Transcriptional regulation of *Escherichia coli* metabolism and engineered metabolic pathways

Dissertation

zur Erlangung des Doktorgrades

der Naturwissenschaften

(Dr. rer. nat)

dem Fachbereich Biologie

der Philipps-Universität Marburg

vorgelegt von

Chun-Ying Wang

aus New Taipei City, Taiwan

Marburg, Juli 2020

Die Untersuchungen zur vorliegenden Arbeit wurden von April 2016 bis Juli 2020 an der Philipps-Universität Marburg unter der Leitung von Dr. Hannes Link am Max-Planck-Institut für terrestrische Mikrobiologie durchgeführt.

Erstgutachter: Dr. Hannes Link

Zweitgutachter: Prof. Dr. Tobias Erb

Weitere Mitglieder der Prüfungskommission: Prof. Dr. Victor Sourjik
Prof. Dr. Michael Bölker

Vom Fachbereich Biologie der Philipps-Universität Marburg als Dissertation am 15.07.2020 angenommen.

Tag der mündlichen Prüfung: 03.09.2020

Acknowledgements

I would like to express my sincere gratitude to people helped me during my doctoral period. Without you, I could not finish this.

First of all, I want to thank you all teachers from Taiwan: Chuan-Mu Chen, Bor-Sen Chen, Jui-Sheng Chang and Chieh-Chen Huang. If you did not give me many recommendation letters, I would not have any chance to get this position in Germany. Thank you, teachers: Ban-Yang Chang and Nien-Taifor Hu correcting my applied forms and proposals so that I got my dream place to do research. Also, thank you my high school teachers Hsin-I Shih and Shi-Hong Lu for supporting and encouraging me all the time.

Thank my Marburg alumni association. Especially, thank you Chiung Wen Chang for helping me to solve plenty of German translation issues and living problems. Of course, my playmates Yu-Hsin Chang, Fang Yu Chang and Yi-Ju Lin. Because you colored my life by going to many German cities and also UK trips. Furthermore, I have to thank you Yu-An Chou, Shao-Min Sun, Felix Lin, and Alex Li for playing basketball, drinking and chatting when I was bored or stressful. I am so lucky to meet you all nice persons here in Marburg.

I would like to thank my colleagues. Dominik Beuter, you are most kind and patient person I met always helping me to deal with troublesome administrative works; Claudia Einloft, thank you, my German teacher always spent time to teach me and shared the German culture and foods; Thank Christina Schmidt for helping the administrative wroks and be super nice to answer my question about German's rules; Dušica Radoš and Peter Temitope Adeboye, you two just like my psychological counselors encouraging me and told me how to face the stress and career issues; Thank you that Janhavi Srirangaraj gave me an unforgettable trip in India for me understanding that I have a wonderful life no matter in Taiwan or in Germany; Thank you Vanessa Pahl and Michelle Kuntz, I really enjoyed the time sharing the experiences of trips where to go/what to eat...and so on. Thank Stefano Donati, Niklas Farke, Martin Lempp, Timur Sander and Thorben Schramm. It is really grateful to have you as my colleagues. Of course, Thank you Dr. Timo Glatter for being patient to measure a plenty of protein samples. Also, thanks Min-Chin Lin and Dominik correcting my thesis. Becasue all of you, my expirements and thesis can be finished.

Of course, thanks to my supervisor-Hannes Link. You are really a trustworthy person who ceaselessly encouraged me when I was stocked in my research, who always patiently corrected my disjointed and unreadable English writing and who endlessly supported me when I needed help. I am extremely happy and grateful to have this chance working with you. Moreover, thanks to the members of my thesis and defense committee: Prof.Dr. Lotte Sogaard-Andersen, Prof. Dr. Tobias Erb and Prof. Dr. Michael Bölker. It is my honor to have you who guided and clarified my research topic in a cool direction. I truly appreciate you so that I have learned a lot from this process.

My best friend and competitor in life-Hui-Chun. Without you, I would not make it. I was the one who was easy to give up, but you made me become stronger. Thanks to my bestie PSQ-Chien-Yun Lee, I am really glad that you also came to Germany to accompany with me. Thanks Yu-Hsuan Chen and Tsun-Ying Yin, even the distance is long, but it is still like we are at the same university. All of you are not just my friends, but like my family who always be there for me and makes joy and fun in my life.

Last but not least, my true family: my mom and two brothers. Because of you, I do not have to worry about other things except for my study.

Thank you all and I love you all!

-This dissertation is dedicated to my angel father-

Chun-Ying, Marburg, 03.09.2020

致謝

俗話說的好：一日為師，終生為父。感謝陳全木、陳博現、張睿昇以及碩士指導教授黃介辰老師們在 2015 年不厭其煩幫我寫多封推薦信，張邦彥以及胡念台老師多次幫我修改申請資料，終於讓我在 2016 年如願以償到德國念書，由衷地謝謝您們的幫助。在德國的生活期間，心靈的支柱當然很重要，但台灣的物資支持也是不容小覷，謝謝石欣怡與呂世宏老師，多次寄台灣補給品給我，讓我可以遠在他鄉仍然享有很多台灣味食物。

在一開始的適應期間莫過於我的玩伴張予馨和張瓊文，想當初我們一起用學生證到多個城市玩耍，這是多麼難得的經驗；另外，值得回味的英國之旅真得很棒，謝謝張芳瑀完美帶路以及林邑孺地理老師的行程導覽，讓我獲益良多。很慶幸在馬堡能夠認識妳們：周予安、孫劭旼、林哲豪以及李楚睿，在我想打球時可以找妳們，在我想放縱時後可以隨時呼喚妳們，很謝謝你們的支持與陪伴。

實驗室的生活，當然要謝謝 Dominik Beuter，你是第一個在德國願意當我是朋友的人，還願意帶我出去騎腳踏車；Dušica Radoš 陪我聊人生黑暗面以及如何應對，你真得很適合當心靈導師；謝謝我的德文老師 Claudia Einloft，您就像是我在德國的媽媽一樣，願意花時間教我德文陪我聊天；感謝 Christina Schmidt 每次我有任何需求都幫我完成；謝謝 Janhavi Srirangaraj 讓我體驗印度生活的樣子，讓我了解我的生活過得很好；謝謝 Peter Temitope Adeboye 除了愛鼓勵我加入宗教以外，你真得很會鼓勵人；我超喜歡跟 Vanessa Pahl 和 Michelle Kuntz 聊要去哪裡玩的，因為妳們讓我知道有些我根本沒想過要去的景點；謝謝實驗室同事們：Stefano Donati、Niklas Farke、Martin Lempp、Timur Sander 以及 Thorben Schramm 在我需要幫忙時後，總是願意幫助我。感謝 Timo Glatter 總是很有效率幫我測量數量龐大的樣本。也感謝林旻瑾和 Dominik 願意幫我改英文，因為你們讓我的實驗和論文順利完成。

博班過程，當然要感謝當初願意收留我的指導老師 Hannes Link，沒有您就不會有今天的我，很謝謝您每次都在我實驗失敗時後，鼓勵我，也沒有因為我英文不好放棄過我，真得很謝謝您。也感謝我的 TAC 和口試委員：Lotte Sogaard-Andersen、Tobias Erb、以及 Michael Bölke 教授，因為你們的直言不諱，讓我在做研究過程可以發現問題並且去解決問題，讓我了解歐洲教育與亞洲教育的差異性。

當然我要感謝一路走來，我得好夥伴與競爭對手-李慧君，總是在我不想念下去的時後刺激或鼓勵我，幫我改論文順投影片等等，讓我重時心情/信心繼續努力。也要謝謝我的夥伴李倩筠，真得很開心你也能來德國陪我；還有謝謝我的前室友們陳宇萱以尹存瑛，能夠來德國探望我，在我心情不好時候陪我聊天，你們是我永遠的心靈支柱。

最後，感謝我的母親，哥哥以及弟弟，讓我無後顧之憂可以到德國念書。

獻給我的父親，我想你可能曾經嚮往過國外的生活，女兒我代替你完成夢想了，謝謝您默默的在天堂幫助守護我！

鈞瑩　謹

03.09.2020

Abstract

A common problem in metabolic engineering projects is to find enzyme levels that enhance productivity and efficiency of synthetic metabolic pathways. This problem is especially important when overexpressing heterologous enzymes that drain metabolites from central metabolism. The additional requirement for precursors and energy causes a growth burden or even leads to cell death (**chapter 1**). Several strategies have been employed to solve this problem and two out of them are addressed in this thesis: i) decoupled overproduction with a two-phase process and ii) growth coupled production with a one-phase process. A two-phase process decouples growth of the host from the production phase, while the one-phase strategy couples production to growth of the host. In this thesis, these two strategies are investigated using overproduction of three chemicals as case studies: glycerol, carotenoids and arginine. To test the two-phase processes, we inserted glycerol genes into the *E. coli* genome and used CRISPR interference (CRISPRi) to down-regulate the expression of RNA polymerase (**chapter 2**). When the CRISPRi system is induced at a relatively high biomass, the glycerol production can be enhanced. This result implies that producing glycerol requires to maintain certain growth instead of draining all metabolites only for overproduction. Therefore, a one-phase process is more feasible for glycerol overproduction.

In **chapter 3**, to use the concept of one-phase processes and combine it with an additional layer of regulation, we designed an artificial feedback circuit using a transcription factor (TF), Cra. To this end, the consensus binding sequence of Cra was inserted directly after the pBAD promoter sequence expressing glycerol genes. This design allowed native transcriptional regulation by Cra to regulate the expression of overproduction pathways. Proteomic and metabolomic data revealed that the Cra-regulation system can overcome growth burden by slowing down enzyme expression and thereby avoid the complete utilization of pathway precursors even before they are replenished. This delayed time enabled an adaptive metabolic response, a gradual increment of product synthesis (i.e. glycerol and carotenoids). The observed adaptive behavior leads to an increase in the expression of glucose transporters and glycolytic enzymes thereby improving host fitness and productivity in glycerol and carotenoid overproduction pathways. Cra-regulation was engineered in multiple types of

promoters including pBAD promoter, pTetR and constitutive promoters (**chapter 4**) which shows that this regulation enables universal dynamical control.

In **chapter 5**, we genomically integrated GFP into SS9 intergenic region and used it as a reporter system to measure burden and fitness defects. In this system, the GFP signal decreases when overproduction pathways drain too many cellular resources and thereby the GFP content per cell reflects cellular fitness. Furthermore, it was sensitive enough to detect the metabolic burden even in the absence of a growth phenotype.

In **chapter 6**, we engineered the arginine pathway with the CRISPRi system to down-regulate the TF of the pathway, ArgR. Our results show that productivity similar to an Δ*argR* overproducing strain can be achieved with a further enhanced growth rate. The result again indicates that the TF-feedback regulated system is capable of altering enzyme expressions for balanced resource utilization.

In conclusion, incorporating TF-based regulation in any designed circuit, can balance host metabolism and the production of several high-value chemicals without retarding the host growth. This is an advantage in comparison to the conventional process that involves complex enzyme screening to obtain optimized enzyme levels. This thesis therefore introduces a new dynamically controlled feedback loop strategy. With its ability to maintain a balance between host metabolism and product formation without causing host cellular burden, this strategy not only can serve as a competitive and facile solution to improve the productivity of a bacterial strain but also can be further expanded in large-scale applications.

Zusammenfassung

Ein häufiges Problem bei Metabolic-Engineering-Projekten ist es, Enzymkonzentrationen zu finden, die die Produktivität und Effizienz von synthetischen Stoffwechselwegen verbessern. Dieses Problem ist besonders wichtig, wenn heterologe Enzyme überexprimiert werden, die Metaboliten aus dem Zentralstoffwechsel abführen. Der zusätzliche Bedarf an Metaboliten und Energie verursacht eine Wachstumsbelastung oder führt sogar zum Zelltod **(Kapitel 1)**. Zur Lösung dieses Problems wurden mehrere Strategien angewandt, von denen zwei in dieser Arbeit behandelt werden: i) die entkoppelte Überproduktion mit zweiphasigen Prozessen und ii) die wachstumsgekoppelte Produktion mit einphasigen Prozessen. Ein zweiphasiger Prozess entkoppelt das Wachstum des Wirts von der Produktionsphase, während die einphasige Strategie die Produktion an das Wachstum des Wirts koppelt. In dieser Arbeit werden diese beiden Strategien anhand der Überproduktion von drei Chemikalien als Fallstudien untersucht: Glycerin, Carotinoide und Arginin. Um die zweiphasigen Prozesse zu testen, haben wir Gene der Glycerin-Biosynthese in das *E. coli*-Genom eingefügt und CRISPR-Interferenz (CRISPRi) zur Herunterregulierung der Expression der RNA-Polymerase verwendet. Wenn das CRISPRi-System bei einer relativ höheren Biomasse induziert wird, kann die Glycerinproduktion gesteigert werden. Dieses Ergebnis impliziert, dass zur Produktion von Glycerin ein bestimmtes Wachstum aufrechterhalten werden muss, anstatt alle Metabolite nur zur Überproduktion zu verwenden **(Kapitel 2)**. Daher ist ein einphasiger Prozess für die Glycerin-Überproduktion vielversprechender.

Um das Konzept der einphasigen Prozesse zu nutzen und mit zusätzlicher Regulierung zu kombinieren, entwarfen wir in **Kapitel 3** einen künstlichen Rückkopplungsmechanismus unter Verwendung des Transkriptionsfaktors (TF), Cra. Zu diesem Zweck wurde die Bindesequenz von Cra direkt nach der pBAD-Promotorsequenz eingefügt, welche verwendet wird um Glycerin-Gene zu exprimieren. Dieses Design ermöglichte die native Transkriptionsregulation durch Cra, um die Expression der Glycerol-Gene zu regulieren. Proteom- und Metabolomdaten zeigten, dass das Cra-Regulationssystem die Wachstumsbelastung überwinden kann, indem es die Enzymexpression verlangsamt und dadurch die vollständige Nutzung von Metaboliten verhindert. Diese verzögerte Zeit ermöglichte eine adaptive Reaktion und eine allmähliche Steigerung der Produktsynthese (d.h. Glycerin und Carotinoide). Das

beobachtete adaptive Verhalten führt zu einer Erhöhung der Expression von Glukose-Transportern und glykolytischen Enzymen, wodurch die Fitness und Produktivität des Wirts in den Glycerin- und Carotinoid-Überproduktionspfaden verbessert wird. Die Cra-Regulierung wurde bei mehreren Arten von Promotoren, einschließlich pBAD-Promotor, pTetR und konstitutiven Promotoren **(Kapitel 4)**, durchgeführt, was zeigt, dass diese Regulierung eine universelle dynamische Kontrolle ermöglicht.

In **Kapitel 5** haben wir GFP genomisch in die intergenische SS9-Region integriert und es als Reportersystem zur Messung von Belastung und Fitnessdefekten verwendet. In diesem System nimmt das GFP-Signal ab, wenn Überproduktionspfade zu viele zelluläre Ressourcen verwenden, weshalb der GFP-Gehalt pro Zelle die zelluläre Fitness widerspiegelt. Darüber hinaus war es empfindlich genug, um die metabolische Belastung auch bei Fehlen eines Wachstumsphänotyps zu erkennen.

In **Kapitel 6** kombinieren wir den Arginin-Syntheseweg mit dem CRISPRi-System, um den TF des Stoffwechselwegs, ArgR, herunter zu regulieren. Unsere Ergebnisse zeigen, dass eine ähnliche Produktivität wie bei einem überproduzierenden ArgR-Deletions-Stamm, aber mit einer weiter verbesserten Wachstumsrate erreicht werden kann. Das Ergebnis zeigt erneut, dass das durch TF-Feedback regulierte System in der Lage ist, die Enzymmenge für eine ausgewogene Ressourcennutzung zu verändern.

Zusammenfassend lässt sich sagen, dass die Einbeziehung der TF-basierten Regulation in jeden entworfenen Kreislauf den Wirtsstoffwechsel und die Produktion mehrerer hochwertiger Chemikalien ausbalancieren kann, ohne das Wirtswachstum zu verzögern. Dies ist ein Vorteil im Vergleich zum herkömmlichen Verfahren, das ein komplexes Enzymscreening umfasst, um optimierte Enzymmengen zu erhalten. In dieser Arbeit wird daher eine neue dynamisch gesteuerte Rückkopplungsschleifenstrategie vorgestellt. Mit ihrer Fähigkeit, ein Gleichgewicht zwischen Wirtsstoffwechsel und Produktbildung aufrechtzuerhalten, ohne die Wirtszellen zu belasten, kann diese Strategie nicht nur als wettbewerbsfähige und einfache Lösung zur Verbesserung der Produktivität eines Bakterienstammes dienen, sondern kann auch in großtechnischen Anwendungen weiter ausgebaut werden.

Contents

List of Abbreviations

TF	transcription factor
pBAD-only	plasmid-based glycerol production strain
pBAD-Cra	plasmid-based glycerol production strain with additional Cra-binding site
G6P	glucose-6-phosphate
Glycerol-P	glycerol-3-phosphate
F6P	fructose-6-phosphate
F1P	fructose-1-phosphate
G1P	glucose-1-phosphate
FBP	fructose 1,6-bisphosphate
6-PG	6-phospho D-glucono-1,5-lactone
1,3-BPG	1,3-bisphosphoglycerate
3-PG	3-phosphoglycerate
2-PG	2-phosphoglycerate
PEP	phosphoenolpyruvate
DHAP	dihydroxyacetone phosphate
Gluconate-6P	D-gluconate 6-phosphate
Ribulose-5P	D-ribulose 5-phosphate
Ribose-5P	D-ribose 5-phosphate
KDGP	2-dehydro-3-deoxy-D-gluconate 6-phosphate
Xylulose-5P	D-xylulose 5-phosphate
Sedoheptulose-7P	D-sedoheptulose 7-phosphate
Erythrose-4P	D-erythrose 4-phosphate
GABA	4-aminobutanoate
SAICAR	5-Phosphoribosyl-4-(N-succinocarboxamide)-5-aminoimidazole
N5-CAIR	5-Carboxyamino-1-(5-phospho-D-ribosyl)imidazole
AIR	5-Amino-1-(5-phospho-D-ribosyl)imidazole
GA3P	glyceraldehyde 3-phosphate
DXP	1-deoxy-Dxylulose-5-phosphate
MEP	2-C-methyl-D-erythritol-4-phosphate
DMAPP	dimethylallyl diphosphate
IPP	isopentenyl diphosphate
GPP	geranyl diphosphate
FPP	farnesyl diphosphate
GGPP	geranylgeranyl diphosphate
gpd1	glycerol-3-phosphate dehydrogenase 1
gpp2	glycerol-3-phosphate phosphohydrolase 2
ptsG	glucose-specific PTS enzyme IIBC component
ptsH	phosphocarrier protein HPr
Crr	enzyme IIA^{Glc}
ptsI	PTS enzyme I
fruA	fructose-specific PTS multiphosphoryl transfer protein
fruB	fructose-specific PTS multiphosphoryl transfer protein
fruK	1-phosphofructokinase
Pgi	glucose-6-phosphate isomerase
ybhA	pyridoxal phosphate/fructose-1,6-bisphosphate phosphatase
glpX	fructose-1,6-bisphosphatase 2
fbp	fructose-1,6-bisphosphatase 1
pfkA	6-phosphofructokinase 1
pfkB	6-phosphofructokinase 2
fbaA	fructose-bisphosphate aldolase class II
fbaB	fructose-bisphosphate aldolase class I
tpiA	triose-phosphate isomerase
gapA	glyceraldehyde-3-phosphate dehydrogenase
pgk	phosphoglycerate kinase
pgmI/gpmM	2,3-bisphosphoglycerate-independent phosphoglycerate mutase
gpmA	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase
eno	enolase

pckA	phosphoenolpyruvate carboxykinase
ppsA	phosphoenolpyruvate synthase
aceE	pyruvate dehydrogenase
aceF	pyruvate dehydrogenase, E2 subunit
lpd	lipoamide dehydrogenase
gltA	citrate synthase
acnA	aconitate hydratase A
acnB	bifunctional aconitate hydratase B and 2-methylisocitrate dehydratase
aceA	isocitrate lyase
aceB	malate synthase A
aceK	isocitrate dehydrogenase kinase / isocitrate dehydrogenase phosphatase
icd	isocitrate dehydrogenase
sucA	2-oxoglutarate decarboxylase, thiamine-requiring
sucB	dihydrolipoyltranssuccinylase
lpdA	lipoamide dehydrogenase
sucA	2-oxoglutarate dehydrogenase E1
sucB	dihydrolipoyltranssuccinylase
sucC	succinyl-CoA synthetase subunit β
sucD	succinyl-CoA synthetase subunit α
fumA	fumarase A
fumB	fumarase B
fumC	fumarase C
fumD	fumarase D
fumE	fumarase E
dxs	1-deoxy-D-xylulose-5-phosphate synthase
dxr	1-deoxy-D-xylulose 5-phosphate reductoisomerase
idi	IPP isomerase
crtB	phytoene synthase
crtI	phytoene dehydrogenase
crtY	lycopene beta-cyclase
crtZ	beta-carotene hydroxylase
crtE	geranylgeranyl diphosphate synthase

1. Chapter 1 – Research background and literature overview

1.1 Motivation

In synthetic bioengineering, circuit-associated burdens are general problems due to severe energy/precursors deficiency resulting from imbalanced enzyme expression levels, accumulation of toxic intermediates and so on[1-5]. These non-preferable conditions eventually affect the cell's fitness and generate heterogeneity, which reduce the product titer and applicability in large-scale fermentations[5, 6]. However, a facile and universal strategy to control and get the optimal enzyme levels in overproduction strains is still very challenging, meaning a fair balance between optimized synthetic enzymes of desired products and sustained the host fitness at an acceptable level. Therefore, in this thesis, we focused on understanding and investigating how *E. coli* adapts to the overexpression of exogenous genes by Cra-regulation and how *E. coli* reacts to the perturbation of endogenous genes by CRISPR interference. Moreover, two general bioprocessing approaches, the two-phase approach (growth switch) and dynamic control (self-regulating feedback loops), are developed to improve productivity. Overall the current thesis is aim to provide universal strategies for rational design in synthetic bioengineering under the premise of enhancing interesting chemicals. The ultimate goal is to improve yield, titer, and productivity of specialized chemicals via the newly-developed strategy in bioprocessing field. In the following paragraphs, the basic concepts of applied methods, literature reviews of transcription factors and mechanism of fine-tuned gene expression are described.

1.2 The switch between growth and production phase in *E. coli*

Several researchers have described the two-phase bioprocessing approach which maintains metabolic activity for desired products with constrained biomass production[5, 7, 8] (**Fig. 1.1**). For example, T. Hanai's group[9] programmed the genetic logic gate to turn off *sucA* and *aceA* genes in TCA cycle and to turn on GABA production simultaneously, achieving 3-fold improvement in total GABA production via the on/off controller. The other study used temperature-sensitive CRISPR inference (CRISPRi) to inhibit the host replication at 37 °C and released the inhibition at 42 °C to recover host cells from the suffering arrest[10]. Besides, downregulation of DNA replication (*dnaA*

and *oriC*) or nucleotide synthesis (*pyrF* or *thyA*) also showed a two-phase process, resulting in 41% increment of mevalonate compared to the counterpart without growth switcher[11]. Moreover, by replacing the native promoter of RNA polymerase (*rpoB*) with an IPTG-inducible promoter, the refined genomic integration will switch the host from growth to production mode[12]. Another two-layer gene regulation (i.e. sensing nutrient and sensing substrate) also decouples growth and production, giving 2.4-fold higher oleic acid production titer [13].

Apart from the above-mentioned genetic programming methods, the switch behavior can also be achieved by changing the surrounding environment. For example, under sulfur or magnesium starvation, 2.3 or 2.0 mM separately higher titer of 3-hydroxypropionic acid (3HP) was detected[14]. In summary, the growth switch has been demonstrated as a promising and universal alternative to optimize productivity in biotechnology and bioengineering[8].

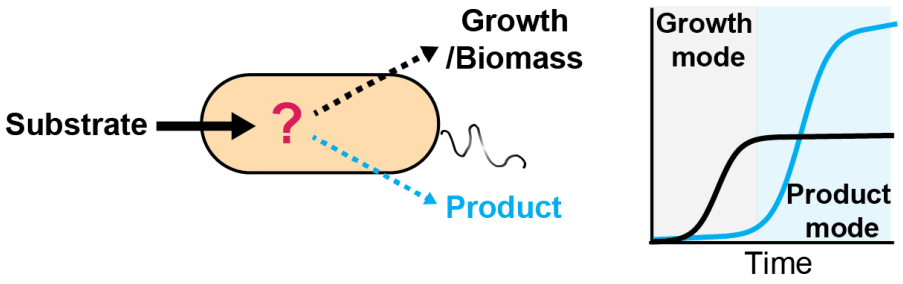


Fig. 1.1 | The concept of the two-phase bioprocess that differentiates growth and production phases. Two-phase bioprocessing to decouple growth from product production. Common ways are through multi-layered regulation to perturb gene expression and switch the cell state from growth to production mode[5, 8, 13, 15].

1.3 Dynamic control of enzyme expression levels by transcriptional feedback loop regulation

In industrial bioprocessing, there is always a struggle between one-stage and two-stage fermentation. The one-stage fermentation is a growth-coupled production; by contrast, like the concept in growth switch, a separated growth and production phase is applied in the two-stage fermentation. Due to the pros and cons in individual systems and diverse constraints, there is actually no strict rules to guarantee a fermentation with the highest productivity[16]. However, it has been shown that growth-coupled

overproduction is highly feasible for five major organisms including *Escherichia coli* (*E. coli*), *Saccharomyces cerevisiae* (*S. cerevisiae*), *Corynebacterium glutamicum* (*C. glutamicum*), *Aspergillus niger* (*A. niger*) and *Synechocystis* sp. PCC 6803[17]. To design a growth-coupled overproducer, how to dictate gene expression through appropriate control is important in designed circuits, and currently, several methodologies have been reported, for example, the dynamic control. Considering the wide usability, dynamic control is regarded as one prevalent and facile strategy for growth-coupled overproducers. Here we review some principles for the achievement of growth-coupled production via dynamical control (**Fig. 1.2**).

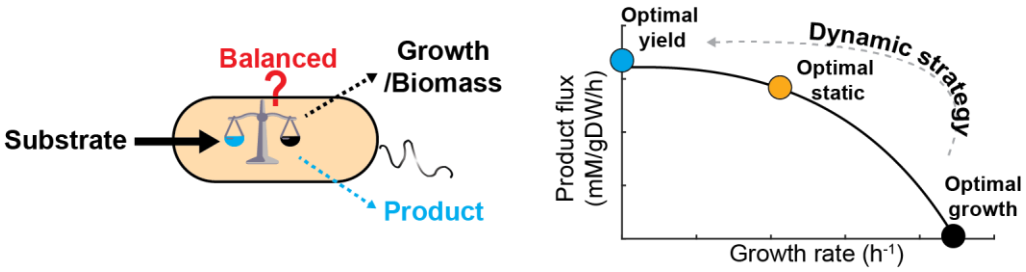


Fig. 1.2 | The concept of the one-phase bioprocess - a dynamic strategy. The dynamic strategy that balances both growth and production and results in the optimal state (yellow circle), bringing the host cells from optimal growth (black circle) to being close to optimal yield (blue circle), is about internal signals (the stress), transcription factors, external environmental changes, etc[18-20].

Natural strategies for dynamic control of homeostasis in the cells have been widely observed through the feedback loop. For example, ArgR, a transcription factor (TF) which inhibits the arginine biosynthetic pathway of *E. coli*, is actually active by the end-product arginine[21-23]. In glycolysis, this highly complex network involves several layers regulation via TF regulators (ex: Cra[24] and Crp[25]), allosteric regulation, posttranslational modifications[26] (ex: phosphorylation of carbon uptake proteins[27]) and oftentimes interplay among each other. Such comprehensive regulation exists even outside the cell to control the cell density by quorum-sensing[28]. Therefore, these well-evolved regulatory mechanisms have been broadly and extensively applied to construct engineered circuits for biosynthetic applications.

Many studies have bloomed up by a combination of native regulated promoters with designed biosynthetic pathways. Native promoters[29] sensing light[30], temperature[31] and substrates[32] are exemplified as a controller responding to

environmental changes. However, when additional integration of engineered circuits is expressed in the host, cells often suffer burden from the toxicity of desired chemicals, protein burden of gene overexpression, oxidative stress of reactive oxygen species (ROS), the variation of metabolites concentration (metabolic burden) and so on. Lots of studies have been successfully applied to solve the extra burden caused by engineered circuits. For example, when the cells face the burden, the burden-driven promoter (P_{htpG1}) controls CRISPR–dCas9-based system to down-regulate gene expression, resulting in higher growth rate and fluorescent protein expression[33]. To avoid the accumulation of harmful compounds, AcuR TF regulator was designed to sense and regulate the concentration of toxic 3-hydroxypropionate, by which the final yield is increased by 23-fold (4.2 g/L) comparing to the previously reported result[34]. Furthermore, the original function of SoxRS regulon sensing the NADP(H) pool to protect the host cells against superoxide[27] and is can be used as an NADPH sensor[35]. Advanced genetic logic gates provide more complex dynamic regulation to sense both environmental or internal varieties and further to achieve applicable digital-like display[36-38].

To sum up, as shown in the mentioned examples, the incorporation of a sensing transcriptional regulation is currently the most prevalent and direct method to enable a dynamic control and to realize balanced growth with improved productivity[39]. To test and expand the usability of this dynamic concept towards more complicated systems, the glycolysis, as a complex and regulatory network, is chosen as the model system in the current thesis, illustrating the efficient resource regulation in the host for balancing between growth and production.

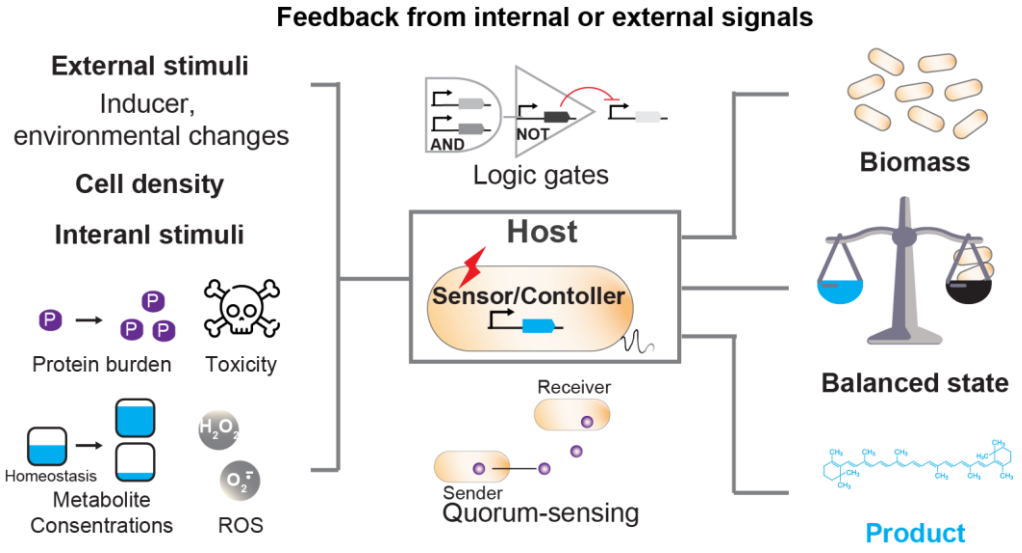


Fig. 1.3 | Feedback loop regulation in the host cells leads to balanced growth and production. A host response circuit typically consists of a biosensor for receiving the internal or external signals and a genetic controller for reacting to the signals. Through environmental responsive promoters, layered logic gate, etc., the cells can control gene expression reaching a balanced state[18, 23, 28, 36, 40, 41].

1.4 Cra-regulator as a flux-sensor to regulate central metabolic pathway

Glycolysis is one of the crucial pathway in almost all organisms, that catalyzes the transformation of carbon molecules into energy sources and building blocks[32]. As a central carbon metabolic pathway, numerous attractive compounds, such as polysaccharides, biofuel, aromatics, and terpenoids, can be derived and converted from glycolytic metabolites (**Fig. 1.4**)[42-44]. According to the concept of "supply and demand"[45, 46], the source of nutrients/glycolysis is defined as the supplier/producer for producing building blocks such as ATP/precursor metabolites, while consuming the goods from the supplier is as demander/consumers such as incorporated synthetic circuits. Under limited nutrients or when the demand is larger than supply, the supplier can no longer meet the requirement of demander. Thus, to compensate for the deficiency, retarded growth of cells is usually detected at the expense of the outcomes between supply and demand. Based on this concept, it is clear that draining too many resources from the supply without proper regulation can collapse the homeostasis in the host cells, restricting the product synthesis reversely. Consequently, a balanced synthetic pathway has to be designed and deployed to coordinate with intrinsic glycolysis under a tolerable disruption between the supplier and the demanding

products. The good news is that *E. coli* has evolved effective regulatory networks in glycolysis via transcription factor-Cra (catabolite repressor/activator).

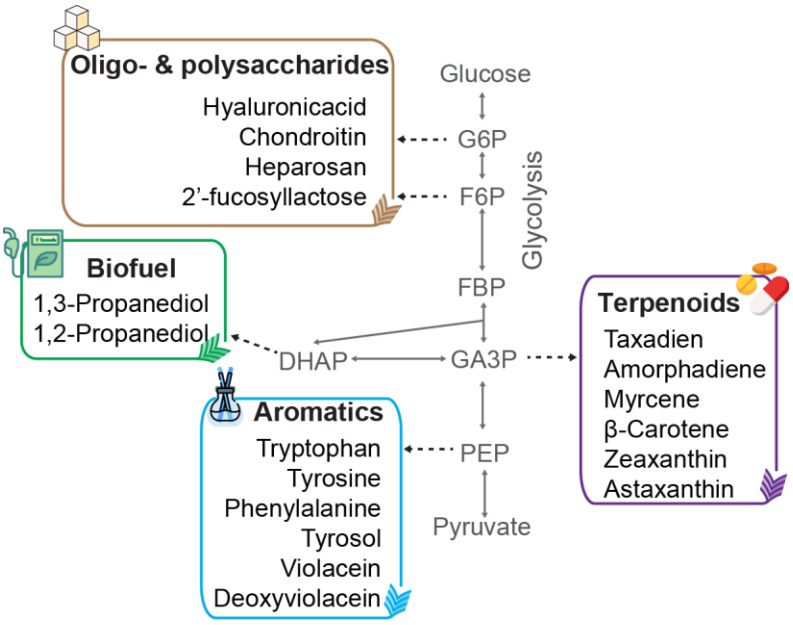


Fig. 1.4 | Bio-based chemicals that are derived from glycolysis metabolites. Glycolysis is important to cell growth. It provides the most of majority precursors for bio-based chemicals like polysaccharides, biofuels, aromatic compounds, and terpenoids[42, 43, 47-50].

Cra has been known as a flux sensor[24, 51, 52] that plays a pleiotropic role to arrange the carbon flow within different metabolic pathways. It acts as an activator of gluconeogenic, TCA and glyoxylate shunt enzymes, and as an inhibitor of Entner-Doudoroff (ED) and glycolytic pathways (Embden–Meyerhof–Parnas pathway) (**Fig. 1.5**). The mechanism of Cra-regulation is at transcriptional level, meaning that the signaling molecule interacts with Cra to prevent Cra binding of its consensus sequences. The distance and the location of the consensus sequence determine whether the Cra acts as an activator/inhibitor of the downstream genes expression[53]. However, the regulatory signal of Cra remains still unclear and has been disputed for a long time[54]. The candidate signaling molecules are fructose-1,6-bisphosphate (FBP) and fructose-1-phosphate (F1P). The individual functional concentration to trigger Cra releasing is millimole (mM) scale in the former and micromole (μM) scale in the latter case[55]. Currently, the studies show that FBP is reflecting the up-glycolytic flux[56, 57]. Although the signaling molecule is ambiguous, many evidences also point out that

depending on the carbon sources, Cra can regulate the central metabolism as either an activator or an inhibitor[52, 56, 58-61].

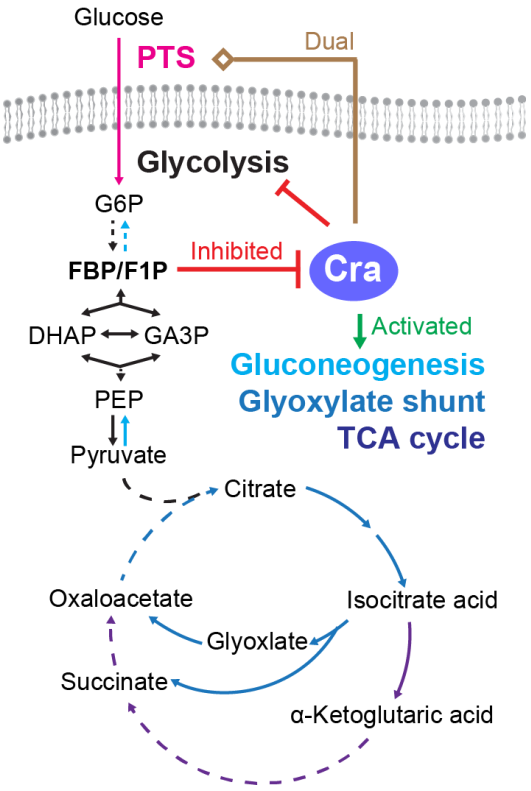


Fig. 1.5 | Schematic illustration of transcriptional regulation by Cra. Cra is a transcription factor that acts as an activator of expression of genes encoding enzymes in gluconeogenic, TCA and glyoxylate shunt enzymes, and as an inhibitor of Entner-Doudoroff (ED) and glycolytic pathways[54-56, 61, 62]. A detail map of Cra-regulated genes and pathways is shown in the **Supplementary figure 1**.

1.5 Current advanced methods for metabolic engineering

1.5.1 CRISPR interference for genetic perturbation

The most widely used of clustered regularly interspaced short palindromic repeats (CRISPR) system comes from *Streptococcus pyogenes*'s type II CRISPR-Cas9 system[63]. This system includes Cas9 endonuclease and mature crRNAs processed by trans-activating crRNA (tracrRNA) with 24-nucleotide complementarity to pre-crRNA[64]. The ternary Cas9-crRNA-tracrRNA complex contributes degrading invading DNA[65]. In genetic engineering of bacteria, crRNA and tracrRNA are synthesized into a single guide RNA (sgRNA)[66-68] which functions similarly to the native crRNA-tracrRNA duplex, guiding Cas9 to a specific locus by dCas9 handle[66]. Generally, the sgRNA is comprised

of 18 to 25 bp sequences of target genes and adjacent to the protospacer adjacent motif (PAM) site according to the applied CRISPR system and the host organism[63]. In *S. pyogenes*, the PAM sequence is NGG or NAG that is more effective in the former than the latter one[66]. The Cas9 and sgRNA complex have been collaterally used in various organisms as an efficient genome-editing tool. Taking advantage of the endonucleolytic activity in Cas9, together with sgRNA, the complex can at first recognize specific genomic sequence, and then directly trigger an on-site double-stranded break (DSB) for homologous recombination[63, 66, 69].

Recently, the nuclease-deficient Cas9 (dCas9) is developed and it contains two-point mutations in RuvC-like (D10A) and HNH nuclease (H840A) domains so that the genomic DNA will not be cleaved after targeting[63, 70]. This deficient dCas9 is so-called CRISPRi system which blocks the transcription of target DNA without cutting the genes [63]. On the other hand, the CRISPR system can also activate gene expression after fusion with functional proteins, such as MS2 coat protein and SoxS protein,[71] giving the so-called CRISPRa system[71, 72]. Thus, the CRISPR system becomes a prevalent way to up and down-regulated genes without the modification of the host's genetic code.

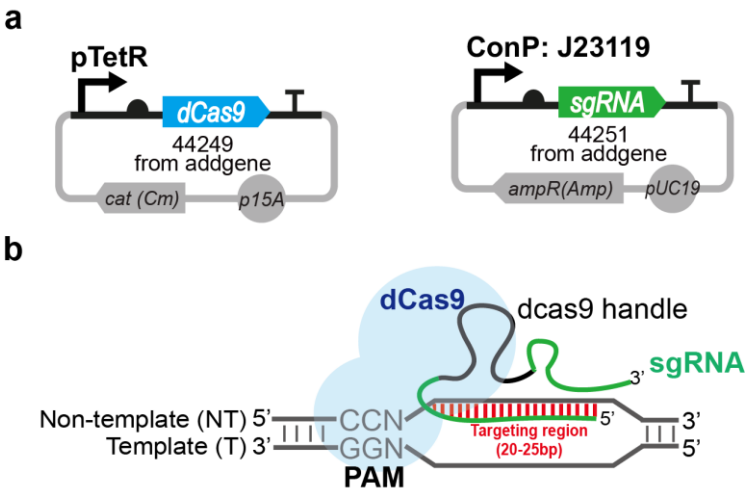


Fig. 1.6 | The CRISPR interference system. a, The CRISPR-Cas9 system includes a dCas9 (44249) and a sgRNA (44251) plasmids. **b**, Schematic illustration shows how dCas9 recognizes the PAM sequence and binds to sgRNA for preventing the transcription of the target gene[68].

1.5.2 Reporter plasmid for metabolic burden quantification

Owing to the improvement of cloning methods, such as error-prone PCR, multiplex automated genome engineering (MAGE)[73], adaptive laboratory evolution, and so on, it becomes easy to create comprehensive strain libraries. It, therefore, raises more attention on how to find a time-saving and cost-effective way to screen the proper strain from libraries. High-throughput screening becomes more important[15]. For example, imaging-based (i.e. microfluidics, single-cell imagination, optical tweezers) and fluorescence-based methods (i.e. fluorescence-activated cell sorting, FACS[74]) are quite common techniques to select the potential strain[75].

We used a fluorescence-based method to monitor the cell state in a real-time manner. The constitutive expression of superfolder GFP was embedded in chromosomal DNA at safe site 9 (SS9)[76]. Following the detection strategy developed by Ellis's group[25], the more stress in the cell is, the less GFP is expressed. Through this method, a series of potential strains can be first evaluated based on the stress response and then sorted for further application and detail analysis.

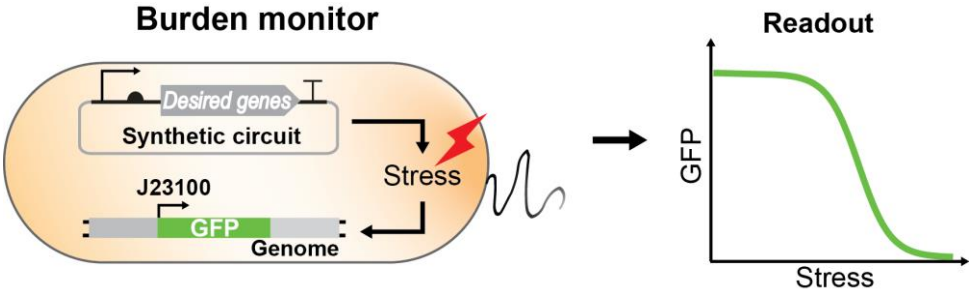


Fig. 1.7 | The mechanism of a burden monitor. The pathway of interest is expressed in the strain with the genomic integration of GFP as a stress-sensing controller. The expression of the synthetic circuit competes for the resources of cells and further causes stress so that the fluorescence level is negatively related to the stress[33].

1.5.3 Cas9-based methods for genome editing

Scarless Cas9 assisted recombineering (No-SCAR)[77] is a method to modify genomic DNA by the combination of CRISPRi and λ-Red recombination system. Combining with the CRISPRi system, sgRNA brings the Cas9 to a specific site to generate double-stranded DNA breaks and then a λ-Red system to recover the break by designed homologous sequences. Bacteriophage λ-Red system is comprised of Gam, Exo, and Bet proteins[78]. Gam prevents RecBCD nuclease from degrading double-stranded linear DNA fragments. Exo has a 5' to 3' double-strand DNA

exonuclease activity which generates 3' linear overhangs DNA, and Bet binds to single-stranded DNA to promote homologous recombination[78]. Compared to the traditional strategy which selecting by antibiotic resistance[79] or TetA-SacB dual-system[80], the No-SCAR system requires no longer both flippase recombination targets (FRT) and dual-selectable marker. The as-designed system provides the advantage of not only the elimination of the antibiotic marker but also short operating time. Through No-SCAR, deletions, point mutations, and replacements can be performed well in target loci without leaving any scar sites[77]. Moreover, the insertion of a larger fragment (>7 kb) becomes highly achievable[65].

