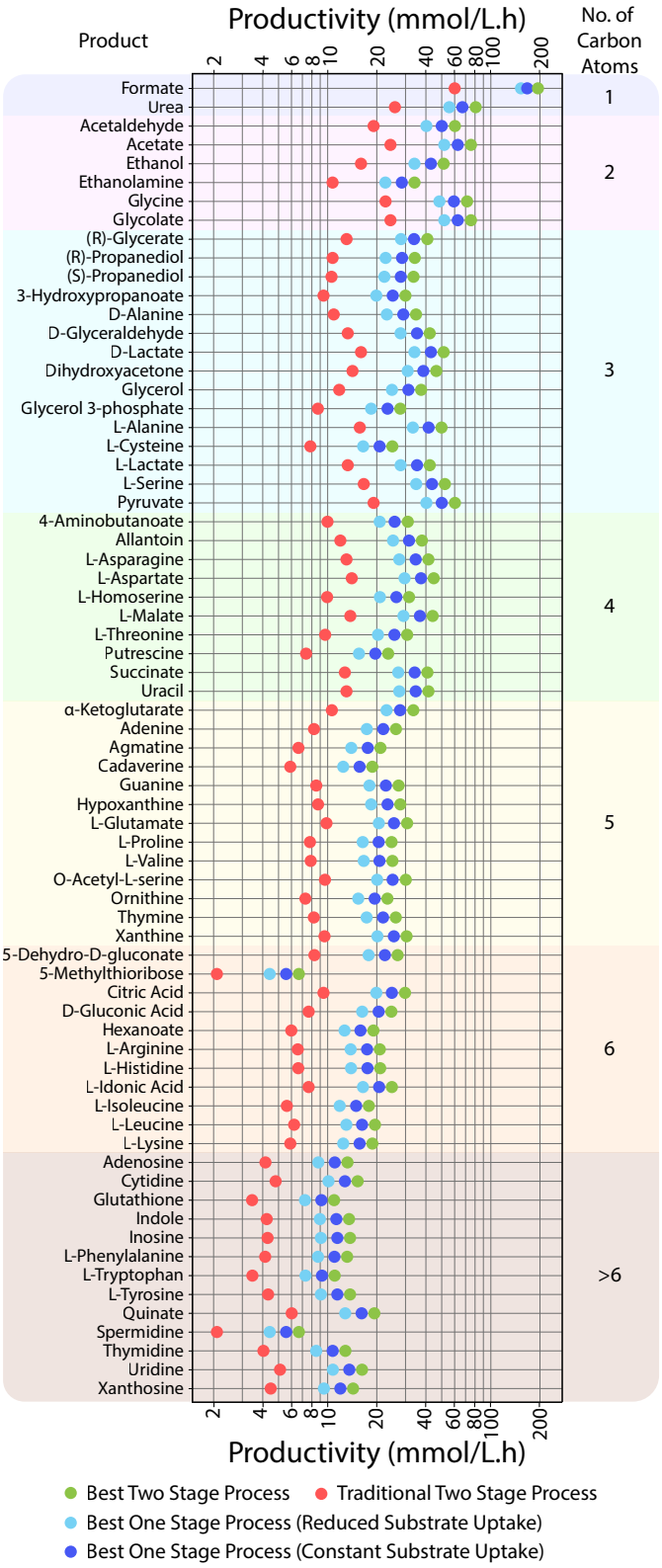

Figure 4: Productivity for exchange metabolite production in *E. coli*: TS-DySScO was used to simulate the production of all exchange metabolites in the iJO1366 reconstruction of *E. coli*, ordered by the number of carbon atoms in the product. The optimized two stage process always results in the highest productivity.

bioRxiv preprint doi: https://doi.org/10.1101/803023; this version posted October 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

later sections. The best TS processes also have a higher yield and end-titer than their OS counterparts (Supplementary Figure S6). Only if the substrate uptake effects of growth rate are neglected, OS processes have a higher yield and end-titer than the best TS processes. The traditional TS processes have very stunted productivities, yields and end-titers for all the exchange metabolites when substrate uptake rates is reduced, consistent with the previous study for D-lactic acid[31].

Upon examining the best TS processes predicted by TS-DySScO, we found that all of them had the same phenotype during the growth stage - wild-type growth. These processes varied in phenotype only during the production stage, where an intermediate growth phenotype resulted in the highest productivity. Similarly, OS processes with the highest productivity are those with an intermediate growth rate and have one growth-coupled production stage. Hence, it is possible to compare the two process types at every operating point in the production envelope during the production stage for all possible substrate uptake rates. This allows us to determine the threshold substrate uptake rate that a strain would need to achieve in order to make the TS process better than the OS process. Since the logistic curve used to model substrate uptake variation is merely an approximation, this analysis would help in determining the best process type under various substrate uptake rate assumptions. If a TS process is compared to an OS process with the same production stage operating point, the TS process has a higher productivity at all substrate uptake rates (Supplementary Figure S7a). Even if we compare the TS processes to the best performing OS process assuming constant substrate uptake, there is a large range of substrate uptake rates where a better performing TS process can be found (Supplementary Figure S7b). The threshold substrate uptake rate that needs to be crossed to make a TS process is given by the contour line that equals zero in Figure S7b. This analysis can be used as a general guideline to decide the process type for production when substrate uptake rates are exactly known for a given production strain.

3.3 Commonalities in flux perturbations hint towards a universal production phenotype

While it is clear that TS processes have optimal productivity, it would be useful to determine how such processes can be physically realized. Hence, we wished to examine the flux perturbations required to achieve the various production strategies for each of the exchange metabolites analyzed in the previous section. This was done by using pFBA to obtain the flux distributions required to achieve the various phenotypes for each of the process types. As mentioned earlier, both the best and traditional TS processes use wild-type growth as the phenotype in the first stage. Therefore, using the wild-type flux distribution as a reference, we examined how many reactions would need to be perturbed in each of the strategies, and classified each perturbed reaction/flux based on whether it was switched on, switched off, upregulated, downregulated or reversed (Supplementary Figures S8,S9, and S10).

A very large number of fluxes (>200) need to be modified to achieve any of the target phenotypes predicted by TS-DySScO. It appears that for all process types, products with fewer carbon atoms require more reactions to be turned on/off compared to larger products. Among these, α-ketoglutarate is an outlier, being a 5-carbon compound requiring an unusually large number of on-off type perturbations (>25) for the OS and best TS processes. Interestingly, the compounds that were determined to have an unusually low yield in the previous section - 5-methylthioribose and spermidine have the largest number of upregulations among all the products. The yield of these products is likely low due to the upregulation of pathways

bioRxiv preprint doi: https://doi.org/10.1101/803023; this version posted October 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