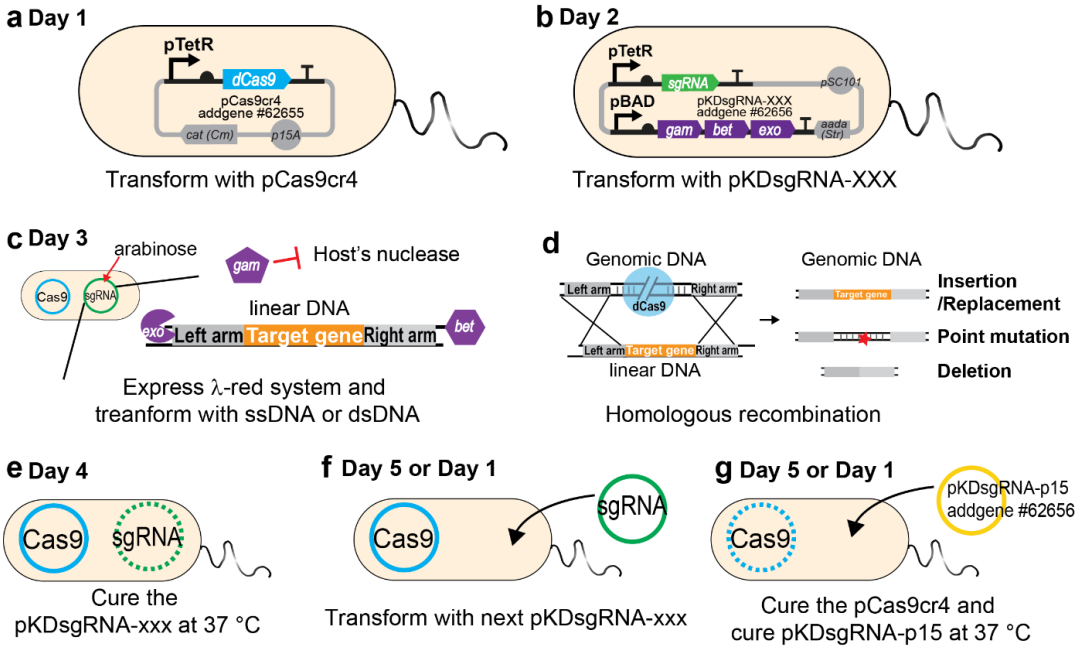


Fig. 1.8 | Genome editing with the scarless Cas9 assisted recombineering (No-SCAR) system. a, On day 1**,** the pCas9cr4 plasmid (addgene #62655) is transformed to the host cell and the cells grow at 37 °C with chloramphenicol (Cm). **b,** On day 2**,** the pKDsgRNA-XXX plasmid is transformed and the cells are plated on LB with spectinomycin (Spec) and Cm at 30 °C. **c,** On day 3, the cells are grown in rich medium (LB or SOB) and is induced with 50 mM arabinose at OD_{600} of 0.5. After ~20 minutes incubation with arabinose, the cells are made electrocompetent and transformed with linear ssDNA or dsDNA. The λ-Red system is expressed to protect the linear fragment from the host's nuclease. **d,** Then the cells are plated on LB with aTc, Spec and Cm and then incubated at 30 °C. At this step, the Cas9 is expressed and makes a double-strand break in the genomic DNA at the designed loci. Through homologous recombination, the chromosomic DNA is replaced/deleted by the linear template DNA. **e,** On day 4, colonies are screened by PCR and the pKDsgRNA-xxx plasmid is cured at 37 °C to get the mutant strain. **f,** On day 5, continue to create another genomic modification with a new pKDsgRNA-xxx plasmid in this mutant strain and repeat the process. **g,** Or cure the pCas9cr4 plasmid with pKDsgRNA-p15 plasmid[77].

2. Chapter 2 – Growth switch via CRISPR Interference targeting RNA polymerase

2.1 Research questions and objectives

Overproduction is one most common and straightforward methods to increase product titer. However, the trade-off between the biomass and targeted products is a restriction for high product titer. A universal strategy to maximize productivity is demanded to balance or even to break such trade-off. Therefore, as shown in **Fig. 2.1**, to test the suitability in our systems, two general approaches- growth coupled (one-phase) or growth switch (two-phase), are demonstrated and evaluated individually for the glycerol overproduction circuits.

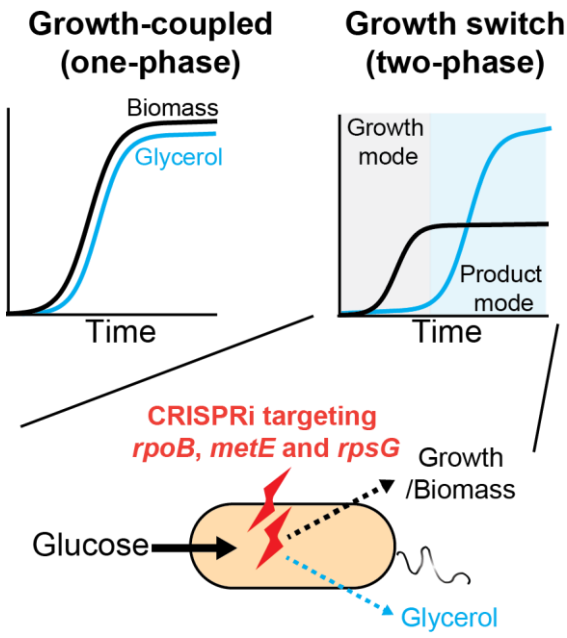


Fig. 2.1 | Illustration of one-phase and two-phase bioprocesses. To understand which processes, growth coupled or growth switch, is more feasible to glycerol overproduction. We applied the CRISPRi system to target *rpoB*, *metE* and *rpsG* genes for growth switch production.

2.2 Results and discussion

2.2.1 Selection for a suitable metabolic pathway

To study the trade-off between growth and production, we selected the glycerol production pathway from *Saccharomyces cerevisiae* SEY6210 as the case study. This pathway consists of two proteins- GPD1 (glycerol-3-phosphate dehydrogenase 1) and GPP2 (glycerol-3-phosphate phosphatase 2) (**Fig. 2.2**). According to previous researches, the GPD1-GPP2 fusion protein, which has 44 bp deletion and one base insertion between GPD1 and GPP2, shows higher catalytic efficiency and lower K_m for DHAP (dihydroxyacetone phosphate) and could form substrate channeling to convert the intermediate glycerol 3-phosphate (glycerol-P) into glycerol efficiently[81]. Consequently, we adapted the GPD1-GPP2 fusion protein into our systems for efficient glycerol production. Additionally, due to the simplicity of the introduced pathway and a less toxic product, the constructed system serves as a good model to get more insight of the trade-off effect between host growth and desired production.

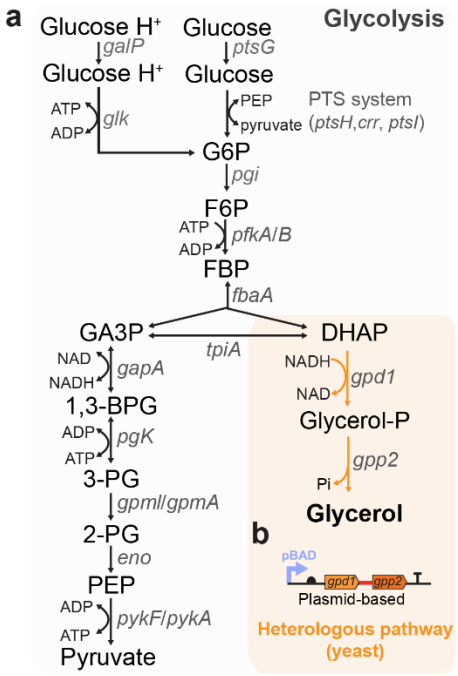


Fig. 2.2 | Glycolysis and a heterologous glycerol production pathway in *E. coli*. **a,** Glycerol production from glucose is a catalyzed process through glycolysis. **b,** The plasmid of GPD1-GPP2 fusion protein was constructed under the control of pBAD promoter. G6P, glucose 6-phosphate; F6P, fructose 6-phosphate; FBP, fructose 1,6-bisphosphate; GA3P, glyceraldehyde 3-phosphate; 1,3-BPG, 1,3-bisphosphoglycerate; 3-PG, 3-phosphoglycerate; 2-PG, 2-phosphoglycerate; PEP, phosphoenolpyruvate; DHAP, dihydroxyacetone phosphate; Glycerol-P, glycerol 3-phosphate.

2.2.2 The selection of suitable targeted genes to achieve growth switch

Once the synthetic pathway was established and functional, we used clustered CRISPRi interference (CRISPRi) system to inhibit growth. First, we excluded that the inducer of dCas9 expression (aTc) caused a defect. Therefore, different amounts of aTc was added and the result shows that wild-type cell can tolerate up to 1000 nM of aTc (**Fig. 2.3a**). We then measured the growth of sgRNA targeting different genes-*rpoB* (RNA polymerase subunit β), *metE* (cobalamin-independent homocysteine transmethylase) and *rpsG* (30S ribosomal subunit protein S7). The results show that, compared to sgRNA-rpoB-1 targeting to template strand (-), the aTc with limited influence on the growth of *E. coli* and sgRNA-rpoB-2, targeting to non-template strand (+), presented actually the strongest effect on the growth (**Fig. 2.3b**). Conversely, the sgRNA of *metE* and *rpsG* genes didn't lead to severe growth effect even under higher induction level (500 nM aTc) (**Fig. 2.3c**). The significantly retarded growth was observed in the sgRNA-rpoB-2 even under lower aTc concentration (62.5 nM). The observed phenomenon is consistent with the previous study, indicating that a sgRNA binding to the non-template strand (+) equips stronger downregulated capability than the one binding to the template strand (-)[68]. Other proofs have provided in studies that design growth switchers by controlling rpoBC (encoded by ββ' subunit of RNA polymerase)[12]. The author replaced the promoter region of *rpoBC* with an IPTG inducible promoter to turn genes on or off via IPTG, revealing that that owing to the properly remained metabolically active of the host, the high product yield is still achievable by simply turning off the *rpoBC*[12]. Based on these results, we chose the sgRNA-rpoB-2 to arrest growth of the glycerol production strain for switching from growth to producing phase.

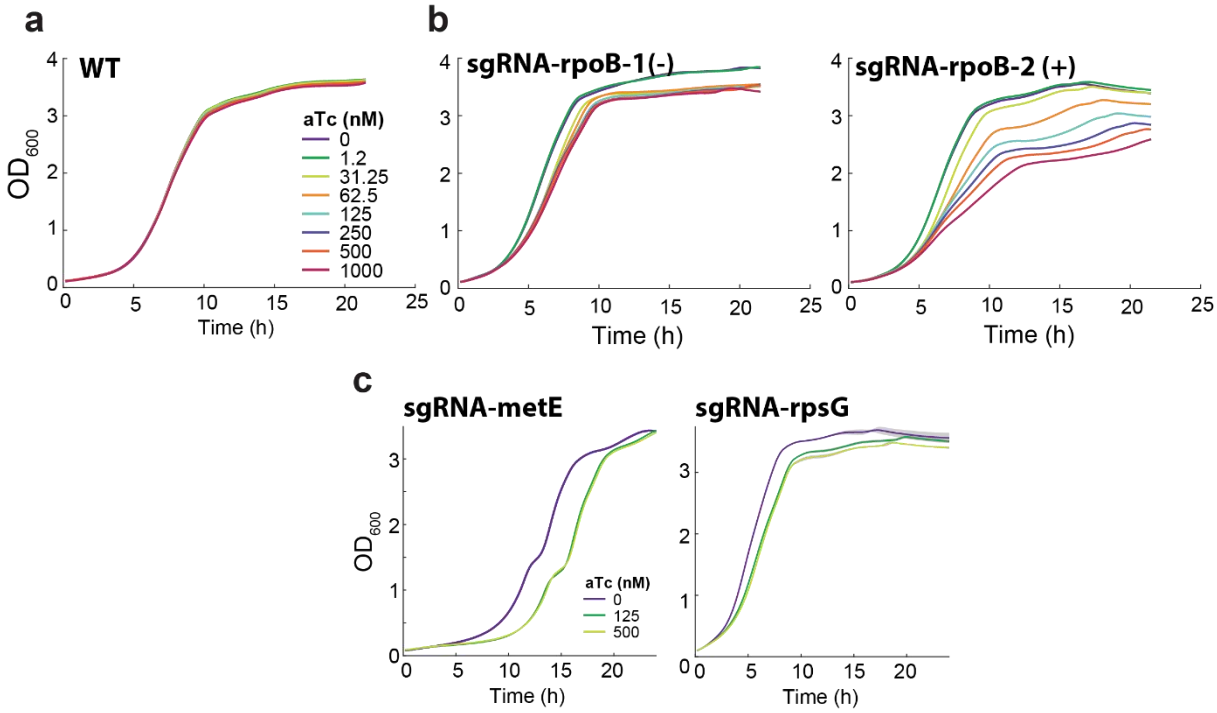


Fig. 2.3 | Growth characteristics of *E. coli* expressing sgRNAs that target different genes. a, Titration of aTc inducer in wild-type cells from 0 to 1000 nM. **b**, The growth of sgRNA-rpoB-1 and sgRNA-rpoB-2 targeting *rpoB* at template (-) and non-template (+) strands. **c**, The growth of sgRNA-metE and sgRNA-rpsG targeting *metE* and *rpsG* separately. All the strains were incubated in a 96-well plate filled with M9 medium containing 0.5% glucose. (n=2).

2.2.3 Using growth-switch process for glycerol production with sgRNA-rpoB-2 targeting RNA polymerase

Base on **chapter 2.2.2**, CRISPRi system to down-regulated RNA polymerase (*rpoBC*) is applied. To reduce the complexity of the expression system, we inserted whole glycerol production cassette in the loci of *lacZ* gene by scarless Cas9 assisted recombineering (No-SCAR), which is a method to modify genomic DNA with the combination of CRISPRi and the λ-Red recombination system[77]. The whole cassette is comprised of a pBAD inducible-promoter driven GPD1-GPP2 fusion protein (4217 bp) and an *rrnB* terminator (**Fig. 2.4**). The as-constructed strain is denoted as pBAD-G in this thesis (**Supplementary Table S1**).

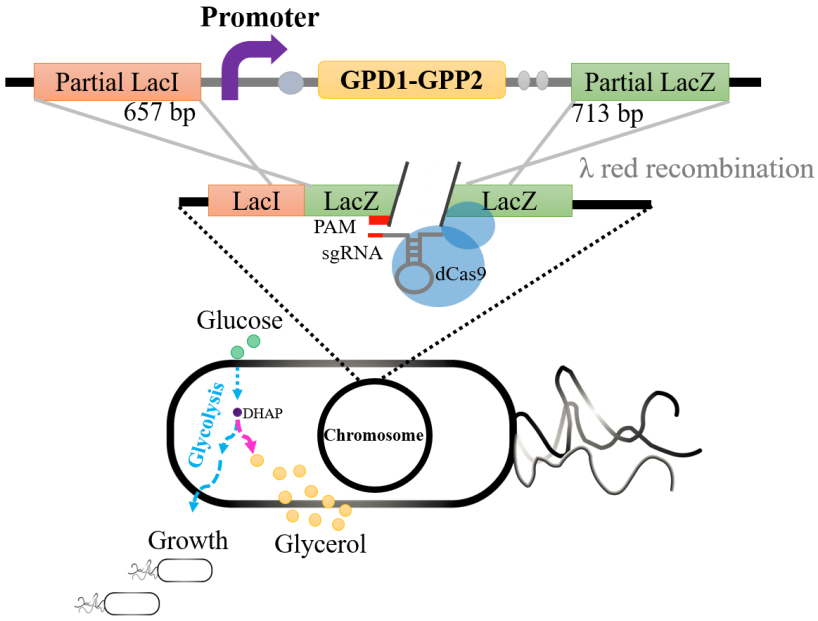


Fig. 2.4 | Insertion of the GPD1-GPP2 fusion protein in the *E. coli* genome. A pBAD inducible promoter driven GPD1-GPP2 fusion protein (4217 bp) and an *rrnB* terminator were integrated to the *E. coli* genome at loci of *lacZ* gene through No-SCAR genome editing method.

The pBAD-G strain was utilized to express the CRISPRi system for downregulation of RNA polymerase (sgRNA-rpoB-2), and we induced glycerol production pathway by adding 0.8% arabinose in the M9 medium. The growth experiment showed that without aTc inducer, the glycerol concentration increased correlative with the growth, and started to decrease within 10 h (**Fig. 2.5a**), indicating that the glycerol pathway is growth coupled production and after 10 h the *E. coli* re-used the synthesized glycerol as a carbon source. **Fig. 2.5b** shows that the *E. coli* suffers severe growth defect especially at the beginning of induction, but the production phase is decoupled from the growth. When the induced production was conducted under OD_{600} at 0.34 (**Fig. 2.5c**), the glycerol concentration was much higher than induction in the beginning. The results point out that, although it is possible to separate the growth and production phase, the titer of glycerol depends mainly on the total biomass. In other words, the more biomass in the system is, the higher glycerol titer can be reached.

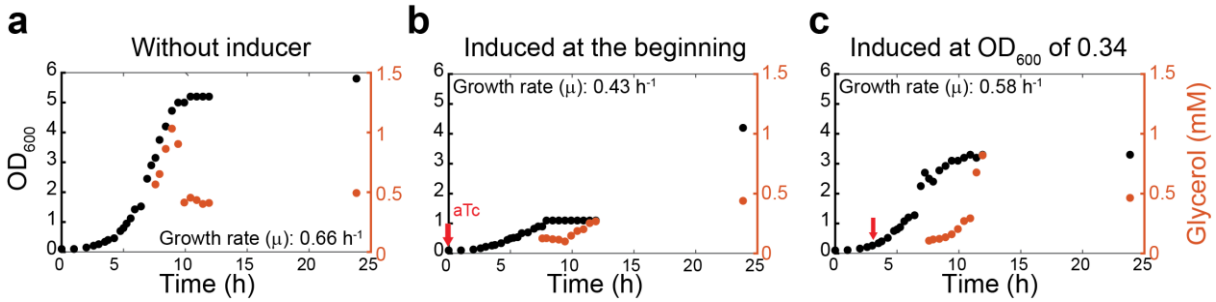
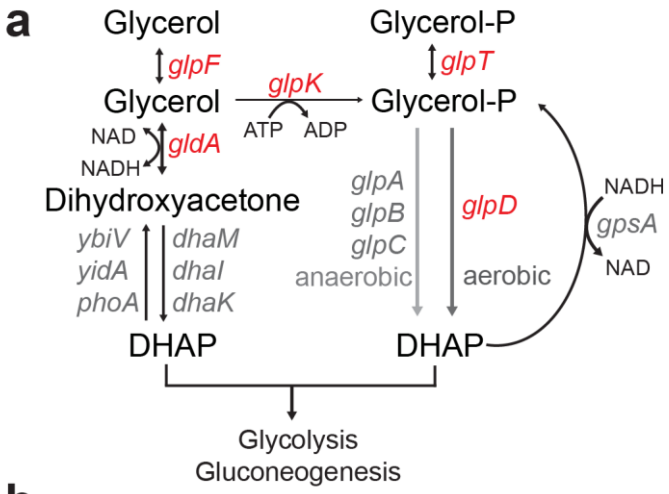


Fig. 2.5 | Growth curves and glycerol production of pBAD (pBAD-G) strain bearing glycerol genes in genomic DNA with the expression of CRISPRi to down regulate RNA polymerase (*rpoB*). a, No aTc to induce CRISPRi system of sgRNA-rpoB-2 targeting RNA polymerase. **b,** 500 nM aTc was added at the beginning of the cultivation (t =0 h) **c,** 500 nM aTc was added when the OD reached 0.34. All the experiments were done in a shake flask with M9 medium containing 0.8% arabinose and 0.5% glucose. Dark dots represent the OD_{600} and brown dots represent the glycerol concentration in mM scale.

2.2.4 Deletion of *glpK* gene to avoid glycerol reutilization

E. coli can use glycerol as a carbon source and to improve production the reuptake glycerol should be considered as well[82]. To decide which gene needed to be removed, we first selected 5 candidate genes (**Fig. 2.6a**) related to glycerol degradation from the Keio collection (*E. coli* BW25113 strain)[83] and cultured them with either 0.5% glucose or 0.5% glycerol as a sole carbon substrate in M9 medium. We found that two genes-*glpK* and *glpD* are highly essential for glycerol utilization in *E. coli* (**Fig. 2.6b**).



b

	Δ glpF	Δ glpK	Δ glpT	Δ gldA	Δ glpD
Glucose (0.5%)	O	O	O	O	O
Glycerol (0.5%)	O	No growth	O	O	No growth

Fig. 2.6 | The impact of gene deletions in the glycerol degradation pathway of *E. coli* on glucose and on glycerol. a, Five genes (*glpF*, *glpK*, *glpT, gldA* and *glpD*) that participate in glycerol degradation were selected and marked in red color. **b,** Essentiality for glycerol utilization. There five strains with deletions of genes in (a) were taken from the Keio collection (*E. coli BW25113* strain) and tested for growth in culture tubes with 0.5% glucose or glycerol as a sole substrate. Hollow circles represent the cells grown. *glpF*, glycerol facilitator; *glpK*, glycerol kinase; *glpT*, sn-glycerol 3-phosphate:phosphate antiporter; *gldA*, glycerol dehydrogenase; *glpD*, aerobic glycerol 3-phosphate dehydrogenase.

To test whether the knockout of *glpK* or *glpD* gene can prevent the re-cycle behavior, instead of using the pBAD-G strain, we cloned two additional glycerol producer strains which were the glycerol plasmid under control of a constitutive promoter J23105 and a native promoter of *epd* gene in a high copy number plasmid to exclude the reuptake of glycerol. As a proof of principle, a higher glycerol concentration is detected in both knockout strains than in wild-type after 24 h cultivation (**Fig. 2.7**), evidencing that reuptake of glycerol is effectively prevented via *glpK* or *glpD* gene deletion.

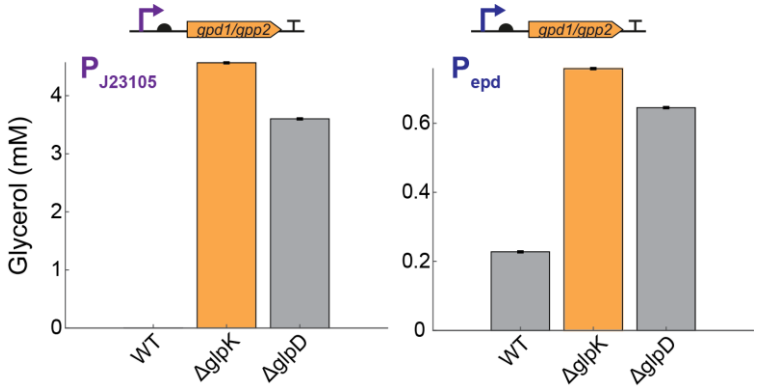


Fig. 2.7 | Glycerol production in a Δ*glpK* or Δ*glpD* strain. The GPD1-GPP2 fusion protein was under control of a constitutive promoter J23105 or a native promoter from *edp* gene in a high copy number plasmid (pSB1A2). The cells were cultured in a 96-well plate and the supernatants were taken for the glycerol measurement after 24 h.

The *E. coli* BW25113 strain of Keio collection is a closely related strain from *E. coli* MG1655 K-12, but it is missing several genes presenting in MG1655[83] and has 29 genetic variations[84]. These genetic differences have been reported that they influence the expression of 17 genes and further affect the transcriptional regulation between the strians[85]. Based on previous studies, we decided to use fewer genetic variations and more wild-type-like strain (MG1655) for our study. As shown in **Fig. 2.7**, compared to the BW25113 wild type and *glpD* system, the *glpK* deletion always presented higher glycerol production regardless of comprised promoter. Thus, we created *glpK* deletion in MG1655, which is unable to use glycerol as a carbon source (**Fig. 2.8**), as a background strain to investigate the glycerol production regulation in the following chapters.

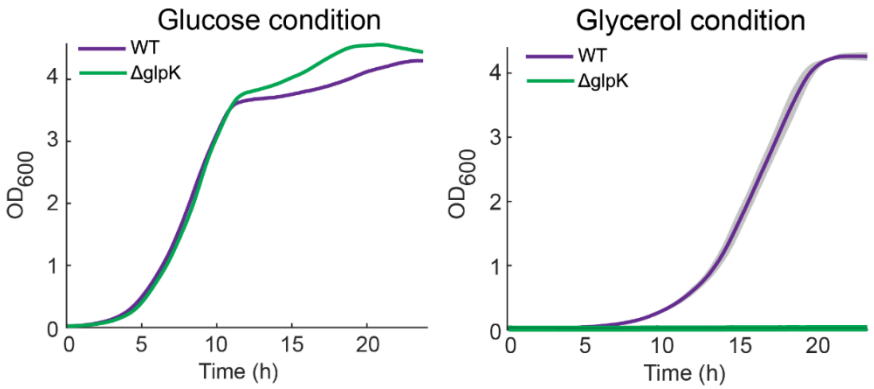


Fig. 2.8 | Growth curves of the Δ*glpK* strain in *E. coli* MG 1655. The *glpK* removed *E. coli* MG1655 was cultured in M9 minimal medium using glucose or glycerol as the sole carbon source.

2.3 Conclusion

We have demonstrated that the growth-switch (two-phase) glycerol production is possible to apply for biosynthetic production. However, in our case study, it suggests that glycerol production in *E. coli* is likely to be a growth-coupled process, suggesting that it would produce more glycerol by maintaining certain growth instead of non-growth or limited growth. Additionally, to avoid the calculations error resulting from host reusing, the deletion of *glpK* gene was executed in the constructed strains. After the preliminary study and preparation, in the next chapter, we attempted to develop a balanced method to reach a higher glycerol productivity.

2.4 Future work

2.4.1 Growth switch achieved by changing cells' metabolic networks or environments

There are other ways to achieve near-zero cell growth but stay metabolic activity, such as downregulation of TCA cycle genes (*sucA* and *aceA*)[9], downregulation of replication as well as the nutrient depletion[14].

2.4.2 Growth switch achieved by replacing proper promoters to control heterogeneous genes

The other strategies achieve the growth switch is by using sensor related promoters, such as temperature sensitive CRISPRi system[10], environment-sensing genetic circutis[13], metabolites-sensing promoters[86] and so on. For this type of strategy, we recommend the review by Ross Kent and coworkers[5].

3. Chapter 3 – Dynamic control of enzyme expression with the transcription factor- Cra

3.1 Research questions and objectives

We wanted to understand how to control the heterologous glycerol pathway and how to balance the trade-off between the growth and glycerol production. According to the results from **chapter 2**, glycerol production was likely to be a growth-coupled pathway. Therefore, it was important to balance both growth and production. To control enzyme expression, we used the transcription factor Cra because it is the main regulator of glycolysis and gluconeogenesis. How to coordinate this Cra-regulation with glycerol overproduction and what the regulatory mechanisms are involved would be discussed in this chapter. Finally, we confirm this Cra-regulation is applicable to other glycolysis-derived product such as carotenoid.

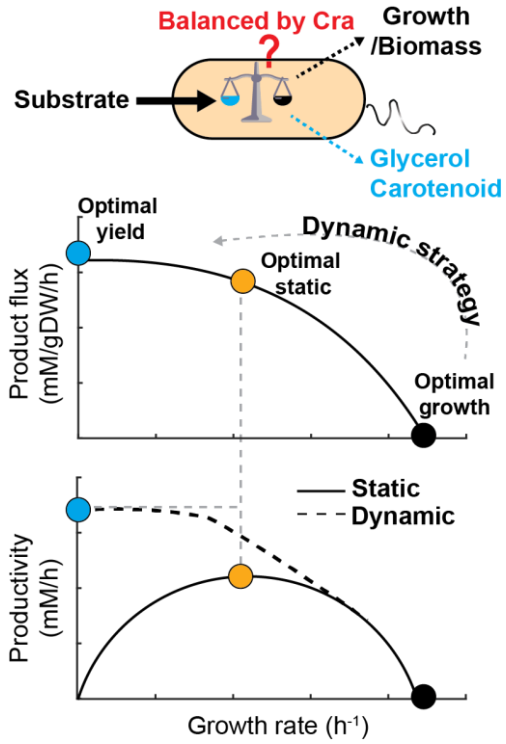


Fig. 3.1 | One-phase bioprecesse. To maximize productivity, integration of Cra-regulation in both glycerol and carotenoid synthetic plasmids can dynamically reach the optimal enzyme levels for increasing product flux from the direction of optimal growth to fairly near an optimal yield.

3.2 Results and discussion

3.2.1 Construction of novel glycerol production strains

To prevent the accumulation of intracellular intermediate (glycerol-3-phosphate, glycerol-P), we separated the fused *gpd1* and *gpp2* genes into two expression segments equipped with the individual promoter. To control flux into the pathway we controlled expression of *gpd1* by an arabinose inducible promoter and used several constitutive promoters to control expression of *gpp2* (**Fig. 3.2**).

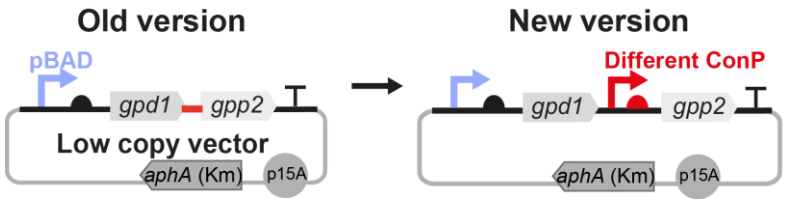


Fig. 3.2 | Two designs of the synthetic glycerol pathway. Left:The GPD-GPP2 fusion protein. Right: the fusion protein was separated and each gene was equipped with a promoter. The arabinose inducible promoter was kept for *gpd1* gene and several different constitutive promoters from the Anderson Collection were used for the *gpp2* gene.

We selected constitutive promoters from Anderson's family promoters[87] based on promoter activity (**Fig. 3.3a**). As demonstrated in **Fig. 3.3a**, a series of constructs with varying promoter activity were designed and test in glucose condition. To realize proper dynamic equilibrium, the systems with similar (J100B34), too strong (J01B34) or too weak promoters activity (J09B34) were first excluded, and all the other five derived promoter-RBS sets were then cloned in front of the *gpp2* gene for the replacement of fused protein of GPD1-GPP2. Apart from using glucose as a carbon source, the consistent promoter activity was observed in both fructose (**Fig. 3.3b)** and mannose (**Fig. 3.3c)** as well, indicating that the promoters maintain relative activities among different conditions[88].

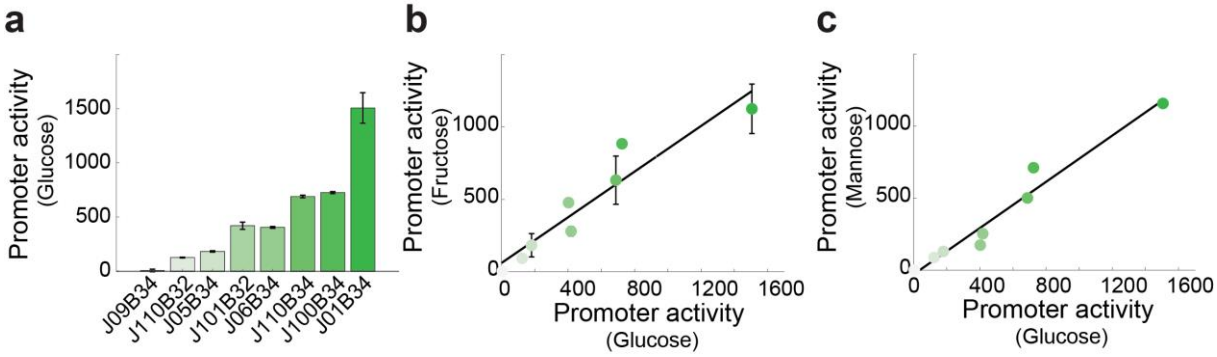


Fig. 3.3 | Promoter activities of the Anderson's promoter library. a, The constitutive promoters under glucose condition. **b,** Promoters display proportional activities under fructose verse glucose conditions. **c**, Same as in (b), for mannose and glucose. For condition-specific analysis, all the strains were grown in 96-well plate and M9 medium with 0.5% of a specific sugar. Error bars show 2 biological replicates.

With these five strains, we first conducted the growth experiment to know how they grew under 3 arabinose concentrations. Compared to the *glpk* deleted strain which is hereafter called wild-type (WT) in this thesis, all five derived strains showed a strongly retarded growth at 0.3% arabinose (**Fig. 3.4a**). Besides, in all induction levels, the intracellular metabolites, glycerol-3-phosphate (glycerol-P), of all five strains were significantly lower than the fused version at 0% arabinose (**Fig. 3.4b**). The blue line in **Fig. 3.4b** represents the reference amount coming from the wild-type strain and points out that J110B32 and J01B32 strains produced relatively less glycerol-P closed to the wild type. However, with similar growth curves, at 0.1% arabinose, the other 3 strains (J05B34, J06B34, and J110B34) showed higher glycerol-P production than the J110B32 and J01B32. Based on these data, it suggests that the growth is not mainly influenced by the accumulation of glycerol-P, but could be severely restrained by draining DHAP due to the imbalanced supply-demand condition. The glycerol concentrations were then normalized to biomass (OD_{600} multiplies parameter 0.37) for comparisons between samples. For better comparisons between samples with distinct growth behavior, the glycerol concentrations were again evaluated after normalized to biomass (OD_{600} multiplies parameter 0.37). Except for the J110B32 with lightly lower productivity, the glycerol productions were actually quite similar within all strains (**Fig. 3.4c**). Taken together, we decided that the J01B32 strain was used for the following process because of the lower glycerol-P accumulation and higher glycerol production.

Therefore, the J01B32 strain was further investigated and this strain is hereafter named as pBAD-only in the following sections.

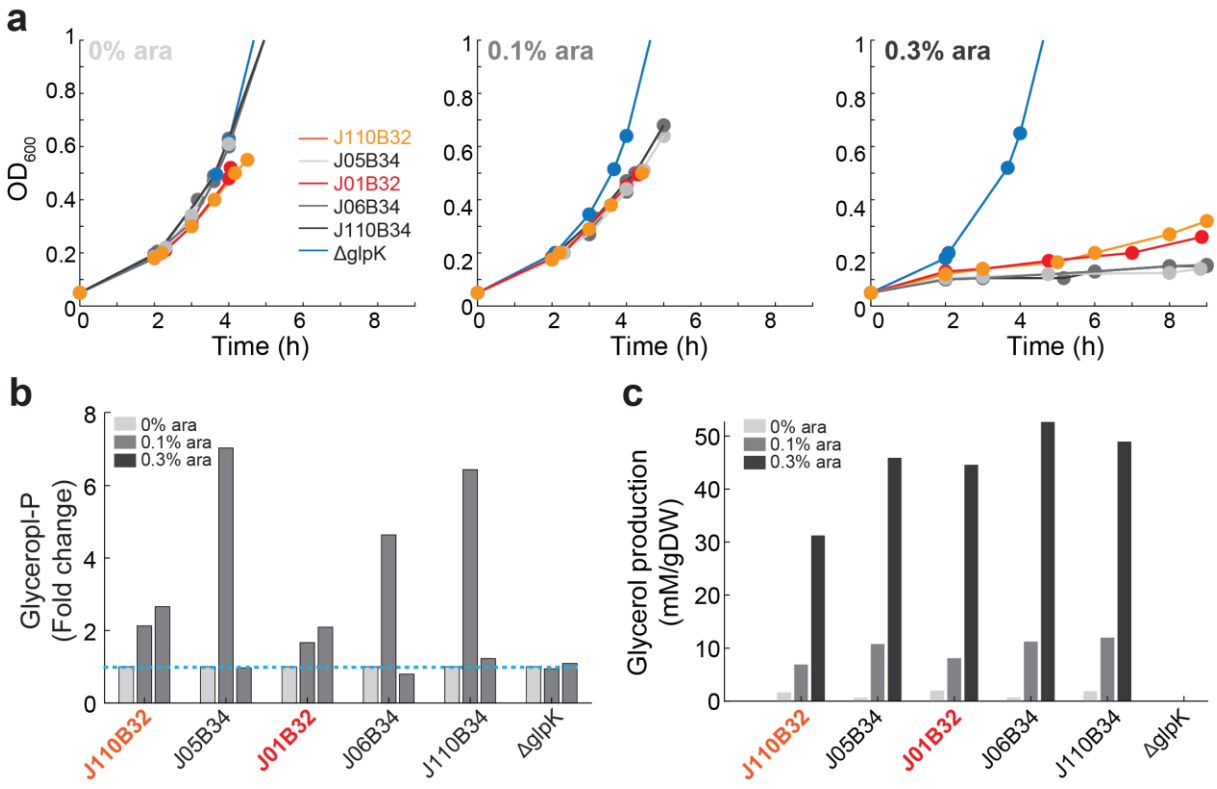


Fig. 3.4 | Characterizations of five strains expressing *gpd1* and *gpp2* genes separately with its own promoters. a, Growth curves of 5 strains bearing the same pBAD promoter driving gpd1 gene and five version of consitutive for *gpp2* gene. The three plots shows 3 different arabinose concentrations (0%, 0.1%, and 0.3%). b, Intracellular concetration of glycerol-3-phosphate (glycerol-P). Intracellular metabolites were taken at OD_{600} of 0.5 or equal biomass amount for slow-growing strains which didn't grow over 9 h cultivation. Blue dashed line is glycerol-P level in *glpk* deleted strain (also called wild-type in this thesis). **c,** Glycerol production was glycerol concentration normalized to biomass. All the strains were grown in 500 mL shake flask with M9 medium supplemented 0.5% glucose. Red and orange color mark the two strains with similar physiological characteristics.

To characterize the pBAD-only strain, we cultivated both glycerol and GFP production strains in 96-well plates with different arabinose concentrations. Distinct from the GFP case, the glycerol-producing strain suffered strong growth reduction so that the glycerol production was not a positive correlation (**Fig. 3.5b**). This experiment

pointed out that glycerol production, as a part of complex networks, is manipulated via other factors within the critical central metabolism.

In general, there were 3 potential reasons which could lead to the constrained glycerol production; that is, the overexpressed exogenous protein, the toxic intermediates, and the overload glycolysis. Referring to the GFP result in **Fig. 3.5a**, the overexpression of exogenous protein was able to be excluded because by using the promoter with the same activity a consistent growth phenotype was presented by the GFP strain (pBAD-only-GFP) even under 2% arabinose treatment. As a result, in the following chapters, the potential two reasons will be discussed. For better understanding, we would like to reemphasize that the constructed GFP strains were used as a non/less-stress experiment reference and *glpk* deletion of *E. coli* MG1655 as a background glycerol production strain in all the experiments of glycerol production.

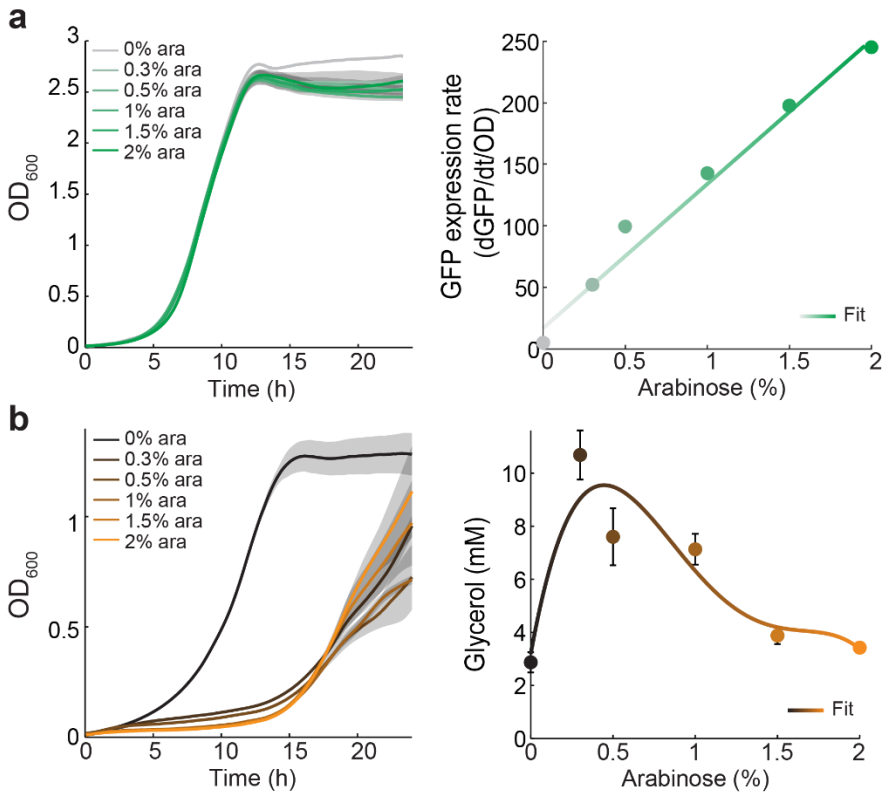


Fig. 3.5 | The correlation level between arabinse and production of GFP or glycerol. a, Growth curves of GFP production strain and GFP expression rate of the pBAD promoter at different arabinose levels. GFP fluorescence and OD_{600} was measured with n = 3 plate reader cultures. **b,** The same as **Fig. 3.5a** for glycerol strain. The glycerol concentrations were measured after 24 h.

3.2.2 Glycerol production pathway overloads host's glycolysis

As **Fig. 3.5** shows that glycerol production is not correlated to protein expression levels, to get more perspective on the dynamic regulation within the glycerol-producing strain, we performed the growth experiment in a shake flask and measured metabolites and proteins at 0.5 of OD_{600}. **Fig. 3.6a** reviews again the comparison between *E. coli*'s native glycolysis and the synthetic glycerol production pathway. We tested how a varying expression-level of *gpd1* affects host growth and glycerol titers by different induction levels with 0%, 0.1%, to 0.5% of arabinose (**Fig. 3.6b**). Under low arabinose (0.1%) induction, the growth curve shows high resemblance to the non-induced one (0% arabinose treatment), but can produce > 5-fold higher glycerol. However, by increasing the induction strength to 0.5%, a strongly reduced growth was detected without a substantial improvement of glycerol production (>7 fold). In other words, with 0.1% arabinose, cells preserve enough capacity of glycolysis (sufficient supply) to sustain basic growth behavior and to conduct the mission of glycerol production (affordable demand). By contrast, with 0.5% arabinose, cells have to struggle for growth or production. Thus, based on the result, we assumed that glycerol production directly affects metabolism in the way of draining either glycolytic metabolites or energy.

To understand how the external glycerol production interconnects with host metabolism, we measured metabolites at 3 different induction levels. Metabolites remained relatively constant at 0.1% inducer, and only the intermediate of the glycerol pathway (glycerol-P) increased ~4-fold compared to the un-induced culture (**Fig. 3.6c**). Nevertheless, at 0.5% inducer, the metabolite levels changed markedly, and all glycolytic metabolites decreased. Among all 96 measured metabolites, the direct glycerol-precursor DHAP showed especially the strongest decrease (**Supplementary figure 2**), suggesting that DHAP consumption is the primary burden resulting from glycerol overproduction. Because glycerol-P levels were similar in both 0.1% and 0.5% cultivation, the result also proves that the accumulation of glycerol-P (phosphate stress) actually not causes Influential disturbance in host's growth. Furthermore, a slightly decreasing in both ATP and NADH levels is observed as well, advocating that instead of the protein accumulation the overloaded glycolysis is the most critical burden for glycerol overproduction in *E. coli*.

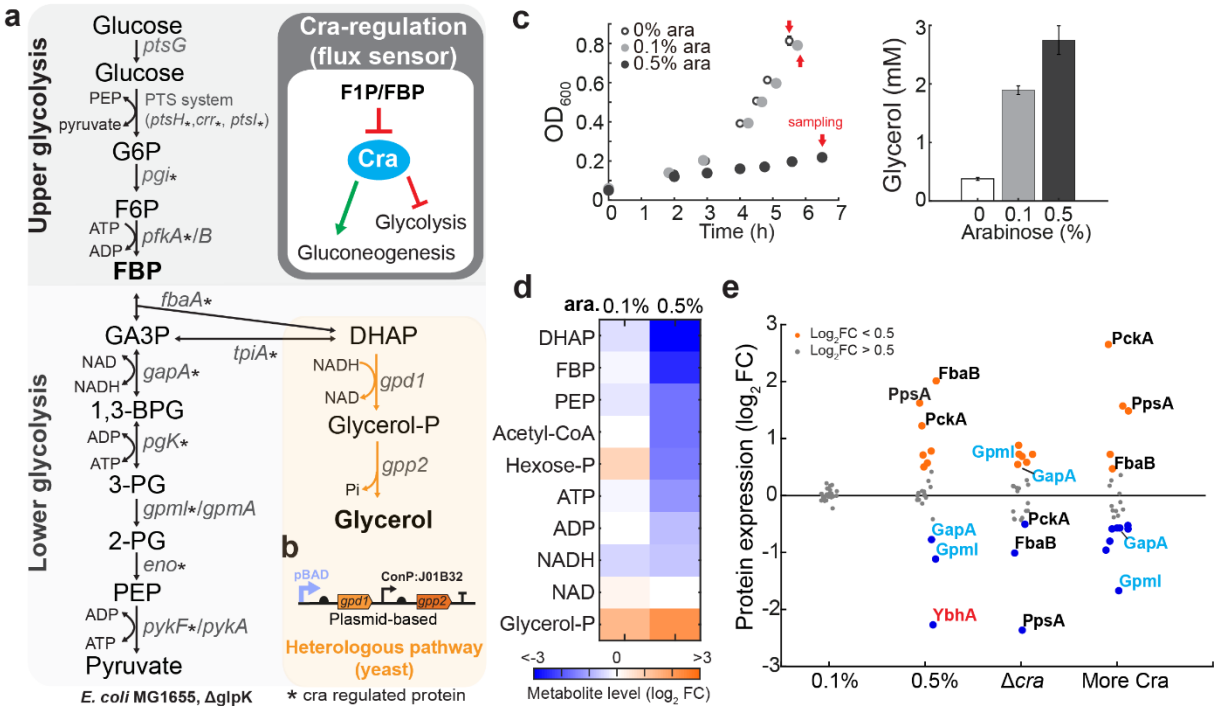


Fig. 3.6 | Growth burden and perturbations in metabolome and proteome caused by glycerol overproduction. a, Schematic overview of the native glycolytic (gray background) and the heterologous glycerol pathway (orange background). The synthetic pathway consists of the two enzymes from yeast-GPD1 and GPP2. A more detailed map is in **Supplementary figure 1**. **b,** Glycerol synthetic pathway comprises of GPD1 and GPP2 form yeast in the plasmid-based expression system which is *gpd1* gene under the control of an arabinose-inducible promoter pBAD and *gpp2* is expressed constitutively with J01B32 promoter. **c,** The growth of pBAD-only strain and glycerol titer at different arabinose levels. The supernatant was taken for glycerol measurement when OD_{600} reached 0.8 or equal biomass for the slow-growing cells (red arrows). **d,** Relative fold-change of major metabolites involved in glycerol production was normalized to 0% treatment. **e,** Relative protein levels of glycolysis and gluconeogenesis. Blue color represents down-regulated proteins at 0.5% arabinose. All strains were grown in a shake flask with M9 medium supplemented with 0.5% glucose and the samples of metabolites and proteins were taken at OD_{600} of 0.5. (n = 3) The abbreviations are in appendices 0 (page xi).

Since the metabolic levels had dramatically changed, the protein levels would have been altered due to the protein-metabolite interaction[89, 90]. As illustrated in **Fig. 3.6e**, the varied protein expression in direct glycerol-producing pathways, including glycolysis and gluconeogenesis, is significantly distinct between 0.5% induction (highly altered) and 0.1% induction (less altered). The result is in line with metabolic results. Intriguingly, the proteins involved in gluconeogenesis (FbaB, PpsA, and PckA) are the top three high values at 0.5% induction and the glycolytic proteins (GapA and GpmI)

decrease conversely by half (**Fig. 3.6e**). The result hinted that the switch between glycolysis and gluconeogenesis is governed by a transcription factor named as Cra (catabolite repressor activator, or fructose repressor, called FruR). As mentioned in **chapter 1.4**, Cra plays as a flux sensor by sensing the concentration of FBP/F1P which reflects the upper-glycolytic flux and distinguishes glycolytic from gluconeogenic substrates[56, 59, 91]. To further confirm these protein changes is attributed to Cra regulation, we compared 0.5% arabinose with *cra* deleted strain (Δ*cra*: no-Cra strain has no growth phenotype under glucose condition, **Supplementary figure 3**) and wild-type strain in gluconeogenic substrate-succinate (more-Cra: Cra-active state[92]). Induced by 0.5% arabinose, the glycerol-producing strain exhibited a pattern similar to Cra-active state, but is reverse to Δ*cra* except ybhA which converts FBP to F6P. This result suggests that glycerol producer attempts to replenish the lower concentration of glycolytic metabolites (FBP, DHAP and PEP) by the upregulation of gluconeogenic proteins.

In previous studies, the Cra activity is adjusted according to surrounding conditions via sensing external parameters, such as glycolytic (glucose) or gluconeogenic (succinate) substrate[51, 92, 93]. Here, we reported the evidence that the changing activity of Cra is subjected to the internal signals instead of outside environmental changes, for example, the shortage of metabolic flux resulting from non-balanced equilibrium between basic growth and product synthesis. Altogether, glycerol overexpression leads to misregulation of *E coli*'s native regulatory networks by activating gluconeogenesis and deactivating glycolysis. The former response causes extra flux burden, and the glycerol overexpression which drains glycolytic metabolites further deteriorates the internal metabolism bias within host. Thereby, the challenge was to coordinate this "glycolytic flux burden" between glycerol overproduction and growth, namely a solution to ease the additional demand of DHAP overloading the glycolysis.

3.2.3 Learning metabolic control from natural regulatory systems

As mentioned before, Cra activity is adjusted according to surrounding conditions[51, 92, 93]. In order to balance both growth and glycerol overproduction, we therefore thought whether we could use Cra regulator to adjust glycerol overproduction.

First, we tested native Cra-activated promoters P_{ppsA} and Cra- inhibited P_{pykF} in shake flasks with either glucose or pyruvate as the only carbon resource. The Cra activity can be altered by the utilization of glucose or pyruvate, giving high or low glycolytic flux to fructose-1,6-bisphosphate (FBP) respectively[51, 93]. The amount of FBP is higher in glucose condition so that the inhibition of Cra-regulator is higher compared to pyruvate conditions (gluconeogenic substrate), and vice versa (**Fig. 3.7a**). Therefore, we demonstrated the "on" and "off" state of Cra-regulator activity by cultivating the cells in pyruvate or glucose M9 medium. Cra-activated promoter (P_{ppsA}) has higher glycerol production in pyruvate conditions (Cra protein at on state); on the opposite, Cra-inhibited promoter (P_{pykF}) has higher glycerol production in glucose conditions (**Fig. 3.7b**). Thus, when glucose is employed as a sole carbon sources, Cra-inhibited promoter is likely to a logical and potential strategy to express glycerol gene.

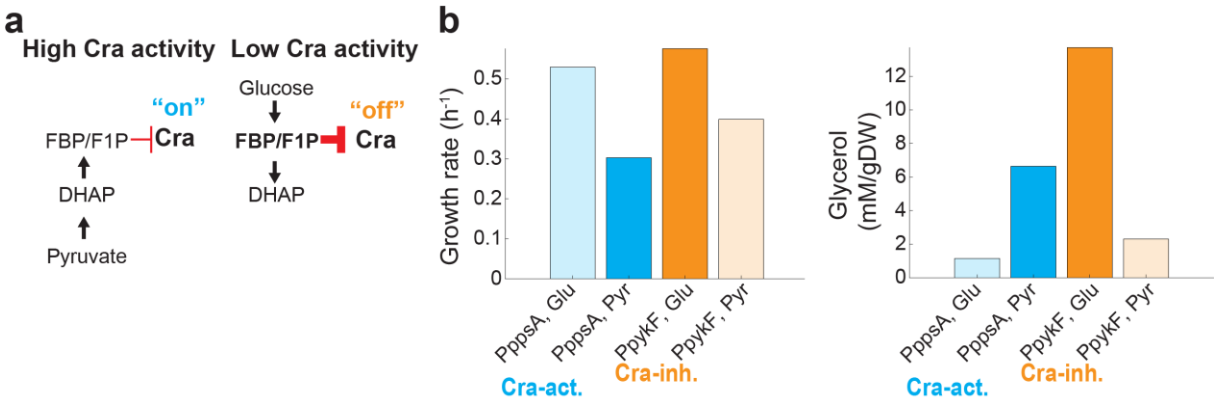


Fig. 3.7 | Glycerol pathway expression with native promoters. a, The logic of Cra-regulation. In pyruvate conditions, the concentration of metabolite-FBP/F1P is lower compared to glucose condition, so the inhibition of the Cra-regulator is lower, which means Cra-regulator is more active in pyruvate than in glucose conditions. The bold letter of FBP represents the concentration is relatively higher. **b**, Two strains with native promoters for glycerol production were grew in shake flasks in M9 medium with 0.07% glucose or pyruvate. The glycerol concentrations were measured when the OD_{600} reached 0.5. FBP, fructose-1,6-bisphosphate; DHAP, dihydroxyacetone phosphate.

Besides, more comprehensive studies regarding native promoters for expressing GFP (**Fig. 5.4**) and glycerol (**Fig. 5.6**) will be discussed later in **chapter 5**. Here, we addressed that only Cra-activated promoters increase GFP production in pyruvate condition and Cra-inhibited promoters increase GFP levels under glucose condition

(**Supplementary figure 4**). These data suggest that Cra-regulator indeed plays a critical role in the gene expression and can be used as an environmental response controller.

3.2.4 The effect of Cra-regulation on GFP expression

The above results inspired us to consider whether any type of promoter could be tuned into a Cra-regulated chimeric promoter. Considering that in most cases, glucose is the primary carbon source provided in the culture medium, the Cra-inhibited version was therefore selected to match the regulatory logic according to **Fig. 3.7a**. Based on the literature survey, the 18 bp of Cra consensus sequence (AGCTGAAGCGTTTCAGTC, **Supplementary Table S2**) from the *epd* gene was chosen for construct due to the reported high affinity toward Cra-regulator[80]. However, to find a suitable position in the desired promoter for Cra recognizing was another unknown question. As a result, we cloned four versions of Cra consensus sequence, namely at -10 region or within the region between -35 and -10 by using two kinds of constitutive promoters, J_{23101} or J_{23105}. Moreover, as a proof-of-principle, we replaced the native Cra-binding site of the P_{pfkA} promoter by a scrambled sequence (CACGAGAGAACAACGTAA, **Supplementary Table S2**) that has the same length as original the 18 bp Cra-binding site but cannot be recognized and regulated by any transcription factors in *E. coli*[93, 94]. As shown in **Fig. 3.7**, constitutive promoter J_{23101} represented 75% inhibition capacity at -10 region and 54% inhibition between -35 and -10 region, while J_{23105} had 64% and 49% inhibition separately. This demonstrates that by adding Cra-binding site, the promoter attracts Cra-regulator to fine-tune the strength of gene expression. Beside this, we figured out that the inhibition effect is more effective at -10 location than between -35 and -10 due to the competition between Cra-regulator and RNA polymerase. When the Cra-binding site was removed, the promoter activity of P_{pfkA} increased by 7-fold. The results confirm that the regulation by Cra can be modified by inserting and removing of the consensus sequence in promoters.

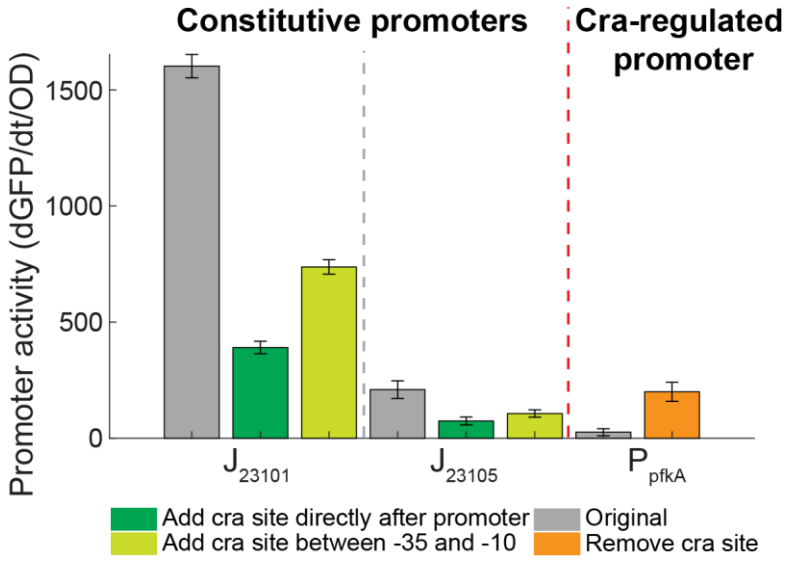


Fig. 3.8 | The GFP expression with synthetic and natural promoters under Cra-regulation. The Cra consensus sequence was inserted to two constitutive promoters (J_{23101} and J_{23105}) or was removed from the native P_{pfkA} promoter which is regulated by Cra. The inhibition effect of the Cra consensus sequence inserted at the loci of -10 or -35 and -10 region of the promoter was measured. On the other hand, the native P_{pfkA} promoter replaced the consensus sequence with the scrambled sequence, the same length of Cra-consensus sequences, was also conducted as a comparison. The consensus sequence used in this thesis can be found in **Supplementary Table S2**.

However, to further demonstrate that the regulation effect was indeed operated by Cra, we deleted the *cra* gene of wild-type MG1655 and *glpK*-deleted strains. The former is used for GFP producers and the latter is used for glycerol producers, by which the reutilization of glycerol can be avoided. These mutant strains show no growth retardation in glucose condition (**Supplementary figure 3**). In △*cra* background, when cells grew in glucose M9 medium, the GFP production was truly reduced in the Cra-activated promoters and increased generally in the Cra-inhibited promoters (**Supplementary figure 5**).

Next, we also used the modified promoters to confirm again the effect of Cra-regulator. As the result of **Fig. 3.8**, under wild-type background, the GFP production was reduced in the modified version of constitutive promoters (J_{23101} and J_{23105}), and native Cra-regulated promoter conversely increase. However, when expressed in Δ*cra* strain, the promoter activity increases a lot in the promoters with additional Cra-binding site (green to light green bar), which means that the inhibition capacity is actually given

by Cra (**Fig. 3.9**). The above results inspired us that Cra-regulator could serve as an universal and facile controller to adjust the overexpression of interested pathways.

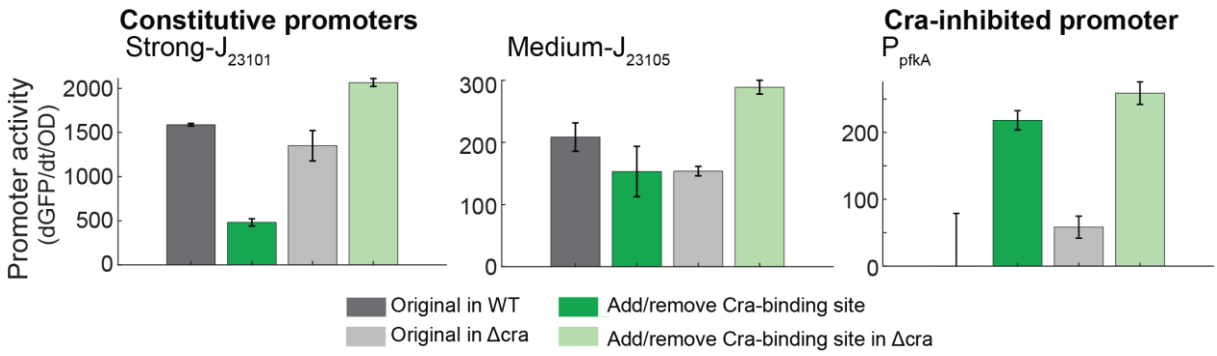


Fig. 3.9 | GFP expression levels with Cra-regulated promoters in wild-type *E. coli* and a Cra deletion strain. The promoters added or removed Cra-binding site were expressed in wild-type (WT) or Δ*cra* strain. (n=2).

3.2.5 Coupling to the Cra-regulation attenuates glycolytic flux burden and stabilizes glycerol overproduction

At first, the GFP was again used to examine the effect of an additional Cra-binding sequence. To confirm and apply Cra-regulation as a controller in glycerol production, unlike constitutive promoters (**Fig. 3.8**), we didn't add Cra-binding site at -10 region of pBAD promoter. Alternatively, we cloned the pBAD-Cra-GFP strains with Cra binding sequence between the pBAD promoter and RBS to avoid the competition between Cra-regulator and RNA polymerase. Compared to original pBAD-driving GFP strain (pBAD-only-GFP), the strain with additional Cra-binding site (pBAD-Cra-GFP) has lower GFP expression as excepted, while the strain produces more GFP in Δ*cra* background (**Fig. 3.10a**). Interestingly, by switching from wild-type to *cra* deleted background, pBAD-Cra-GFP strain regained the GFP expression level (slop increases from 47 to 110) closed to pBAD-only-GFP strain's performance (slop is 117). The results indicate that the Cra-regulation system is achieved only when Cra exists in the cells. Otherwise, it should perform like the scrambled sequence, meaning that the promoter remains always higher activity no matter which background strain is used (**Fig. 3.10b**). That is, Cra-regulator has to bind first to its consensus sequence and then represses the downstream gene expression.

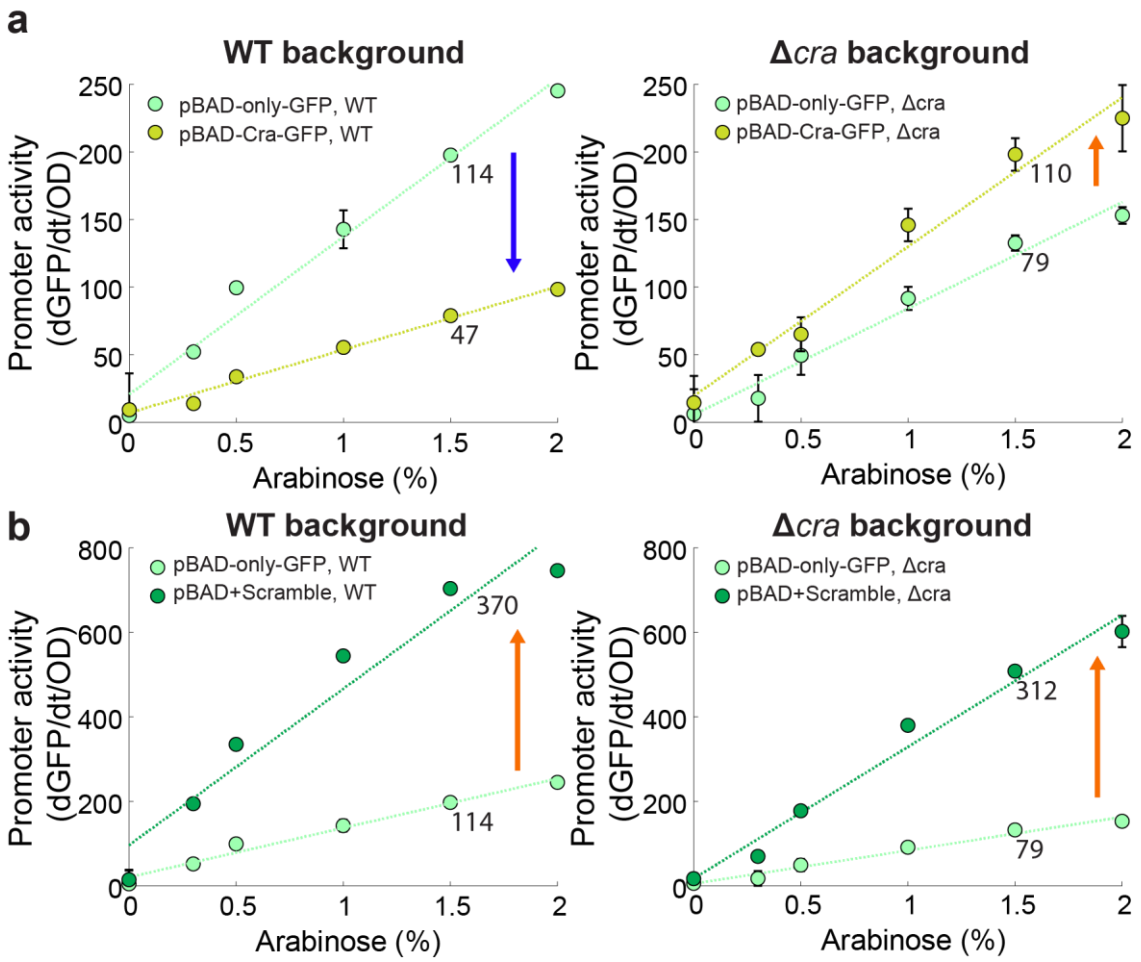


Fig. 3.10 | Activities of Cra regulated promoters in WT and Δ*cra* background. a, Comparison between the strains without or with Cra-binding site (the former is pBAD-only-GFP and the latter is pBAD-Cra-GFP). **b,** The pBAD-only-GFP strain compared to pBAD+scramble strain. The number closed to the line is the slop. Arrows mean GFP production level compared to original pBAD-only-GFP strain. (n=2)

So far, we knew that Cra regulator can recognize the Cra-binding site and repress GFP expression which is not a harmful pathway for *E. coli*. We then applied the same approach to the glycerol production pathway, a more metabolism-involved and resource-draining system. We cloned in the first step two pBAD promoters, without (named pBAD-only) and with Cra-binding site (named pBAD-Cra), to drive glycerol genes. These two plasmids were then transformed into *glpK* and *cra* deleted strains separately, forming 4 different kinds of constructs. Following the strategy in GFP expression, we also tried to construct scrambled sequences in the current systems, but we could not get any colony. Referring to the GFP results in **Fig. 3.10b**, it implied

that the promoter with scrambled sequences could be too strong to produce glycerol due to the resulting internal stress.

Among these four strains, only pBAD-Cra strain with functional Cra-regulator maintained suitable growth together with enriched glycerol titers; on the other hand, the strongest growth retardation and less production occurred in pBAD-only system expressing in WT background (**Fig. 3.11**). The growth of the pBAD-Cra system was stable even at high arabinose-induction (3.5%), and glycerol titers were more than 2-fold higher (53.91 mM) than with the pBAD-only system (22.28 mM) (**Supplementary figure 6**). Thereby, integration of native Cra-regulation achieves higher glycerol titers within a wide range of inducing concentrations; and by contrast, a narrower functional window for glycerol production is traced in the pBAD-only system including both WT or Δ*cra* background strains (pBAD-only, WT and pBAD-only, Δcra). Interestingly, the result also shows that when the pBAD-only system expressed in Δ*cra* instead of WT strain, this strain increases the stability of growth and production as the previous study showed that using Δ*cra* as a vector can alter metabolic gene expression and further diminish the burden from designed plasmid[95]. Conversely, the improved growth by Δ*cra* does not be observed in the pBAD-Cra system, indicating the balanced effect is truly regulated by Cra and reaches maximized stabilization in the WT background. In summary, the result shows that the combination of transcription factor Cra with glycerol synthetic pathway can improve not only internal adaptability but also robustness against variation of GPD1 expression and is able to extend the tolerance range of glycerol production.

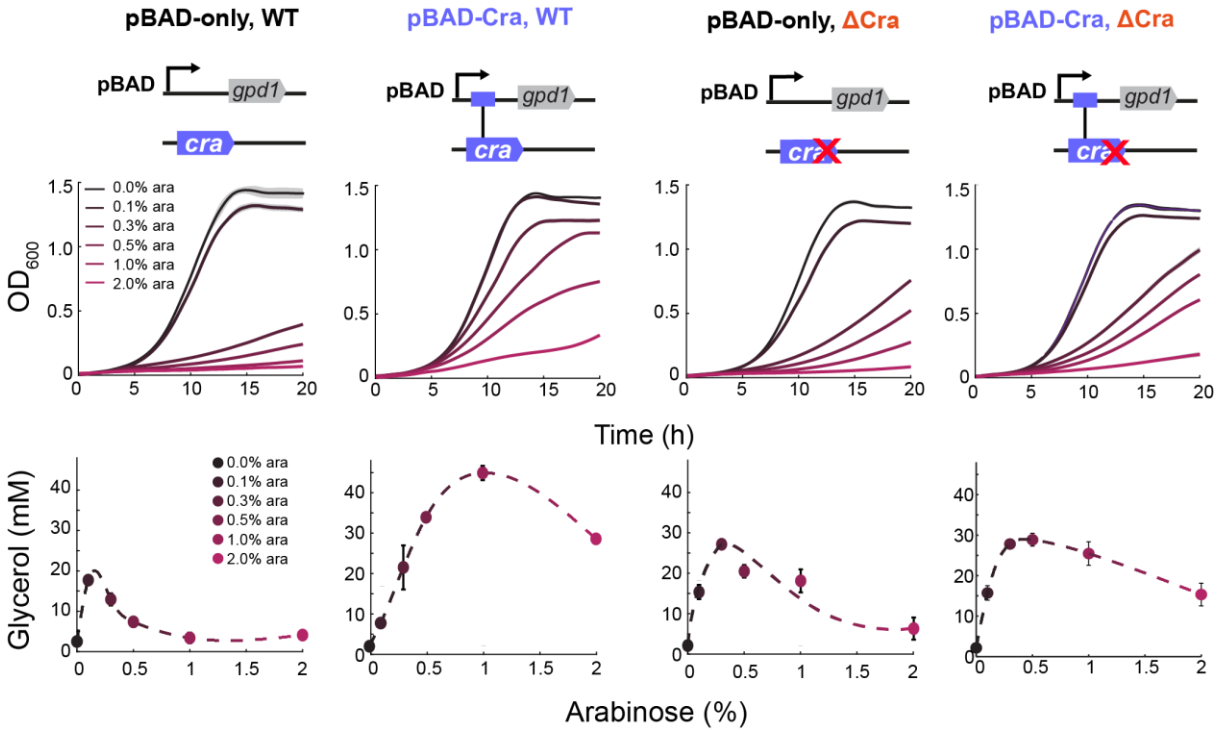


Fig. 3.11 | Growth and gylcerol production with a Cra-regulated promoter. Growth and glycerol titer at different inducer concentrations. The pBAD-only and pBAD-Cra plasmids were expressed in both Δ*glpk* (or called wild-type strain referred in glycerol producer) and Δ*cra* background separately. All the strains were cultivated in M9 medium with 0.5% glucose and glycerol concentration was measured after 24 h in a 96-well plate. (n=2).

3.2.6 Confirmation about Cra involving in fine-tuning gene expression

The question was raised from why pBAD-Cra strain can stabilize the glycerol overproduction but pBAD-only strain cannot. For the observed phenomena, we proposed three hypotheses: 1) the observed balanced effect is achieved under the regulation of additional Cra-binding site; 2) there is no/less crucial competition which divides the Cra-regulator pool from intrinsic Cra-regulated genes in the host; 3) the population heterogeneity was existed in pBAD-only strain due to the overloaded glycolysis.

To address the stabilized effect exclusively via Cra-binding site, another native transcription factor (TF) binding sites were constructed and these TF regulators participate in glucose catabolism (Crp[96] and Mlc[97]), the mediate at the transition of O_2 level (Fnr[98]), maintenance of nucleoid relative cellular processes (Fis[99]) and iron homeostasis (Fur[100]). In the same construct of pBAD-Cra strain, we inserted the

consensus DNA binding sequences (**Supplementary Table S3**) of each TF at the position between the pBAD promoter and the ribosomal binding site (RBS) of *gpd1* gene (**Fig. 3.12a**) and expressed these five plasmids to obtain five TF-coupling strains. The growth and 24 h glycerol titers at different arabinose-levels of the TF-coupling strains are shown in **Fig. 3.12b**, revealing that the improved growth and glycerol titers are only achievable in the Cra-regulated system. Besides, these five TF-coupling strains show similar characteristics as the pBAD-only strain that maintains a proper growth rate at 0.1% induction but reduces growth rate over 0.5% arabinose induction. The result confirms that the Cra-regulation phenomenon indeed originates from the inserted Cra-binding site, by which the as-designed Cra-regulated system can sense the glycolytic flux and adjust downstream gene expression.

Next, as the figure shown (**Fig. 3.12c)**, to rule out the effect from competition of Cra regulator within the host cell due to additional Cra-binding site, we constructed 1 or 2 additional Cra-binding site in both systems. Namely, if the observed balanced effect is merely resulted from enhanced glycolytic flux due to additional Cra-binding site which competes for the intrinsic pool of Cra-regulator, the pBAD-only system with 1 or 2 Cra sites should improve glycerol titer. However, the experimental results declined this explanation, because no matter how many binding sites was added there is no significant change in glycerol production. Therefore, the dynamic control via Cra is the more convincing and suitable explanation for the glycerol overproduction in host with improved stress adaptability and production.

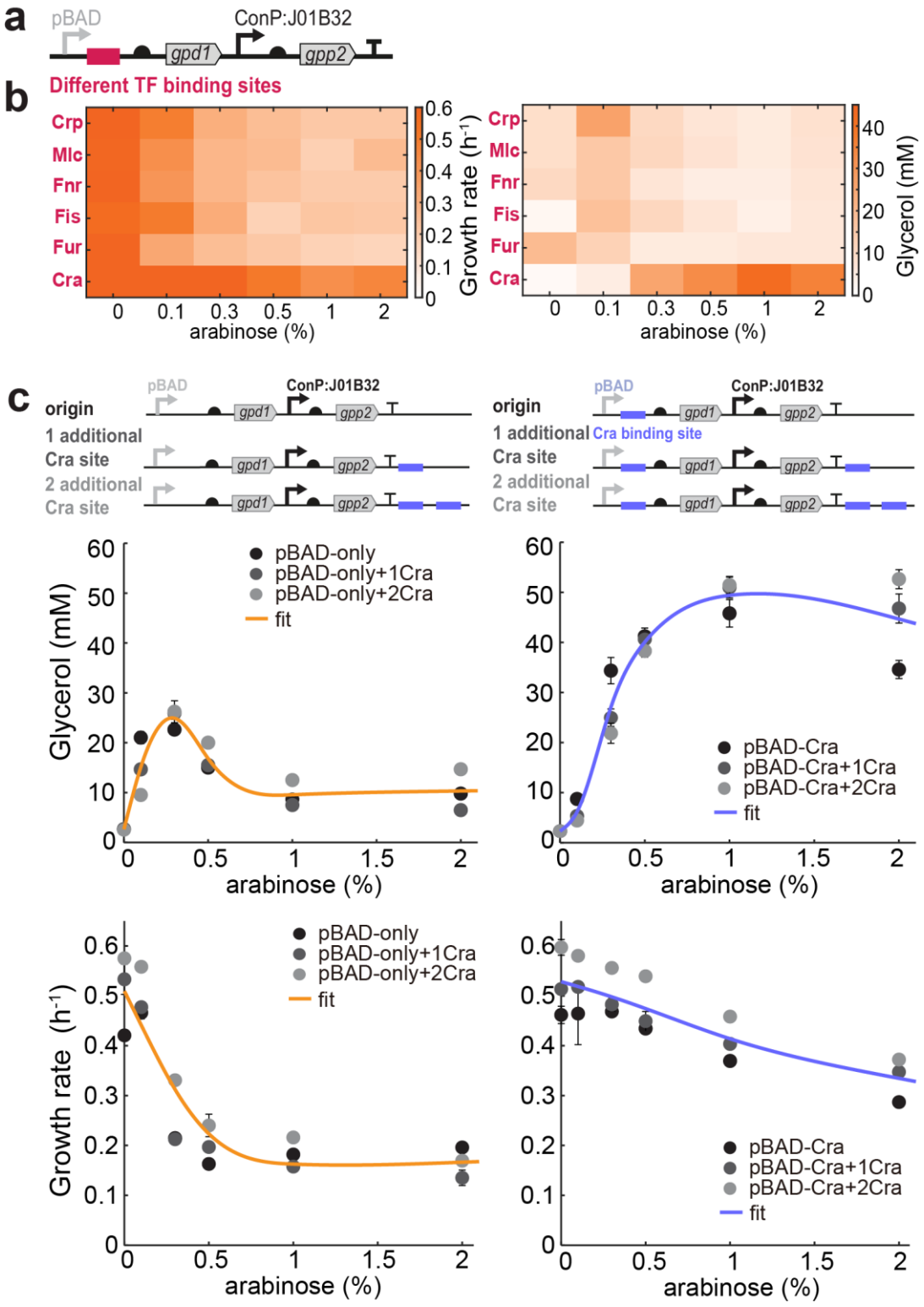


Fig. 3.12 | The fine-tuning gene expression by Cra. a, Different TF-binding sites were inserted between the promoter and the RBS of the *gpd1* gene to get 5 types of TF-coupling systems. **b,** Growth rates (left) and glycerol titers (right) of each TF-coupling strain were cultivated with different arabinose levels in a 96-well plate. **c,** The additional Cra-binding site were cloned separately with original pBAD-only and pBAD-Cra systems. The glycerol titers and growth rates of pBAD-only (left) and pBAD-Cra series (right) strains were measured and calculated after 24 h cultivation. (n=2).

Moreover, population heterogeneity is also a common problem in bioprocesses and leads to lower productivity[6]. To detect whether subpopulation was induced in the pBAD-only strain, the GFP sequence was fused directly to *gpd1* gene as a reporter protein (**Fig. 3.13a**). Among all induction levels, both glycerol producers (**Fig. 3.13b**) displayed single peaks not overlapped or two separated peaks suggesting that the population is homogeneous in these two strains. Therefore, we confirmed that the pBAD-only strain could not produce as much glycerol as the pBAD-Cra strain is not due to the population heterogeneity. Apart from the detection of population heterogeneity, the GFP signal can also be utilized for tracing the GPD1 expression level. By merging the signals in **Fig. 3.13b** and **Fig. 3.13c** shows that over 0.1% induction, the GFP expression levels of pBAD-only were higher than pBAD-Cra strain, which implied the GPD1 overabundance is the culprit for lower production in the pBAD-only strain.

Considering the distinct physiological reactions presented by two glycerol producers, pBAD-Cra and pBAD-only strains, responding to glycerol production, series of experiments were conducted to elucidate the detail mechanism. We had excluded the competition effect of internal Cra with glycerol production plasmid and confirmed that only Cra-binding site at Cra-inhibited position between the promoter and RBS region can improve product titer without losing growth rate. As a result, through cooperating with Cra, the pBAD-Cra strain maintains balanced metabolism by adjusting growth and production dynamically, but the pBAD-only strain with misregulation in glycolysis suffers severe growth impact resulting from GPD1 overabundance.

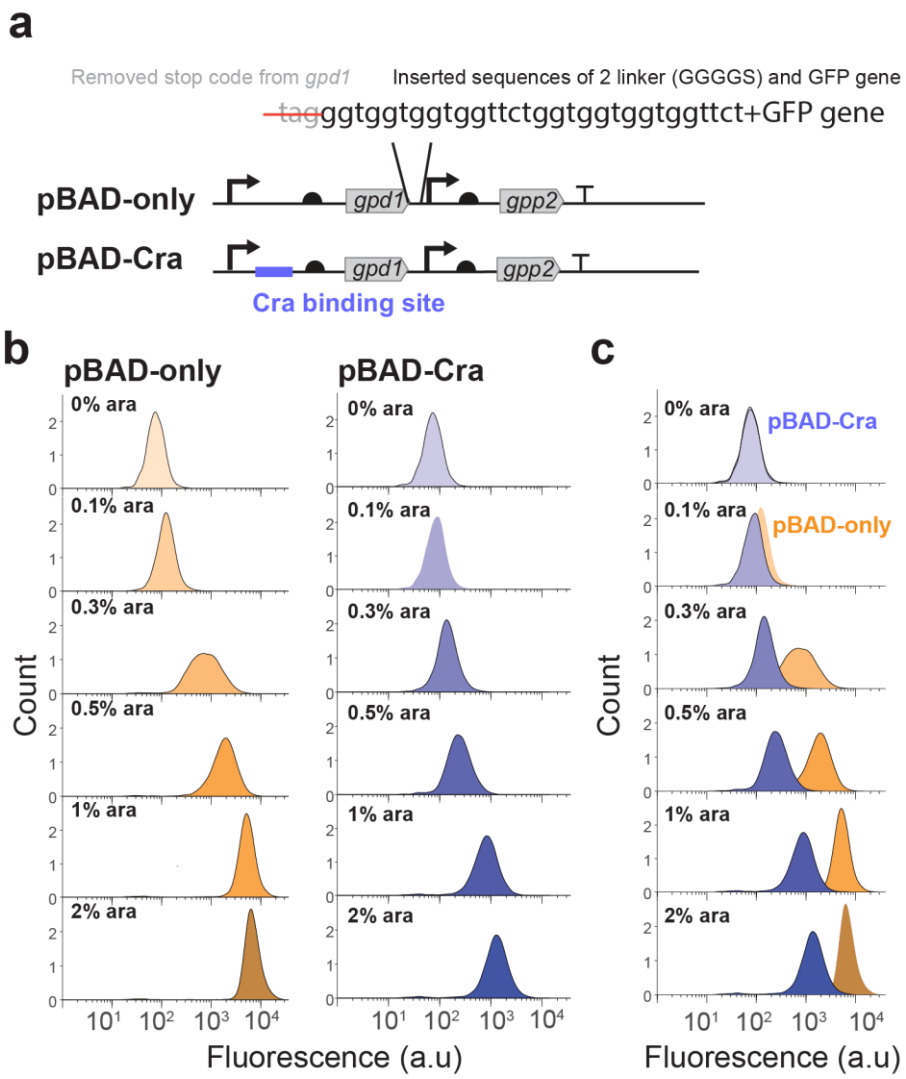


Fig. 3.13 | FACS analysis of the glycerol producer that expresses a GFP-GPD1 fusion protein. a, Fusion GFP was cloned after *gpd1* gene by removing stop code and adding 2 linker connected to GFP gene. **b,** Fluorescent levels of pBAD-only and pBAD-Cra systems were induced at various arabinose concentrations from 0% to 2%. **c,** Merged these two strains shows pBAD-only strain expresses more GFP than pBAD-Cra strain. All the strains were cultivated in culture tubes with 5 mL M9 medium supplemented with various arabinose concentrations and measured the fluorescent levels at exponential phases.

3.2.7 Proteome and metabolome analysis of non-regulated and Cra-regulated strains

To further investigate the distinct physiological behavior between pBAD-Cra and pBAD-only strains, we conducted the time-dependent experiment to carefully explore when they changed metabolites and how they reacted to overloaded glycolysis. As shown in **Fig. 3.14a**, both strains were grown separately without inducer for 3 h and

then turned on glycerol genes with 0.5% arabinose. The GPD1 protein expressed immediately after the induction (**Fig. 3.14b**), and the glycerol titers remained comparable in the first hour (**Fig. 3.14c**). However, the glycolytic metabolites (FBP, DHAP and PEP) dropped quickly within 1 h in the pBAD-only strain (**Fig. 3.14d**).

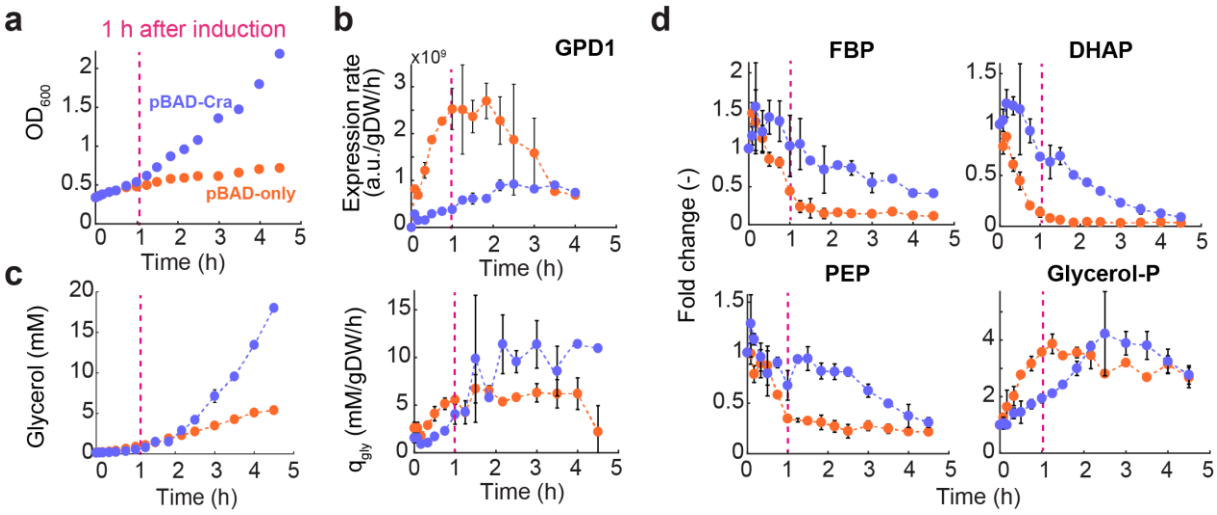


Fig. 3.14 | Dynamic analysis of the two glycerol producing *E. coli* strains. a, Growth curves of pBAD-only (orange) and pBAD-Cra strains (blue). The pink lines represent the timing of pBAD-only strain shows reducing growth. **b,** The GPD1 protein expression rate. **c,** The specific glycerol production rates were calculated from glycerol titers according to Materials and Methods. **d,** The fold changes of FBP, DHAP, PEP and glycerol-P were normalized to the uninduced state of each system. FBP, fructose 1,6-bisphosphate; PEP, phosphoenolpyruvate; DHAP, dihydroxyacetone phosphate; glycerol-P, glycerol 3-phosphate. (n=2)

Unlike pBAD-only strain, the pBAD-Cra strain consumed these metabolites slowly with a gradually increasing glycerol titer and such a higher production rate (10-12 mM glycerol/gDW/h) was sustained almost 4.5 h till the experiment terminated. The intermediate glycerol-P could be toxic to the cells[101]. However, considering that both strains reached a similar level of glycerol-P, the toxicity of intermediate should not be the main stress in our case as mentioned above (**Fig. 3.4b**). The other concern was that the glycerol pathway drains NADH converted to glycerol-P for glycerol production, but the energy (NADH, NAD and ATP) alters in a similar trend between these two strains (**Supplementary figure 7**), showing energy imbalance is also not the major cause of the observed distinct controllability. Intriguingly, except that FBP was higher

in pBAD-Cra than pBAD-only strain, both DHAP and PEP eventually reached the same level, which implies the FBP could be one signaling molecule for running Cra-regulated system. Due to the huge difference of GPD1 expression level between these two strains, it suggests that pBAD-Cra strain can maintain lower GPD1 via Cra regulation so that the deficiency of metabolites cannot be as serious as pBAD-only strain.

To quickly compare these two strains, the protein amount of pBAD-Cra, in arbitrary units, was divided by pBAD-only strain and transformed to a $\log_2$ scale. From this transformed heatmap (**Fig. 3.15a**), the relative protein expression didn't change much within 1 h but diverted obviously after 1 h induction. The proteins involved in glycolysis and the phosphoenolpyruvate-dependent sugar phosphotransferase system (PTS) are higher in pBAD-Cra strain, but the proteins involved in TCA, glyoxylate and gluconeogenesis conversely are relatively lower in pBAD-Cra strain. In other words, in the pBAD-only strain, the Cra-activated proteins are up-regulated which is consistent with **Fig. 3.6e** that gluconeogenic proteins increase expression under glucose condition causing misregulation. The misregulation in pBAD-only strain further aggravates the growth retardation. In the pBAD-Cra strain, it seems that the Cra regulator provides the host 1 h preparation time before boosting glycerol overproduction.

Furthermore, the upregulation of the PTS system (**Fig. 3.15a** and **Supplementary figure 8**) could benefit the pBAD-Cra strain to sustain such high glycerol production rate. To compare whether the glucose uptake rate was increased in pBAD-Cra strain, the GFP plasmid driven by pBAD inducible promoter was expressed in the same background strain (*glpK* deleted strain) as a reference because pBAD-only cannot uptake glucose that will be discussed later. As mentioned in **subchapter 3.2.5**, the plasmids of pBAD-only-GFP and pBAD-Cra-GFP were expressed in both wild-type and Δ*cra* strains separately, the results show the strain harboring pBAD-only-GFP produces higher GFP than pBAD-Cra-GFP in both different background strains; moreover, no growth effect is detected between harboring pBAD-only-GFP or pBAD-Cra-GFP plasmid (**Supplementary figure 9**). That was the reason pBAD-only-GFP plasmid was chosen for the comparison if PTS protein increased, the glucose uptake rate will also enhance. As shown in **Fig. 3.15b** both glycerol and GFP producers were grown in M9 medium provided with 0% or 0.5% arabinose. In the pBAD-only-GFP

strain, limited change in both glucose uptake rate (6.91 and 6.27 mM/gDW/h) and acetate production rate (3.32 and 3.58 mM/gDW/h) were detected regardless of the induction by 0% or 0.5% arabinose, indicating neither *glpK* deletion nor arabinose degradation will affect glucose uptake and acetate production. Distinct from the pBAD-only system, an enhanced glucose uptake rate was observed in the pBAD-Cra strain even without inducer (8.65 mM/gDW/h), and was further enhanced after 0.5% induction (10.93 mM/gDW/h). This result explains why pBAD-Cra strain can sustain higher glycerol production rate but still keep growth via up-regulating PTS proteins to increase glucose uptake.

Usually, the fast-growing cells show higher acetate overflow[102-104], meaning the GFP producer strain should produce more acetate than glycerol producer due to the higher growth rate. However, the acetate was above 2-fold higher in glycerol producer and surprisingly reduced at 0.5% induction (from 9.40 to 7.03 mM/gDW/h) (**Fig. 3.15b**), suggesting the recycling/reducing of acetate happens when glycerol genes are turned on. A study has shown that *tktB* or *talA* mutants have lower acetate yield compared to wild-type[105]. Based on the current result, the glycerol producer (pBAD-Cra) increased TktB and TalA protein expression resulting in higher acetate production than GFP producer. A suitable explanation and mechanism of acetate secretion still need to be further investigated.