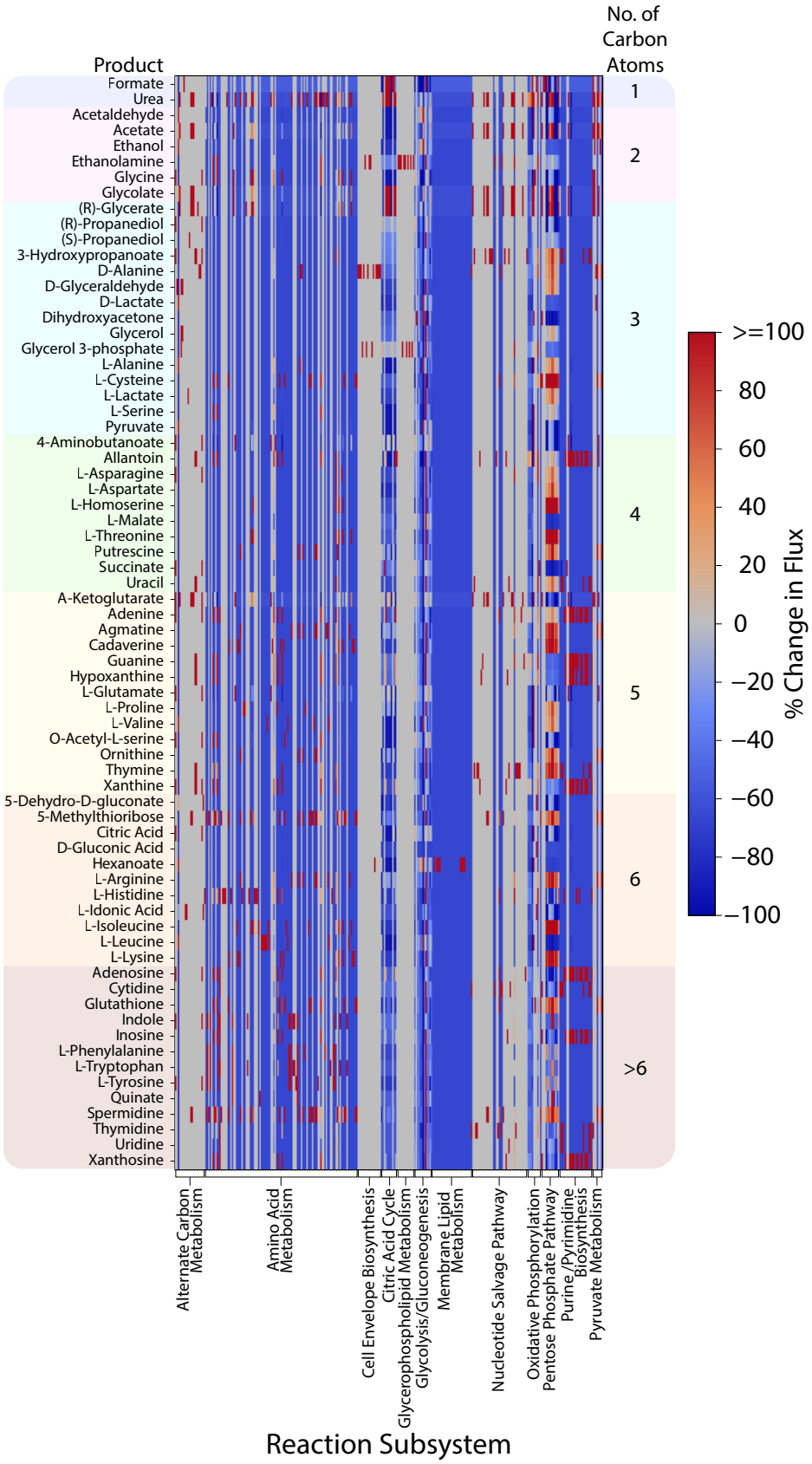

Figure 5: Flux perturbations required for best TS processes: Percent change in flux through all reactions compared to wild-type flux distribution required to achieve the best two-stage processes for native exchange metabolites in *E. coli*. Due to the large number of reactions involved, only the reaction subsystems are shown.

bioRxiv preprint doi: https://doi.org/10.1101/803023; this version posted October 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

that result in yield losses during production and the absence of alternative pathways that conserve yield. All process types involve a large number of downregulations. However, the number of downregulations for the OS and best TS processes are very high while the number of reactions turned off is very high for the traditional TS processes.

To further understand the production phenotypes for each of the process types, we examined the magnitude of flux changes from the wild-type flux distribution for every product and arranged the reactions based on the subsystem in which they occur (Figure 5, Supplementary Figure S11). As noted previously, most perturbations for the best TS process are downregulations. It appears that the reactions which require downregulation are concentrated in specific subsystems. Notably, all reactions involved in membrane lipid metabolism are downregulated for all but one of the products. Similarly, many reactions involved in glycolysis, the citric acid cycle, amino acid metabolism and nucleotide biosynthesis are downregulated for most products. Not surprisingly, reactions involved in the production of biomass are downregulated for all products. The magnitude of flux changes is also very similar across the entire product range. Reactions in amino acid metabolism and nucleotide biosynthesis are upregulated/switched on only for the production of these products. Interestingly, reactions in the pentose phosphate pathway are equally divided between being upregulated or downregulated together for different products. Hence, this subsystem appears to act as a key node that controls precursor availability to manufacture metabolites within the cell. Also, it appears that the same fluxes need to be perturbed for the OS process as well, with the only difference being the magnitude of the flux change required (Supplementary Figure S11). These results suggest that it is possible to engineer an *E. coli* strain with a universal production phenotype that maximizes productivity, where flux perturbations that appear for all the products can be dynamically controlled and flux through them throttled to reduce the growth rate to a certain value. Then, depending on the class of product required - amino acid, nucleotide, central carbon metabolite, etc, those reactions that require upregulation can be dynamically expressed. Therefore, it could be possible to create productivity maximizing platform strains for each class of product.

3.4 Perturbations increasing phosphoenolpyruvate and NADPH availability are enriched

In order to identify key control reactions, we analyzed which reactions are enriched in production strategies for the exchange metabolites previously analyzed. We did this by looking at the number of products for which each reaction appeared as a perturbation and classified them based on the type of perturbation - on, off, upregulation, downregulation or reversal (Figure 6). Only non-transport reactions involved directly in metabolism were retained for the final analysis. The full names of reactions have been provided in supplementary table S1.

Among the reactions that are switched on for TS strategies (Figure 6), HEX1 (ATP dependent hexokinase) occurs in more than half of the products. This reaction serves as an alternative to the phosphotransferase system that is used by wild-type *E. coli* to transport and phosphorylate glucose. Other modes of glucose phosphorylation such as the XYLI - HEX7 (Xylose isomerase - Hexokinase) system that converts glucose to fructose and then phosphorylates it, also appears in many products. These reactions differ from the conventional phosphotransferase system, in that they use ATP for phosphorylation as opposed to phosphoenolpyruvate (PEP). PEP is a key precursor to many products and therefore, alternative

bioRxiv preprint doi: https://doi.org/10.1101/803023; this version posted October 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