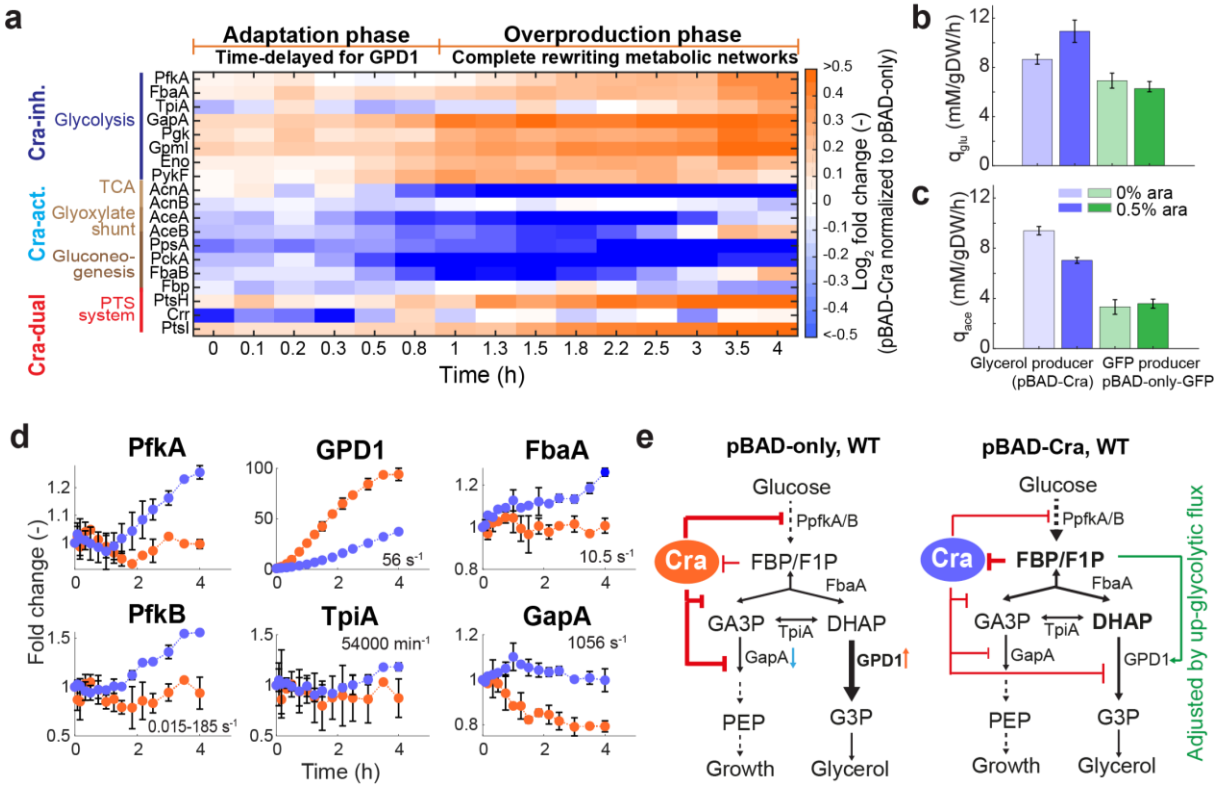


Fig. 3.15 | The dynamic control via Cra that enables the cells to adapt to the overloaded glycolysis due to glycerol overproduction. a, The protein amount of pBAD-Cra was normalized to pBAD-only strain to get relative protein expression with $\log_2$ transformation. **b,** The glucose uptake rate and acetate production rate of pBAD-Cra and pBAD-GFP strain. As an experimental control, the pBAD-GFP strain was also expressed in *glpK* deleted background to avoid the effect of *glpK* deletion. **c,** The fold change of protein expression was normalized to uninduced state of each strain. The number inside each subfigure is the turnover number (K_{cat}) of each enzyme. PfkA has the same K_{cat} as PfkB. All the K_{cat} values are taken from the BRENDA enzyme database. **d,** Schematic view of different regulatory mechanisms via Cra in pBAD-only and pBAD-Cra strains. The thickness of the lines represents the relative expression by comparing between these two strains in glucose cultivation. The color of the Cra-regulator indicates its activity. The Cra-activity is comparatively higher in pBAD-only (orange color) than in pBAD-Cra strain (blue color). (n=2)

Increasing glucose uptake rate provides the pBAD-Cra strain with incrasing capacity to ease the over-demanding stress from glycerol production. On the other side, pBAD-only strain faces a shortage of glycolytic metabolites within 1 h induction and it could neither increase glucose uptake nor upregulate glycolysis promptly to balance the demand, which eventually causes metabolites drainingout and serious growth reduction. When looking closer to the proteins involved in the glycerol production pathway (**Fig. 3.15c**), we observed the GPD1 was unlimited expressed in pBAD-only

strain but much slowly increased in pBAD-Cra strain (**Fig. 3.14b and Fig. 3.15c**). Such slow expression of GPD1 actually gives the pBAD-Cra strain a crucial buffer time to adjust the whole metabolic pathways, and to properly respond to the external demand of overloaded glycolysis (as shown in **Fig. 3.15a**).

From the turnover number (K_{cat}) of these enzymes near the glycerol synthetic pathway, it seems that up-glycolysis could be the bottleneck restricting the glycerol production, because the K_{cat} of up-glycolytic enzymes, such as PfkA/B and FbaA, are much lower (0.015~185 s^{-1} and 10.5 s^{-1} separately) than lower-glycolytic enzymes, such as TpiA and GapA (54000 min^{-1} and 1056 s^{-1} separately). Considering that the K_{cat} of FbaA (10.5 s^{-1}) producing DHAP and GPD1 (56 s^{-181}) consuming DHAP, higher GPD1 expression, in theory, will cause a shortage of DHAP. From practical results (as shown in **Fig. 3.14Fig. 3.15d**), the strong shortage effect is observed in pBAD-only strain after 1 h induction. Moreover, these enzymes (PfkA/B, FbaA, TpiA and GapA) are belonging to the Cra-inhibited group (which will be inhibited by Cra) and are fewer in pBAD-only strain (**Fig. 3.15c)**. The results suggest that the pBAD-only strain represses these glycolytic genes and trigger the gluconeogenesis, causing the misregulation in the central metabolism and performing like a starvation-like cell even though enough glucose presents in the surrounding environment. Besides, considering the higher K_{cat} of lower-glycolytic enzymes, we can speculate that pBAD-only strain tending to use gluconeogenesis is strongly favored by thermodynamics.

Combined with all the experimental results, schematic view of Cra-regulation was drew in **Fig. 3.15d**. Taken together, it suggests that how to increase the reaction-determined up-glycolytic flux and keep glycolysis running normally (without activating gluconeogenesis) are two major strategies to maintain high glycerol production rate. The as-designed pBAD-Cra strain can meet these two requisites by upregulation of both glycolytic and PTS proteins. By contrast, the pBAD-only strain further misregulates itself by repressing PTS/glycolysis and activating gluconeogenesis together with increased expression of several stress-responsive proteins, such as sigma factor RpoS for general stress response[106] and CstA for starvation[107] (PTS system proteins are shown in **Supplementary figure 8**, all measurable Cra-regulated enzymes are provided in **Supplementary figure 10** and **Supplementary figure 11**). With the Cra-binding site, the pBAD-Cra strain can repress the GPD1 via Cra first

(adaptation phase) and in meanwhile adjust the associated pathways for the preparation of glycerol production later on (overproduction phase) (**Fig. 3.15a**). Thus, Cra can either balance both growth and production or deteriorate the cells depending on how and where the Cra-regulation is located (as shown in **Fig. 3.12c**). Without creating a small set library for various enzyme expression levels such as different promoter strengths[108] or promoter-RBS combination[109, 110], we here demonstrated a simple but efficient way to get improved productivity and balanced growth synchronously.

3.2.8 The postponeed GPD1 expression level in a time-delayed manner increases glycerol production

We have observed that dynamic control by Cra-regulation can serve as a facile and universal valve to balance central metabolism and glycerol production. The Cra-regulation adjusts gene expression depending on the concentration of FBP/F1P which is changeable all the time. We then hypothesized if the way of Cra-regulation is like time-interval mode (time-delayed), by titrating gene expression levels, could we mimic Cra-regulation and further improve pBAD-only strain with improved growth and production?

Because pBAD-only strain shows strong growth reduction after direct 0.5% arabinose induction, to reduce the metabolic burden resulting from GPD1 expression, we tried to titrate pBAD-only and pBAD-Cra strains interval induction to the mimic time-delayed effect of the pBAD-Cra strain. Unlike common one-time induction which is added fixed concentration of inducer at once, we split the concentration of one-time induction into 10 times and added the divided concentration to the culture medium every 30 minutes for 4.5 h. We titrated the pBAD-only strain with two different final concentrations of 0.3% and 0.5% arabinose due to strong growth reduction, while pBAD-Cra strain was only titrated to a final concentration of 0.5% arabinose. By the interval induction, the pBAD-only strain grown much better than the common one-step induction, especially at 0.3% inducer (**Fig. 3.16 a** and **b**). Owing to the improvement of growth, glycerol production increases together with the accelerated glucose consumption from 4.87 to 8.11 mM/gDW/h at 0.3% induction and 4.27 to 6.37 mM/gDW/h at 0.5% induction. Intriguingly, the pBAD-Cra strain, however, does not

change significantly in both growth and glucose consumption, but the glycerol production rate is slightly higher by one-time (9.84 mM/gDW/h) than interval induction (7.02 mM/gDW/h) (**Fig. 3.16c**). This result points out that unlike pBAD-only strain, the pBAD-Cra strain can reach a dynamic equilibrium promptly, and the intrinsic equilibrium capacity is less affected by the external induction way.

Some synthetic pathways are difficult to find a native regulator for auto-regulating gene expression. In these cases, an induction strategy caused delaying protein expression is an alternative way to earn a responding time in which the host is able to adjust the combined metabolic networks and adapt to the encountered stress. Induction strategy has played a huge role in product yield and burden reduction, and several processes, such as the induction time[111] (early exponential growth[112]), the length of induction[113] and the inducer concentration, are developed and currently applied especially for industrial scale-up[112]. Here, we demonstrated that via an interval-induction (fine-tuned the protein expression levels by time) the titer and growth of pBAD-only strain can be refined, mimicking the Cra-regulation. However, considering the multistep induction and less efficiency, by incorporating the dynamic control of Cra-regulation, the designed overproduction system can synthesize target molecules (i.e. glycerol in this thesis) in a more cost-effective way without complicated titration to find a proper induction strategy.

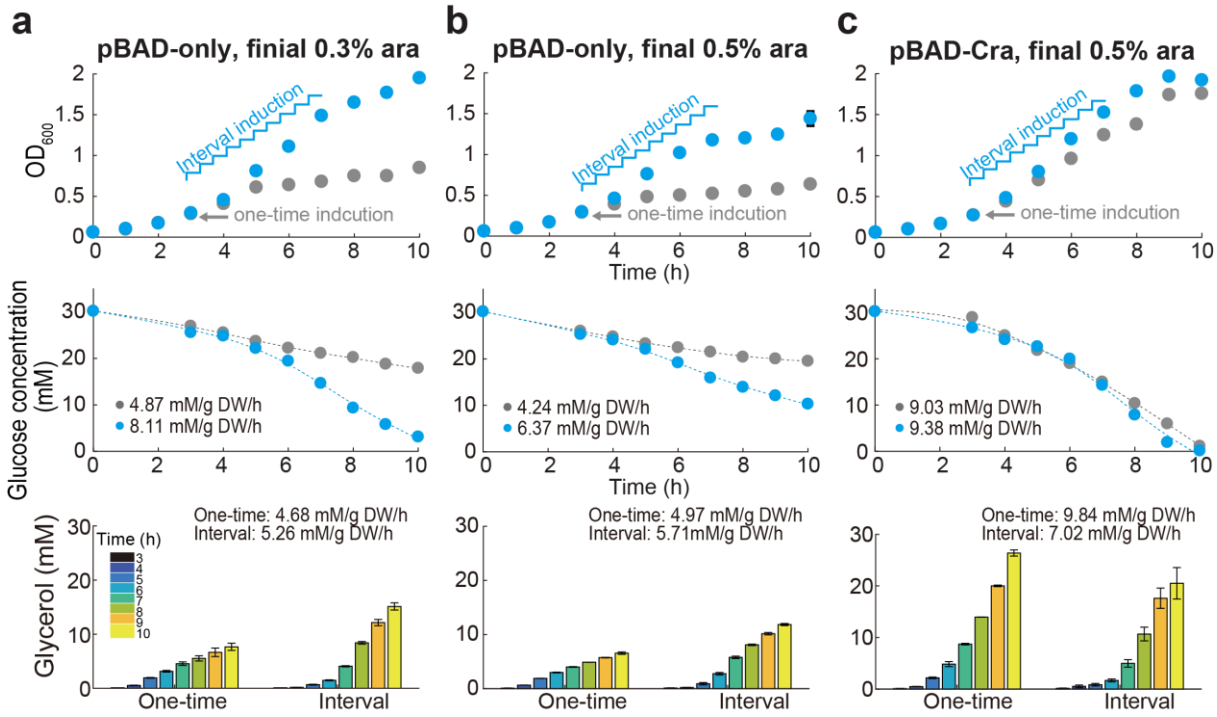


Fig. 3.16 | The comparison between one-time and interval induction of arabinose for glycerol production. a, Titration of pBAD-only strain to a final concentration of 0.3% arabinose. **b,** Titration of pBAD-only strain to a final concentration of 0.5% arabinose. **c,** Titration of pBAD-Cra strain to a final concentration of 0.5% arabinose. The gray arrow marks the induction time of arabinose. The blue ladder shows the 30 min-interval of arabinose induction from 3 to 7.5 h for total 4.5 h. The numbers represent glucose consumption rates and glycerol production rates.

3.2.9 Auto-regulation via Cra in carotenoid biosynthetic pathway ensures the general applicability of Cra-regulation

To test whether Cra-regulation is generally applicable for biosynthetic pathways, we chose carotenoid pathways as an extended study because this pathway also drains central metabolites as precursors (glyceraldehyde 3-phosphate and pyruvate) (**Fig. 3.18a**). It has been shown that overexpression of *dxs* (1-deoxy-D-xylulose-5-phosphate synthase) and *dxr* (1-deoxy-D-xylulose 5-phosphate reductoisomerase) can enhance the yield of carotenoid-relative compounds such as lycopene[73], β-carotene[114] and zeaxanthin[115]. Therefore, genes of *dxs* and *dxr* were amplified from *E. coli* MG1655 and then integrated into a pBAD controlled plasmid to get a pController-Non plasmid. Then, following the same strategy in glycerol production, we inserted Cra consensus sequence to form a pController-Cra plasmid. Besides, another plasmid

provided by Prof. Dr. Victor Sourjik (MPI, Marburg) is named as pCarotenoid (**Fig. 3.17a**) which harbors five genes (*crtE*, *crtB*, *crtI*, *crtY*, and *crtZ*) from *Pantoea ananatis* for synthesizing carotenoids and two genes (*dxs* and *idi*) from *E. coli* for methylerythritol 4-phosphate (MEP) pathway. The dxs and idi have been reported as rate-limiting steps in the MEP pathway and this pathway can be further improved by overexpressing these two genes[115, 116].

Different from glycerol production, the wild-type BW25113 was chosen due to its lack of arabinose degradation genes. Since pCarotenoid plasmid includes seven genes, the plasmid was solely expressed in wild-type and *cra* deleted strain separately to verify whether the plasmid was a burden for cells or not. Various concentrations of arabinose were added to the medium, and the results showed limited changed in growth but the titers of carotenoid were slightly higher in wild-type than △*cra* strain (**Fig. 3.17b**). Overall, overexpression of pCarotenoid is unlikely to create stress for the cells.

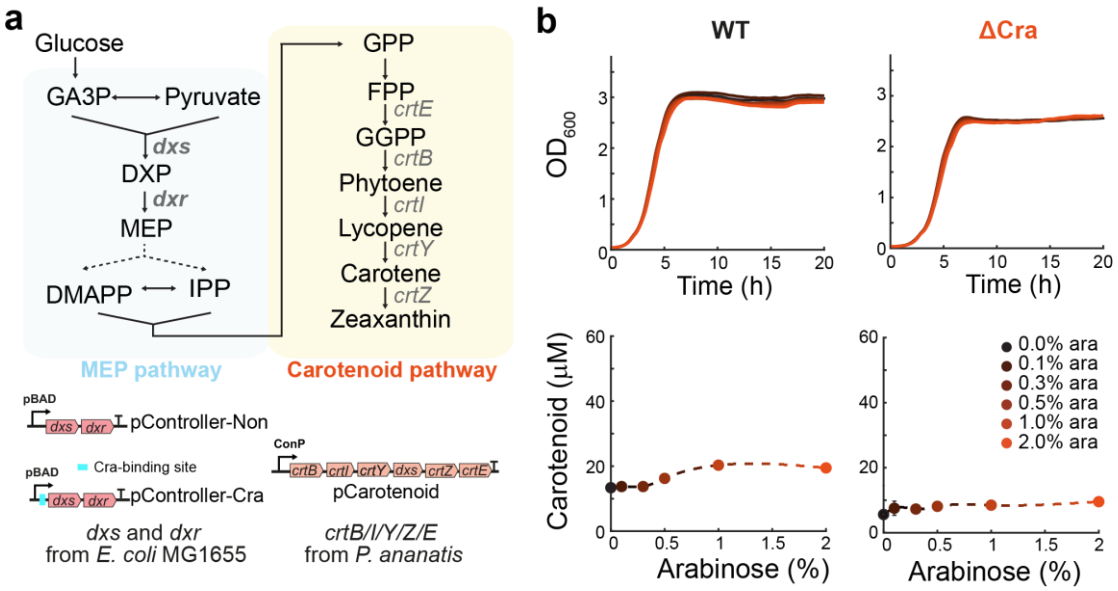


Fig. 3.17 | Overexpression of pCarotenoid plasmid alone in wild-type and Δ*cra* BW25113 strains. a, Biosynthesis of carotenoid through the MEP pathway and the construct of pController-Non, pController-Cra (integrated with consensus sequence) and pCarotenoid plasmids. **b,**The pCarotenoid plasmid expressed sole in both wild-type and *cra* knowkout *E. coli* BW25113 strains in M9 medium supplemented with 0.5% glucose and an additional 20% LB. The concentration of carotenoid was measured after 24 h. GA3P, glyceraldehyde 3-phosphate; DXP, 1-deoxy-Dxylulose-5-phosphate; MEP, 2-C-methyl-D-erythritol-4-phosphate; DMAPP, dimethylallyl diphosphate; IPP, isopentenyl diphosphate; GPP, geranyl diphosphate; FPP, farnesyl diphosphate; GGPP, geranylgeranyl diphosphate.

Besides, producing carotenoids shows the same phenomenon as in glycerol production (**Fig. 3.11**) that is, pBAD-only strain has instable growth and limited window to maximize product titer. Conseverly, only pBAD-Cra strain with functional Cra can stabilize both growth and production; moreover, this strain can maintain high titer of carotenoids throughout a wider range of induction, confirming again the Cra-regulation can fine-tune gene expression and enable high adaptability against overloaded glycolysis. By exploiting the Cra-regulation as a glycolytic flux controller, the synthetic circuits are equipped with a tunable and auto-regulated feedback device to boost the fabrication of desired products derived from glycolysis.

Next, we coexpressed the pController (either pController-Non or pController-Cra) and pCarotenoid in a given background strain at the same time. Using the same naming method, the pBAD-only is the strain bearing pController-Non and pCarotenoid plasmids; on the other hand, the pBAD-Cra is a coexpression of pController-Cra and pCarotenoid plasmid. As demonstrated in **Fig. 3.18**, compared to the strain with only pCarotenoid plasmid, these four strains exhibited enhanced yield of carotenoids, suggesting that providing enough precursors is critical for overproduction, but over a certain threshold a damage will eventually emerge to the cells. Besides, producing carotenoids shows the same phenomenon as in glycerol production (**Fig. 3.11**) that is pBAD-only strain has instable growth and limited window to maximize product titer. Conseverly, only pBAD-Cra strain with functional Cra can stabilize both growth and production; moreover, this strain can maintain high titers of carotenoids throughout a wider range of induction, confirming again the Cra-regulation can fine-tune gene expression and enable high adaptability against overloaded glycolysis. By exploiting the Cra-regulation as a glycolytic flux controller, the synthetic circuits are equipped with a tunable and auto-regulated feedback device to boost the fabrication of desired products derived from glycolysis.

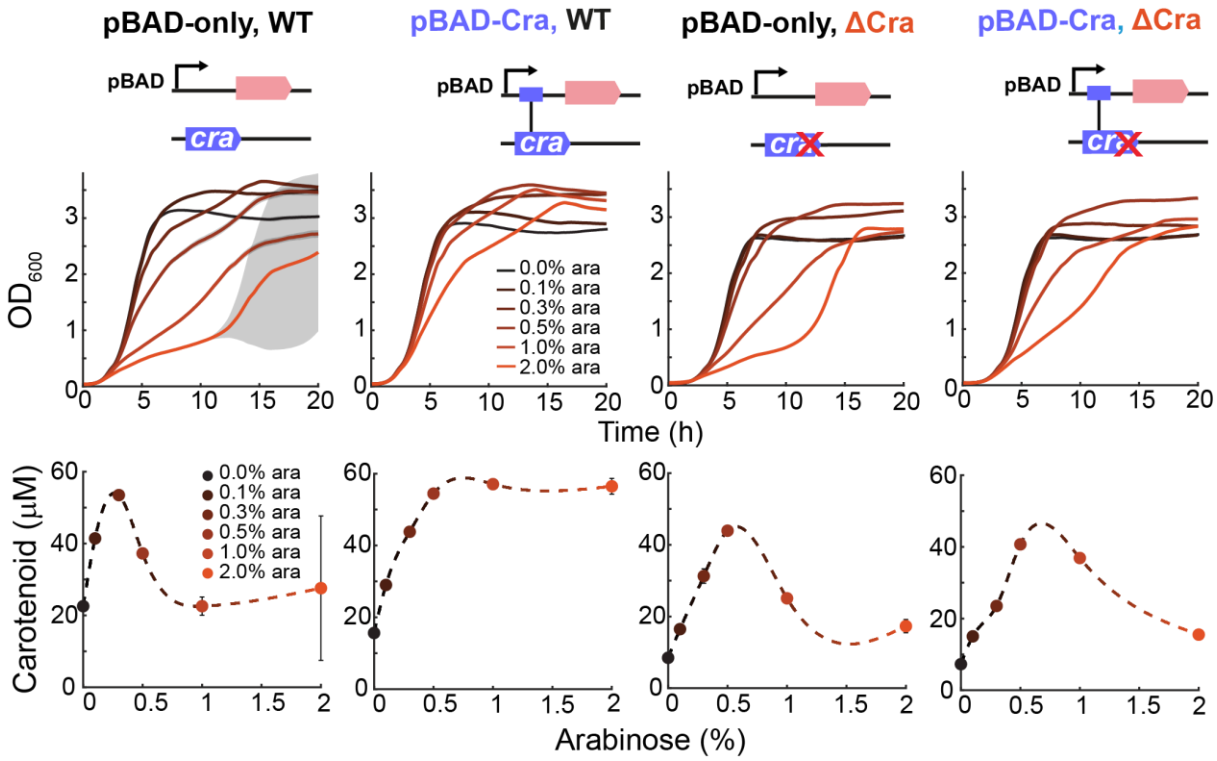


Fig. 3.18 | The general applicability of integrated Cra-regulation to dynamically control carotenoid biosynthetic pathway. Pasmids of pController (either pController-Non or pController-Cra) and pCarotenoid were coexpressed in wild-type and Δ*cra* BW25113 strains. The pBAD-only strain represents the coexpression of pController-Non and pCarotenoid. The pBAD-Cra represents the coexpression of pController-Cra and pCarotenoid. The concentration of carotenoid was measured after 24 h.

3.3 Conclusion

Cells have evolved tightly and comprehensively regulatory networks for sustaining homeostasis. Additional implementation of designed pathways always disrupts the homeostasis by unbalancing precursors, energy, co-factors, and so on. Especially in the current study, a strong growth impairment is resulting from draining precursors from glycolysis over a certain threshold. As demonstrated in the synthesis of glycerol and carotenoids, to overcome this impairment, we take advantage of native Cra-regulation to precisely express desired genes and to achieve stabilized growth as well as improved productivity. In principle, the time-interval induction of designed genes offers multistep protein expression which eases the instant draining-out of glycolytic metabolites so that the cells obtain auxiliary time to adjust/prepare the metabolic pathways for product overproduction. However, considering the high controllability,

one-step process and high applicability, the incorporation of Cra-regulation is a cost-effective operation and requires no additional injection device when scale-up in a bioreactor. In summary, this Cra-regulation can be theoretically incorporated into many other high-value products that rely on central metabolism and save time to find optimal fermentation conditions.

3.4 Future work

3.4.1 Can the pBAD-Cra strain further increase productivity via a two-stage process?

It has been suggested that enhancing the substrate uptake rate is the key requirement for the two-stage process otherwise one-stage process would be more appropriate for the highest product yield[16]. In other words, the pBAD-Cra could tend to perform much well in the two-stage process since this strain has improved the glucose uptake rate.

3.4.2 Rewriting metabolic pathway is a solution to tackle overloaded glycolysis

The pBAD-only strain has strong growth retardation when the pathways are turned on. However, we never measure the physiological changes after 24 h. The cells may evolve new metabolic networks to conquer the overloaded glycolysis or it just mutates the designed genes to release the burden. In this thesis, we have noticed that glycerol genes driven by constitutive promoters (J_{23101}) show insertion sequence (IS) inserts in the promoter region to block *gpd1* gene expression after 24 h cultivation, but we did not find IS sequence in pBAD controlled system. Hence, we speculate that pBAD-only strain can rewrite the metabolic pathways to escape from overloaded glycolysis.

3.4.3 Modification of genetic circuits is not always mandatory for high product yield

Without embedded Cra-consensus sequences, we improved glycerol production by interval titration. This indicates the fermentation techniques are also important for product titers. As mentioned before, there are many strategies to do it. One of the

common methods is called auto-induction which is using the hierarchy of catabolite repression that glucose depletes first for growth and the following lactose consumes for desired protein expression[117]. By auto-induction, there is no need to monitor the cell density or selection of time to add inducer. What if we combine both Cra-regulation with auto-induction for designed circuits, will we improve much higher titer?

3.4.4 The Cra-regulation mechanism is a not dynamic control instead but a thermodynamic reaction

What if all the dynamic regulation is just a basic chemical reaction. The additional protein expression interferes with the internal homeostasis. Therefore, the inside pH, ionic strength and catalytic capacity of enzymes change to meet the imbalanced. When the unfavorable reaction can return to proper chemical equilibrium via time-interval titration, the cells regain growth again and again between each time-interval induction.

4. Chapter 4 – The studies about Cra-regulation

4.1 Research questions and objectives

Besides the major finding of dynamic control in **chapter 3**, we also found some interesting results relative to the Cra-regulated response. For example, by additional providing extra phosphate sugar rescues, the burden from overloaded glycolysis in the pBAD-only strain can be released. The more detail will be described in the following subchapters.

4.2 Results and discussion

4.2.1 Integration of Cra-regulation enhances glycerol titer in other types of promoter systems

We tested different promoter systems for the integration of Cra-regulation, including anhydrotetracycline (aTc) inducible and constitutive promoters. These two systems show the same advantages that Cra-regulator improves both growth and glycerol production.

4.2.1.1 Integration of Cra-regulation via aTc inducible system

Anhydrotetracycline (aTc) inducer is a derivative of tetracycline but it exhibits no antibiotic activity. We created two strains, namely without (pTetR-only) and with the Cra-binding site (pTetR-Cra) as the pBAD-inducible system (**Fig. 3.11**). Integration of Cra-regulation improves both growth and glycerol production in the presence of aTc (**Fig. 4.1**). The pTetR-only strain didn't produce more glycerol even under higher glucose conditions (1%), which points out the restrictions of glycerol production is not lack of carbon sources. Inspired by the conclusion in the previous chapter, the obstructed metabolism, such as precursor lacking or the unaffordable uptake rate, is the main reason why the consumption rate/glycerol production rate is halted. Unlike pTetR-only strain, the pBAD-Cra strain further enhances the glycerol production when the glucose concentration increases, indicating that pTetR-Cra adapts to glycerol overproduction by increasing glucose uptake that is in line with the result of pBAD-Cra strain (**Fig. 3.15a** and **b**). Due to the transcription factor Cra, the pTetR-Cra strain

reduces growth impairment of overloaded glycolysis and then boosts glycerol production.

Basically, aTc is a bright yellow chemical and its wavelength varies in the presence of magnesium (Mg^{2+}) or calcium cations (Ca^{2+}). The absorbance ranges from 430 nm to 436 nm and the emission wavelength changes from 532 to 502 nm when the excitation wavelength (Ex) set at 305 nM[118]. The compounds in biological samples are more complicated than composites of just aTc and Mg^{2+} or Ca^{2+}. Since the glycerol concentration is determined by glycerol kit which is actually an absorption-based measurement (absorbance, 570 nm or fluorescence (Ex/Em: 535/587 nm), the absorption spectrum could be disturbed by the metal ion-associated biological samples. In case the glycerol measurement was influenced by the addition of aTc, we decided not to use this aTc system for deep study.

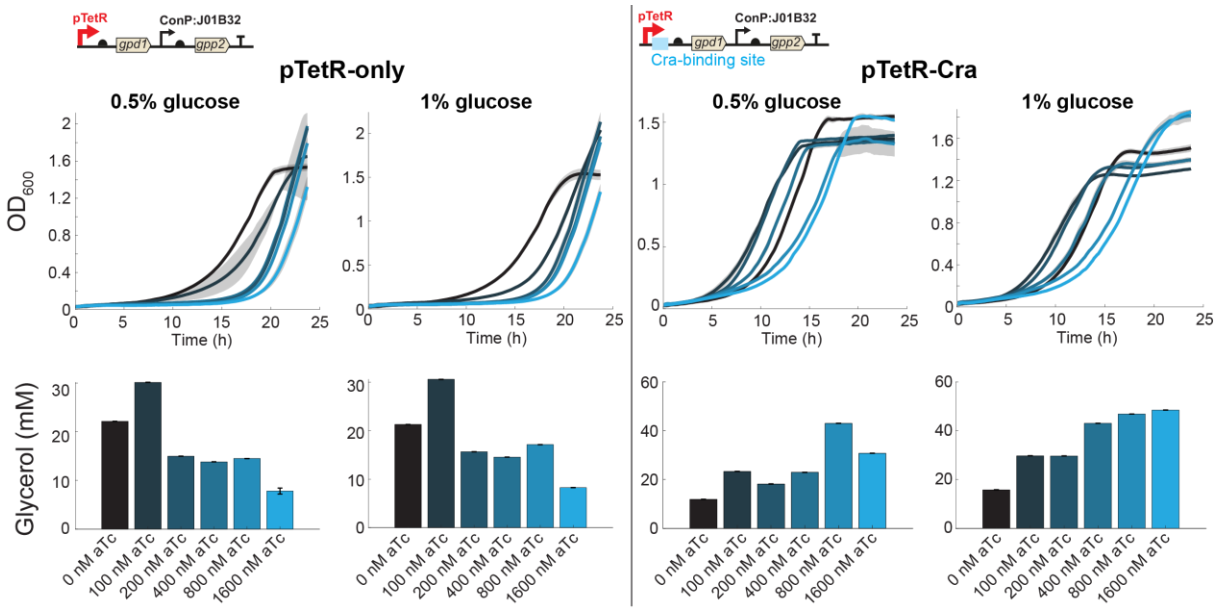


Fig. 4.1 | Engineering Cra-regulation in the pTetR promoter. Cra binding site was inserted directly after pTet promoter as a Cra-regulated system (pTetR-Cra). The strains without (pTetR-only) and with Cra-binding site (pTetR-Cra) were cultivated in 0.5% and 1% glucose M9 medium supplemented with various aTc concentrations. (n=2)

4.2.1.2 Integration of Cra-regulation with constitutive promoter system

Because of the hierarchical behavior of carbon usage[119-121], the usability pBAD promoter cannot grow in different carbon sources. To test the environmental response

by supplement with various carbon sources, we created several constitutive promoters with the integration of Cra-regulationa and also tested native Cra-inhibited promoter-P_{pfkA}. As illuminated in **Fig. 4.2**, a balanced effect that growth and production maintain in a synchronized way can only be achieved when the Cra-binding site and the Cra-regulator are coexisted in the strain. Especially, under less-nutrient succinate conditions, producing glycerol becomes more difficult for cells, but the Cra-regulated strain stands out no matter which promoters were used. Although, the glycerol titer is much lower than glucose condition, the strains with Cra-integration still manage to grow. However, there was one exception that the P_{pfkA} strain in ∆cra background did not lose growth and production. We assumed that other regulators recognize the promoter of *pfkA* and participate in the regulation of overproduction. Despite this, the results confirm again that Cra play an significant role to regulate glycerol gene expression and to stabilize host growth.

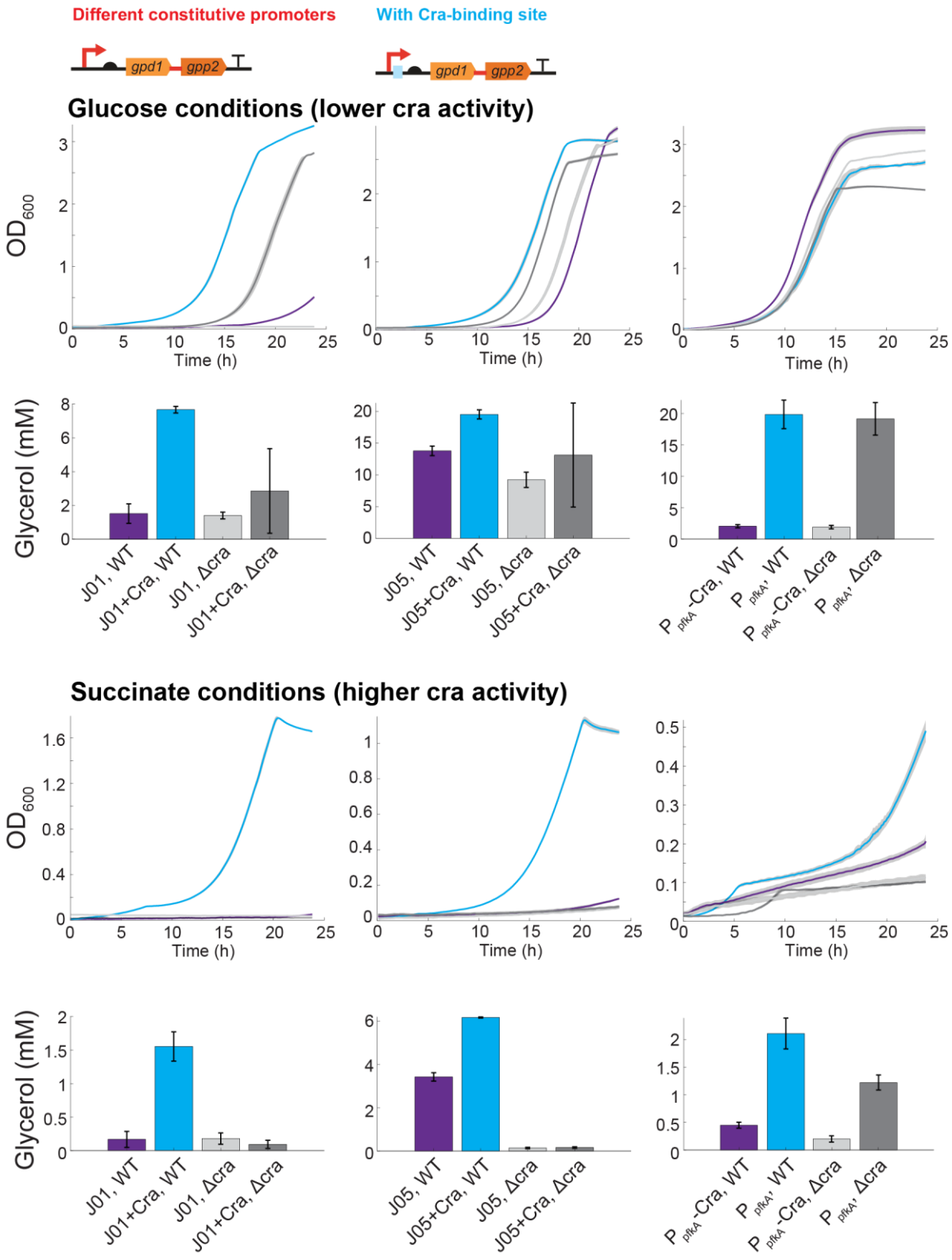


Fig. 4.2 | Engineering Cra-regulation in a constitutive promoter. Under lower (glucose conditions) and higher Cra activity (succinate conditions), the growth and glycerol titer were measured. The strains were expressed in the ∆*glpK* and ∆*cra* background. (Fusion version of glycerol genes)

As shown in **Fig. 4.2**, to evaluate the improvement stemmed from the Cra-integration, we measured the host growth and glycerol production in various carbon substrates. Generally, compared to the strain without the Cra-binding site, the growth of +Cra strain (with the Cra-binding site) is usually faster with shorter lag phase and only in the fructose condition a relatively slow growth is observed. The season is that fructose is directly transferred to F1P/FBP, and the F1P/FBP is the inhibitor of Cra-regulator so that it shuts down the benefit of Cra-regulation. From the data, we also observed that to maintain the growth, the host has somehow restrained the glycerol production as a sacrifice. In a nutshell, the Cra-regulated system helps host cells to face environmental changes and balance both growth and production.

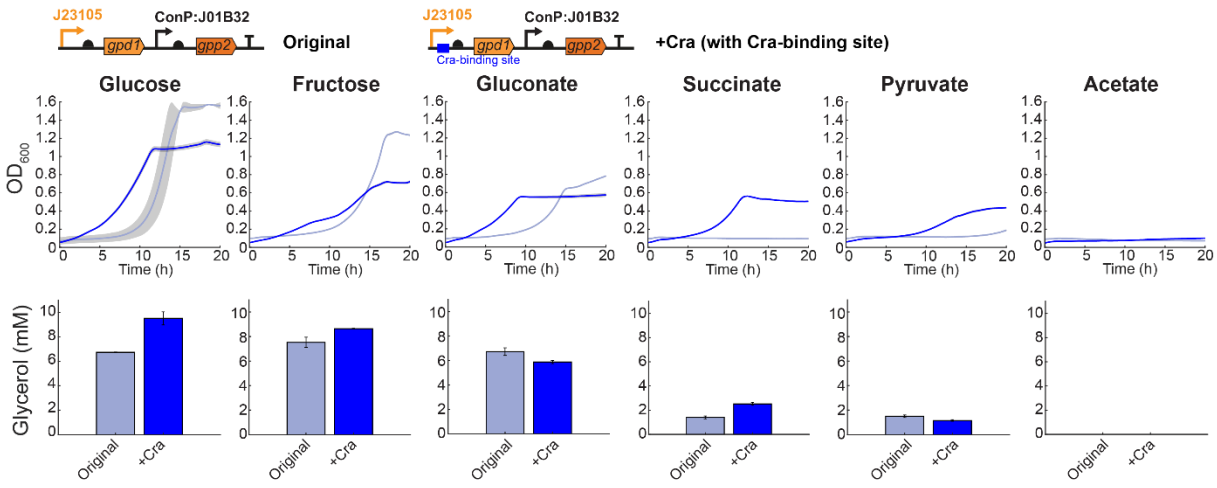


Fig. 4.3 | Growth and glycerol production with 6 different carbon sources. Two strains with Cra-regulation (blue) and without (grey) were grown in 3 gluconic (glucose, fructose, and gluconate) and 3 gluconeogenic carbon substrates (succinate, pyruvate, and acetate) and the glycerol titers were measured after 24 h cultivation.

4.2.2 Supplement of phosphate sugars can rescue overloaded glycolysis

According to **chapter 3**, we know that producing glycerol can overload intrinsic glycolysis and then lead to misregulated central metabolism in the host. The results inspired us whether this misregulation effect can be recused by providing enough precursor for glycerol synthesis or by directly adding the signaling molecule FBP of Cra (**Fig. 4.4a**). Six compounds (G6P, F6P, FBP, F1P and G1P) that can be converted to DHAP were tested (**Fig. 4.4b**). Intriguingly, after 24 h, three different five-carbon

sugars (F6P, FBP and F1P) improved the growth at 0.1% induction of arabinose but not at 0.5% (**Fig. 4.4c**). This implies that the five-carbon sugars are analogous to FBP and can indeed alleviate the misregulation by obtaining more precursors and bypassing the PTS system or regulate by the signaling molecule FBP but the alleviation depends on the draining speed of the precursors and the permeability of the sugars. The later parameter might be the reason why the pBAD-only strain remains ailing although the environment is fulfilled with sugars.

Remarkably, the pBAD-Cra strain did not show a significant change in growth, but the highest glycerol titer was in the presence of FBP (**Fig. 4.4d**). This result suggests that the cells might harbor FBP transporter which permeates FBP (FBP→DHAP→Glycerol-P) for direct transformation FBP into glycerol in a short cut. However, compared to without any phosphates sugar, the growth was slightly reduced by adding FBP that implies slight misregulation exists due to FBP-Cra interaction or just increasing cells' GPD1 capacity. Thus, it is hard to distinguish the major effect from the speed of permeating FBP, FBP-Cra interaction or enough reservoir for GPD1 using.

Unlike six-carbon sugars of G6P and G1P, five-carbon sugars of F1P and F6P are a one-step reaction to FBP. This could be one of the possible explanations why five-carbon sugars are better than six-carbon for pBAD-only strain to overcome the misregulation. However, neither the transporter existing nor the permeability of each compound is considered or experimental measured here. This is only an evidence that overloaded glycolysis can be compensated by supplementation of FBP, F1P and F6P. Overall, these results demonstrate that pBAD-only strain can be rescued by supplying five-carbon sugars while pBAD-Cra strain already reaches the maximal balance so there is less additional benefits from these sugars.

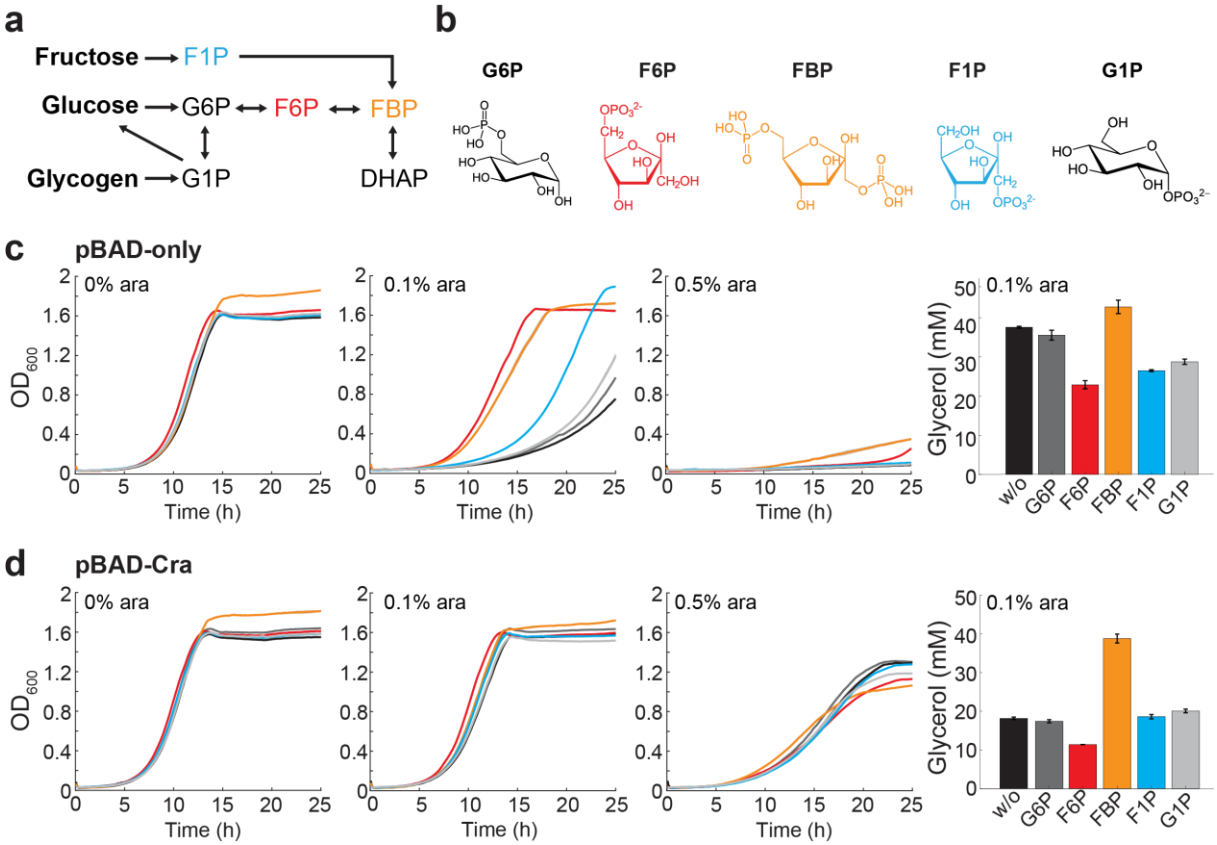


Fig. 4.4 | The growth and glycerol titer of pBAD-only and pBAD-Cra glycerol producers under 6 different phosphate sugars and their chemical structures. a, The relationship between DHAP precursor for glycerol production and the other phosphate-sugars. **b**, Structures of 6 phosphate-sugars that are involved in DHAP synthesis; Sugars of five-carbon ring are colored in red, orange, and blue. **c,** Growth and glycerol titer of pBAD-only strain. **d,** Growth and glycerol titer of pBAD-Cra strain. The pBAD-only and pBAD-Cra strains were grown in M9 medium provided with 0.3 mM individual phosphate sugar and were induced at the beginning with three arabinose levels (0, 0.1 and 0.5%). G6P, glucose-6-phosphate; F6P, fructose-6-phosphate; FBP, fructose 1,6-bisphosphate; F1P, fructose-1-phosphate; G1P, glucose-1-phosphate.

4.2.3 Comparison between the wild-type and Δ*cra* strains confirms the impact of Cra-regulation in cell's pathway selection

Glycolysis is a ubiquitous route for living organisms to convert glucose into pyruvate and release energy during the processes. One well-known glycolytic route is the Embden–Meyerhof–Parnas (EMP pathway) and the other is the Entner-Doudoroff Pathway (ED Pathway). To address the significant role of Cra-regulation in responding to both environmental changes and glycerol production, we selected 3 representative carbon sources, glucose, gluconate and fructose, and utilized them as the only carbon

source separately. In principle, glucose is catabolized by both EMP and ED pathways; gluconate is utilized by ED and PPP pathways, while fructose is only employed for EMP pathway (**Fig. 4.5a**).

To understand the effect of Cra regulator, we first compared the wild-type and Δ*cra* mutant under glucose condition (no Cra exists). **Fig. 4.5b** illustrates that when Cra did not exist in the cells, 73% (11 out of 15) of Cra-activated proteins showed a reduced expression levels, and an increased expression was observed in 74% (35 out of 47) of Cra-inhibited proteins. Distinct from glucose condition, under succinate conditions (Cra-activating state), Cra activity is expected higher than in glucose conditions. However, the results turned out that 66% (10 out of 15) of Cra-activated proteins enhanced the expression levels and 50% of Cra-inhibited proteins decreased expression levels. Although the variation of percentage seems to be not high enough, when narrowing down the proteins only regulated by Cra (orange font), 11 out of 13 Cra-regulating proteins (EnvC and Glk slightly high) under succinate condition were altered, corresponding to the logic of Cra-regulation. Moreover, all the protein changes shown in the inactivating Δ*cra* system agree with the previous transcriptional and protein study that Δ*cra* down regulates gluconeogenesis, TCA cycle and anaplerotic pathways as well as upregulation of fructose operon and glycolysis[95, 122, 123].