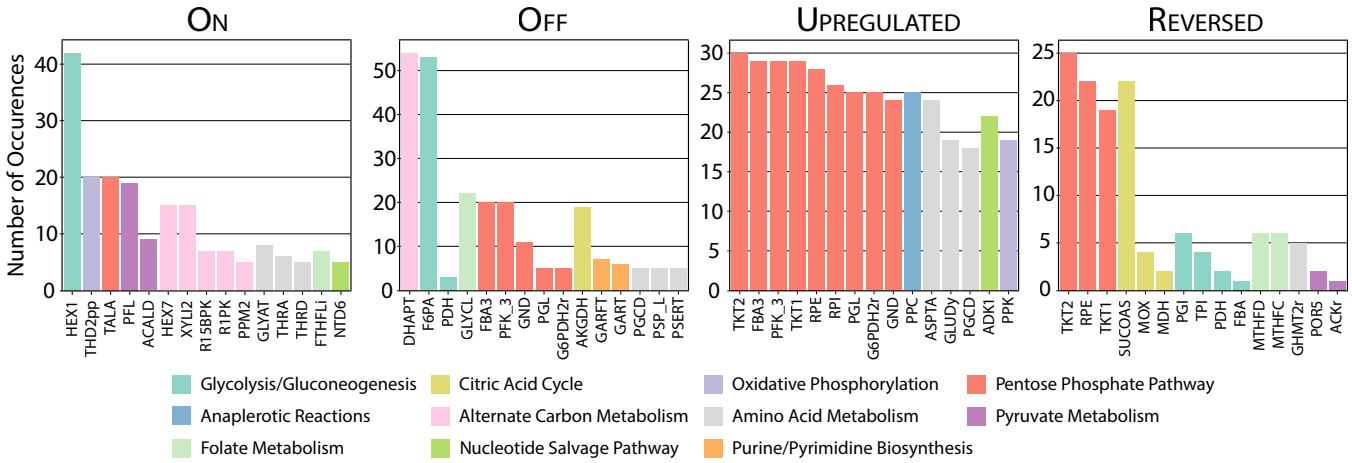

Figure 6: Key reactions in perturbations for best TS processes: Reactions most frequently occuring in the best two-stage production strategies for native exchange metabolites in *E. coli* were obtained and classified based on the type of perturbation.

means of glucose usage that use less PEP are enriched in the 'on' type perturbations. Consequently, the HEX1 glucose transport system has previously been studied for its importance in creating platform strains for microbial chemical production[42]. Furthermore, extra usage of ATP has been explored as a means to increase substrate uptake rates, potentially increasing product flux under low growth conditions[43], suggesting the importance of studying alternatives to the native phosphotransferase system. The reaction THD2pp (NADP transhydrogenase) is also required to be switched on for many products. This likely serves to increase NADPH availability to cater to production pathways. Many reactions that need to be switched on are concentrated in the alternate carbon metabolism subsystem.

Among other types of perturbations, DHAPT (Dihydroxyacetone phosphotransferase) and F6PA (Fructose 6-phosphate aldolase) are reactions that need to be turned off for most products. These also likely serve to collectively increase PEP availability for production reactions since DHAPT utilizes PEP. Many upregulated reactions are from the pentose phosphate pathway subsystem, serving to increase the availability of NADPH and pentose sugars for products. The transketolases and RPE (Ribulose 5-phosphate epimerase) need to be reversed in direction for many of the products to cater to varying precursor requirements for different product classes. In summary, reactions from the pentose phosphate pathway occur very frequently in the best TS strategies. Also, over 70% of the reactions involved in membrane lipid metabolism, cofactor biosynthesis, lipopolysaccharide biosynthesis and glycerophospholipid metabolism, which are used for biomass production need to be downregulated (Supplementary Figure S13).

The best OS and TS strategies have a lot of commonly required perturbations (Figure 6 and Supplementary Figure S12). Notably, the PEP conserving strategies and the pentose phosphate pathway reactions are enriched for either process type. This bolsters the importance of PEP as a key bow-tie metabolite in making metabolism modular.

4 Conclusions

We have seen that the choice of process type influences the process metrics and therefore the profitability of a microbial chemical production process to a great extent. Furthermore, strain design choices are also

bioRxiv preprint doi: https://doi.org/10.1101/803023; this version posted October 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

influenced by process choice. One-stage processes require static genetic intervention strategies that couple growth and production whereas, TS processes require dynamic intervention strategies where gene expression is temporally controlled. Recent advances in CRISPR[44], transcriptional switches[45], riboswitches[46], and other gene regulatory elements present an exciting outlook for the experimental implementation of such intervention strategies. There has also been interest in computational algorithms that predict dynamic control strategies which begin with high growth and switch over to growth-coupled production as required by the best TS production strategies predicted in this study[47].

The flux perturbations necessary to achieve the target production phenotypes predicted for these strategies is daunting due to the sheer number of perturbations required. However, it is important to note that this analysis does not take into account the fact that many pathways are linear and sequential. Therefore, it would not be necessary to actively perturb all fluxes predicted in this study. A reduction of the metabolic network would help to identify key control reactions that actually need to be perturbed. Furthermore, algorithmic approaches can be used to predict the genetic perurbations required to achieve target phenotypes given the constraints predicted by TS-DySScO. Most reactions that need to be downregulated are those that direct flux towards biomass production. It is interesting to note that the production of most products involves enhancing PEP conserving and NADPH overproducing strategies. The emergence of PEP as a key precursor indicates its importance as a bow-tie metabolite, funneling flux into different pathways[2,48]. Furthermore, reactions in the pentose phosphate pathway seem to work in unison to increase precursor availability for number of products by being upregulated or downregulated, alluding to their importance in making metabolic networks malleable and robust to perturbations. These common features in strains with enhanced production of a wide range of metabolites give rise to the idea of a universal production strain that could be used to maximize productivity in a TS process by redirecting flux from growth to production related processes for various classes of products. Such a platform strain that maximizes productivity could be realized by placing a minimal number of control reactions under dynamically repressible/inducible promoters to throttle biomass production flux. This is similar to the concept of a modular cellular chassis for the production of many different compounds, that has gained interest recently[49,50].

From our results, TS processes do seem to outperform OS processes in terms of productivity for native exchange metabolites in *E. coli*, and since the production characteristics can be expected to be similar for other organisms and non-native products too, this conclusion can be extended to products in other hosts as well. We found that this conclusion holds true over a wide range of industrially relevant fermentation start parameters. The overall yield and end-titer of such a TS process is also higher than an OS process selected for high productivity. However, the model for substrate uptake rate variation used in this study is only an approximation. In future, better substrate uptake rate measurements for combinatorial deletion mutants will help in making more accurate predictions of process performance. While it is true that the process metrics depend on the substrate uptake rate of the mutant strain, we have shown that a TS process can outperform an OS process at every production phenotype, regardless of substrate uptake rate. Further improvements in substrate uptake rates through various strategies[51] will improve productivity even further. The software framework presented here - TS-DySScO, has the ability to determine the effectiveness of each process type and predict optimal hypothetical phenotypes for experimental evaluation. It also provides information about the fermentation conditions under which each process type would perform better. We anticipate that TS-DySScO and the findings obtained in this study will be very valuable to make process

bioRxiv preprint doi: https://doi.org/10.1101/803023; this version posted October 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

and strain design decisions for industrial scale production of chemicals using microorganisms. Furthermore, the concept of a universal production strain that has the same growth phenotype and several common flux perturbations required to switch on the production of many chemicals in a flexible manner may provide a paradigm shift in the way chemical production processes are designed in the future.

Acknowledgements

This work was supported by the Natural Sciences and Engineering Research Council (NSERC), the NSERC Industrial Biocatalysis Network, the Ontario Ministry of Research and Innovation, and Genome Canada.

References

[1] E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A.-L. Barabási. Hierarchical organization of modularity in metabolic networks. *Science*, 297(5586):1551–1555, 2002.

[2] Tamar Friedlander, Avraham E. Mayo, Tsvi Tlusty, and Uri Alon. Evolution of bow-tie architectures in biology. *PLOS Computational Biology*, 11(3):1–19, 03 2015.

[3] Kayla Nemr, Jonas E.N. Müller, Jeong Chan Joo, Pratish Gawand, Ruhi Choudhary, Burton Mendonca, Shuyi Lu, Xiuyan Yu, Alexander F. Yakunin, and Radhakrishnan Mahadevan. Engineering a short, aldolase-based pathway for (R)-1,3-butanediol production in *Escherichia coli*. *Metabolic Engineering*, 48:13 – 24, 2018.

[4] Kaushik Raj, Siavash Partow, Kevin Correia, Anna N. Khusnutdinova, Alexander F. Yakunin, and Radhakrishnan Mahadevan. Biocatalytic production of adipic acid from glucose using engineered *Saccharomyces cerevisiae*. *Metabolic Engineering Communications*, 6:28 – 32, 2018.

[5] Stephanie Galanie, Kate Thodey, Isis J. Trenchard, Maria Filsinger Interrante, and Christina D. Smolke. Complete biosynthesis of opioids in yeast. *Science*, 349(6252):1095–1100, 2015.