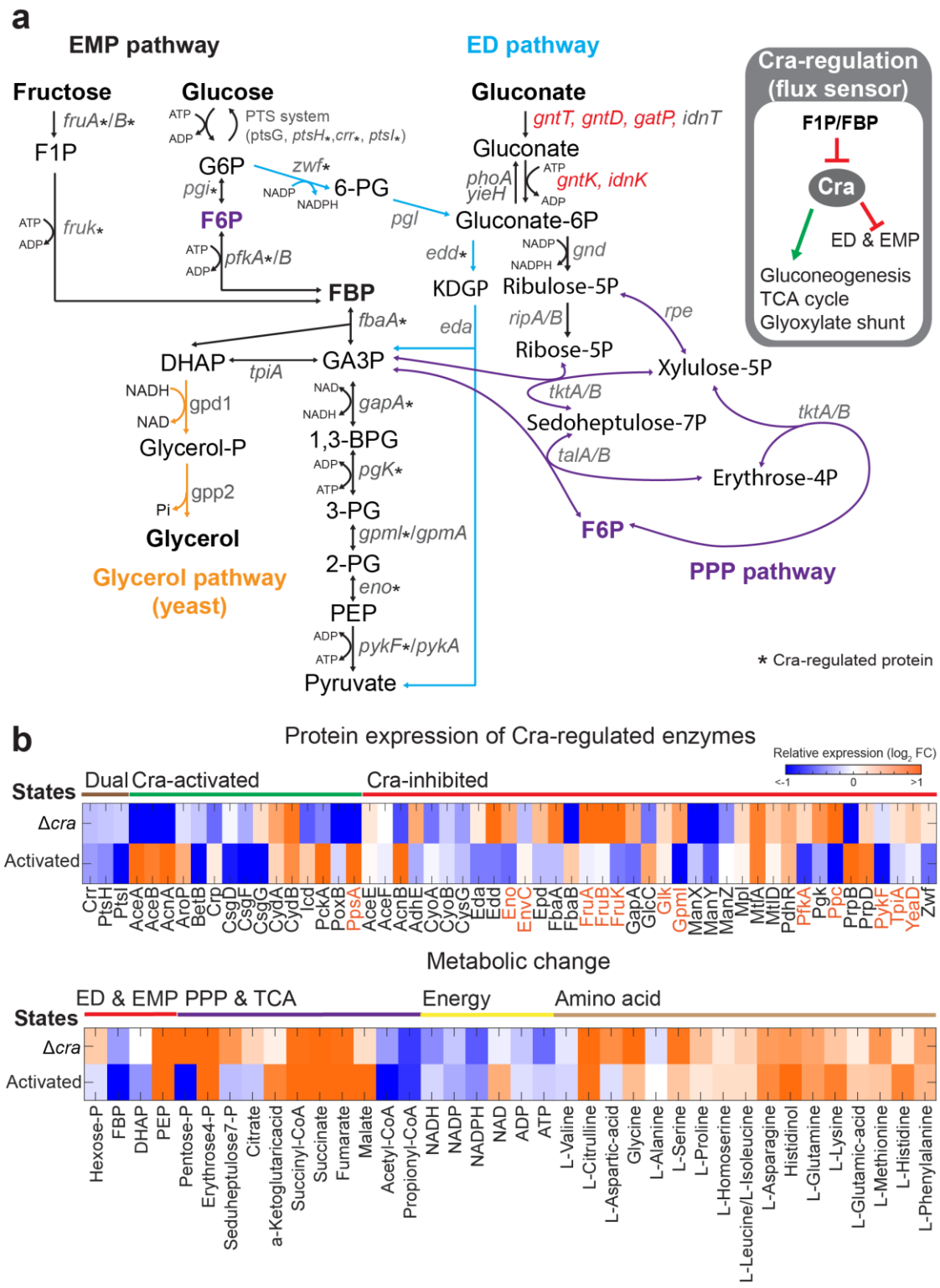


Fig. 4.5 | Proteome data of the Δ*cra* strain and the WT grown on succinate. a, Cells choose different bioprocess pathways according to their carbon sources (fructose, glucose, or gluconate). **b,** The proteins and metabolites were measured under glucose or succinate environment. Δ*cra* strain was grown in glucose, representing no Cra in the cell. The WT was grown in succinate, representing higher Cra activity, named activated. 54 proteins out of 79 known Cra-regulated enzymes and metabolites were detected. The data were normalized to wild-type strain cultivated in the glucose M9 medium. 13 proteins whose names are colored in orange are only regulated by Cra.

Aligning with the result of **Fig. 4.5**, instead of grouping the proteins in Cra-regulated order, we classified the protein by bioprocessing function as shown in **Fig. 4.6a**. As expected, the protein expression of glycolysis was up-regulated and the TCA cycles, glyoxylate and gluconeogenesis were down-regulated in △*cra* and vice versa the wild-type in the succinate condition (Cra-activated state). Although the protein expression level changes lots between these two different Cra-states, most metabolites maintained a similar pattern (**Fig. 4.6b**) that is the metabolites of the TCA cycle is increased. Intuitively, the lower protein expression should produce fewer metabolites of the TCA cycle, but interestingly in △*cra* state, the metabolites were increasing even though the protein expression was relatively lower than the wild-type strain. There are two feasible reasons that leads to a similar pattern of metabolic change. First, the ∆*cra* strain loses Cra-regulator to inhibit glycolytic enzymes so these enzymes may facilitate glycolysis without any limitation. Due to the higher speed of glycolysis, the concentration of PEP increases, and the metabolites in TCA also increase to match the chemical equilibrium. The other explanation is that the catalytic capacity of the enzyme is higher enough to execute the metabolic reactions. Namely, even though the protein expression is down-regulated in △*cra* state, the overabundance enzyme is still enough to run the TCA cycle. All the hypotheses here are needed to be further investigated. To sum up, the change of FBP concentration influences the Cra activity and then triggers protein and metabolic variations, revealing Cra-regulator indeed has a global effect on the regulation of central metabolism (**Fig. 4.6c**).

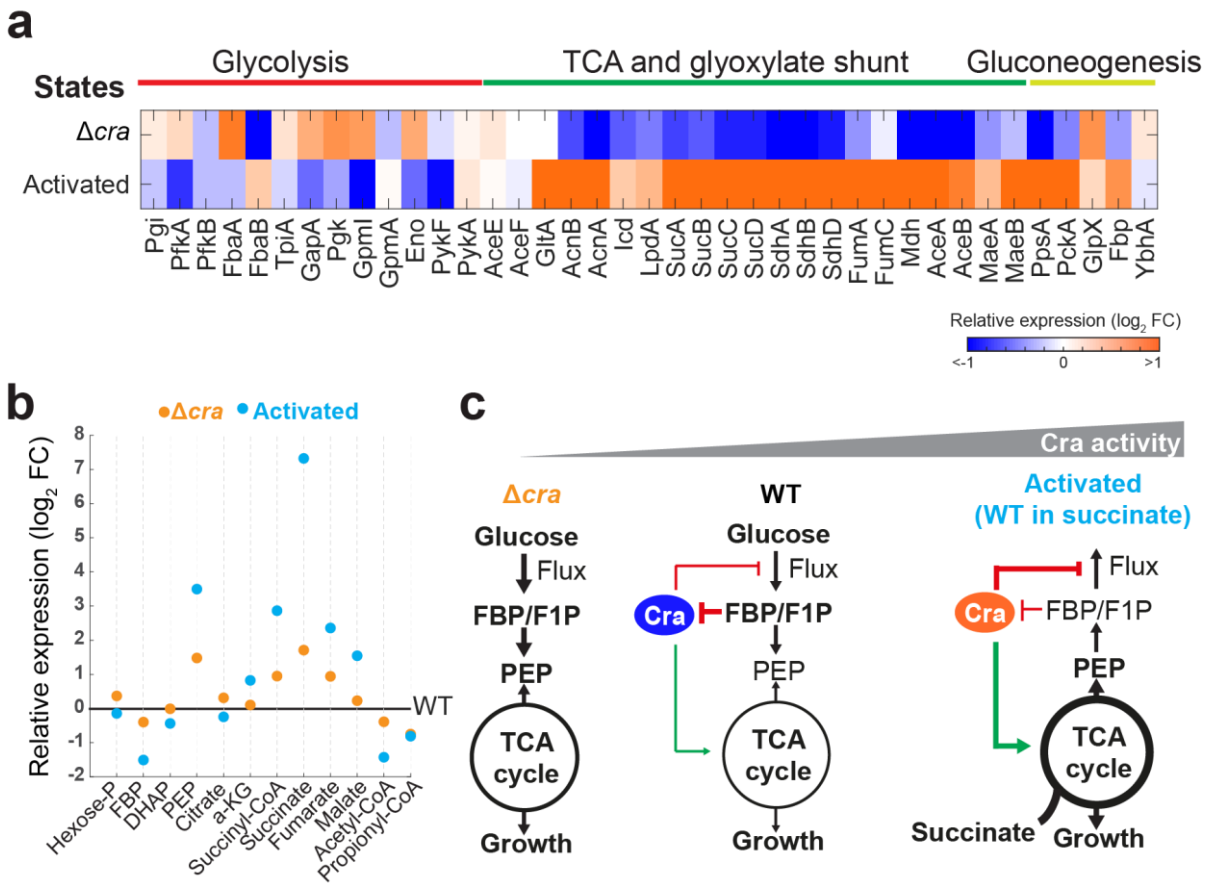


Fig. 4.6 | The comparison of cell Cra activities between no-Cra and increased Cra activity. a, Protein levels of glycolytic, gluconeogenetic, glyoxylate and TCA cycles. The Δ*cra* strain was grown in glucose medium as a state of no Cra activity while the wild-type strain was cultivated in succinate conditions as a higher Cra activated state. The data of these two strains were normalized to WT strain cultivated in glucose conditions. **b,** The metabolites of glycolytic and TCA cycles. **c,** The Cra-activity from low to high is Δ*cra*, WT and activated state. The bold letters represent higher metabolites concentration from the result of **Fig. 4.6a**. The thickness of the lines represents the relative expression by comparing with WT strain cultivated in glucose conditions.

From **chapter 3**, we already knew the misregulation of Cra due to glycerol overproduction in glucose conditions. Here, we tested other carbon sources- gluconate and fructose to reveal the importance of Cra-regulator. When gluconate was applied as a sole carbon substrate, only the wild-type strain harboring pBAD-Cra plasmid could endure the overloaded glycolysis (**Fig. 4.7a**) and produce glycerol (**Fig. 4.7b**). As mentioned in other studies[90], providing gluconate in the medium, the FBP concentration of wild-type strain is 25% lower than glucose medium. Considering the results of **Fig. 4.5b** and **Fig. 4.6b**, under glucose condition, the △*cra* strain already

faces 24% FBP reduction compared to the wild-type strain. Combined of 25% lower FBP on gluconate condition and 24% lower FBP due to *cra* deletion, it suggests the intrinsic shortage of FBP is enhancing when cells are cultivated on gluconate than on glucose substrate. Such insufficient situation is further aggravated both *cra*-deleted strains bearing pBAD-only or pBAD-Cra plasmid once the glycerol pathway is turned on, retarding the host growth severely. Moreover, without the Cra-regulator, the uptake rate of gluconate might reduce, because Crp, as a Cra-activating protein, will remain less active and unable to activate gluconate transporters (**Supplementary figure 1**) causing strong stress to the cells. In sum, compared to glucose condition, there are two feasible reasons for the strong growth impacts, no adequate replenishment of FBP (gluconate only metabolized by ED pathway) and inefficient gluconate uptake due to lack of Cra.

On the other hand, when fructose was used, the wild-type strain harboring pBAD-Cra plasmid reduced both growth and production compared to glucose cultivation, presenting the same misregulation of the wild-type strain bearing pBAD-only plasmid. The misregulation is due to the tug-of-war of FBP which on the one hand inhibits Cra-regulator and on the other hand is continuously drained by GPD1. When cells grow on glucose or fructose condition, host's ED and PPP pathway (as shown in **Fig. 4.5a**) serve as branch points for the stronger cycling of carbon back to central metabolic pathways. That's why *cra* deleted strain bearing pBAD-only or pBAD-Cra plasmid still keeps growth dissimilarity from gluconate conditions. Taken together, all the results indicate the Cra-regulation is indispensable for the host to respond to various environmental changes and by providing fructose the best producer in this thesis can be deceived and lose the balance between growth and production.

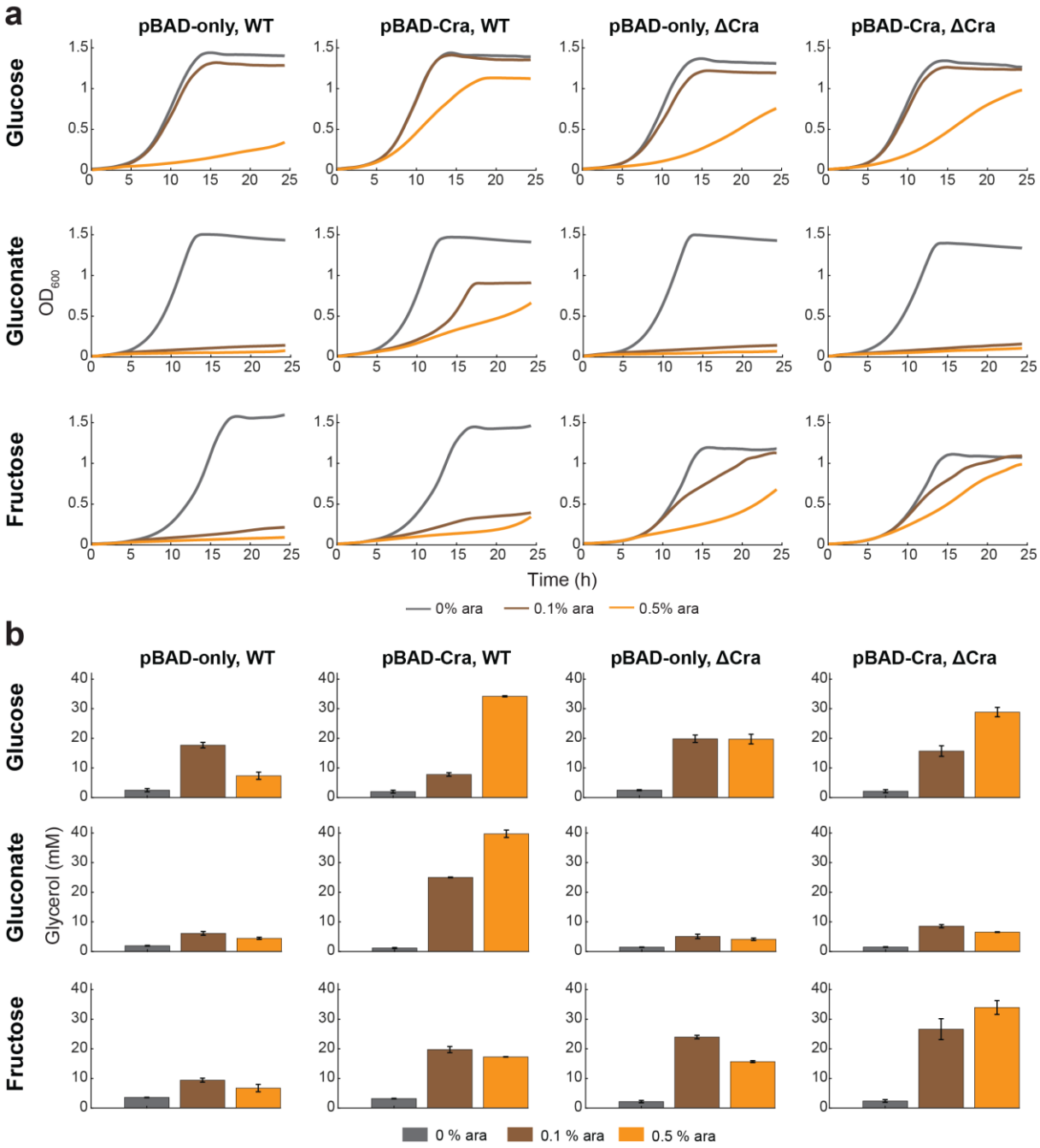


Fig. 4.7 | A comparison of growth and glycerol production in four strains were cultivated separately in three different carbon sources (glucose, gluconate and fructose). a, Growth curves, OD_{600}. **b,** Glycerol titers measured after 24 h. The pBAD-only and pBAD-Cra plasmids were separately expressed in wild-type and *cra* deleted backgrounds to get four strains and these strains were grown separately in 0.5% of different carbon substrates. The glycerol production pathway was induced by 0, 0.1 and 0.5% arabinose.

4.2.4 Cra-regulated system makes cells robust against osmotic and oxidative stresses

We knew how indispensable Cra-regulation is for the host's to properly respond to carbon source changes. Moreover, one research group has shown that deletion of Cra causes down-regulation of BeA/B/T which protect cells from osmotic stress[124]. Inspired by this resulting, we are wondering what could happen when the Cra-regulating hosts are exposed to osmotic (NaCl) and oxidative stresses (H_2O_2). It turned out that only pBAD-Cra with functional Cra was able to reach stationary phase within 24 h (**Fig. 4.8a** and **b**). Although this is not a fair comparison, *E. coli* reduces growth at 0.5% arabinose induction (**Fig. 4.7a**). It still reveals how vital Cra-regulated needs for cells to fight both intracellular (glycolytic overloaded) and extracellular stresses. Although this is not a fair comparison, wild-type strain harboring pBAD-only plasmid already reduces growth at 0.5% arabinose induction due to misregulation. It still reveals how vital Cra-regulation is needed for cells to fight against both intracellular (glycolytic overloaded) and extracellular stresses (NaCl or H_2O_2).

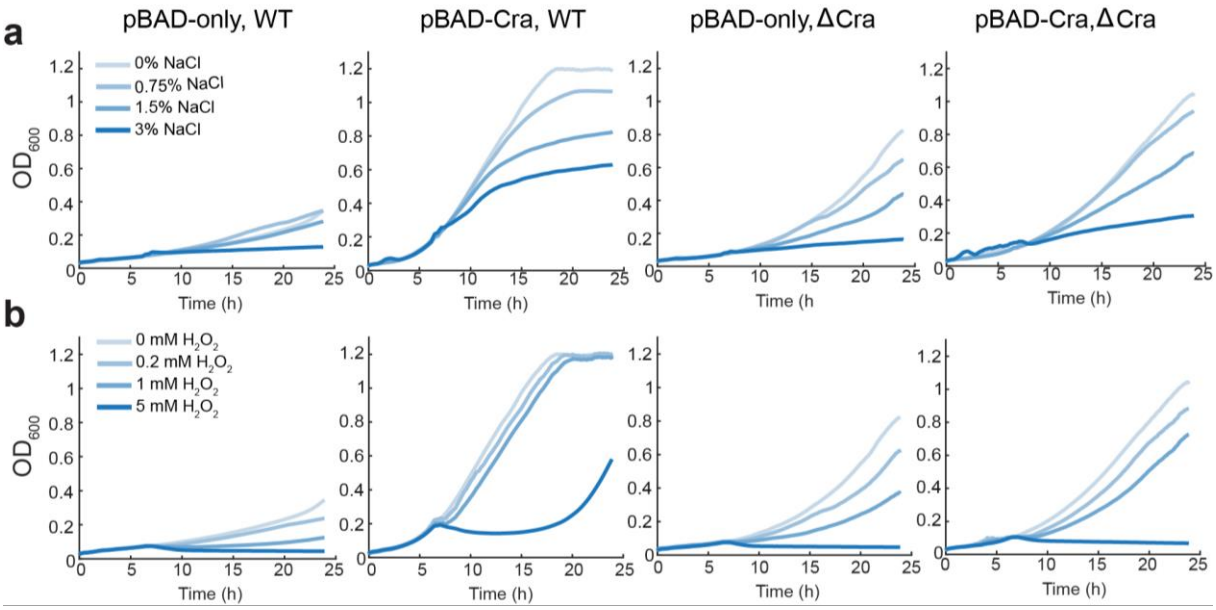


Fig. 4.8 | Increment cell tolerance toward stresses (NaCl and H_2O_2) in functional Cra-regulation strains. a, Osmotic stress (NaCl) **b,** Oxidative stress (H_2O_2). Strains bearing specific glycerol production plasmid were cultivated in 0.5% glucose M9 medium supplemented with 0.5% arabinose.

4.2.5 Reduced carbon dioxide emissions during glycerol production

Usually, the growth behavior is linearly coupled to CO_2 production. Nevertheless, we accidentally observed that the CO_2 production was reduced when the glycerol pathway was turned on in the applied bioreactor. From metabolic map in **Supplementary figure 1**, we postulated that *E. coli* may try to reroute the metabolic pathway by up-regulating glyoxylate proteins (AceA and AceB) so that the cells can bypass two steps of TCA cycle which produce CO_2 (from isocitrate acid to succinyl-CoA). Previous researches have shown that few key enzymes (PckA, Ppc, MaeA and MaeA) involved CO_2 fixation in *E. coli*, and the PckA account for CO_2 fixation under gluconeogenesis or high hydrogen carbonate (HCO^{3-}) concentration[125, 126]. In this thesis, we found the expression of PckA was increased throughout the time. Therefore, increasing the glyoxylate shunt and gluconeogenesis compensate for glycerol overproduction and base on this rerouting pathway, the cells reduce CO_2 production at a certain time.

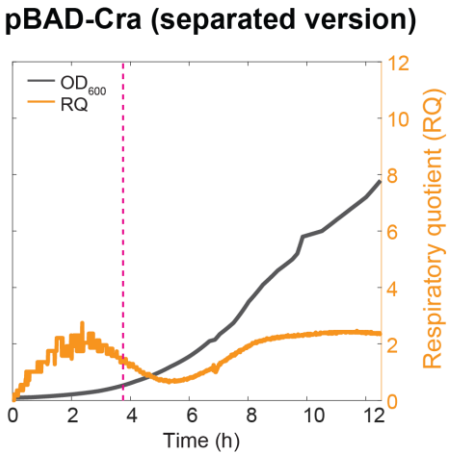


Fig. 4.9 | The growth and respiratory quotient of pBAD-Cra strains were measured in a batch reactor. The pBAD-Cra strains bearing fused *gpd1-gpp2* or separated *gpd1-gpp2* genes show a dropping in CO_2 production at the time when arabinose was added indicated by the pink dashed line. The black line is OD_{600}. The orange line is respiratory quotient.

4.3 Conclusion

The Cra-regulation system can be further applied to various types of promoters with highly-applicable controllability. Although the chosen promoters still have a dominant impact, such as the leakiness (exemplified by aTc inducible promoter) or the strength of promoter, on the transformer's behavior, the incorporated Cra-regulation system can fine-tune the metabolic networks to optimize internal defects. On the other hand, the five-carbon sugars (F6P, FBP and F1P) can be utilized to rescue the overloaded glycolysis in pBAD-only strain. Even though we had not considered the permeability of sugar transporters, this still gives the evidence that Cra-regulator is regulated by FBP/F1P as literature has demonstrated[24, 54, 55, 61]. Overall, the Cra-regulation system is vital not only for balancing intracellular burden as the trade-off between growth and production but also for overcoming extracellular stresses/changes such as osmotic stress, oxidative stress and different carbon sources.

4.4 Future work

4.2.1 Why cells use FBP/F1P as a flux sensor?

When fructose is used, the pBAD-Cra strain loses the balanced benefit of Cra-regulation. Does it mean fructose and fructose-transformed five-carbon sugars (F6P, FBP and F1P) are a signal factor for the cells or does it have other physiological meaning?

4.2.2 The viability of using glycerol production strains for CO_2 fixation.

The result here implies producing glycerol leads to rewrite the metabolic pathway and reduce CO_2 emissions. However, it has been known that metabolic repair through the emergence of new pathways in *E. coli*[127]. If our engrained cells also emerge new pathways and can bypass CO_2 production, can we combined the glycerol pathway with additional CO_2-fixing enzyme such as RuBisCO[128]? By doing so, do the cells improve CO_2 fixation?

5. Chapter 5 – A burden monitor to trace cells' physiological states

5.1 Research questions and objectives

In bioengineering, the goal is to understand the trade-offs among gene expression, cell growth, and production. We designed a simple glycerol pathway in *E. coli* as a case study. The GFP gene was integrated into the genome, which we called it a burden monitor, to detect and reflect the physiological states of the cells. Moreover, the burden monitor can be widely used for the screening overproduction.

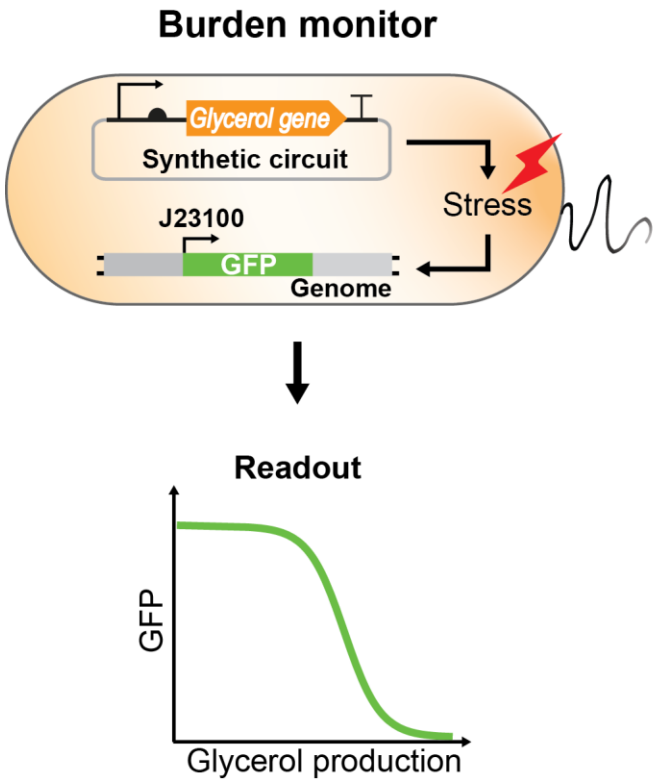


Fig. 5.1 | Application of the burden monitor, a GFP gene integrated in *E. coli* genome. The glycerol pathway was expressed in the strain with genomic integration of GFP as a stress-sensing controller (burden monitor). The expression of the glycerol pathway competes for the same resources in host cells and further causes stress so that the fluorescence level is negatively related to glycerol production.

5.2 Results and discussion

5.2.1 The selection of a suitable expression system for glycerol production

To study which promoters and plasmid backbones can maintain both growth and production, we cloned parallel total 92 (23 x 2 x 2) plasmids including 23 various promoters, 2 different copy-number vectors and 2 kinds of genes (GFP and glycerol) as shown in **Fig. 5.2**. As mentioned many times in precious chapters, the GFP protein is used as a reference because unlike glycerol production pathway draining DHAP from glycolysis, GFP protein does not directly participate in the host's metabolic networks. Basically, the expression level of GFP can represent not only the promoter strengths but also the protein burden in the host cells. On the other hand, the expression of GPD1-GPP2 fusion protein reports pressures resulting from both protein burden due to expression of heterologous pathway and an internal metabolic burden due to GPD1 draining glycolytic metabolites. All the plasmids are listed in **Supplementary Table S1**.

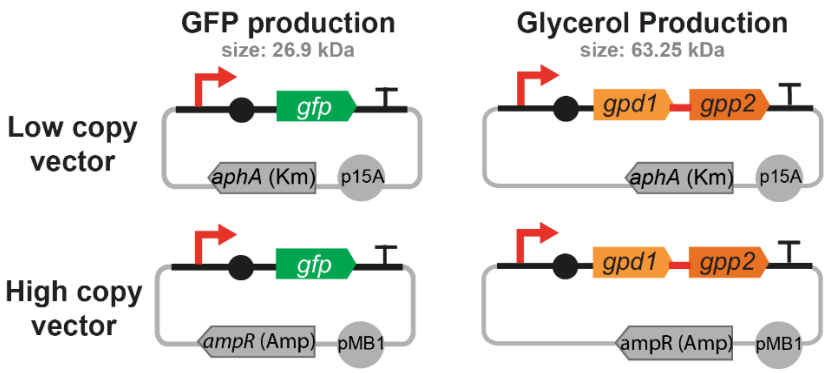


Fig. 5.2 | Four types of plasmid-based systems for selecting a suitable expression system. The GFP or glycerol gene was driven by different promoters and cloned to low copy vector- p15A or high copy number vector- pMB1 separately.

Specifically, we cloned 4 constitutive promoters (different strength from high to low promoter activity: J_{23101}, J_{23100}, J_{23105} and J_{23109} from iGEM repository[129-134] and the abbreviation of these promoters is J01, J100, J05 and J09 in this chapter) and 19 native promoters from *E. coli*. The 19 native promoters include 5 Cra activated promoters (P_{icd}, P_{ppsA}, P_{cydA}, P_{crp} and P_{pck}), 7 Cra inhibited promoters (P_{pykF}, P_{pfkA}, P_{gapA}, P_{fbaB}, P_{gpmM}, P_{ppc} and P_{epd}), and 7 non-Cra-responsive promoters (P_{soxS}, P_{acs}, P_{galSP}, PptsGp1, P_{galEp1} and P_{gltB}).

First, we measured the GFP production and compared the fluorescent level at OD_{600} reaching 0.5. **Fig. 5.3a** shows that the GFP signal of 22 different promoters is mostly higher, except P_{ppc} and P_{epd}. Second, to test the promoter activity according to carbon source, total 46 different strains (23 x 2, high or low-copy vector) were grown in M9 medium containing 0.5% glucose or 0.5% pyruvate. The result is consistent with the study from Leeat et al.[88] that promoters maintain relative expression values across different conditions (**Fig. 5.3b**).

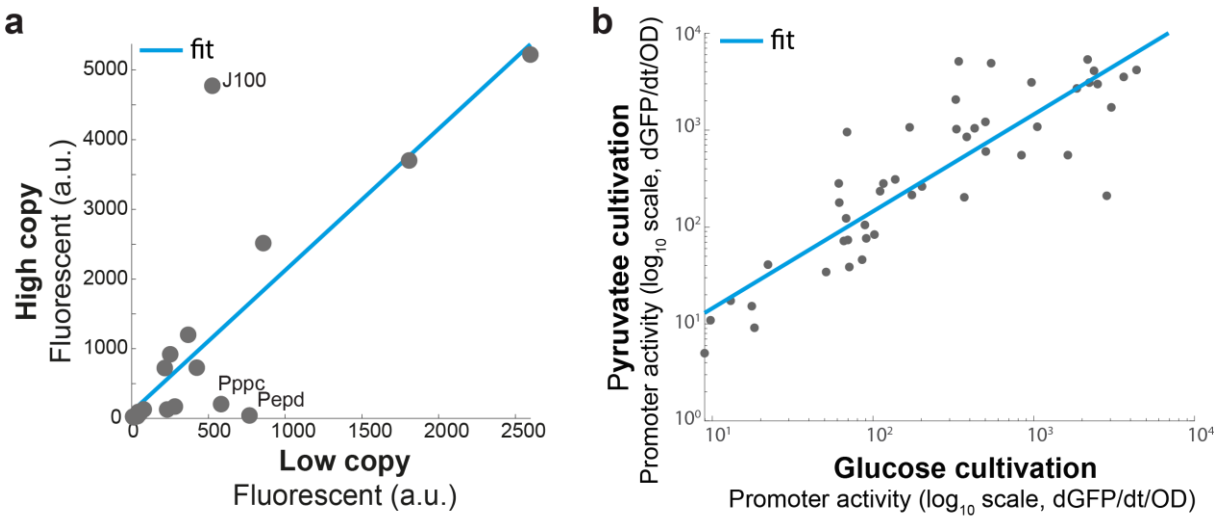


Fig. 5.3 | The characteristic of GFP production under plasmids with different copy number and carbon sources. a, Fluorescence was measured at OD_{600} of 0.5. P_{gap} was removed due to no GFP signal in the high copy number (total 22 promoters). **b**, Promoters maintain relative activity under glucose and pyruvate conditions. The cells harboring plasmid with different promoters were cultured in M9 medium with 0.5% glucose or pyruvate. (n=2)

When we tested the stability of the plasmid-based system, the results show that, in general, the GFP expression levels in the low-copy system (**Fig. 5.4a**) is more stable than in the high-copy system (**Fig. 5.4b**). The combination of the high-copy system with the constitutive promoter is especially unstable; for example, the constructs are comprised of high-strength promoters, J01 and J100.

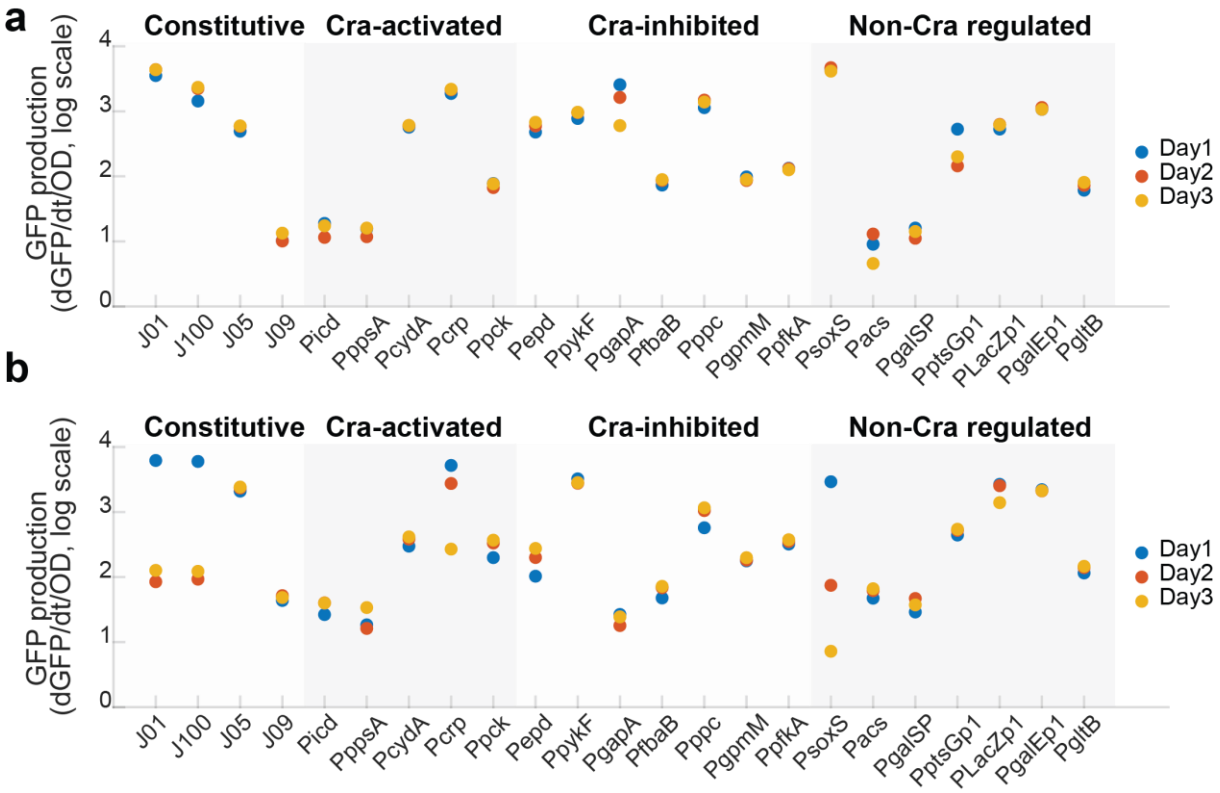


Fig. 5.4 | **Stability of plasmid-based GFP prodcution system. a,** The GFP production of the low-copy plasmid. **b,** The GFP production of the high-copy plasmid. All the experiments were conducted in 96-well plates, and the overnight cultured cells were re-diluted every day with fresh M9 medium containing 0.5% glucose. The OD_{600} and GFP were measured with 5 min interval for 24 h.

To elucidate the unstable GFP production of the constitutive promoter, we furthermore investigated the difference between high and low copy plasmids. The strongest promoter (i.e. J01) and the mild promoter (i.e. J05) were taken and measured by fluorescence-activated cell sorting (FACS). Within the high-copy system, there are two distinct peaks that represent sub-populations in the strongest promoter as J01(**Fig. 5.5**). We then sequenced the cultures and found there was insertion-sequence (IS) as IS10 or IS5 inserting between the promoter and RBS region. Therefore, although the high-copy system can provide more GFP, namely higher product yield, it can be problematical[135] for product overproduction due to sub-populations/heterogeneity from mutation[136, 137].

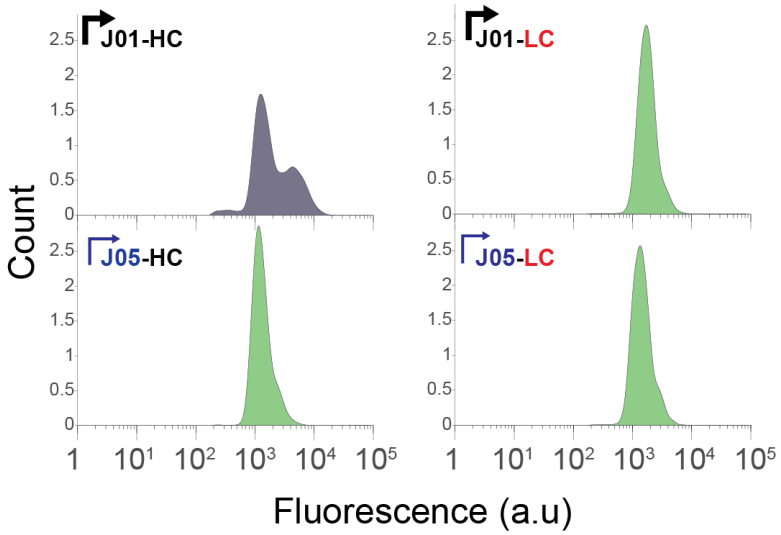


Fig. 5.5 | FACS analysis of GFP production under the control of different copy number plasmids. The cells were cultivated in M9 medium with 0.5% glucose and were measured at the exponential phase by FACS. HC: high-copy number plasmid; LC: low-copy number plasmid. J01 and J05 two constitutive promoters with strong and medium activity, respectively.

5.2.2 Various promoters for glycerol overexpression

Since heterogeneity is one of the hard challenges for industrial-scale biomanufacturing[6], to solve this problem, we decided to use the low-copy system for a glycerol overproduction with suitable stability. According to previous results in **chapter 2.2.4**, to prevent the glycerol re-utilization, the Δ*glpK* strain of *E. coli* MG1655 was used for all the following experiments in this section. Therefore, we screened at first these 23 various promoter systems for glycerol production and expressed each system in Δ*glpK* strain. Generally, the constitutive versions were less reproducible, and the native regulated promoters mostly could maintain consistent glycerol production (**Fig. 5.6**). Here, the two Cra-inhibited promoters- P_{gpmM} and P_{pfkA} reached relatively higher glycerol concentration among all the promoters, giving a hint that Cra-inhibited promoters can be applied for glycerol production.

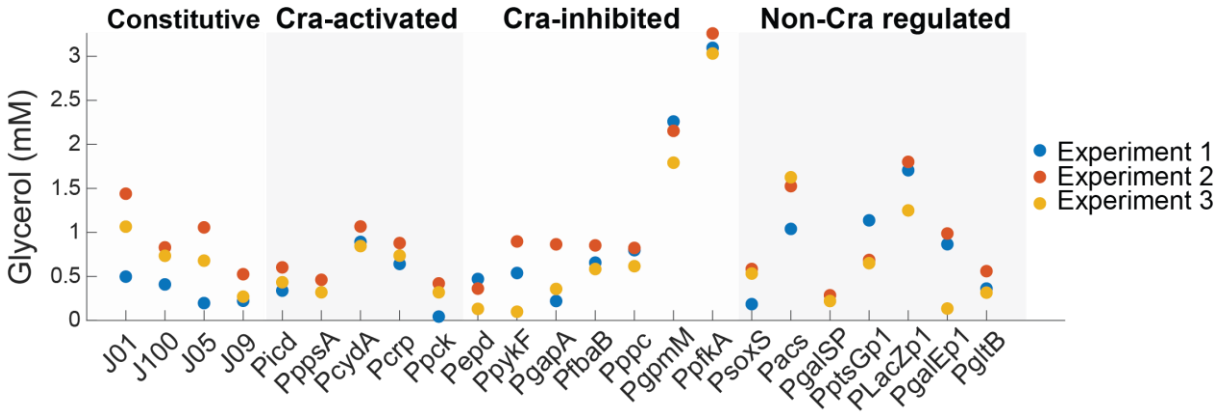
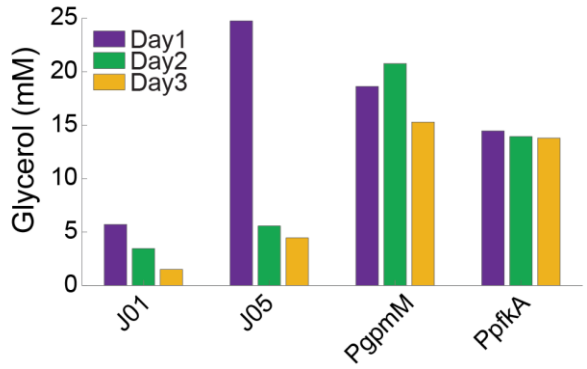


Fig. 5.6 | The reproducibility of glycerol production under the control of various promoters. The glycerol production genes were driven by constitutive, Cra-regulated and non-Cra regulated promoters. All strains were cultured in M9 medium with 0.07% glucose and the glycerol concentration was measured from 3 independent experiments after 24 h cultivation.

To test the stability of as-designed glycerol production systems, like in the GFP production (as shown in **Fig. 5.4**), the 2 strains, P_{gpmM} and P_{pfkA}, were selected and compared to constitutive promoters by growing them in culture tubes. The supernatants were harvested after 24 h for continuous three days. Glycerol production strains under the control of constitutive promoters (J01 and J05) declined with the day, while regulated promoters (P_{gpmM} and P_{pfkA}) could keep consistent glycerol titer throughout the 3 day-experiment (**Fig. 5.7a)**. Interestingly, the IS1 was only found in constitutive promoters (**Fig. 5.7b)**, meaning that stress is induced in the host cells. Because the IS element is usually activated to maintain the host's fitness[136, 138], by blocking the glycerol genes expression which is regarded as the sources of stress, the hosts can be released from the burden and able to grow normally again.

a



b

	Day1	Day2	Day3
J01	-	O	O
J05	-	-	O
PgpmM	-	-	-
PpfkA	-	-	-

Fig. 5.7 | The stability of glycerol production by daily measurement. a, The glycerol pathway was expressed from different promoters (J01, J05, P_{gpmM} and P_{pfkA}). Glycerol concentrations were measured every day at the same time. **b,** An insertion-sequence (IS) element was detected by sequencing. The open circle presents IS element was detected. The dashed line presents no IS1 element was found. All the strains were cultivated in culture tubes with M9 medium for continuous three days.

According to this result, we conclude that, unlike GFP production (**Fig. 5.4a**), producing glycerol makes systems more unstable even under low-copy plasmid. The reason for such sensitive instability could be that producing glycerol is not only taking common resources from the fundamental host building block causing protein burden but also draining glycolytic metabolites causing metabolic burden which overloads the central metabolism. Overall, the synthetic glycerol pathway in the *E. coli* is successfully introduced into the plasmid-based systems with well function.

However, when we plotted the growth rates and glycerol productions of all 23 promoters, insignificant relevance is observed (**Fig. 5.8**). This phenomenon could be attributed to the general effect due to the applied promoters with comprehensive characteristics and strength. Considering that the high diversity of promoters may complicate a systematical study, we decided to unify the system via the arabinose-inducible promoter owing to the dose-dependent property and the non-toxic arabinose inducer[37, 139, 140].

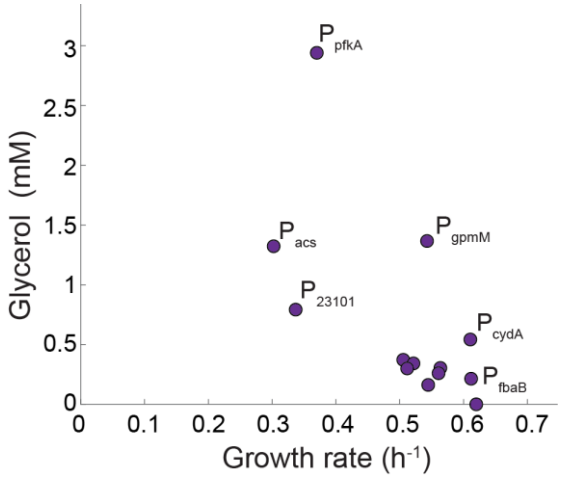


Fig. 5.8 | Growth rate versus glycerol production on a low-copy plasmid. The cells were cultured in M9 medium supplemented with 0.07% glucose and the growth rates were calculated at OD_{600} around 0.2. The glycerol concentrations were measured after 24 h.

5.2.3 The most stable plasmid-based system for the glycerol overexpression

Based on **chapter 5.2.2**, we decided to use the pBAD promoter to explore glycerol production systematically. As a starting point, we cultivated wild-type *E. coli* in various arabinose levels to confirm that arabinose is not toxic for the host cells. The highest concentration tested here was 2% arabinose, and there was difference in growth among all these tested levels (**Fig. 5.9**), confirming applicable arabinose is harmless for the host cells.

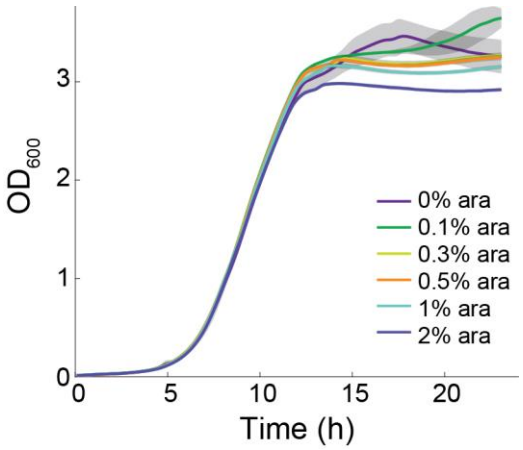


Fig. 5.9 | Arabinose titration of wild-type *E. coli*. Wild-type *E. coli* strain was grown in M9 medium supplemented 0.5% glucose and the growth curves were monitored for 24 h in a 96-well plate. Arabinose levels were 0, 0.1, 0.3, 0.5, 1, 2%. (n=2).

The GFP gene was implemented on high and low-copy plasmids as a reference. In both backbones, the promoter activity of **Fig. 5.10a** increases proportionally with arabinose concentration but the high-copy backbone shows a higher basal expression even before induction as shown in other's study[141]. In spite of higher basal expression, the GFP production is correlated linearly to induction levels and also copy numbers (**Fig. 5.10b**) which is consistent with the results of Kafri's group[142].

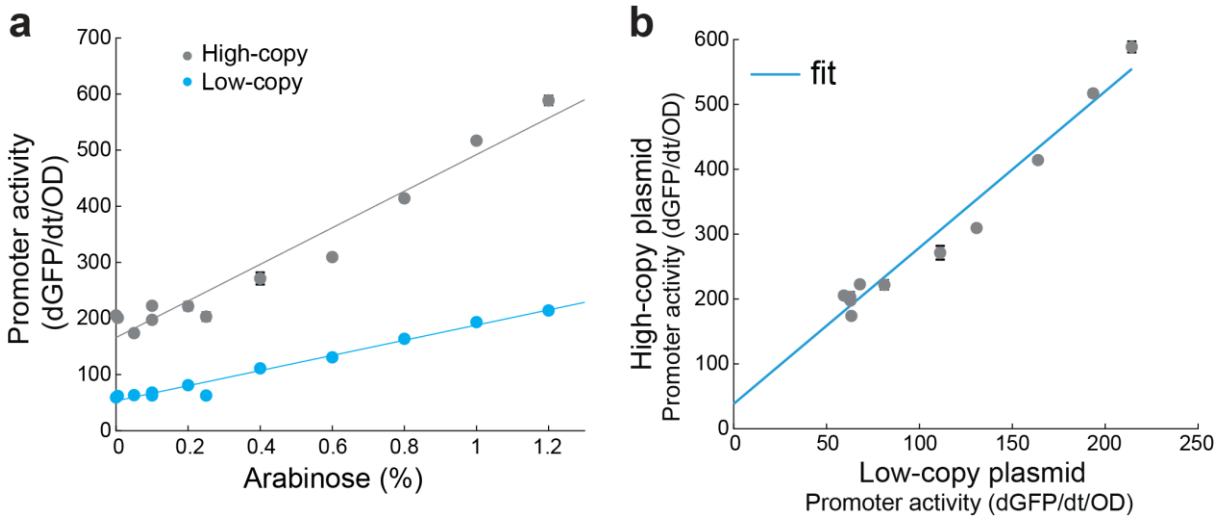


Fig. 5.10 | The promoter activity of high- and low-copy number plasmid under different arabinose concentrations. a, Promoter activity of high or low copy vector increases with arabinose induction levels. **b,** Promoter activity of high-copy number plasmid is well correlated with low-copy number plasmid. Arabinose induction levels are 2-fold serial dilution from 1.2% arabinose. The cells were grown in M9 medium supplemented with 0.5% glucose and inoculated in a 96-well plate. (n=2).

Next, we cloned the pBAD-driven GPD1-GPP2 genes in 3 different copy numbers that a genome integration strain represents one copy number and both high and low-copy number plasmids are 10-12 and 100-300 per cell separately[143, 144]. As expected, glycerol production is a linear correlation to arabinose concentrations except the high-copy vector that glycerol production decreased at 1.2% induction (**Fig. 5.11**). Unlike genomic integration and low-copy system, the high-copy plasmid produced less glycerol at 1.2% arabinose that is attributed to the side effect of high-copy backbone, such as protein aggregation or non-functional misfolded protein[87]. Thus, different from simple GFP production, glycerol production as a more complicated synthetic loop interconnected to central metabolism is more sensitive to the variation of host's

resource, so finding a suitable expression level of protein is critical to realize and improve glycerol production.

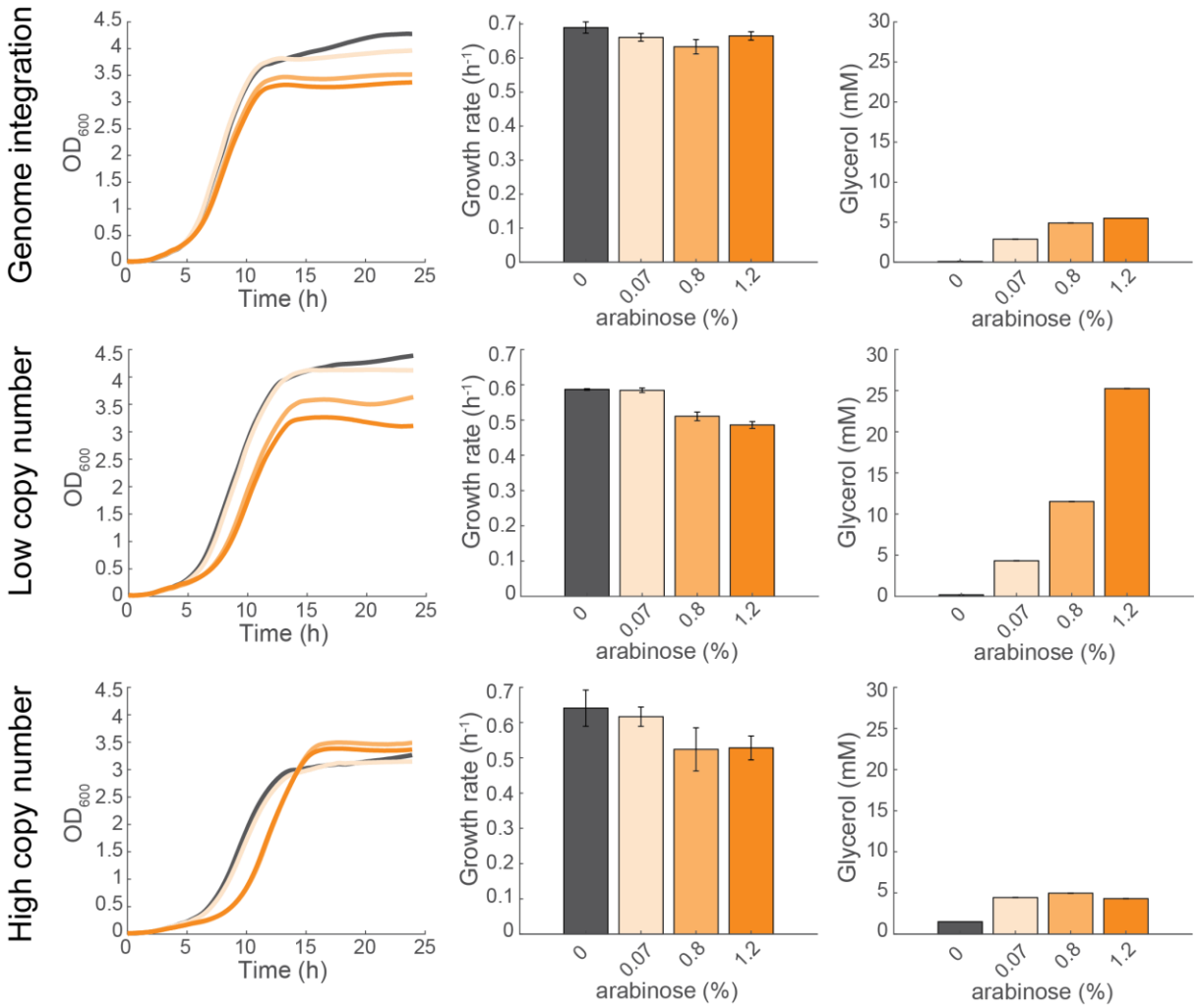


Fig. 5.11 | Observations of glycerol gene expression in 3 different backbones (genomic integration, low- and high-copy number) driven by pBAD promoter. All the strains were grown in M9 glucose (0.5%) medium and the glycerol concentrations were detected after 24 h. (n=2).

5.2.4 Burden monitor was designed to detect the physiological state of glycerol producer

We attempted to develop a facile strategy, by which we could monitor the burden instantly inside the cells. In most cases, when the host cells encounter stresses which can come from both inside, such as protein overexpression and metabolic flux burden,

and outside like environmental changes, a reduced growth is commonly observed in response to the stresses. Sometimes, there is no growth phenotype, but the cells already suffer from stresses. The benefit of the real-time detection is that the internal condition of a target host can be revealed, and then the culture conditions or induction concentration can be fine-tuned in a timely manner based on the newly-obtained signals. Therefore, we inserted GFP sequence at intergenic site named as safe site 9 (SS9) from previous study[76] to clone a genome-based GFP strain called "burden monitor" strain through amplifying coding region of a burden-monitor plasmid from Ellis's group[9].

To demonstrate the concept of real-time detection, the genome-based glycerol producer was chosen because of no/less-growth effect after induction (**Fig. 5.12a** and **Fig. 5.11**). There was only 9% reduction of growth rate at 1.2% arabinose (**Fig. 5.12b**), but 26% reduction in GFP production (**Fig. 5.12c**). This proofs that the burden monitor indeed reflects the cells' states and can be applied as a more sensitive way than common optical density (OD) measurement of bacterial culture.

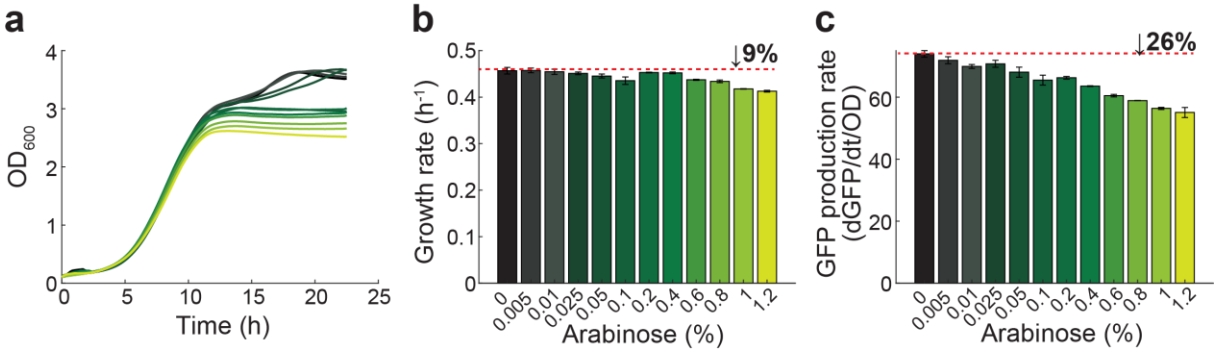


Fig. 5.12 | The detection result of the burden monitor in a genome-based glycerol production strain (pBAD-G). a, Growth curves were measured with 5-min intervals for 24 h. **b,** The growth rates among different inducers were calculated according to the appendices of materials and methods. **c,** The GFP production rates of the burden monitor are dropped with increasing arabinose concentrations. Arabinose induction levels are: 0, 0.005, 0.01, 0.025, 0.1, 0.2, 0.4, 0.6, 0.8, 1, and 1.2%. The numbers above the red lines were the comparison of 1.2% induction with the uninduced condition.

5.3 Conclusion

It is hard to find an optimal enzyme level in synthetic pathways. Through trial and error, we realized that using strong promoters or high-copy number plasmids for glycerol production is not a guaranteed approach to boost product titer. Although stronger promoters enhance the protein production[145-147] such as GFP in this thesis, a higher protein amount does not always refer to a higher product titer. Sometimes, under higher copy plasmid, sub-populations in the cell culture reduces the productivity and is a critical obstacle for production at an industrial scale. In general, for the engineering of desired chemicals, the bottleneck usually coming from the shortage of precursors[21, 148], co-substrates or energy[149]. To ease these deficient issues, the whole metabolic networks have to be taking into consideration for a better-designed and well-balanced pathway.

GFP production is relatively stable than glycerol production among different conditions and constructed backbones because the synthesis of GFP consumes only secondary metabolites (amino acid) instead of draining primary metabolites (DHAP). That's why GFP can be utilized as a burden monitor to reflect the internal stress from the cells' resource competition.

5.4 Future work

5.4.1 The approach for increasing production titer

Many other options have already mentioned in **chapter 2.4.2**.

5.4.2 Burden monitor for high throughput screening production strains

With the combination of burden monitor and fluorescence-activated cell sorting (FACS), it provides a time-saving method to screen the potential overproduction strains from a library.

6. Chapter 6 – CRISPR interference for gene perturbation

6.1 Research questions and objectives

In this chapter, we aimed to investigate the tightness and the potential application of modified CRISPR interference (CRISPRi) system into overproduction systems. We chose the CRISPRi system from Qi's research[68] which is a two-plasmid based system consisting of an aTc inducible promoter expressing dCas9 and a constitutive promoter for sgRNA. To improve the controllability of the CRISPRi system, various modified systems were constructed as shown in scheme 6.1. Then, the refined systems were applied to target a specific gene in the chosen biosynthetic pathway.

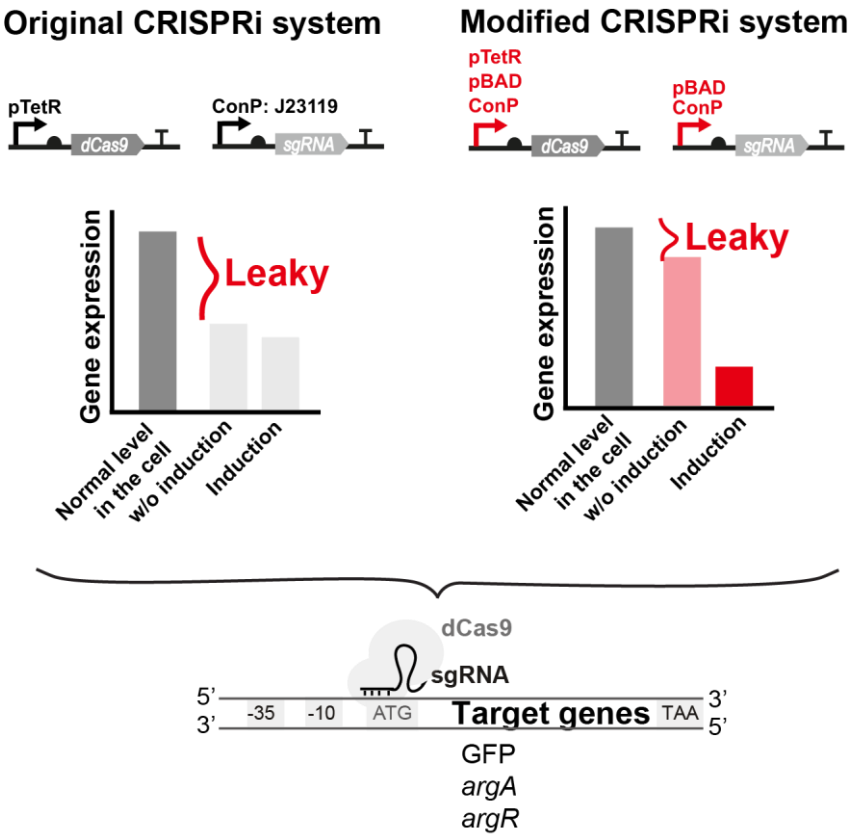


Fig. 6.1 | The difference between original and modified CRISPRi systems. We tried to reduce the leakiness of original CRISPRi system from Qi's research[68] by several different promoter combination and texted them in different target genes such as GFP, *argA* and *argR.*

6.2 Results and discussion

6.2.1 Characterization of CRISPRi systems by targeting GFP

Before jumping into the specific metabolic pathways, GFP was constructed as a reference and further for a fair comparison. As **chapter 5**, we had created a GFP genome-based system named as “burden monitor” to detect the real-time state of host cells. Herein, by using this burden monitor strain as a GFP reporting signal, we could compare and screen suitable CRISPRi systems with enhanced tightness simply by monitoring the change of GFP expression. In principle, once the CRISPRi system is turned on, the GFP level producing by the burden monitor will decrease, and furthermore, a comparison of the sensitivity between on/off switch can reflect the tightness of individual CRISPRi system. To develop an inducible CRISPRi system with high sensitivity, series of combinations were formed via introducing varied promoter and ribosomal binding site (RBS) region, the as-designed CRISPRi systems were evaluated and compared with the original version developed by Qi’s research[6]. I In order to estimate the fluorescence decrement arising from targeted GFP, an empty plasmid, named as CTRL, was utilized as an experimental control. The applied CTRL was lacking 20 bp sgRNA sequences to target the GFP gene, but in the experiment systems the gRNA-GFP-2 with 20 bp sgRNA was used to target the region of GFP transcriptional initiation.

Fig. 6.2 illustrates that, even without aTc, a serious leaky inhibition of GFP expression was presented by the original system. The original system is comprised of two plasmids, that is, dCas9 driven by tetR promoter and sgRNA driven by a constitutive promoter. To verify such major leaky effect was caused by which plasmid, the other two systems consisted of either constitutive dCas9 or dual-regulated were constructed. The constitutive system is that the original TetR promoter of dCas9 was replaced with a series of constitutive promoters with increasing promoter activity from J05-73, J05-37, J114 to J05. The modified constitutive system shows that in the control group without GFP sgRNA, the promoter activity is negatively correlated to the dCas9 expression, suggesting the increasing promoter activity is mainly attributed to the growth retardation due to protein burden from dCas9 overexpression. When the GFP sgRNA is expressed, the strong knockdown of GFP can be observed and shows no difference between various dCas9 expression. The results also point out that dCas9

can serve as a powerful controller because with the existence of the GFP sgRNA slight dCas9 expression is already enough to knockdown GFP expression but a high amount of dCas9 like J05 group cause strong growth effect even without targeting GFP genes. Based on the observation, we realized that more dCas9 expression can only impose a severe burden in the host cells instead of repressing gene expression. Thus, the lower level of dCas9 is enough to perform repressing targeted gene expression and a high amount of dCas9 expression will cause protein burden and damage the host cells.

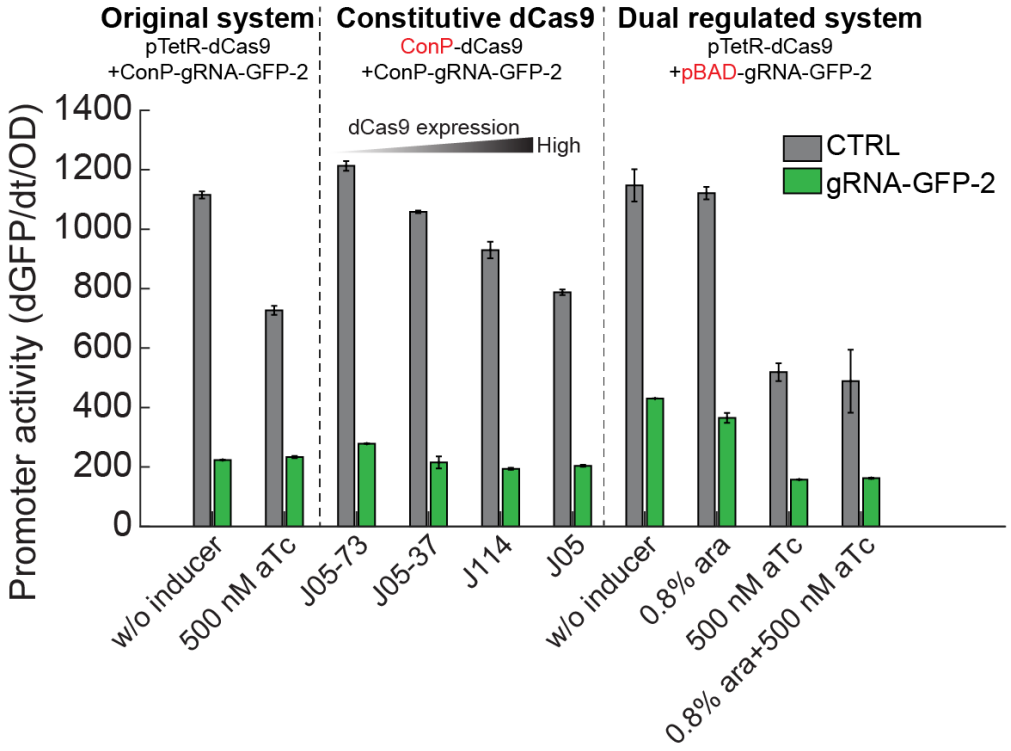


Fig. 6.2 | Characterization of 3 different CRISPRi systems. We expressed 3 CRISPRi systems (original, constitutive and dual regulated systems) in the burden monitor strain and measured the promoter activity with specific inducers induced at the beginning. The original CRISPRi system is a two-plasmid system comprised of one TetR promoter driving dCas9 and one constitutive promoter driving sgRNA. Constitutive dCas9 system replaced the TetR promoter of original dCas9 by different constitutive promoters. The promoter activity from weak to strong is J05-73, J05-37, J114 and J05. The dual-regulated system replaced the constitutive promoter of the original sgRNA by arabinose inducible promoter. CTRL represents the plasmid without 20 bp sgRNA targeting the GFP gene as an experimental control. The gRNA-GFP-2 represents the plasmid with 20 bp sgRNA complementary to the GFP coding region. All the plasmids used are in **Supplementary Table S1**.

To confirm that the stress is mainly coming from dCas9 rather than sgRNA, the dual-regulated system was cloned by replacing the original constitutive promoter of sgRNA plasmid with an inducible pBAD promoter. The control group (CTRL) showed that once the aTc was added, the dCas9 was expressed, and then the growth effect was induced so as to restrain the promoter activity. On the contrary, once the arabinose was added, the sgRNA was expressed but there was no growth effect so as to maintain the promoter activity as the strain without inducer. The results explain the burden of a two-plasmid based CRISPRi system is mainly from dCas9 expression. When the GFP sgRNA was expressed, compared to the original and constitutive systems, an improved tightness can be achieved by the dual-regulated system. Taken together, the result evidences that the overexpression of dCas9 is the main cause leading to a significant protein burden in the host cells, which is consistent with the result of constitutive dCas9 expression.

6.2.2 Characterization of CRISPRi systems in the arginine pathway

As illustrated in the previous section (**chapter 6.2.1)**, we first applied the original CRISPRi system to control GFP expression and determined that the major burden is from the overexpression of dCas9. To verify the concept and further investigate the interference effect in biological physiology, we then designed CRISPRi plasmids to target the *argA* gene in the arginine biosynthetic pathway. As expected, dCas9 expression is a negative correlation to growth rate, and the strongest promoter (J05) showed a restrained growth by even 61% reduction (**Fig. 6.3a**). Once again, it proves that the strong expression of dCas9 is not a proper way for genetic perturbation.

Considering the eminent controllability of the dual-regulated system and the reduced leakage compared to the original system, we also constructed the dual-regulated system to target *argA* gene. As shown in **Fig. 6.3b**, the growth remained similar to the non-induced condition by expressing only the sgRNA, but a 38%-reduced growth was detected even by only the dCas9 expression. When both plasmids were expressed, a slightly-degreasing growth rate from 38% to 42% was observed. The result is corresponding to the observed phenomenon in GFP, advocating that only a few amounts of dCas9 expression are enough to switch off gene expression and the more dCas9 reversely stresses the cell. Besides, when this dual-regulated system was

turned on, it showed a much longer lag phase compared to the original CRISPRi system (**Fig. 6.3c**). Taken together with the results from **chapter 6.2.1** and **Fig. 6.3b**, we once again confirmed that dCas9 is a major burden for the host cells, and only sgRNA expression is less effective to the growth; moreover, a combination of dCas9 and sgRNA will trigger more severe growth retardation.

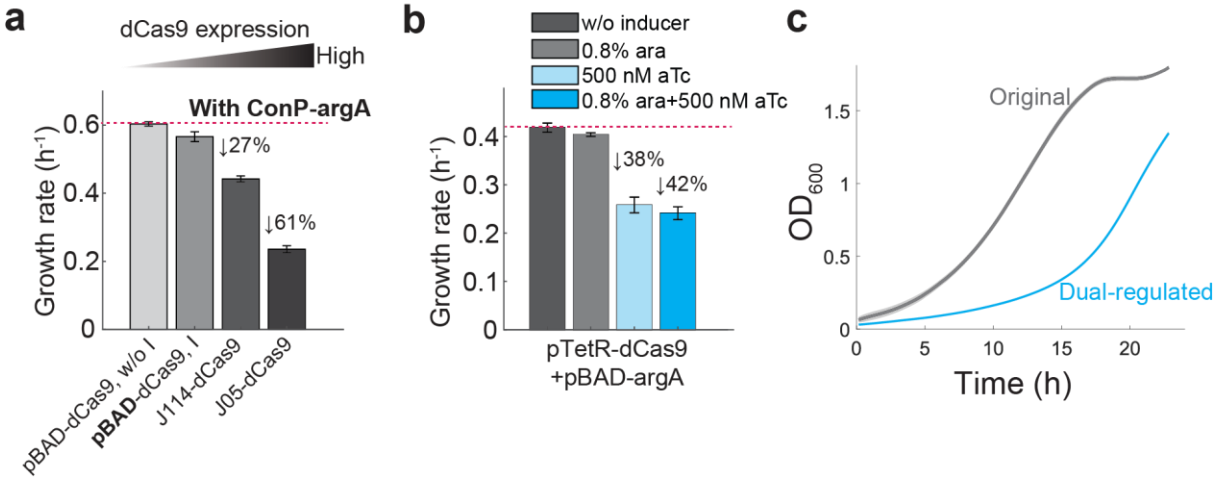


Fig. 6.3 | Different CRISPRi systems targeted *argA* gene of the arginine pathway. a, The effect of dCas9 expression levels. The growth rates of original CRISPRi perturbation. Different strength of dCas9 expressed under 3 different promoters (the promoter strengths from weak to strong are pBAD, P_{J23114} and P_{J23105}). **b,** The dual-regulated CRISPRi system consists of aTc induced promoter for dCas9 and arabinose induced promoter for expression of argA-sgRNA. **c,** Growth curves of original CRISPRi system (from Qi's research[6]) and constructed dual-regulated system from (b).

Accordingly, we realized that the proper quantity of dCas9 is important to execute the CRISPRi system properly. Furthermore, surpassing the conventional system the dual-regulated system is demonstrated as a promising and competitive alternative to provide synthetic pathways with improved controllability and reduced leakage. However, considering the complexity of 2 inducers and strong growth impact (i.e. long lag phase) at the controlled experiment, the dual regulated system is still insufficient and has to be further improved. Generally, we tightened the expression through the dual-regulated system and found out that the leaky expression of dCas9 is a critical problem in the CRISPRi system as pointed out in other studies[10].

In a practical way, using the dual-regulated system was actually under the sacrifice of a much longer lag phase; therefore, considering the productivity and analytic property, we still decided to use the original CRISPRi system to study the genetic perturbation. To investigate the efficiency of CRISPRi system, by knocking down the *argA* gene in arginine biosynthetic pathway (**Fig. 6.4a**), we continuously evaluated the CRISPRi system in the term of growth rate measurement and analysis of metabolites. When the plasmids of the CRISPRi system were expressed separately in *E. coli* MG1655, the growth rate of only dCas9 plasmid was reduced by 8.5% after aTc induction, and no significant change was detected by only the sgRNA plasmid. However, when both plasmids were coexpressed in the cell, a 10%-less growth rate was first detected even before induction, and more serious growth retardation by 41% was detected once the inducer, aTc, was added (**Fig. 6.4b**). By simple calculation, the result points out that the CRISPRi system itself without induction leads to around 10% growth retardation and the down-regulated the *argA* gene further deteriorates the host growth by about 30%.

To verify if the CRISPRi system expressed the designed functions properly, we induced the system at 6-h and traced the time-dependent metabolite changes. The metabolites of the arginine pathway (L-citrulline and L-arginine) dropped after the activation of the CRISPRi system, except N-acetyl-L-glutamate (**Fig. 6.4c**). There are three possible reasons to explain why after induction the concentration of N-acetyl-L-glutamate rose in the first 30 min and then started to decreasing like other metabolites: i) the replenishment of N-acetyl-L-glutamate from ornithine pathways, ii) time-delayed expression of CRISPRi system and iii) overabundance of ArgA enzyme or combined effects from these three. Overall, although the original CRISPRi system is a leaky system, for the downregulation of genes, it still functions well as a feasible method to study how cells respond to genetic perturbation.

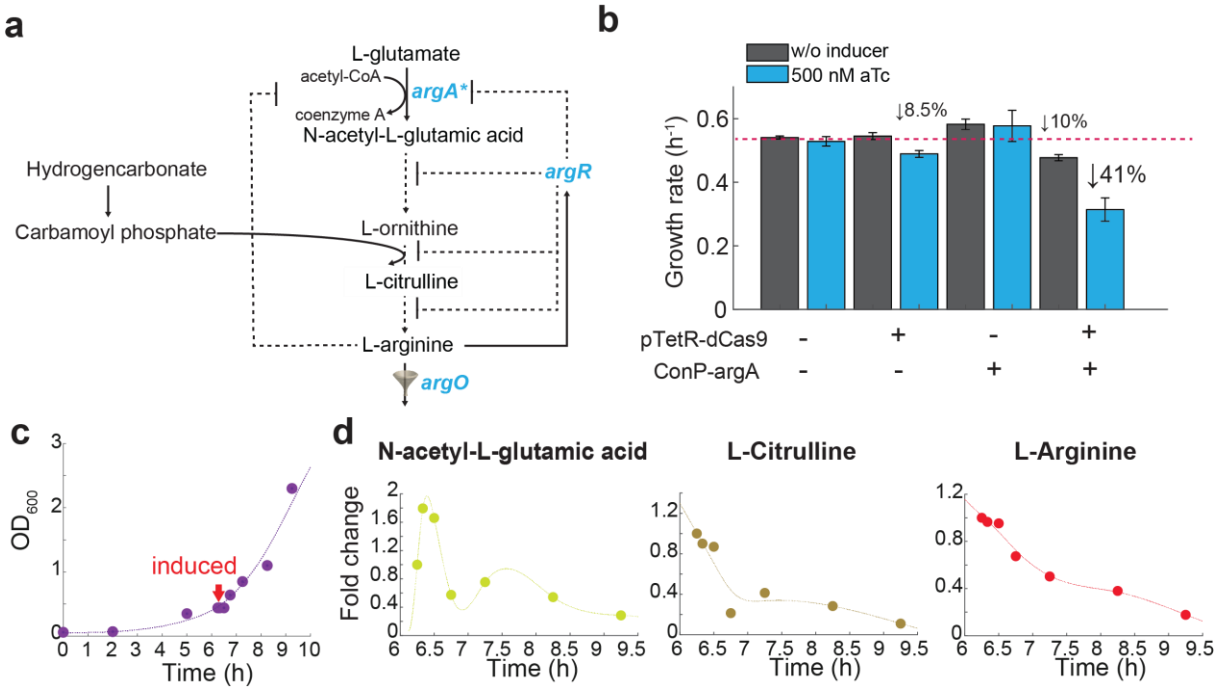


Fig. 6.4 | **Characterization of CRISPRi system in the arginine pathway. a,** The arginine pathway. **b,** Growth rates of bacterial strains with different CRISPRi systems were taken from 96-well plate experiment and all the strains were induced at the beginning. The dCas9 (pTetR-dCas9) and sgRNA plasmid (ConP-argA) were expressed separately or co-expressed in wild type MG1655. **c,** Growth curve of the *argA* knockdown was measured in a shake flask. **d,** Metabolites were taken directly after induction and normalized to 0% aTc (before induction).

6.2.3 Application of CRISPRi system for arginine overexpression

Although we hadn't found an optimized CRISPRi system perfectly sealed (**Fig. 6.2** and **Fig. 6.3a**), we used the original system without adding an inducer as an alternative way to reduce the burden from dCas9 overexpression. Here, instead of manipulating metabolic genes such as *argA*, we targeted a transcription factor ArgR, a DNA-binding transcriptional dual regulator that negatively regulates transcription of arginine biosynthetic genes as shown in **Fig. 6.4a**[22, 150, 151].

Another study has indicated that expression of feedback-resistant ArgA* protein increase arginine production[11]; moreover, the double mutant strain (allosteric mutant ArgA (H15Y) and deletion of the *argR* gene named as KO strain) can further improve the arginine production under the sacrifice of an obvious growth retardation[12]. Therefore, by using the CRISPRi system to knockdown the *argR* gene, we aimed to reduce the burden from the deletion of the *argR* gene

Fig. 6.5a). We used the feedback-resistant ArgA* as a basic strain to express the CRISPRi system, and two knockdown strains with different interfering strengths (#7 and #10) were created for the downregulation of the *argR* gene.

Compared the wild-type *E. coli* and the double-mutant strain (KO), the results indicated that the knockdown strains present a growth rate higher than the double mutant strain (**Fig. 6.5b**)[13] and show similar specific arginine production rate (2.2 and 2.3 mmol gDW^{-1} h^{-1}, not shown in this dissertation). To understand why the knockdown strains can reach the same productivity without severe growth retardation, the proteomics was analyzed. Among 156 proteins in amino acid metabolism, the double mutant strain had more protein changes in comparison with knockdown strains (**Fig. 6.5c**). Such variation was remarkablely presented in nine proteins participating in the arginine pathway (orange color), resulting in higher protein levels in the trend of double mutant strain > knockdown strain (#7) > knockdown strain (#10) > wild-type. The observed phenomenon indicates that the deletion of ArgR causes more global changes in amino acid pathways than downregulation of ArgR; besides, targeting the *argR* at the position near the start codon (#7) brings a stronger effect than near the stop codon (#10) as other studies have shown[68, 152]. As a result, the protein overabundance could be the main reason why the host with the deletion of ArgR suffered from more serious growth retardation than the knockdown ones.

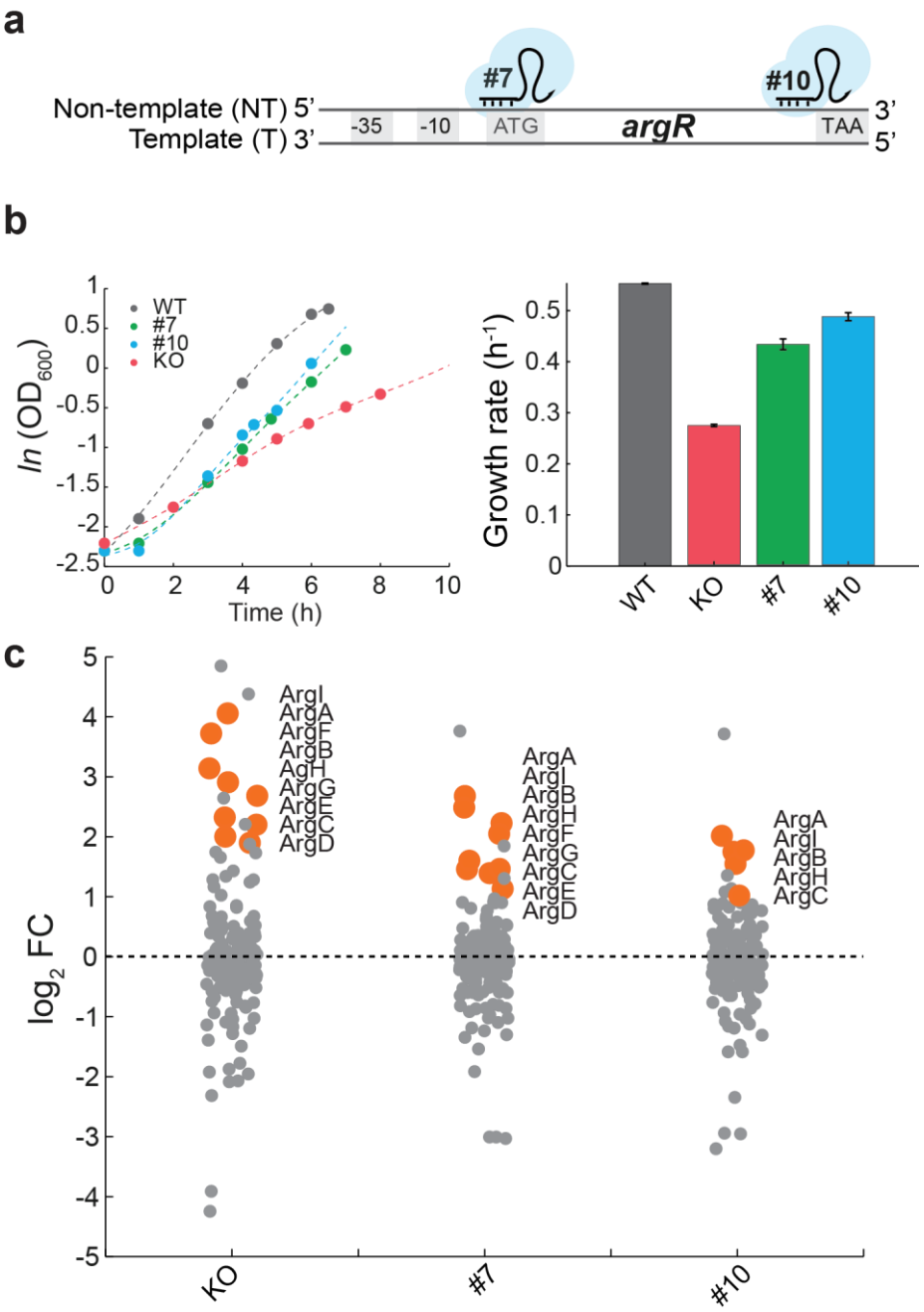


Fig. 6.5 | Protein expression in arginine overexpression strains. a, The designed sgRNA that targets the sequence near the initial site (#7) or at the end of the coding region (#10) of the argR gene. **b,** Growth curve of arginine production. Four strains, the downregulation of ArgR (#7 and #10) by the CRISRPi system, double mutant (KO) and the wild-type strains, were cultivated in the M9 medium without adding aTc inducer. Except wild-type strain, all the other strains harbor ArgA (H15Y) mutation. **c,** Protein expression of 156 proteins, involving in currently known amino acid metabolism, was normalized to wild-type and transferred to $\log_2$ scale. The orange color dots are the protein in the arginine biosynthetic pathway.

6.3 Conclusion

In this thesis, we have demonstrated that the controllability and leaky effect of an original CRISPRi system can be refined by replacing the intrinsic sgRNA plasmid with arabinose inducible promoter, forming a dual-regulated system. The modified version functions well in GFP expression and in the arginine pathway. Furthermore, we also proved that through targeting a transcription factor-ArgR by the CRISPRi system, the concept of dynamic control can be again observed in *E. coli* arginine biosynthesis, which not only maintains the production rate but also improves the host growth rate. In summary, our results advocate that the CRISPRi system is a promising technique that can expand to lots of potential applications in bioengineering and fundamental research.

6.4 Future work

6.4.1 Improvement of CRISPRi system

To create a tight and inducible CRISPRi system, the dCas9 integrated into genomic DNA would be a better way to prevent the burden from overexpression of dCas9[153, 154].

6.4.2 Other CRISPRi system

There are many types of CRISPRi system have found and been applied in lots of experimental and clinical purposes[155, 156], such as Cpf1[157], Cas12a/Cas12a-like[158], MAD7(Inscripta[159]), and so on. These nucleases have similar functions and mechanisms but the size is smaller than the yeast's Cas9 so that the additional burden or off-target effect of original Cas9 can be reduced[160]. Besides, instead of conserve PAM sequence CCN in Cas9 orthologs, the PAM sequence among these systems using TTTN or TTN has a higher frequency in the genome, giving the applicability become much wide[157, 158, 161, 162].

7. Chapter 7 – Conclusion and outlook

7.1 Summary of all projects

To find optimized enzyme expression levels is a typical challenge in designed circuits for cost-effective bioprocessing and industrial-scale fermentation. The knowledge of multiple gene interactions and competitions between hosts with their common resources in the designed circuits is still insufficient; therefore, how synthetic biology can make an improvement in facilitating host cells to adapt to genetic perturbations and to balance the perturbations are covered in this thesis.

In **Chapter 1**, the utilization of two-phase (growth decoupled) and one-phase processes (growth coupled) systems was reviewed as two common strategies for production of desired compounds. Although nowadays biotechnology such as CRISPRi and CRISPRi-associated genomic editing has made constructing synthetic circuits more convenient, stabilizing synthetic circuits with host-resource is still a challenge. To overcome the burden arising from the external circuits, native transcription factor-Cra could be a potential regulator to adjust central glycolytic networks and to coordinate the synthetic circuits.

For achieving high-titer production, there is always a debate between two-phase and one-phase processes. A two-phase process was first introduced in the glycerol-producing system which competes the same precursor in the glycolysis (**Chapter 2**), but the results indicated that producing glycerol is more likely to be a growth-coupled product. Therefore, in order to implement the one-phase processes, we inserted the Cra-consensus binding sequence directly after pBAD promoter so that this circuit can be recognized and regulated dynamically by Cra (**Chapter 3**) depending on the metabolic concentration of FBP/F1P in the glycolysis. Compared with the strain with no consensus sequence, the Cra-integrated strain is more fitness and reach a balance between both host growth and product synthesis at wider functional ranges. Through time-course proteomic and metabolomic characterization, we observed that no Cra-misregulates glycolysis which further worsens the host growth. In other words, when the draining flux coming from GPD1 overexpression is higher over a certain threshold, the host cells switch the glycolysis to gluconeogenesis even enough glucose presents in the medium. On the contrary, the Cra-regulation strain shows a time-delayed

enzyme expression that slows down the *gpd1* expression level and upregulates glucose transporters and glycolytic enzymes for adapting to the higher demanding of glycerol precursors. Apart from glycerol production, carotenoid synthetic pathway was applied under the same regulation, and the result confirms that Cra-driven regulation is generally applicable for synthesizing glycolysis-derived chemicals. Moreover, this Cra-driven regulation is not promoter specific and it can be successfully incorporated into any types of promoters, such as constitutive or pTetR inducible promoters, to achieve stabilized growth and production (**Chapter 4**). By simply adding Cra binding sequence in designed circuits, compared to Δ*cra* strains, the Cra-regulated strains are more robustness against environmental changes caused by carbon sources, oxidative damage and osmotic stresses. Thus, our studies prove that one-phase growth coupled strategy is a more feasible and effective way to reorganize and redirect the glycolytic flux into desired valuable compounds.

When synthetic circuits are expressed in the host cells, a common phenomenon is growth retardation. Apart from the lower growth, by reporting the host physiological state timely, a burden monitor GFP signal (**Chapter 5**) provides an additional perspective to understand the resource allocation in the host. In this thesis, the burden monitor was not only used as a reporter but also applied for screening proper CRISPRi systems with improved tightness (**Chapter 6**). Furthermore, by down-regulation of transcription factor-ArgR, both enhanced growth and product yield are observed in the host compared to the Δ*argR* strain. From **Chapters 3** and **6,** it suggests that the transcription factor is indeed an effective and promising controller to synchronize host's growth and production either by cooperation with it (e.g. Cra-regulation) or by downregulation of it (e.g. ArgR), and eventually to enhance product titer.

In a nutshell, it is clear that there will be no higher productivity when enzyme expression is not optimized that means too high or too low enzyme expression cannot guarantee a higher productivity. As a result, to find an optimized enzyme level in specific host cells and individual pathway is still very challenging and cumbersome in bioengineering. Therefore, instead of the case-specialized screening, establishing an auto-regulated loop in the host is a facile and compatible strategy to increase the host fitness. For example, as previously demonstrated in this thesis, the integration of Cra-regulation which controls the on/off valve of the synthetic pathway is a cost-effective, time-saving and universal approach to optimize the production of glycolysis-derived

chemicals. Via simply insertion of Cra-binding sequence, the native regulated system in the host can auto-adjust the enzyme expression levels for that it requires no further adjustment or fermentation strategies. All the studies mentioned in this thesis were aimed to provide strategies for rational design in synthetic bioengineering and to further enhance the production of valuable chemicals.

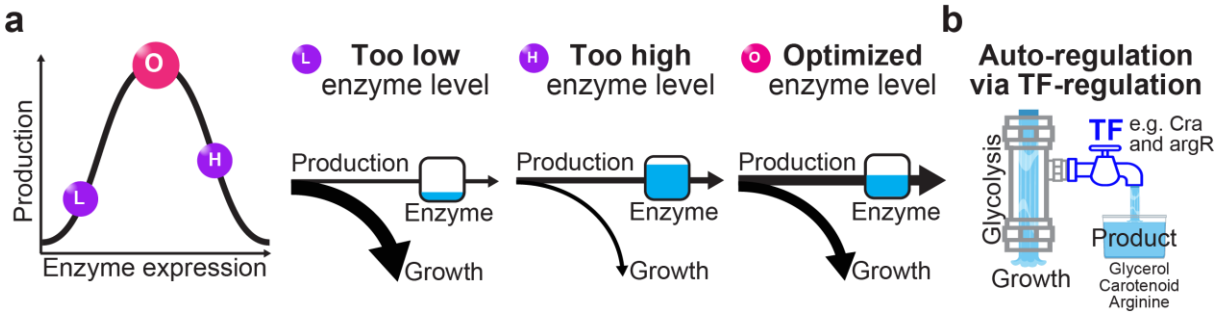


Fig. 7.1 | Optimized enzyme levels in bioengineering. a, Enzyme expression levels determined the product production. **b**, In this thesis, we demonstrated that native transcription factors such as Cra and ArgR, can act as a valve-like device to fine-tune gene expression depending on the concentration of glycolytic metabolites and further accomplish stabilized growth and increased productivity of glycerol, carotenoid and arginine.