[6] Yanran Li, Sijin Li, Kate Thodey, Isis Trenchard, Aaron Cravens, and Christina D. Smolke. Complete biosynthesis of noscapine and halogenated alkaloids in yeast. *Proceedings of the National Academy of Sciences*, 115(17):E3922–E3931, 2018.

[7] Weerawat Runguphan and Jay D. Keasling. Metabolic engineering of saccharomyces cerevisiae for production of fatty acid-derived biofuels and chemicals. *Metabolic Engineering*, 21:103 – 113, 2014.

[8] Andreas Schirmer, Mathew A. Rude, Xuezhi Li, Emanuela Popova, and Stephen B. del Cardayre. Microbial biosynthesis of alkanes. *Science*, 329(5991):559–562, 2010.

[9] Sang Yup Lee, Hyun Uk Kim, Tong Un Chae, Jae Sung Cho, Je Woong Kim, Jae Ho Shin, Dong In Kim, Yoo-Sung Ko, Woo Dae Jang, and Yu-Sin Jang. A comprehensive metabolic map for production of bio-based chemicals. *Nature Catalysis*, 2(1):18–33, jan 2019.

[10] Anthony Burgard, Mark J. Burk, Robin Osterhout, Stephen Van Dien, and Harry Yim. Development of a commercial scale process for production of 1,4-butanediol from sugar. *Current Opinion in Biotechnology*, 42:118 – 125, 2016.

bioRxiv preprint doi: https://doi.org/10.1101/803023; this version posted October 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

[11] C J Paddon, P J Westfall, D J Pitera, K Benjamin, K Fisher, D McPhee, M D Leavell, a Tai, a Main, D Eng, D R Polichuk, K H Teoh, D W Reed, T Treynor, J Lenihan, M Fleck, S Bajad, G Dang, D Dengrove, D Diola, G Dorin, K W Ellens, S Fickes, J Galazzo, S P Gaucher, T Geistlinger, R Henry, M Hepp, T Horning, T Iqbal, H Jiang, L Kizer, B Lieu, D Melis, N Moss, R Regentin, S Secrest, H Tsuruta, R Vazquez, L F Westblade, L Xu, M Yu, Y Zhang, L Zhao, J Lievense, P S Covello, J D Keasling, K K Reiling, N S Renninger, and J D Newman. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature*, 496(7446):528–532, apr 2013.

[12] David Jullesson, Florian David, Brian Pfleger, and Jens Nielsen. Impact of synthetic biology and metabolic engineering on industrial production of fine chemicals. *Biotechnology Advances*, 33(7):1395 – 1402, 2015.

[13] Nelson R. Barton, Anthony P. Burgard, Mark J. Burk, Jason S. Crater, Robin E. Osterhout, Priti Pharkya, Brian A. Steer, Jun Sun, John D. Trawick, Stephen J. Van Dien, Tae Hoon Yang, and Harry Yim. An integrated biotechnology platform for developing sustainable chemical processes. *Journal of Industrial Microbiology & Biotechnology*, 42(3):349–360, Mar 2015.

[14] Sang Yup Lee and Hyun Uk Kim. Systems strategies for developing industrial microbial strains. *Nature Biotechnology*, 33(10):1061–1072, 2015.

[15] Stephen Van Dien. From the first drop to the first truckload: commercialization of microbial processes for renewable chemicals. *Current Opinion in Biotechnology*, 24(6):1061 – 1068, 2013.

[16] Jens Nielsen and Jay D. Keasling. Engineering cellular metabolism. *Cell*, 164(6):1185 – 1197, 2016.

[17] Jeremy S. Edwards, Rafael U. Ibarra, and Bernhard O. Palsson. In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data. *Nature Biotechnology*, 19(2):125–130, feb 2001.

[18] Steffen Klamt and Radhakrishnan Mahadevan. On the feasibility of growth-coupled product synthesis in microbial strains. *Metabolic Engineering*, 30:166–178, 2015.

[19] Naveen Venayak, Kaushik Raj, Rohil Jaydeep, and Radhakrishnan Mahadevan. An optimized bistable metabolic switch to decouple phenotypic states during anaerobic fermentation. *ACS Synthetic Biology*, 7(12):2854–2866, 2018.

[20] Apoorv Gupta, Irene M Brockman Reizman, Christopher R Reisch, and Kristala L J Prather. Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. *Nature Biotechnology*, 35(3):273–279, 2017.

[21] Peng Xu, Lingyun Li, Fuming Zhang, Gregory Stephanopoulos, and Mattheos Koffas. Improving fatty acids production by engineering dynamic pathway regulation and metabolic control. *Proceedings of the National Academy of Sciences*, 111(31):11299–11304, 2014.

[22] Nikolaos Anesiadis, William R. Cluett, and Radhakrishnan Mahadevan. Dynamic metabolic engineering for increasing bioprocess productivity. *Metabolic Engineering*, 10(5):255–266, 2008.

bioRxiv preprint doi: https://doi.org/10.1101/803023; this version posted October 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

[23] Naveen Venayak, Nikolaos Anesiadis, William R. Cluett, and Radhakrishnan Mahadevan. Engineering metabolism through dynamic control. *Current Opinion in Biotechnology*, 34:142–152, 2015.

[24] Victor Chubukov and Uwe Sauer. Environmental dependence of stationary-phase metabolism in bacillus subtilis and escherichia coli. *Applied and Environmental Microbiology*, 80(9):2901–2909, 2014.

[25] Markus W. Covert, Eric M. Knight, Jennifer L. Reed, Markus J. Herrgård, and Bernhard O. Palsson. Integrating high-throughput and computational data elucidates bacterial networks. *Nature*, 429:92–96, 2004.

[26] Eliane Fischer, Nicola Zamboni, and Uwe Sauer. High-throughput metabolic flux analysis based on gas chromatography–mass spectrometry derived 13C constraints. *Analytical Biochemistry*, 325(2):308–316, 2004.

[27] Uwe Sauer, Fabrizio Canonaco, Sylvia Heri, Annik Perrenoud, and Eliane Fischer. The Soluble and Membrane-bound Transhydrogenases UdhA and PntAB Have Divergent Functions in NADPH Metabolism of *Escherichia coli*. *Journal of Biological Chemistry*, 279(8):6613–6619, 2004.

[28] Annik Perrenoud and Uwe Sauer. Impact of global transcriptional regulation by arca, arcb, cra, crp, cya, fnr, and mlc on glucose catabolism in *Escherichia coli*. *Journal of Bacteriology*, 187(9):3171–3179, 2005.

[29] Stephen S Fong, Annik Nanchen, Bernhard O Palsson, and Uwe Sauer. Latent pathway activation and increased pathway capacity enable *Escherichia coli* adaptation to loss of key metabolic enzymes. *The Journal of biological chemistry*, 281(12):8024–33, 2006.

[30] Bart R B Haverkorn van Rijsewijk, Annik Nanchen, Sophie Nallet, Roelco J Kleijn, and Uwe Sauer. Large-scale 13c-flux analysis reveals distinct transcriptional control of respiratory and fermentative metabolism in *Escherichia coli*. *Molecular Systems Biology*, 7(1):477, 2011.

[31] Steffen Klamt, Radhakrishnan Mahadevan, and Oliver Hädicke. When do two-stage processes outperform one-stage processes? *Biotechnology Journal*, 13(2):1700539, 2018.

[32] Kai Zhuang, Laurence Yang, William R Cluett, and Radhakrishnan Mahadevan. Dynamic strain scanning optimization: an efficient strain design strategy for balanced yield, titer, and productivity. DySScO strategy for strain design. *BMC biotechnology*, 13(1):8, jan 2013.

[33] Annik Nanchen, Alexander Schicker, and Uwe Sauer. Nonlinear dependency of intracellular fluxes on growth rate in miniaturized continuous cultures of *Escherichia coli*. *Applied and Environmental Microbiology*, 72(2):1164–1172, 2006.

[34] Nobuyoshi Ishii, Kenji Nakahigashi, Tomoya Baba, Martin Robert, Tomoyoshi Soga, Akio Kanai, Takashi Hirasawa, Miki Naba, Kenta Hirai, Aminul Hoque, Pei Yee Ho, Yuji Kakazu, Kaori Sugawara, Saori Igarashi, Satoshi Harada, Takeshi Masuda, Naoyuki Sugiyama, Takashi Togashi, Miki Hasegawa, Yuki Takai, Katsuyuki Yugi, Kazuharu Arakawa, Nayuta Iwata, Yoshihiro Toya, Yoichi Nakayama, Takaaki Nishioka, Kazuyuki Shimizu, Hirotada Mori, and Masaru Tomita. Multiple high-throughput analyses monitor the response of e. coli to perturbations. *Science*, 316(5824):593–597, 2007.

bioRxiv preprint doi: https://doi.org/10.1101/803023; this version posted October 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

[35] Anke Kayser, Jan Weber, Volker Hecht, and Ursula Rinas. Metabolic flux analysis of *Escherichia coli* in glucose-limited continuous culture. i. growth-rate-dependent metabolic efficiency at steady state. *Microbiology*, 151(3):693–706, 2005.