7.2 Outlook

7.2.1 What is a good strategy for bioprocessing- two-phase or one-phase processes for synthetic circuit?

One study suggests that if the substrate uptake rate is high, a two-phase process is appropriate and vice versa[16]. The other study points out that growth-coupled (one-phase) is feasible for almost metabolites in five different organisms[17]. In this thesis, dynamic control of glycerol production surpasses a two-phase (growth-switch) process. Does this represent a general rule for the chemicals draining from glycolysis? Because many studies using two-phase production for amino-acid derived chemicals[5, 8], this implies sub-branch pathways (not central metabolism) could get a higher yield by two-phase and conversely central metabolism derived chemicals by dynamic control. Can we combine both strategies? For example, at first the one-phase is applied to both produce chemicals and maintain growth dynamically to a certain amount of biomass until the substrate depleted in the system, and then automatically to switch from growth to production phase with a limited growth rate that recycling all the left extracellular or Intracellular resource to further enhance product titer.

7.2.2 Improving analysis of the time-dependent data

There is more information can be discussed in the time-dependent experiments. How do proteins and metabolites interact? When does the regulation be triggered in the host and how regulatory networks response to the regulation? Do the cells reprogram the metabolic networks under the external or internal burden? What is the unknown proteins pumping out in the pBAD-only strain (GlcG, YdfZ, YjbJ, etc.) and conversely in pBAD-Cra strain (YaaA, YbiC, YgA, etc.)? Are they replicating the other proteins' function (misregulated by Cra as an instance)? There are so many questions waiting for model-based investigation.

7.2.3 Why fructose-1,6-bisphosphate is the signal for glycolytic flux across organisms?

In eukaryotic cells, it has been identified four key committed steps for controlling glycolytic flux: substrate import, phosphorylation, fructose-1,6-bisphosphate (FBP) production, and lactate export[163]. In particular, FBP participates in yeast proliferation through interacting with a regulator-Ras[59] and also serves as a flux sensor[164]. On the other hand, in prokaryotes, FBP/F1P has long been known as a glycolytic flux sensor and influences Cra activity[55, 165]. This attractive FBP regulates glycolysis not only in eukaryotes[164] but also in prokaryotes[56]. However, why cells use FBP as a signaling molecule instead of other metabolites? Does fructose play another role more than just a substrate in life because FBP can be converted from fructose? Studies have shown that fructose induces hepatic lipogenesis in the liver[166, 167]. Almost all the known information about FBP regulation is mainly from artificially cultured conditions not from the bacteria clustering real life. There is one study pointing out that fructose is not just a nutrient but can be as a signal molecule stimulating bacterium to express phosphatase genes[168]. Therefore, FBP/fructose must have evolved meaning for bacteria living in the ecosystem.

7.2.4 Does draining glycolytic metabolites trigger overflow metabolism or rewriting metabolic networks?

Overflow metabolism refers to the cells which obtain energy through a seemingly wasteful strategy, namely fermentation, instead of the more efficient respiration and

usually occurs in fast-growing cells[103, 169]. Overflow metabolism is also known as the Warburg effect in cancer cells[89, 163]. The characteristic of overflow metabolism is that cells excrete metabolites, such as lactate, acetate and ethanol. The interesting thing was both glycerol producer (pBAD-only and pBAD-Cra) can produce higher acetate than the GFP producer, suggesting that competing with glycolysis seems to trigger acetate overflow. The reason for the pBAD-Cra strain is understandable because the cells try to maintain relative faster-growing resulting in overflow metabolism to overcome the overloaded glycolysis. However, the contradictory phenomenon was found in the pBAD-only strain which shows higher acetate production but slow growth.

Apart from overflow metabolism, our result also implies that when glycolysis is overloaded, the cells could try to overcome this stress via bypassing the EMP pathway and going through the ED pathway instead. There are many studies pointing out that the cells can reroute or bypass the metabolic pathways when facing the internal or external burden[49, 127, 170]. However, further study is required to understand the detailed regulatory mechanism in both pBAD-only and pBAD-Cra strains.

Appendices

I. Materials and Methods

1. Strains and plasmids

All the strains are derivatives of *Escherichia coli* K-12 MG1655 (wild type, DSMZ No. 18039) and plasmids constructed in this thesis are listed in **Supplementary Table S1**. The consensus sequences of different transcription factors are given in **Supplementary Table S2**. The sequences for glycerol production constructs are provided in **Supplementary Table S3**. For cloning used, the primers synthesized by Eurofins Genomics are listed in **Supplementary Table S4**. The primers were used to get linear target genes separately by polymerase chain reaction (PCR) of Q5® High-Fidelity DNA Polymerase (M0491L, BioLabs). Restriction digestion, circular polymerase extension cloning (CPEC) methods[171] and Gibson assembly (E2611S, Biolabs) were used for cloning. No-SCAR system[77] for genome editing was bought from addgene (https://www.addgene.org/). Constitutive promoters J-series was selected from iGEM (http://parts.igem.org/Main_Page).

2. Bacterial cultivations

Frozen bacterial stocks were plated out on Luria-Bertani (LB) with specific antibiotics overnight. Next day, single colonies were picked from the LB agar plate and transferred to 5 mL LB broth medium at 30 °C or 37 °C depending on the origin harboring in the plasmid. Fresh recovered cells were cultivated in M9 minimal medium with glucose or glycerol as a sole carbon substrate (5 g L^{-1}) depend on experiment designed. M9 medium was composed by (per liter): 42.2 mM Na_2HPO_4, 22 mM KH_2PO_4, 11.3 mM $(NH_4)_2SO_4$, 8.56 mM NaCl, 1 mM $MgSO_4 \cdot 7H_2O$, 100 μM $CaCl_2 \cdot 2H_2O$, 60 μM $FeCl_3$, 7.6 μM $CoCl_2 \cdot 6H_2O$, 7.1 μM $MnSO_4 \cdot 2H_2O$, 7 μM $CuCl_2 \cdot 2H_2O$, and 6.3 μM $ZnSO_4 \cdot 7H_2O$. Antibiotics of kanamycin (50 μg mL^{-1}), ampicillin (100 μg mL^{-1}) and spectinomycin (100 μg mL^{-1}) were added when necessary depending on the construct of the plasmid.

3. Growth characterization and fluorescence assay of engineered strains in a 96-well plate

Engineered *E. coli* cells were first grown in 5 mL LB for 6 h and diluted to a ratio of 1:500 in fresh M9 media supplemented with 0.5% (w/v) glucose at 37 °C for overnight cultivation under shaking at 220 rpm. The following day, M9 pre-cultures were diluted to a ratio of 1:50 with 150 μL M9 media and incubated in 96-well plates. Optical density (OD) at 600 nm was measured every 5 min using a plate reader (Epoch, Biotek) and

GFP fluorescence (excitation 490 nm, emission 530 nm) was measured with 5-min intervals continuously shaking in a plate reader (Synergy, Biotek). Growth rates (μ) were calculated as dln(OD)/dt by linear regression over the indicated time windows. The promoter activities were calculated as dGFP/dt/OD during the exponential phase.

4. Determination of glycerol, glucose and carotenoid concentration

Single colonies that produced glycerol were transferred from agar plates and cultured in 5 mL LB for approximately 6 h. The cells were then diluted to a ratio of 1:500 in fresh M9 medium with 0.5% (w/v) glucose and cultured at 37 °C overnight.

For 96-well plate experiments, 3 μL overnight cultured cells were harvested and diluted cells to a value of 0.05 at OD_{600} in a 147 μL of fresh M9 medium added with 0.5% glucose. After cultivating for 24 h, the cells were harvested by centrifuging at 4000 rpm in a centrifuge (5810R, Eppendorf). The concentration of glycerol was measured using the glycerol assay kit (MAK117-1KT, Sigma) according to the manufacturer's description manual. Briefly, 10 μL supernatant was mixed with 100 μL reaction buffer for 20 min incubation and analyzed the absorbance at 570 nm.

For shake flask experiments, overnight cultured *E. coli* were inoculated in M9 medium supplied with 0.5% (w/v) glucose to obtain an initial value of 0.05 at OD_{600} and cultured at 37 °C in a incubator with a shaking frequency of 220 rpm. 500 μL of this cultivation was taken at defined intervals and the OD values recorded for calculating the growth rate. The cells were spun down and the supernatant was used to measure the amount of glycerol. The metabolic samples were extracted and measured when the growing cells reached 0.5 of OD_{600}. Metabolic sampling method is further described later when discussing the metabolic analysis.

Samples for glucose measurement were collected in a similar process used for glycerol measurement. Glucose concentrations were measured with the D-glucose assay kit (K-GLUC, GOPOD Format, Megazyme). Modified from the manufacturer's description, 10 μL supernatant was taken mixed with 190 μL reaction buffer for 20 min incubation at 42 °C and then measured at an absorbance of 510 nm. Both glycerol or glucose measurements needed to be executed in a dark environment.

To quantify the concentration of carotenoids, carotenoid producers were cultivated in 96-well plates with M9 minimal medium containing 0.5% glucose and 20% LB at 37 °C with continuous shaking. After cultivating 24 h, the cells were collected and the supernatant was discarded by centrifugation at 4000 rpm. Then, repeat the centrifugation again to remove the remaining supernatant as clear as possible. The remaining cell pellets were resuspended in 120 μL DMSO[172] and sonicated for 30 s to disrupt the cell walls for the extraction of carotenoids. Samples were centrifuged again and the red-orange supernatants (carotenoids) were collected into new PCR tubes. 50 μL of collected supernatants were placed in a 384-well pate to quantified

the carotenoids by measuring the absorbance at 470 nm. For quantification, the standard solutions of β-carotene (C4582-25MG, Sigma) were prepared in a range of 5, 10, 25 and 50 mg mL^{-1}. The stock solution is 5g mL^{-1} in THF (Tetrahydrofuran) solution.

5. Metabolic analysis

Freshly transformed strains were plated on LB plates and two individual colonies were picked, inoculated into 5 mL LB and cultured for 6 h. This pre-culture was then diluted with 5 mL M9 medium in a ratio of 1:500. This was then left to culture overnight at 37 °C. The following day, these overnight cultures were then inoculated into a 50 mL M9 medium to obtain OD_{600} value of 0.05. When the cultivated cells reached OD_{600} of 0.5, 2 mL cultured aliquots was harvested by using 0.45 µm filter paper (HVLP02500, Merck Millipore). This filter paper was then transferred into a vial containing 1 mL quenched solution consisting of acetonitrile/methanol/water in the ratio of 40:40:20 and incubated at -20 °C for 15 min. After cooling, the tube was centrifuged for 20 min at -9 °C and 13,000 rpm to remove the cell debris. Supernatant was added with 10 µL ^{13}C-labeled internal control in each well of a 96-well plate. The 96-well plate was then placed in a liquid chromatography (HILIC) LC-MS (Agilent 1290 Infinity II UHPLC system, Agilent Technologies) machine for metabolite measurement. A volume of 3 µL was injected into the columns made of Acquity BEH Amide (30 x 2.1 mm, 1.7 µm) and iHILIC-Fusion (50 x 2.1 mm, 5 µm) were used at 30 °C in a column oven. Running buffer A was water with 10 mM ammonium formate and 0.1% formic acid (v/v) and running buffer B was acetonitrile with 0.1% formic acid (v/v). A gradient elution method was used with 90% B at $T_0 = 0$; 40 % B at $T_1 = 1.3$ mins, 40 % B at $T_2 = 1.5$ min, 90 % B at $T_3 = 1.7$ min and 90 % B $T_4 = 2$ min. The $^{13}C/^{12}C$ ratio was determined by peak heights.

6. Proteomics

Cultivations were performed as described above. Culture aliquots were transferred into 2 mL reaction tubes and washed two times with PBS buffer (0.14 mM NaCl, 2.7 mM KCl, 1.5 mM KH_2PO_4 and 8 mM Na_2HPO_4). After washing, cell pellets were resuspended in 200 µL of lysis buffer containing 100 mM ammonium bicarbonate and 0.5 % sodium laroyl sarcosinate. Cells were again incubated for 15 min with 5 mM Tris(2-carboxyethyl) phosphine (TCEP) at 95 °C followed by alkylation with 10 mM iodoacetamide for 30 min at 25 °C. We used SP3 bead method[173, 174] for a large number of samples. Fixed protein amount of 50 µg measured by BCA assay (23225, Thermo Fischer) and mixed with 4 µL SP3 beads stock (10 µg $µL^{-1}$ stock) in 96-well high-volume v-bottom plate (710879, Biozym Scientific GmbH). To initiate peptide binding to the beads, 75 µL of 100% ethanol were added with the mixture of protein

and beads for 5 min at room temperature. Tubes were placed in a magnetic rack for 5 min. The supernatant was discarded and the beads were rinsed two times with 200 µL of 70% ethanol and then 180 µL of 100% ethanol on a magnetic rack. For elution tubes were removed from the magnetic rack, and the beads were reconstituted in 28 µl 10% acetonitrile/10 mM NH_4HCO_3 with 1 µg trypsin (Promega) incubated overnight at 30 °C at 1000 rpm. After incubation, the tubes were sonicated for 30 s and placed on a magnetic rack, and the supernatant was recovered as the purified peptides and transferred to new tubes. Recovered peptides were acidified by adding trifluoroacetic acid (TFA) to a 1.5% final concentration for 10 min. The tubes were centrifuged 10 min at 10,000 rpm and the supernatant was then purified through C18 microspin column (Harvard Apparatus) according to the manufacturer's instruction. The eluted peptides were dried and resuspended in 0.1 % TFA for analysis of peptides. Analysis of peptides was performed by a Q-Exactive Plus mass spectrometer connected to an Ultimate 3000 RSLC nano with a Prowflow upgrade and a nanospray flex ion source (Thermo Scientific) as previously described in the study of Sander et. al[21, 151]. Briefly, peptides were separated by a reverse-phase HPLC column (75 µm x 42 cm) packed with 2.4 µm C18 resin (Dr. Maisch GmbH, Germany) at a flow rate of 300 nl/min by gradient model which is from 98% solvent A (0.15% formic acid) and 2% solvent B (99.85% acetonitrile, 0.15% formic acid) to 25% solvent B over 105 min and to 35% solvent B for additional 35 min. The data acquisition was set to obtain one high resolution MS scan at a resolution of 70,000 full width at half maximum (at m/z 200) followed by MS/MS scans of the 10 most intense ions. Label-free quantification (LFQ) of the data acquired from mass spectrometry was processed with Progenesis QIP (Waters), and MS/MS search was performed in MASCOT (v2.5, Matrix Science). The following search parameters were used: full tryptic search with two missed cleavage sites, 10 ppm MS1 and 0.02 Da fragment ion tolerance. Carbamidomethylation (C) as fixed, oxidation (M) and deamidation (N,Q) as variable modification. Progenesis outputs were further processed with SafeQuant.

7. Interval-induction

Single colonies of pBAD-only and pBAD-Cra strains were transferred from agar plates and inoculated in 5 mL LB and incubated for 6 h. The cells were then diluted to a ratio of 1:2000 with 10 mL fresh M9 medium with 0.5% (w/v) glucose and grown at 37 °C overnight. The following day, the cells were re-inoculated in 50 mL M9 medium with kanamycin (50 µg/mL) and diluted to an OD_{600} of 0.06 and grown at 37 °C with vigorous shaking for 3 h. The cells were then added with arabinose depending on the way of induction. For one-time induction, the cells were induced once with 0.3% and 0.5% arabinose. For ramp-up time induction, the cells were gradually added arabinose with 30 min interval till 4.5 h reached a final concentration of 0.3% or 0.5% arabinose. The supernatants were taken before and after induction. Glucose concentrations were analyzed using D-glucose assay kit (K-GLUC, gopod format, Megazyme) as

mentioned in the paragraph of “Determination of glycerol, glucose and carotenoid concentration”.

8. Bioreactor cultivation

Bioreactor cultivation was performed in a BioFlo 120 system (Eppendorf) with a culture volume of 0.5 L equipped with pH and dissolved oxygen sensors (Mettler Toledo). Exhaust gases were analyzed with a DasGip GasAnalyser. The culture was inoculated in M9 minimal media with 15 g/L glucose to obtain an initial OD_{600} of 0.1. At approximately OD_{600} of 2, production was induced with 0.5 % arabinose. Additionally, 10 g/L glucose was added to the culture. Supernatant and biomass were sampled over time with higher focus after induction. Glucose was measured using LC-MS/MS with 13C glucose standards. Glycerol was measured with a glycerol assay kit.

9. Calculation of specific production rates

Usually, when *E. coli* produce higher interesting products, it reduces the growth rate. Therefore, the specific production rate is defined as

$$q_p = \frac{1}{X} \times \frac{dP}{dt} \text{ (equation 1)}$$

Where p represents the concentration of product, X is the amount of biomass and q_p is the rate of product formation (mmole product /g biomass/h). The value of OD_{600} is multiplied by coefficient 0.37 gDW/L to get the biomass in a unit of gDW/L. It represents the cell-growth negatively correlates with the production. Multiply the numerator and the denominator of equation 1 by dX to get

$$q_p = \frac{1}{X} \times \frac{dP}{dt} \times \frac{dX}{dX} \quad \text{(equation 2)}$$

The equation of equation 1 can be written as

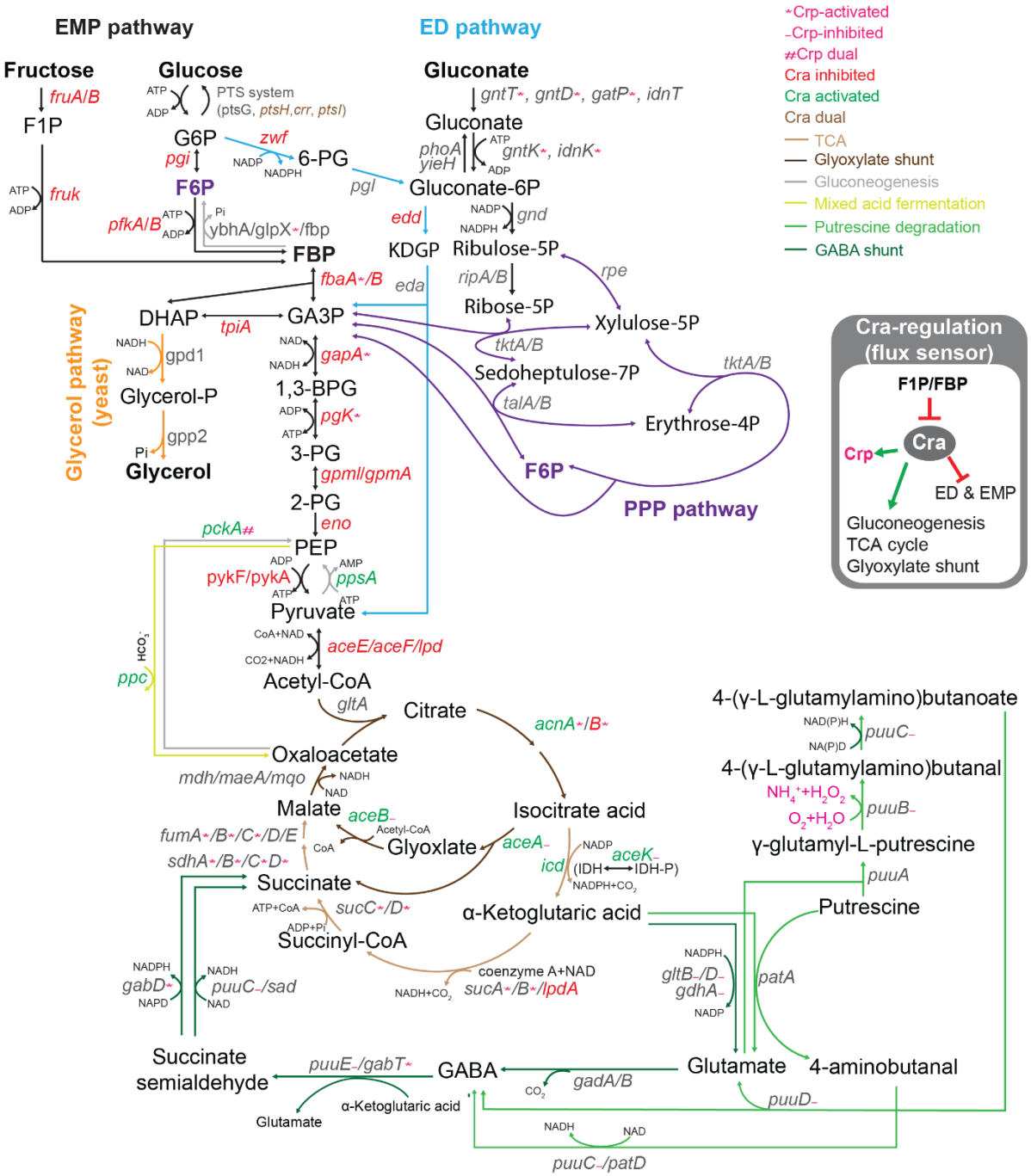
$$q_p = \frac{dP}{dX} \times \mu \text{ (equation 3)}$$

Finally, the specific production rate can also be calculated as equation 3, indicting the slope from product and biomass is multiplied by growth rate (μ).

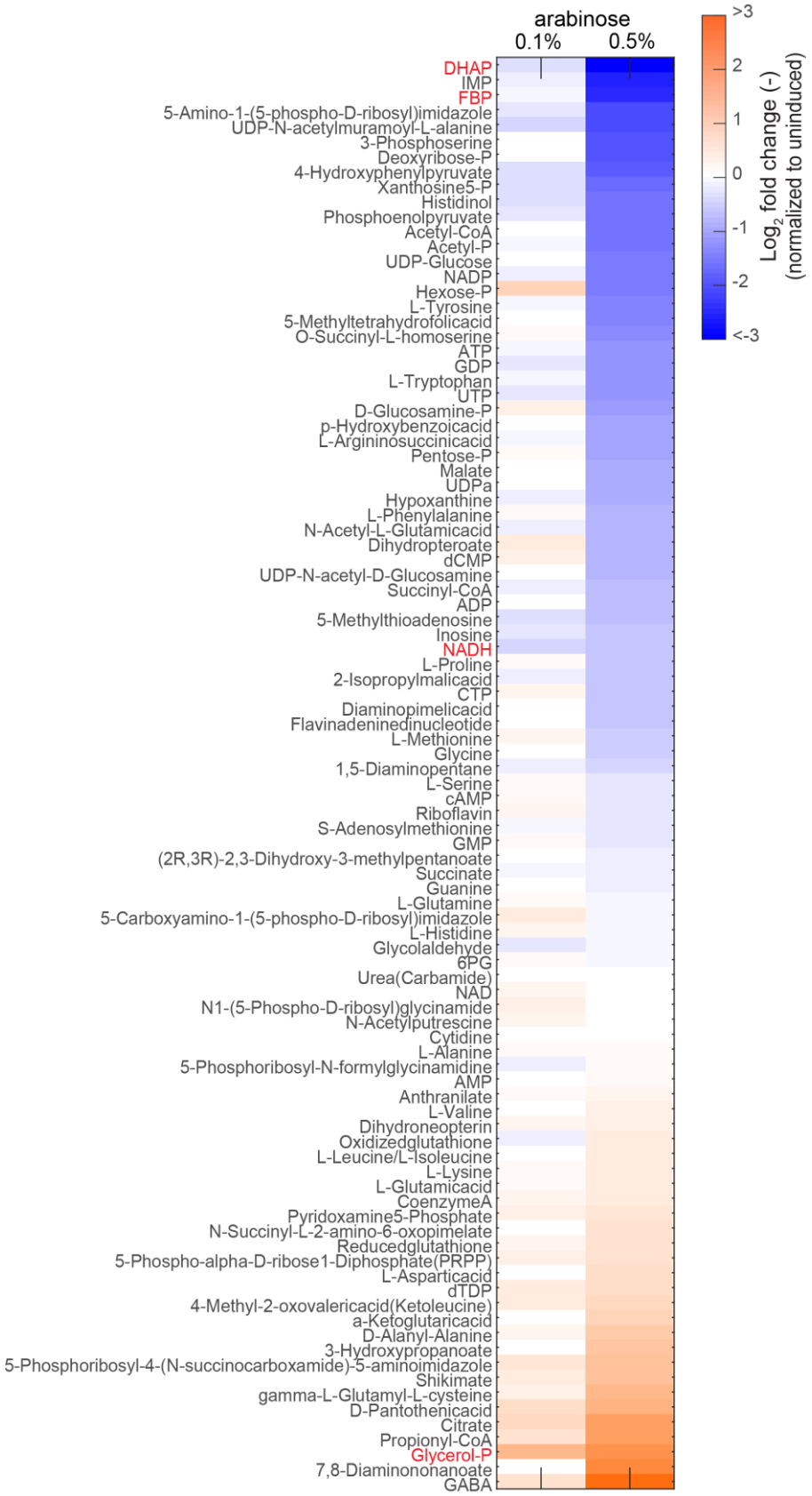
10. Statistical analysis

Statistical analysis was taken under Matlab software (R2018b) and performed with custom Matlab scripts. The number of replicates (n) of each experiment can be found in the respective figure caption. For proteomics and metabolomics n represents the number of independent shake flask cultures. In growth assays, n represents the number of independent microtiter plate cultures. The gene/protein lists were chosen according to gene ontology resources[175] and their functions in specific pathways.

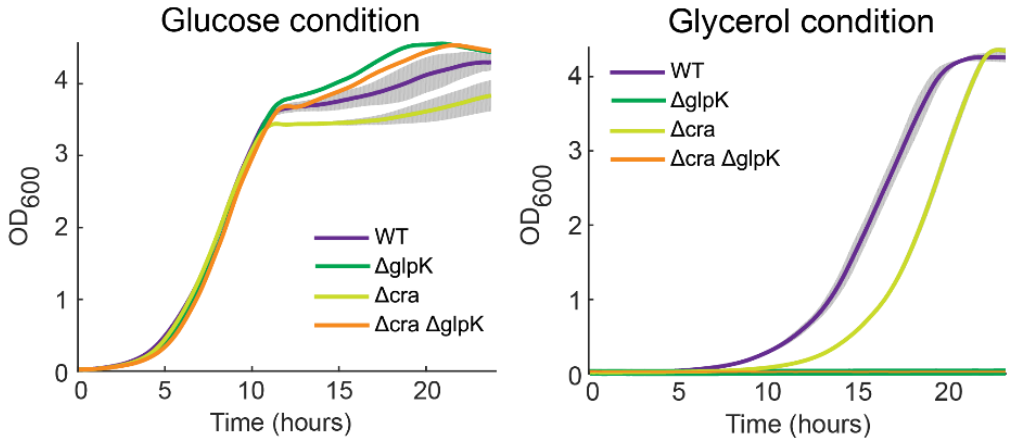
II. Supplementary figures



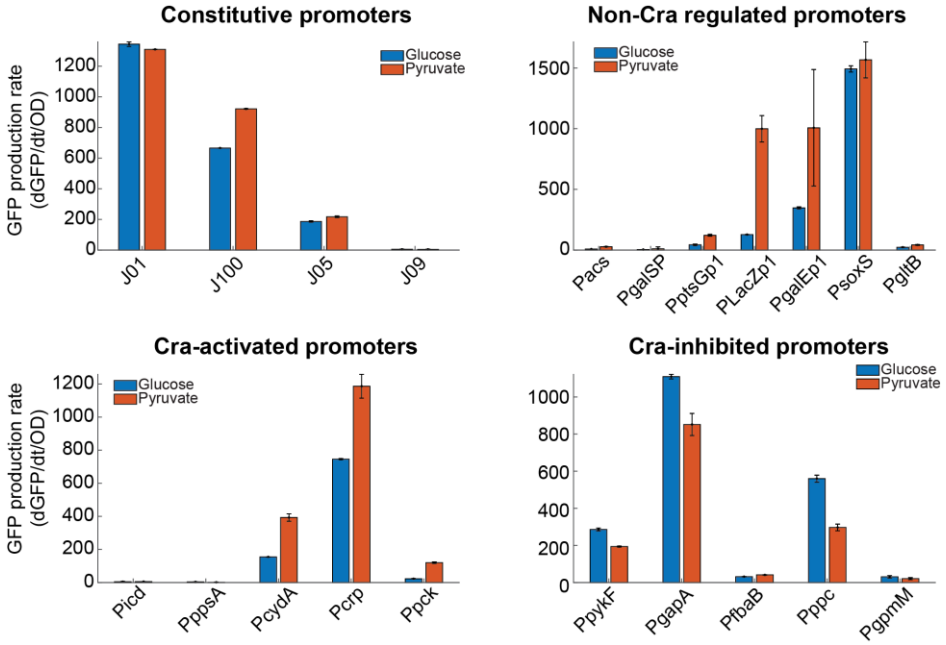
Supplementary figure 1 | The detailed map of glycerol synthetic pathway from yeast and the central metabolic pathway pathways in *E. coli*. A complete map involved *E. coli* exogenous glycerol pathway, glycolysis, gluconeogenesis, glyoxylate, PPP, TCA cycle and GABA shunt.



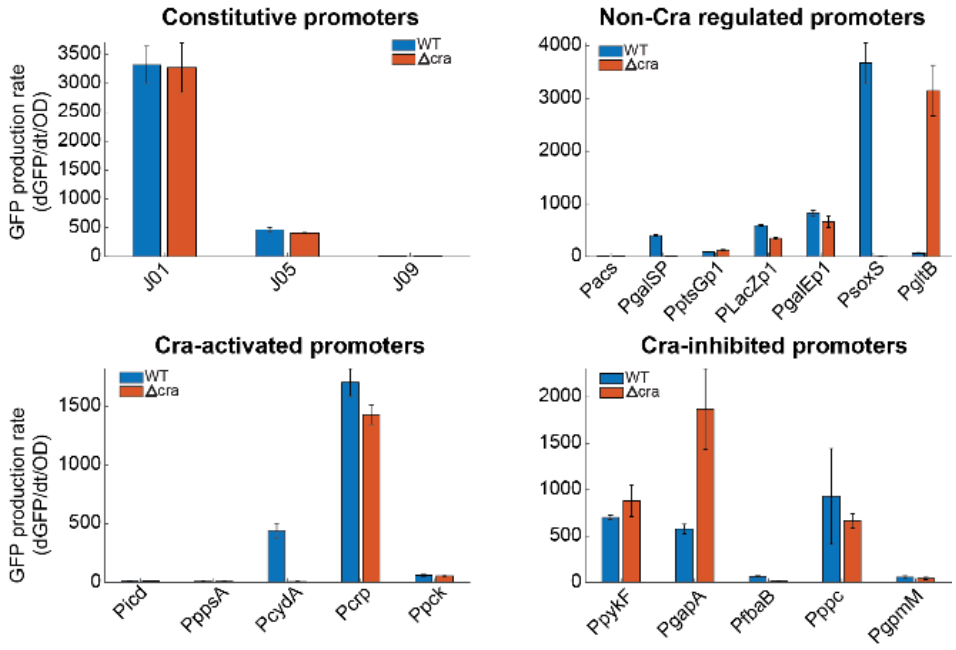
Supplementary figure 2 | The ninety six measured intracellular metabolites. The pBAD-only strain was grown in shake flasks under 0, 0.1% and 0.5% arabinose induction. The data were normalized to the uninduced condition (before induction).



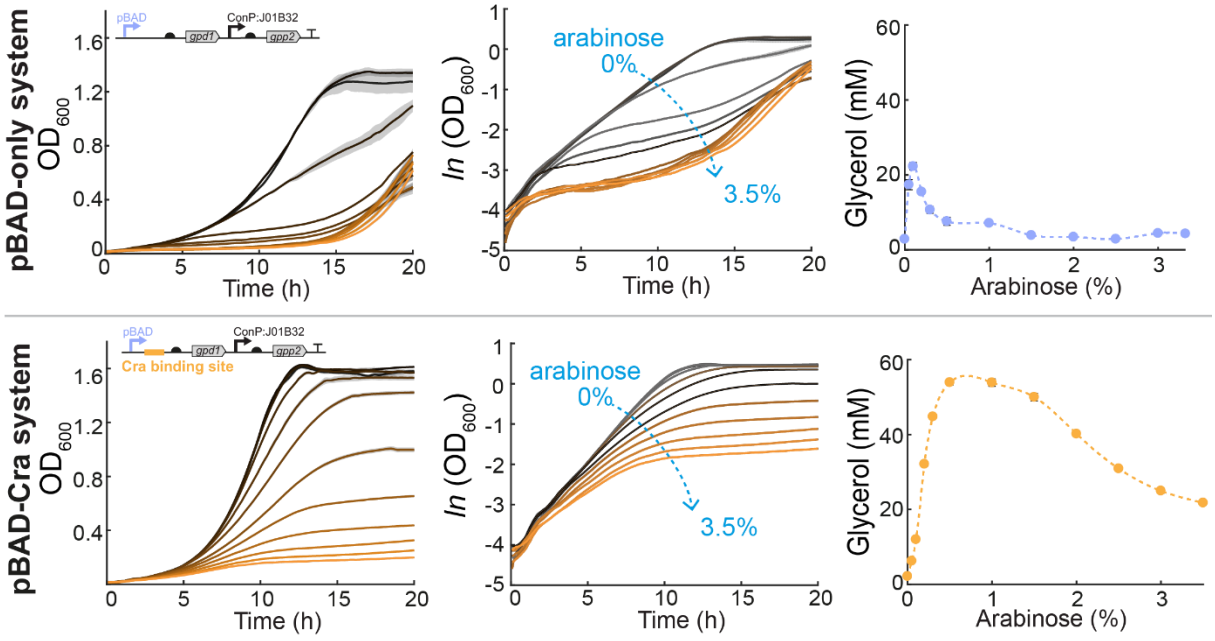
Supplementary figure 3 | Growth curves of Δ*cra* and *Δglpk* strains in glucose and glycerol cultivation. Strains cultivated in a 96-well plate with 0.5% glucose M9 medium.



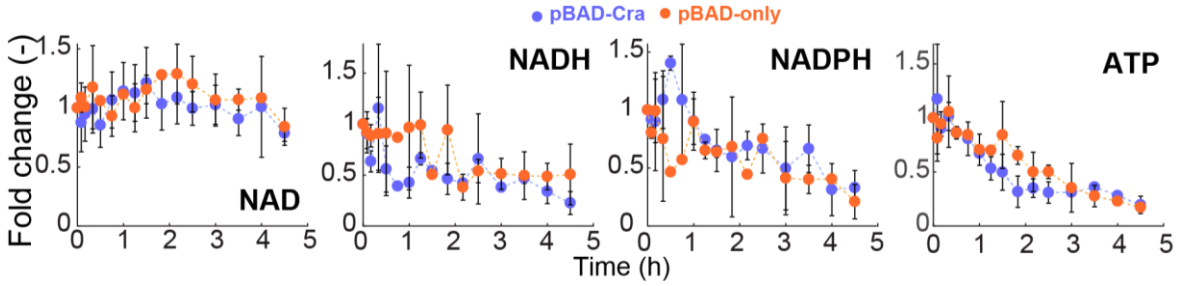
Supplementary figure 4 | The GFP expression under four types of promoters in different conditions. Twenty-one strains bearing different promoters driving GFP expression were grown in glucose or pyruvate conditions. P_{epd} and P_{pfkA} were removed because they did not grow in pyruvate conditions.



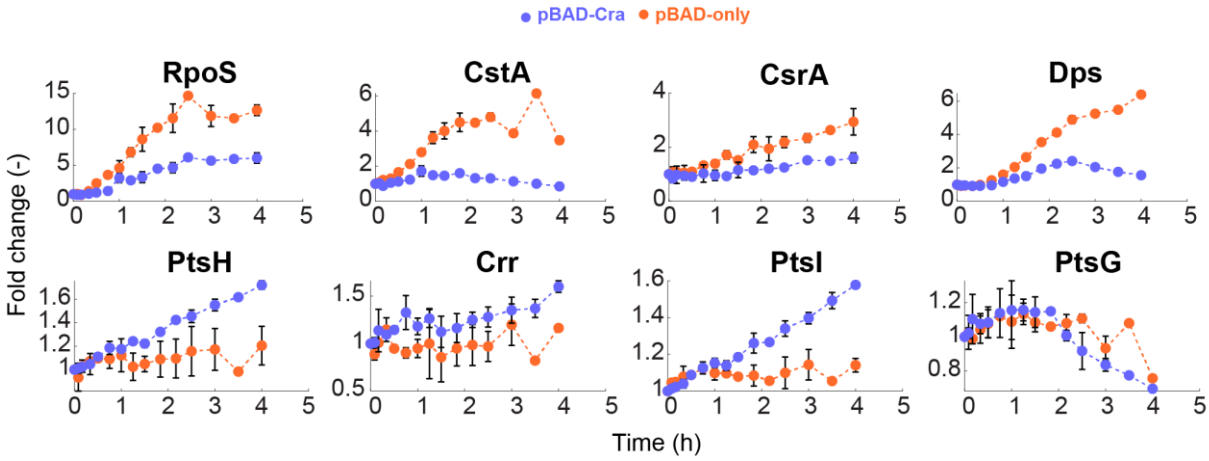
Supplementary figure 5 | The GFP expression in WT and Δ*cra* background. Twenty strains of GFP production were cultivated in a 96-well plate with M9 medium containing 0.5% glucose and measured the GFP levels within 5 min interval. P_{epd}, P_{pfkA} and J100 were removed because they did not grow well in Δ*cra* background.



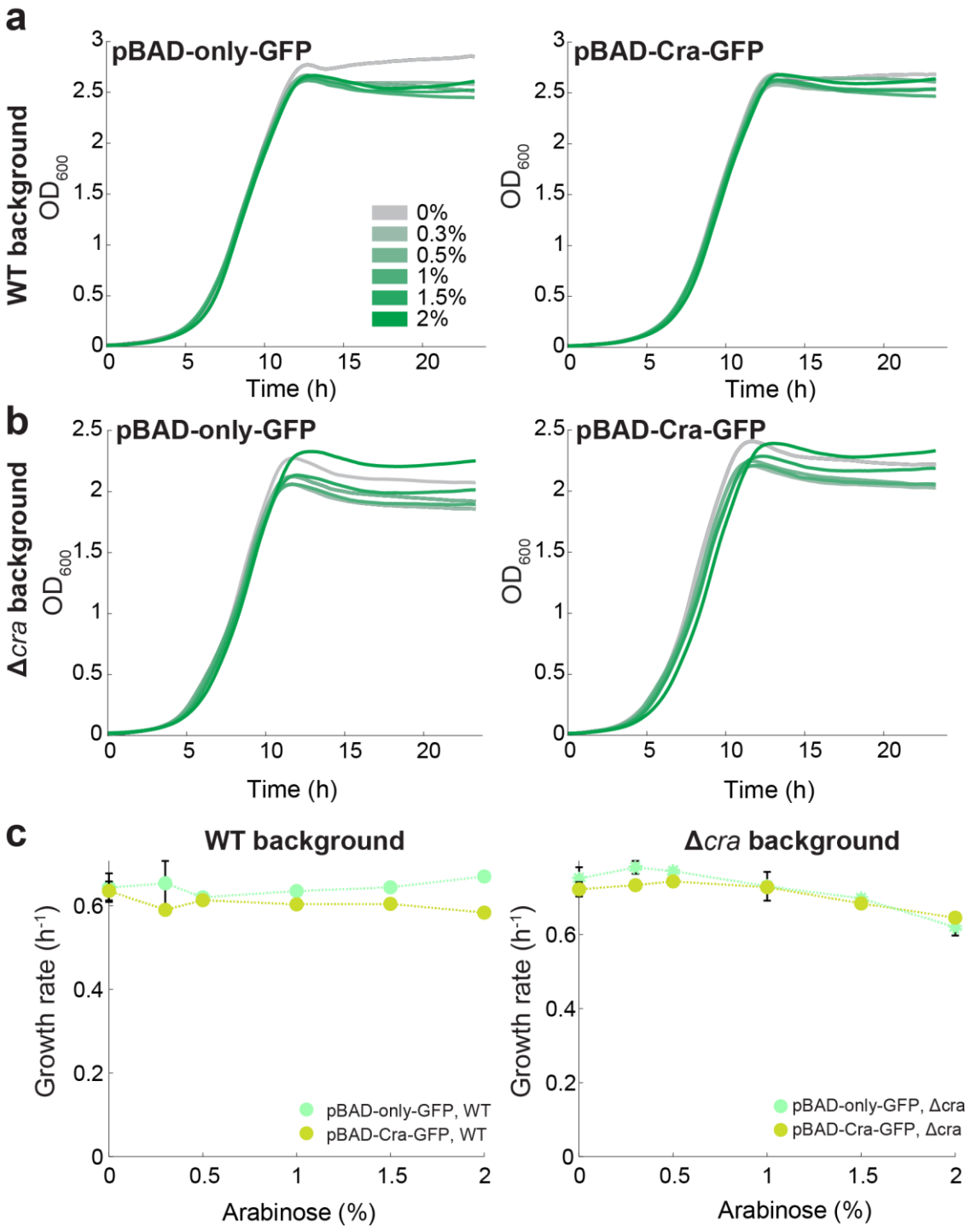
Supplementary figure 6 | Glycerol titers and growth of pBAD-only and pBAD-Cra in the *glpK* background. The same experimental approach as shown in **Fig. 3.11** was conducted but the titration concentration was up to 3.5% instead of 2% arabinose.



Supplementary figure 7 | Energy changes of pBAD-only and pBAD-Cra strain over time. The protein expression of both pBAD-only and pBAD-Cra glycerol producers.

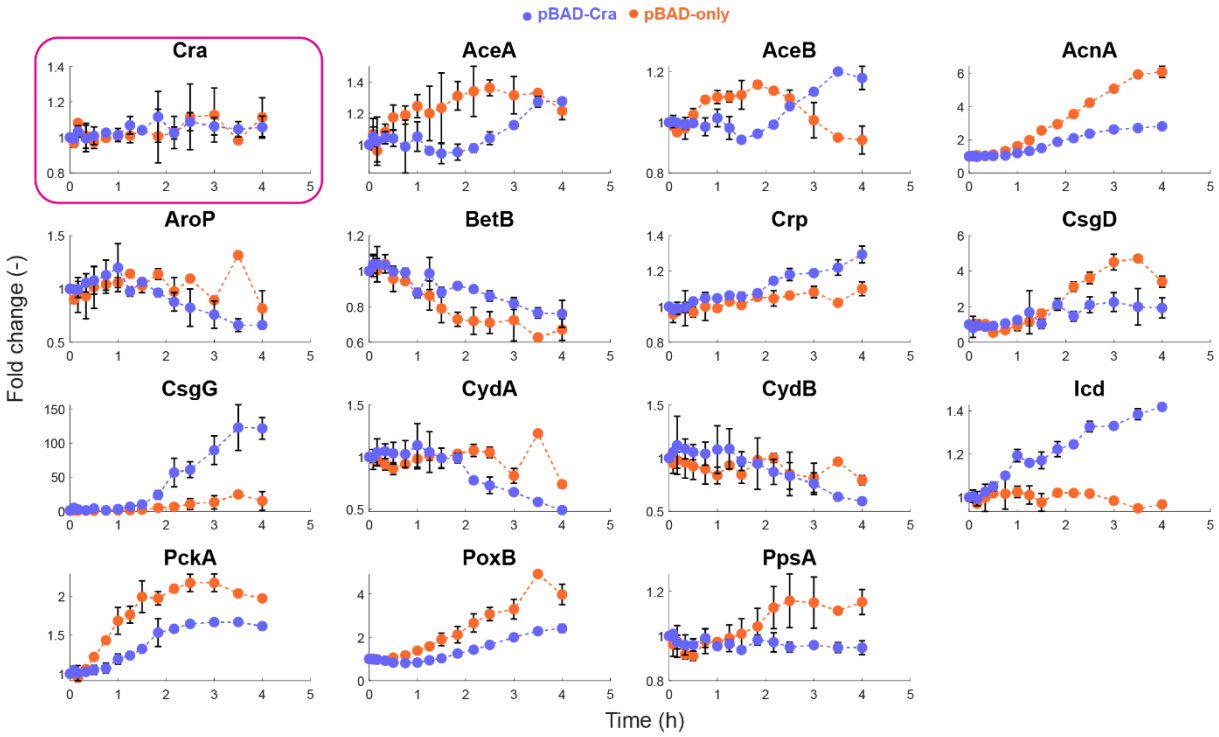


Supplementary figure 8 | Protein expression. The upper four proteins (RpoS, CstA, CsrA and Dps) are involved in starvation and the button four proteins are phosphoenolpyruvate-carbohydrate phosphotransferase system (PTS) enzymes.