[36] Annik Nanchen, Alexander Schicker, Olga Revelles, and Uwe Sauer. Cyclic amp-dependent catabolite repression is the dominant control mechanism of metabolic fluxes under glucose limitation in *Escherichia coli*. *Journal of Bacteriology*, 190(7):2323–2330, 2008.

[37] Radhakrishnan Mahadevan, Jeremy S. Edwards, and Francis J. Doyle. Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophysical Journal*, 83(3):1331 – 1340, 2002.

[38] Ali Ebrahim, Joshua A Lerman, Bernhard O Palsson, and Daniel R Hyduke. COBRApy: COnstraints-Based Reconstruction and Analysis for Python. *BMC Systems Biology*, 7(1):74, 2013.

[39] João G. R. Cardoso, Kristian Jensen, Christian Lieven, Anne Sofie Lærke Hansen, Svetlana Galkina, Moritz Beber, Emre Özdemir, Markus J. Herrgård, Henning Redestig, and Nikolaus Sonnenschein. Cameo: A python library for computer aided metabolic engineering and optimization of cell factories. *ACS Synthetic Biology*, 7(4):1163–1166, 2018.

[40] Jeffrey D Orth, Tom M Conrad, Jessica Na, Joshua a Lerman, Hojung Nam, Adam M Feist, and Bernhard Ø Palsson. A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism–2011. *Molecular systems biology*, 7(535):535, 2011.

[41] Nathan E Lewis, Kim K Hixson, Tom M Conrad, Joshua A Lerman, Pep Charusanti, Ashoka D Polpitiya, Joshua N Adkins, Gunnar Schramm, Samuel O Purvine, Daniel Lopez-Ferrer, Karl K Weitz, Roland Eils, Rainer König, Richard D Smith, and Bernhard Ø Palsson. Omic data from evolved e. coli are consistent with computed optimal growth from genome-scale models. *Molecular Systems Biology*, 6(1):390, 2010.

[42] Verónica Hernández-Montalvo, Alfredo Martínez, Georgina Hernández-Chavez, Francisco Bolivar, Fernando Valle, and Guillermo Gosset. Expression of galp and glk in a *Escherichia coli* pts mutant restores glucose transport and increases glycolytic flux to fermentation products. *Biotechnology and Bioengineering*, 83(6):687–694, 2003.

[43] Simon Boecker, Ahmed Zahoor, Thorben Schramm, Hannes Link, and Steffen Klamt. Broadening the scope of enforced atp wasting as a tool for metabolic engineering in escherichia coli. *Biotechnology Journal*, 14(9):1800438, 2019.

[44] Alexander C. Reis, Sean M. Halper, Grace E. Vezeau, Daniel P. Cetnar, Ayaan Hossain, Phillip R. Clauer, and Howard M. Salis. Simultaneous repression of multiple bacterial genes using nonrepetitive extra-long sgRNA arrays. *Nature Biotechnology*, pages 1–8, oct 2019.

[45] Jeong Wook Lee, Andras Gyorgy, D. Ewen Cameron, Nora Pyenson, Kyeong Rok Choi, Jeffrey C. Way, Pamela A. Silver, Domitilla Del Vecchio, and James J. Collins. Creating Single-Copy Genetic Circuits. *Molecular Cell*, 63(2):329–336, jul 2016.

bioRxiv preprint doi: https://doi.org/10.1101/803023; this version posted October 13, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

[46] Abigail N. Leistra, Nicholas C. Curtis, and Lydia M. Contreras. Regulatory non-coding sRNAs in bacterial metabolic pathway engineering. *Metabolic Engineering*, 52:190–214, mar 2019.

[47] Naveen Venayak, Axel von Kamp, Steffen Klamt, and Radhakrishnan Mahadevan. MoVE identifies metabolic valves to switch between phenotypic states. *Nature Communications*, 9(1):5332, 2018.

[48] Jing Zhao, Hong Yu, Jian-Hua Luo, Zhi-Wei Cao, and Yi-Xue Li. Hierarchical modularity of nested bow-ties in metabolic networks. *BMC Bioinformatics*, 7(1):386, aug 2006.

[49] Sergio Garcia and Cong T. Trinh. Multiobjective strain design: A framework for modular cell engineering. *Metabolic Engineering*, 51:110 – 120, 2019.

[50] Brandon Wilbanks, Donovan S. Layton, Sergio Garcia, and Cong T. Trinh. A prototype for modular cell engineering. *ACS Synthetic Biology*, 7(1):187–199, 2018.

[51] Qiang Yan and Stephen S Fong. Increasing carbon source uptake rates to improve chemical productivity in metabolic engineering. *Current Opinion in Biotechnology*, 53:254 – 263, 2018.