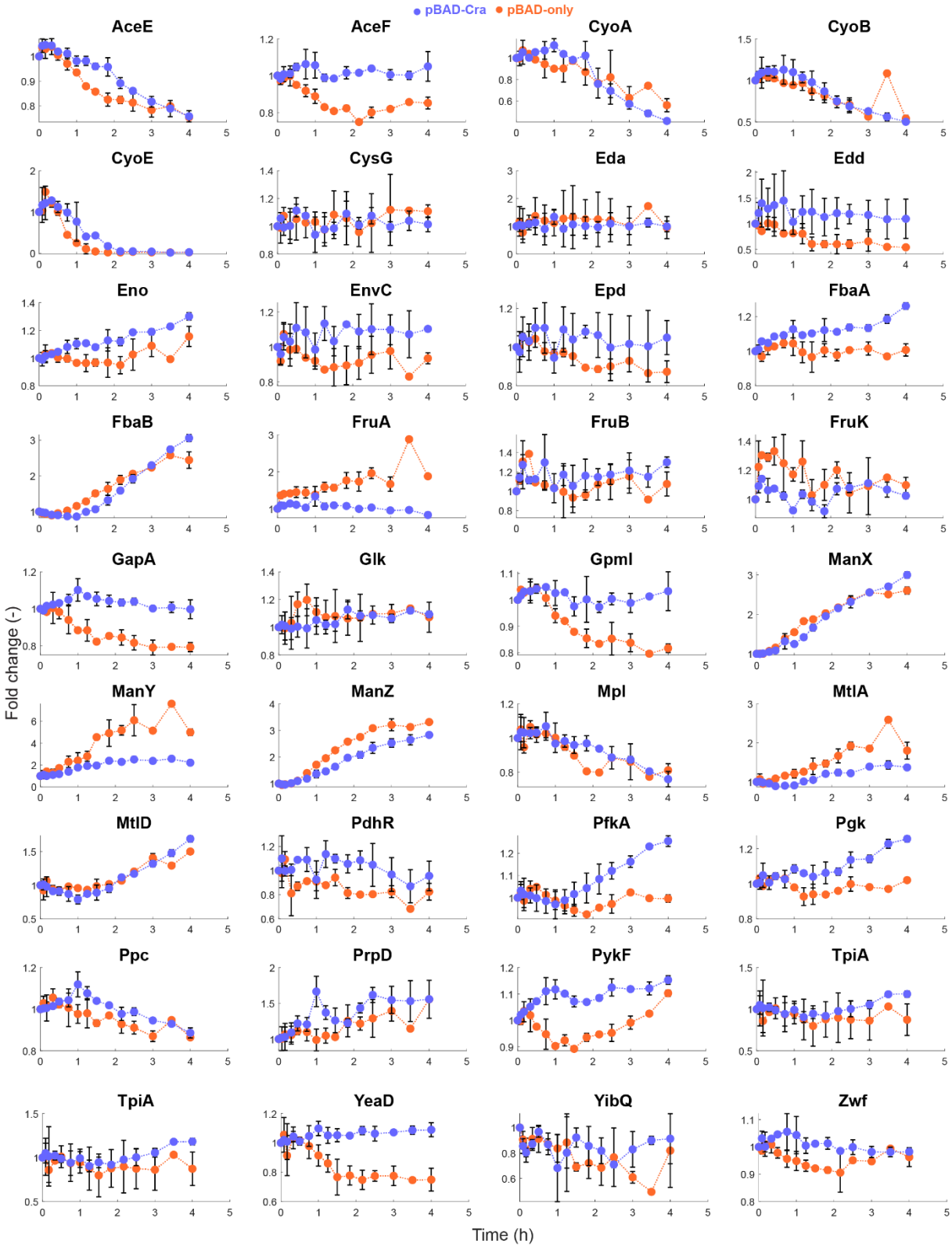


Supplementary figure 9 | The GFP plasmids under the control of pBAD inducible promoter. Two versions of the GFP expression system with or without Cra-binding site were transformed in WT or Δ*cra* background separately. Growth rates of the pBAD-only-GFP and pBAD-Cra-GFP in the wild-type background (**a**) and Δ*cra* background (**b**). **c**, Calculated growth rates of with and without Cra-binding site are compared in

WT or Δ*cra* background.



Supplementary figure 10 | The Cra-regulator and all measurable Cra-activated proteins. The protein expression of both pBAD-only and pBAD-Cra glycerol producers. Pink frame marks the Cra regulator itself. The change of proteins was measured after 0.5% arabinose induction and was normalized to uninduced state.



Supplementary figure 11 | All measurable Cra-inhibited proteins. The protein expression of both pBAD-only and pBAD-Cra glycerol producers. The change of proteins was measured after 0.5% arabinose induction and was normalized to uninduced state.

III. Supplementary Tables

Supplementary Table S1 | Strains and plasmids

Strains/plasmid	Genotype or description	Host	References
MG1655	Wild-type *E. coli* K-12: F-, lambda-, rph-1	-	DZMS-German Collection Of Microorganisms and Cell Cultures (No. 18039)
E. coli DH5α	Used during cloning	-	Cat#18265017, Invitrogen, Thermo Fischer Scientific
BW25113	Derived from MG1655; F-, lambda-, rph-1, Δ*lacZ*4787(::rrnB-3), Δ(*araB–D*)567, Δ(*rhaD–B*)568, hsdR514	-	Keio Collection[83]
GFP (pBAD-only)	P_{BAD} – BBa_B0034 – BBa_E0040 (GFP) – BBa_B0015, p15A, Kanamycin, pSB3K3	MG1655	Prof. Bor-Sen Chen[176]
Genomic DNA	*Saccharomyces cerevisiae* SEY6210	-	Prof. Dr. Victor Sourjik
Chapter 2 - Growth switch			
dCas9	$P_{LtetO-1}$–pCas9, Chloramphenicol, p15A	MG1655	Addgene plasmid #44249
sgRNA	P_{BBa_J23119}–sgRNA, Ampicillin, pUC19	MG1655	Addgene plasmid #44251
pKD46	For homologous recombination, Ampicillin, pMB1	-	Wanner group[79]
pKDsgRNA-ack	P_{J23119} – gRNA, Spectinomycin, pSC101	MG1655	Addgene plasmid #62654
pCas9cr4	P_{tetR} – dCas9, Chloramphenicol, p15A	MG1655	Addgene plasmid #62655
pKDsgRNA-p15	P_{BAD} – dCas9 sgRNA, Spectinomycin, pSC101	MG1655	Addgene plasmid #62656
sgRNA-metE	pdCas9 + metE sgRNA	MG1655	This thesis
sgRNA-rpsG	pdCas9 + rpsG sgRNA	MG1655	This thesis
sgRNA-rpoB-1	pdCas9 + rpoB sgRNA (target template strand, -)	MG1655	This thesis
sgRNA-rpoB-2	pdCas9 + rpoB sgRNA (target non-template strand, +)	MG1655	This thesis
pKDsgRNA-glpK	P_{J23119} – glpK gRNA, Spectinomycin, pSC101	MG1655	This thesis
sgRNA-metE	pdCas9 + metE sgRNA	MG1655	This thesis
pBAD-HC	P_{BAD}+ BBa_B0034+GPD1/GPP2+rrnB terminator, Ampicillin, pMB1	MG1655	This thesis
pBAD-G	Δ *lacZ*::P_{BAD}+ BBa_B0034+GPD1/GPP2+rrnB terminator	MG1655	This thesis
ΔglpF	Δ*glpF*786::kan, Δ(*araD-araB*)567, Δ*lacZ4787*(::rrnB-3), λ-, rph-1, Δ(*rhaD-rhaB*)568, *hsdR*514, F-, BW25113 (JW3898)	BW25113	Keio collection
ΔglpK	Δ*glpK*785::kan, Δ(*araD-araB*)567, Δ*lacZ4787*(::rrnB-3), λ-, rph-1, Δ(*rhaD-rhaB*)568, *hsdR*514, F-, BW25113 (JW3897)	BW25113	Keio collection

ΔglpT	Δ*glpT*720::kan, Δ(*araD-araB*)567, Δ*lacZ4787*(::rrnB-3), λ-, rph-1, Δ(*rhaD-rhaB*)568, *hsdR*514, F-, BW25113 (JW2234)	BW25113	Keio collection
ΔgldA	Δ*gldA*732::kan, Δ(*araD-araB*)567, Δ*lacZ4787*(::rrnB-3), λ-, rph-1, Δ(*rhaD-rhaB*)568, *hsdR*514, F-, BW25113 (JW5556)	BW25113	Keio collection
ΔglpD	Δ*glpD*759::kan, Δ(*araD-araB*)567, Δ*lacZ4787*(::rrnB-3), λ-, rph-1, Δ(*rhaD-rhaB*)568, *hsdR*514, F-, BW25113 (JW3389)	BW25113	Keio collection
ΔglpK	Δ*glpK*, base strain of the study	MG1655	This thesis
Chapter 3 - Dynamic control			
J09B34	P_{J23109} – BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	MG1655	This thesis
J110B32	P_{J23110} – BBa_B0032 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	MG1655	This thesis
J05B34	P_{J23105} – BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	MG1655	This thesis
J01B32	P_{J23101} – BBa_B0032 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	MG1655	This thesis
J06B34	P_{J23106} – BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	MG1655	This thesis
J110B34	P_{J23110} – BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	MG1655	This thesis
J100B34	P_{J23100} – BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	MG1655	This thesis
J01B34	P_{J23101} – BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	MG1655	This thesis
J23101B0032	P_{BAD}– BBa_B0034 – GPD1 – P_{J23110} – BBa_B0032 – GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	This thesis
J23105-B0034	P_{BAD}– BBa_B0034 – GPD1 – P_{J23105} – BBa_B0034 – GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	This thesis
J23101-B0032	P_{BAD}– BBa_B0034 – GPD1 – P_{J23101} – BBa_B0032 – GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	This thesis
J23106-B0034	P_{BAD}– BBa_B0034 – GPD1 – P_{J23106}–BBa_B0034 – GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	This thesis
J23110-B0034	P_{BAD}– BBa_B0034 – GPD1 – P_{J23110} – BBa_B0034 – GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	This thesis
J23101+Cra(-10)	J23101-GFP plasmid, add Cra-binding site at promoter -10 region	MG1655	This thesis
J23101+Cra(-35/-10)	J23101-GFP plasmid, add Cra-binding site between promoter region of -35 and -10	MG1655	This thesis
J23105+Cra(-10)	J23105-GFP plasmid, add Cra-binding site at promoter -10 region	MG1655	This thesis
J23105+Cra(-35/-10)	J23105-GFP plasmid, add Cra-binding site between promoter region of -35 and -10	MG1655	This thesis
P_{pfkA}-Cra	PpfkA-GFP plasmid, remove native Cra-binding site	MG1655	This thesis
pKDsgRNA-cra	P_{J23119} – cra gRNA, Spectinomycin, pSC101	MG1655	This thesis
Δcra	Δ*cra*	MG1655	This thesis

Δcra/ΔglpK	*Δcra ΔglpK*	MG1655	This thesis
J23101	P_{J23101} – BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	Δcra	This thesis
J23105	P_{J23105} – BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	Δcra	This thesis
J23109	P_{J23109} – BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	Δcra	This thesis
Picd	P_{icd} – BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	Δcra	This thesis
PppsA	P_{ppsA} – BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	Δcra	This thesis
PcydA	P_{cydA} – BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	Δcra	This thesis
Pcrp	P_{crp} – BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	Δcra	This thesis
Pepd	P_{epd} – BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	Δcra	This thesis
PpykF	P_{pykF} – BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	Δcra	This thesis
PgapA	P_{gapA} – BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	Δcra	This thesis
PfbaB	P_{fbaB} – BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	Δcra	This thesis
Pppc	P_{ppc} – BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	Δcra	This thesis
PgpmM	P_{gpmM} – BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	Δcra	This thesis
PsoxS	P_{soxS} – BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	Δcra	This thesis
Pacs	P_{acs}– BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	Δcra	This thesis
PgalSP	P_{galSP}– BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	Δcra	This thesis
PptsGp1	P_{ptsGp1}– BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	Δcra	This thesis
PLacZp1	P_{LacZp1}– BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	Δcra	This thesis
PgalEp1	P_{galEp1}– BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	Δcra	This thesis
PgltB	P_{gltB}– BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	Δcra	This thesis
J01+Cra(-10), Δcra	Plasmid J23101+Cra(-10) expressed in Δcra background	Δcra	This thesis
J05+Cra(-10), Δcra	Plasmid J23105+Cra(-10) expressed in Δcra background	Δcra	This thesis
P_{pfkA}+Cra, Δcra	Plasmid P_{pfkA}-Cra expressed in Δcra background	Δcra	This thesis
pBAD-only-GFP, WT	P_{BAD}– BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	MG1655	This thesis
pBAD-Cra-GFP, WT	Added Cra-binding site in pBAD-GFP plasmid	MG1655	This thesis
pBAD+ scramble, WT	Added scrambled sequence CACGAGAGAACAACGTAA in pBAD-GFP plasmid	MG1655	This thesis
pBAD-only-GFP, Δcra	Plasmid of pBAD-GFP, WT expressed in Δcra background	Δcra	This thesis

pBAD-Cra-GFP Δcra	Plasmid of pBAD+Cra-GFP, WT expressed in Δcra background	Δcra	This thesis
pBAD+ scramble Δcra	Plasmid of pBAD+Scr-GFP, WT expressed in Δcra background	Δcra	This thesis
pBAD-only, WT	Plasmid J23101-B0032 expressed in *ΔglpK* strain	ΔglpK	This thesis
pBAD-only, Δcra,	Plasmid J23101-B0032 expressed in Δ*cra* strain	Δcra, ΔglpK	This thesis
pBAD-Cra, WT	Plasmid J23101-B0032 added Cra-binding site and then expressed in *ΔglpK* strain	ΔglpK	This thesis
pBAD-Cra, Δcra,	Plasmid J23101-B0032 added Cra-binding site and then expressed in Δ*cra* strain	Δcra, ΔglpK	This thesis
pBAD-only-1 Cra	Modified pBAD-only by adding additional 1 Cra consensus sequences after terminator, pSB3K3	ΔglpK	This thesis
pBAD-Cra-1 Cra	Modified pBAD-Cra by adding additional 1 Cra consensus sequences after terminator, pSB3K3	ΔglpK	This thesis
pBAD-only-2 Cra	Modified pBAD-only by adding additional 2 Cra consensus sequences after terminator, pSB3K3	ΔglpK	This thesis
pBAD-Cra-2 Cra	Modified pBAD-Cra by adding additional 2 Cra consensus sequences after terminator, pSB3K3	ΔglpK	This thesis
pBAD-only-fused to GFP	P_{BAD}– BBa_B0034 – GPD1 –Linker*2 – BBa_E0040– P_{J23110} –BBa_B0032 – GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	This thesis
pBAD-Cra-fused to GFP	P_{BAD} – Cra-binding site – BBa_B0034 – GPD1 – Linker*2 – BBa_E0040– P_{J23110} –BBa_B0032 – GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	This thesis
pBAD-only-GFP, ΔglpK	P_{BAD}– BBa_B0034 – BBa_E0040 – BBa_B0015, Kanamycin, p15A	ΔglpK	This thesis
pCarotenoid	Native promoter and RBS from *Pantoea ananatis* – *crtBIY* – *dxs* – *crtZ* – *idi* – *crtE*, Streptomycin, pMB1	–	Prof. Dr. Victor Sourjik
pCarotenoid	Native promoter and RBS from *Pantoea ananatis* – *crtBIY* – *dxs* – *crtZE*	BW25113	This thesis
pController_Non	P_{BAD}– BBa_B0034 – *dxs* – BBa_B0034 – dxr – BBa_B0015, Kanamycin, p15A	BW25113	This thesis
pController_Cra	P_{BAD}— Cra-binding site – BBa_B0034 – *dxs* – BBa_B0034 – dxr – BBa_B0015, Kanamycin, p15A	BW25113	This thesis
Chapter 4 - Studies about Cra-regulation			
pTet-only	P_{J23105}– BBa_B0034 –BBa_C0040 –BBa_B0015 – BBa_R0040 – BBa_B0034 – GPD1 – P_{J23106}– BBa_B0034 – GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	This thesis
pTet-Cra	P_{J23105}– BBa_B0034 –BBa_C0040 –BBa_B0015 – BBa_R0040 – Cra site –BBa_B0034 – GPD1 – P_{J23106}–BBa_B0034 – GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	This thesis
J01, Δ*cra*	Plasmid of J01 expressed in a double mutation strain (Δ*cra* and ΔglpK)	Δcra, ΔglpK	This thesis
J05, Δ*cra*	Plasmid of J05 expressed in a double mutation strain (Δ*cra* and ΔglpK)	Δcra, ΔglpK	This thesis
J01+Cra, WT	P_{J23101} – Cra-binding site – BBa_B0034 – GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	This thesis
J05+Cra, WT	P_{J23101} – Cra-binding site – BBa_B0034 – GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	This thesis
J01+Cra, Δ*cra*	Plasmid of J01+Cra, WT expressed in a double mutation strain (Δ*cra* and ΔglpK)	Δcra, ΔglpK	This thesis

J05+Cra, Δ*cra*	Plasmid of J05+Cra, WT expressed in a double mutation strain (Δ*cra* and ΔglpK)	Δcra, ΔglpK	This thesis
J05-spe	P_{J23105}– BBa_B0034 – GPD1 – P_{J23101} – BBa_B0032 – GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	This thesis
Chapter 5 - Burden monitor			
J23100	P_{J23100} – BBa_B0034 – BBa_E0040 – BBa_B0015, Ampicillin, pMB1	MG1655	This plasmid (pSB1A2) was replaced ampicillin and pMB1 with kanamycin and p15 as a low copy number plasmid (pSB3K3).
J23101	P_{J23101} – BBa_B0034 – BBa_E0040 – BBa_B0015, Ampicillin, pMB1	MG1655	
J23105	P_{J23105} – BBa_B0034 – BBa_E0040 – BBa_B0015, Ampicillin, pMB1	MG1655	
J23109	P_{J23109} – BBa_B0034 – BBa_E0040 – BBa_B0015, Ampicillin, pMB1	MG1655	
Picd	P_{icd} – BBa_B0034 – BBa_E0040 – BBa_B0015, Ampicillin, pMB1	MG1655	
PppsA	P_{ppsA} – BBa_B0034 – BBa_E0040 – BBa_B0015, Ampicillin, pMB1	MG1655	
PcydA	P_{cydA} – BBa_B0034 – BBa_E0040 – BBa_B0015, Ampicillin, pMB1	MG1655	
Pcrp	P_{crp} – BBa_B0034 – BBa_E0040 – BBa_B0015, Ampicillin, pMB1	MG1655	
Pepd	P_{epd} – BBa_B0034 – BBa_E0040 – BBa_B0015, Ampicillin, pMB1	MG1655	
PpykF	P_{pykF} – BBa_B0034 – BBa_E0040 – BBa_B0015, Ampicillin, pMB1	MG1655	
PgapA	P_{gapA} – BBa_B0034 – BBa_E0040 – BBa_B0015, Ampicillin, pMB1	MG1655	
PfbaB	P_{fbaB} – BBa_B0034 – BBa_E0040 – BBa_B0015, Ampicillin, pMB1	MG1655	
Pppc	P_{ppc} – BBa_B0034 – BBa_E0040 – BBa_B0015, Ampicillin, pMB1	MG1655	
PgpmM	P_{gpmM} – BBa_B0034 – BBa_E0040 – BBa_B0015, Ampicillin, pMB1	MG1655	
PpfkA	P_{pfkA} – BBa_B0034 – BBa_E0040 – BBa_B0015, Ampicillin, pMB1	MG1655	
PsoxS	P_{soxS} – BBa_B0034 – BBa_E0040 – BBa_B0015, Ampicillin, pMB1	MG1655	
Pacs	P_{acs}– BBa_B0034 – BBa_E0040 – BBa_B0015, Ampicillin, pMB1	MG1655	
PgalSP	P_{galSP}– BBa_B0034 – BBa_E0040 – BBa_B0015, Ampicillin, pMB1	MG1655	
PptsGp1	P_{ptsGp1}– BBa_B0034 – BBa_E0040 – BBa_B0015, Ampicillin, pMB1	MG1655	
PLacZp1	P_{LacZp1}– BBa_B0034 – BBa_E0040 – BBa_B0015, Ampicillin, pMB1	MG1655	
PgalEp1	P_{galEp1}– BBa_B0034 – BBa_E0040 – BBa_B0015, Ampicillin, pMB1	MG1655	
PgltB	P_{gltB}– BBa_B0034 – BBa_E0040 – BBa_B0015, Ampicillin, pMB1	MG1655	
J100	P_{J23100} – BBa_B0034 – GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	This plasmid (pSB3K3 was replaced with ampicillin and
J01	P_{J23101} – BBa_B0034 – GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	

J05	P_{J23105} – BBa_B0034 – GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	pMB1 as a high-copy number plasmid (pSB1A2).
J09	P_{J23109} – BBa_B0034 – GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	
Picd	P_{icd} – BBa_B0034 –GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	
PppsA	P_{ppsA} – BBa_B0034 –GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	
PcydA	P_{cydA} – BBa_B0034 – GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	
Pcrp	P_{crp} – BBa_B0034 – GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	
Pepd	P_{epd} – BBa_B0034 – GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	
PpykF	P_{pykF} – BBa_B0034 – GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	
PgapA	P_{gapA} – BBa_B0034 – GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	
PfbaB	P_{fbaB} – BBa_B0034 – GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	
Pppc	P_{ppc} – BBa_B0034 – GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	
PgpmM	P_{gpmM} – BBa_B0034 – GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	
PpfkA	P_{pfkA} – BBa_B0034 – GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	
PsoxS	P_{soxS} – BBa_B0034 – GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	
Pacs	P_{acs}– BBa_B0034 – GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	
PgalSP	P_{galSP}– BBa_B0034 – GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	
PptsGp1	P_{ptsGp1}– BBa_B0034 – GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	
PLacZp1	P_{LacZp1}– BBa_B0034 – GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	
PgalEp1	P_{galEp1}– BBa_B0034 – GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	
PgltB	P_{gltB}– BBa_B0034 – GPD1/GPP2 – rrnB terminator, Kanamycin, p15A	ΔglpK	
pBAD-HC	Δ*glpK*, P_{BAD}+ BBa_B0034+GPD1/GPP2+rrnB terminator, Ampicillin, pMB1	ΔglpK	This thesis
pBAD-LC	Δ*glpK*, P_{BAD}+ BBa_B0034+GPD1/GPP2+rrnB terminator, Kanamycin, p15A	ΔglpK	This thesis
pBAD-G	Δ *lacZ::*P_{BAD}+ BBa_B0034+GPD1/GPP2+rrnB terminator	ΔglpK	This thesis
	P_{J23100}+ superfolder GFP+ BBa_B1002, Gentamycin	MG1655	Addgene plasmid #66074
Burden monitor (GFP monitor)	Δ*SS9*^{76}*::* P_{J23101}-superfolder GFP^{177}	MG1655	This study
pBAD-G/ΔglpK	Δ*glpK* Δ*lacZ::*P_{BAD}+ BBa_B0034+GPD1/GPP2+rrnB terminator	ΔglpK	This thesis
Chapter 6 - CRISPRi system			
pKDsgRNA-glpK	P_{J23119} – glpK sgRNA, Spectinomycin, pSC101	MG1655	This thesis

pKDsgRNA-Cra	P_{J23119} – Cra sgRNA, Spectinomycin, pSC101	MG1655	This thesis
pTetR-dCas9	P_{tetR}-dCas9, Spectinomycin, pSC101	MG1655	This thesis
ConP-gRNA-GFP-2	P_{BBa_J23119} – GFP sgRNA, Ampicillin, pUC19	MG1655	This thesis
ConP-dCas9	$P_{J23105-73}$ – dCas9, Spectinomycin, pSC101	MG1655	This thesis
	$P_{J23105-37}$ – dCas9, Spectinomycin, pSC101	MG1655	This thesis
	P_{J23114} – dCas9, Spectinomycin, pSC101	MG1655	This thesis
	P_{J23105} – dCas9, Spectinomycin, pSC101	MG1655	This thesis
pBAD-gRNA-GFP-2	P_{J23105} – BBa_B0034 – BBa_C0080 (araC) – BBa_B0015 – P_{BAD}- GFP sgRNA, Ampicillin, pUC19	MG1655	This thesis
ConP-argA	P_{BBa_J23119} – argA sgRNA, Ampicillin, pUC19	MG1655	This thesis
pBAD-argA	P_{J23105} – BBa_B0034 – BBa_C0080 (araC) – BBa_B0015 – P_{BAD}- argA sgRNA, Ampicillin, pUC19	MG1655	This thesis
pBAD-dCas9	P_{BAD}-dCas9, Spectinomycin, pSC101	MG1655	This thesis
argA*	$argA^{H15Y}$, feedback resistant	MG1655	This thesis
ΔargR	ΔargR	MG1655	[21, 151]
Double knockout	ΔargR, argA*	MG1655	[21, 151]
Knockdown #7	pdCas9 + argR sgRNA in argA*	MG1655	[21, 151]
Knockdown #10	pdCas9 + argR sgRNA in argA*	MG1655	[21, 151]

Supplementary Table S2 | Consensus sequences of transcription factors

pBAD promoter with different TF binding sites				
pBAD+TFs(Cra)+scar sequence+B0034 (RBS)	AGAAACCAATTGTCCATATTGCATCAGACATTGCCGTCACTGCGTCTTTTACTGGCTCTTCTCGCTAACCAAACCGGTAACCCCGCTTATTAAAAGCATTCTGTAACAAAGCGGGACCAAAGCCATGACAAAAACGCGTAACAAAAGTGTCTATAATCACGGCAGAAAAGTCCACATTGATTATTTGCACGGCGTCACACTTTGCTATGCCATAGCATTTTTATCCATAAGATTAGCGGATCCTACCTGACGCTTAGCTGAAGCGTTTCAGTCTACTAGAGAAAGAGGAGAAA			
TFs	**Sequence of TF binding sites**	**Numbers of binding sites**	**Numbers of targets**	**Taken from which gene**
Crp	TTAAATTGATCACGTTTTAGAC	283[178]	*522*[179]	*cyoA*
Mlc	TTTTTTTAAAGCTCGTAATTAAT			*ptsG*
Fnr	TTGATATTTATCAAtgta	300–700[98, 180]	465[98]	*cydA*
Fis	ttGCGCAATTTGTCAACg	1269[178]	231[99]	*gltX*
Fur	TGAGATAATGCGTATCATT	134-144[181]	12-39[181]	*fur*
AcrR	TACATACATTTGTGAATGTATGTA			*acrA*
Cra	AGCTGAAGCGTTTCAGTC	164[182]	79[179]	*epd*

Supplementary Table S3 | Sequences of glycerol production construct

Genes	Sequence (5'→3')
pBAD-Cra strain: pBAD+Cra-binding site+B0034(RBS)+GPD1+J23101+B0032(RBS)+GPP2+terminator **pBAD-only strain:** All the sequences are exactly same as pBAD-Cra strain except without Cra-binding site.	AGAAACCAATTGTCCATATTGCATCAGACATTGCCGTCACTGCGTCTTTTA CTGGCTCTTCTCGCTAACCAAACCGGTAACCCCGCTTATTAAAAGCATTCT GTAACAAAGCGGGACCAAAGCCATGACAAAAACGCGTAACAAAAGTGTCT ATAATCACGGCAGAAAAGTCCACATTGATTATTTGCACGGCGTCACACTTT GCTATGCCATAGCATTTTTATCCATAAGATTAGCGGATCCTACCTGACGCT TAGCTGAAGCGTTTCAGTCTACTAGAGAAAGAGGAGAAAtactagatgTCTGC TGCTGCTGATAGATTAAACTTAACTTCCGGCCACTTGAATGCTGGTAGAAA GAGAAGTTCCTCTTCTGTTTCTTTGAAGGCTGCCGAAAAGCCTTTCAAGGT TACTGTGATTGGATCTGGTAACTGGGGTACTACTATTGCCAAGGTGGTTG CCGAAAATTGTAAGGGATACCCAGAAGTTTTCGCTCCAATAGTACAAATGT GGGTGTTCGAAGAAGAGATCAATGGTGAAAAATTGACTGAAATCATAAATA CTAGACATCAAAACGTGAAATACTTGCCTGGCATCACTCTACCCGACAATT TGGTTGCTAATCCAGACTTGATTGATTCAGTCAAGGATGTCGACATCATCG TTTTCAACATTCCACATCAATTTTTGCCCCGTATCTGTAGCCAATTGAAAG GTCATGTTGATTCACACGTCAGAGCTATCTCCTGTCTAAAGGGTTTTGAAG TTGGTGCTAAAGGTGTCCAATTGCTATCCTCTTACATCACTGAGGAACTAG GTATTCAATGTGGTGCTCTATCTGGTGCTAACATTGCCACCGAAGTCGCT CAAGAACACTGGTCTGAAACAACAGTTGCTTACCACATTCCAAAGGATTTC AGAGGCGAGGGCAAGGACGTCGACCATAAGGTTCTAAAGGCCTTGTTCC ACAGACCTTACTTCCACGTTAGTGTCATCGAAGATGTTGCTGGTATCTCCA TCTGTGGTGCTTTGAAGAACGTTGTTGCCTTAGGTTGTGGTTTCGTCGAA GGTCTAGGCTGGGGTAACAACGCTTCTGCTGCCATCCAAAGAGTCGGTTT GGGTGAGATCATCAGATTCGGTCAAATGTTTTTCCCAGAATCTAGAGAAG AAACATACTACCAAGAGTCTGCTGGTGTTGCTGATTTGATCACCACCTGC GCTGGTGGTAGAAACGTCAAGGTTGCTAGGCTAATGGCTACTTCTGGTAA GGACGCCTGGGAATGTGAAAAGGAGTTGTTGAATGGCCAATCCGCTCAA GGTTTAATTACCTGCAAAGAAGTTCACGAATGGTTGGAAACATGTGGCTCT GTCGAAGACTTCCCATTATTTGAAGCCGTATACCAAATCGTTTACAACAAC TACCCAATGAAGAACCTGCCGGACATGATTGAAGAATTAGATCTACATGAA GATtagGCATGCGAGCTCCATATGACTAGTtttacagctagctcagtcctaggtattatgcta gcTACTAGAGtcacacaggaaagtactagatgGGATTGACTACTAAACCTCTATCTTT GAAAGTTAACGCCGCTTTGTTCGACGTCGACGGTACCATTATCATCTCTCA ACCAGCCATTGCTGCATTCTGGAGGGATTTCGGTAAGGACAAACCTTATT TCGATGCTGAACACGTTATCCAAGTCTCGCATGGTTGGAGAACGTTTGAT GCCATTGCTAAGTTCGCTCCAGACTTTGCCAATGAAGAGTATGTTAACAAA TTAGAAGCTGAAATTCCGGTCAAGTACGGTGAAAAATCCATTGAAGTCCC AGGTGCAGTTAAGCTGTGCAACGCTTTGAACGCTCTACCAAAAGAGAAAT GGGCTGTGGCAACTTCCGGTACCCGTGATATGGCACAAAAATGGTTCGA GCATCTGGGAATCAGGAGACCAAAGTACTTCATTACCGCTAATGATGTCA AACAGGGTAAGCCTCATCCAGAACCATATCTGAAGGGCAGGAATGGCTTA GGATATCCGATCAATGAGCAAGACCCTTCCAAATCTAAGGTAGTAGTATTT GAAGACGCTCCAGCAGGTATTGCCGCCGGAAAAGCCGCCGGTTGTAAGA TCATTGGTATTGCCACTACTTTCGACTTGGACTTCCTAAAGGAAAAAGGCT GTGACATCATTGTCAAAAACCACGAATCCATCAGAGTTGGCGGCTACAAT GCCGAAACAGACGAAGTTGAATTCATTTTTGACGACTACTTATATGCTAAG GACGATCTGTTGAAATGGtaagatggtagtgtggggtctccccatgcgagagtagggaactgcc aggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttcgttttatctgttgtttgtcggtgaacgct ctcctgagtaggacaaatccgccgggagcggatttgaacgttgcgaagcaacggcccggagggtggcgg gcaggacgcccgccataaactgccaggcatcaaattaagcagaaggccatcctgacggatggccttttgc gt

Supplementary Table S4 | Primers used in this study

Oligonucleotide	Sequence (5'→3')	Purpose
Chapter 2 - Growth switch		
sgRNA-GFP-1-F (+)	GGAGGTAAGATAATGCGTAAgttttagagctagaaatagcaagttaaaataaggc	Clone GFP sgRNA from pgRNA plasmid
sgRNA-GFP-2-F (-)	CCGTCCAGCTCGACCAGAATgttttagagctagaaatagcaagttaaaataaggc	
sgRNA-RNAP(rpoB)-1-F (-)	TCATTGCGATTTACACAGAAgttttagagctagaaatagcaagttaaaataaggc	Clone rpoB sgRNA from pgRNA plasmid
sgRNA-RNAP(rpoB)-2-F (+)	CTGTGTAAATCGCAATGAAAgttttagagctagaaatagcaagttaaaataaggc	
sgRNA-metE-F	CAGGCCAACGCGAGGGAAACCGgttttagagctagaaatagcaagttaaaataaggc	Clone metE sgRNA from pgRNA plasmid
Ec-R	actagtattatacctaggactgagctagc	Reverse primer for amplification of customized CRISPRi plasmids
pBAD-HC plasmid: pBAD inducible glycerol producer- GPD1-GPP2 fusion version (plasmid-based)		
pBAD-F	AATTCTCATGTTTGACAGcttcctcgctcactgactcgc	Amplification of pBAD promoter from GFP (pBAD-only) plasmid
pBAD-R	cagcagcagacatctagtaTTTCTCCTCTTTCTCTAGTAGCTAGCC	
GPD1-F	AGGAGAAAtactagatgTCTGCTGCTGCTGATAGATTAAACTTAAC	Amplification of *gpd1* gene from yeast genomic DNA
GPD1-R	GTCAATccccatataactaCAATCATGTCCGGCAGGTTCTTC	
GPP2-F	ACATGATTGtagttatatggggATTGACTACTAAACCTCTATCTTTGAAAGTTAACGC	Amplification of *gpp2* gene from yeast genomic DNA
GPP2-R	cacactaccatcttaCCATTTCAACAGATCGTCCTTAGC	
pKD-13-rrnB-F	CTGTTGAAATGGtaagatggtagtgtggggtctcc	Amplification of pBAD promoter from pKD46 plasmid as a backbone
pKD-13-rrnB-R	gaggaagCTGTCAAACATGAGAATTAATTCC	
LacZ/pBAD-F	AGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGGACCCCCGGCTAGAAACCAATTGTCCATATTGCATC	Construction of homologous arms from *E. coli*'s *lacZ* gene
LacZ/pKD-R	AAGACTGTTACCCATCGCGTGGGCGTATTCGCAAAGGATCAGCGGGCGCGCTGTCAAACATGAGAATTAATTCCG	
Seq-lacZ-F	GAATTCCTGGCACGACAGG	For sequencing
Seq-lacZ-R	CTGCTGGTGTTTTGCTTCC	
pBAD-G: pBAD inducible glycerol producer- GPD1-GPP2 fusion version (genome-based)		
lin-gly-lacZ-F	GTCTCTGACCAGACACCCATCAAC	Amplification of pBAD-HC plasmid to get a linearized fragment
lin-gly-lacZ-R	GATGATGCTCGTGACGGTTAACG	
LacZ-sgRNA-F	GACCATGATTACGGATTCACgttttagagctagaaatagcaag	Clone sgRNA for No-SCAR genomic editing from pKDsgRNA-ack plasmid
SCAR-sgRNA-R (LacZ)	GTGAATCCGTAATCATGGTCgtgctcagtatctctatcactga	
Forward001	tttataacctccttagagctcga	
Reverse001	ccaattgtccatattgcatca	
colony-SCAR-LacZ-F	CAGTGGGCTGATCATTAACTATCC	For screening genomic integration by colony PCR
colony-SCAR-LacZ-R	CGTAATCACCCGAGTGTGATC	
Deletion of *glpK* gene in MG1655		

glpk-sgRNA-F	CGTGATCCATTACGACCGCGgttttagagctagaaatagcaag	Clone of *glpK* sgRNA by inverse PCR from pKDsgRNA-ack plasmid
SCAR-sgRNA-R	gtgctcagtatctctatcactga	
Knockout-glpk	CTACGGGACAATTAAACATGACTGAAAAAAAATATATCGTTGCGCTCGACCAGGGCACCACCAGCTCGGTTAAACGCGCGATGGCGTGGGAAGAACACGACGAAtaaTGTAAATGCCGAATG	Template for homologous recombination
colony-glpK-F	GTCCTAACCATTTGTAGCCAGCCAG	Colony PCR for checking *glpK* deletion
colony-glpK-R	CATTTGCCTACCGCAAACTGATTG	
Chapter 3 - Dynamic control		
Construction of separated *gpd1* and *gpp2* genes		
Gly-B-F	gatgGGATTGACTACTAAACCTCTATCTTTG	Backbone
Gly-B-R	GTAGATCTAATTCTTCAATCATGTCCGGCAGGTTC	
J23110-B0032	CCGGACATGATTGAAGAATTAGATCTACATGAAGATtagGCATGCGAGCTCCATATGACTAGTtttacggctagctcagtcctaggtacaatgctagcTACTAGAGtcacacaggaaagtactagatgGGATTGACTACTAAACCTCTATCTTTG	Construction of separated *gpd1* and *gpp2* genes and the gpp2 gene was cloned under different promoters
J23105-B0034	CCGGACATGATTGAAGAATTAGATCTACATGAAGATtagGCATGCGAGCTCCATATGACTAGTtttacggctagctcagtcctaggtactatgctagcTACTAGAGAAAGAGGAGAAAtactagatgGGATTGACTACTAAACCTCTATCTTTG	
J23101-B0032	CCGGACATGATTGAAGAATTAGATCTACATGAAGATtagGCATGCGAGCTCCATATGACTAGTtttacagctagctcagtcctaggtattatgctagcTACTAGAGtcacacaggaaagtactagatgGGATTGACTACTAAACCTCTATCTTTG	
J23106-B0034	CCGGACATGATTGAAGAATTAGATCTACATGAAGATtagGCATGCGAGCTCCATATGACTAGTtttacggctagctcagtcctaggtatagtgctagcTACTAGAGAAAGAGGAGAAAtactagatgGGATTGACTACTAAACCTCTATCTTTG	
J23110-B0034	CCGGACATGATTGAAGAATTAGATCTACATGAAGATtagGCATGCGAGCTCCATATGACTAGTtttacggctagctcagtcctaggtacaatgctagcTACTAGAGAAAGAGGAGAAAtactagatgGGATTGACTACTAAACCTCTATCTTTG	
J23110-B0032	CCGGACATGATTGAAGAATTAGATCTACATGAAGATtagGCATGCGAGCTCCATATGACTAGTtttacggctagctcagtcctaggtacaatgctagcTACTAGAGtcacacaggaaagtactagatgGGATTGACTACTAAACCTCTATCTTTG	
Seq-conP-F	TCGGTTTGGGTGAGATCATC	Check insertion of promoters between GPD1 and GPP2
Colony-ConP-R	CAATGGCTGGTTGAGAGATG	
Adding Cra-binding site for GFP production plasmids		
J-series-Cra-10-site-F	AGCTGAAGCGTTTCAGTCTACTAGAGAAAGAGGAGAAATACTAGATGCGTAAAG	Add Cra consensus sequence at -10 site
J01-Cra-10-R	GCTAGCATAATACCTAGGACTGAGCTAGCTG	
J05-Cra-10-R	GCTAGCATAGTACCTAGGACTGAGCTAGC	
J01-E40-Cra-35/10-R	ACTGAAACGCTTCAGCTTGTAAACTCTAGAAGCGGCCGC	Add Cra consensus sequence between -35 and -10 site
J05-E40-35/10-R	ACTGAAACGCTTCAGCTCGTAAACTCTAGAAGCGGCCG	
PpfkA-NO-Cra-F	CACGAGAGAACAACGTAAGGAAGTGATTGTTATACTATTTGCACATTCG	Replace native Cra consensus sequence with

PpfkA-NO-Cra-R	CAGGCCAAATGGCGGTATTTTATAC	CACGAGAGAACAACGTAA
Construction of *cra* deletion		
Cra-sgRNA-F	ACTGGATGAAATCGCTCGGCgttttagagctagaaatagcaag	Clone of cra sgRNA by inverse PCR with SCAR-sgRNA-R
Knockout-Cra	GATCTCAATGCGCAATTTACAGCCCAACATGTCACGTTGGGCCGCGCCAGGTGAATTTCCCTCTGGCGCGTAGAGTACGGGACTGGACATC	Template for homologous recombination
colony-Cra-F	TAGCAGGCACCAGCGGTAAGC	Colony PCR for genomic modification of glpK
colony-Cra-R	GCAATTCATCGCCGCATGTAAC	
Adding Cra-binding site for pBAD-GFP plasmids		
IF-backbone-F	TACTAGAGAAAGAGGAGAAAtactagatgcgtaaa	Add scrambled sequence in pBAD-GFP
pBAD_ scr_E40-R	TTACGTTGTTCTCTCGTGGCTAGCCCAAAAAAACGGTATGGAG	
pBAD-R	GCTAGCCCAAAAAAACGGTATGGAG	
TF-F	CTTCTAGAGCAACGCAATTAATGTGAG	Change promoter region from pBAD-Cra-gly
TF-R	ctttacgcatctagtaTTTCTCCTCTTTCTCTAGTA	
IF-backbone-F	TACTAGAGAAAGAGGAGAAAtactagatgcgtaaaggagaagaacttttcactggag	
IF-backbone -R	CTCACATTAATTGCGTTGCTCTAGAAGCGGCCGCGAATTC	
Cra(-10)-F	CTCTAGTAGACTGAAACGCTTCAGCTGCTAGCCCAAAAAAACGGTATGG	Add Cra binding site for pBAD-Cra-gly
Cra(-10)--R	AAAGAGGAGAAAtactagatgTCTGCTGC	
Template+Cra	AGCTGAAGCGTTTCAGTCTACTAGAGAAAGAGGAGAAAtactagatgTCTGCTGC	
Construction of diffent TFs-binding sites of pBAD-only plasmid		
TFL-backbone-F	GAAAGAGGAGAAAtactagatgTC	Amplification from pBAD-only plasmid
TFL-backbone-R	GTATGGAGAAACAGTAGAGAGTTG	
TFL-F	caaCTCTCTACTGTTTCTCCATAC	Amplification of linear template sequence into double stranded DNA
TFL-R	GAcatctagtaTTTCTCCTCTTTC	
ArcR_2_template	CAACTCTCTACTGTTTCTCCATACCGTTTTTTTGGGCTAGCTACATACATTTGTGAATGTATGTAAGAAAGAGGAGAAAtactagatgTC	Template sequence for the construction of differnt TF binding site.
Crp_2_template	CAACTCTCTACTGTTTCTCCATACCGTTTTTTTGGGCTAGCTTAAATTGATCACGTTTTAGACAGAGAAAGAGGAGAAAtactagatgTC	
FadR_2_template	CAACTCTCTACTGTTTCTCCATACCGTTTTTTTGGGCTAGCAGGAGGTCTGACCACTTgTACTAGAGAAAGAGGAGAAAtactagatgTC	
Fis_2_template	CAACTCTCTACTGTTTCTCCATACCGTTTTTTTGGGCTAGCttGCGCAATTTGTCAACgTACTAGAGAAAGAGGAGAAAtactagatgTC	
Fnr_2_template	CAACTCTCTACTGTTTCTCCATACCGTTTTTTTGGGCTAGCTTGATATTTATCAAtgtaTACTAGAGAAAGAGGAGAAAtactagatgTC	
Fur-Fe_2_template	CAACTCTCTACTGTTTCTCCATACCGTTTTTTTGGGCTAGCTGAGATAATGCGTATCATTACTAGAGAAAGAGGAGAAAtactagatgTC	
Mlc_2_template	CAACTCTCTACTGTTTCTCCATACCGTTTTTTTGGGCTAGCTTTTTTTAAAGCTCGTAATTAATGAGAAAGAGGAGAAAtactagatgTC	
Additional Cra-binding site		
Cra_3_F	CTCGAGTCCCGTCAAGTCAG	Amplification from pBAD-only or pBAD-Cra plasmid
Km-R	ctgcctcggtgagttttctc	

Cra_3_1Temp	gctcactcaaaggcggtaatAGCTGAAGCGTTTCAGT Cctcgagtcccgtcaagtcag	Add 1 additional Cra binding site
Cra_3-2Temp	gctcactcaaaggcggtaatAGCTGAAGCGTTTCAGT CGCGATCTTGTCTTTAACCCTAAGCCAGGTT GGCGCTTTTTTTCTAGCTGAAGCGTTTCAGTC ctcgagtcccgtcaagtcag	Add 2 additional Cra binding site
Glycerol producer with fused GFP		
fGFP_ggggs_F	ggtggtggtggttctcgtaaaggagaagaacttttcactg	Add 2 linkers (GGGGS) in pBAD-only and pBAD-Cra plasmid
fGFP_ggggs_R	actagtcatatggagctcgcatgcttattatttgtatagttcatccatgc catg	
fGFP_B_F	GCATGCGAGCTCCATATGA	
fGFP_B2_R	cctttacgagaaccaccaccaccagaaccaccaccaccATCT TCATGTAGATCTAATTCTTCAATCATGTC	
VR	attaccgcctttgagtgagc	Add additional Cra-binding site of both pBAD-only and pBAD-Cra plasmid by gibson assembly
Carotenoid producer		
pBAD-Bdxs-F	gTACTAGAGccaggcatcaaataaaacgaaag	Make backbone with or without Cra-binding site
pBAD-Bdxs-R	ctagtaTTTCTCCTCTTTGGTACCGCTAGCCCAA AAAAACGGTATGGAGAAAC	
pBAD-Cra-Bdxs-R	ctagtaTTTCTCCTCTTTGGTACCGACTGAAAC	
dxs-F	GGTACCAAAGAGGAGAAAtactagatgAGTTTTG ATATTGCCAAATACCCGAC	Amplify *dxs* gene from MG1655
dxs-R	ctagtaTTTCTCCTCTTTCATATGttaTGCCAGCC AGGCCTTGATTTTG	
dxr-F	cataaCATATGAAAGAGGAGAAAtactagatgAAGC AACTCACCATTCTGGGCTC	Amplify *dxr* gene from MG1655
dxr-R	tttgatgcctggCTCTAGTActgcaGtcaGCTTGCGAG ACGCATCA	
Seq-dxs-F	GATGAAGCGTTAATTCTGGAAATG	Sequencing and checking
colony-dxs-R	GTTCTTTAGCGTGGTGATAAG	
colony-dx-F	GATGAAGCGTTAATTCTGGAAATG	
colony-dx-R	AACGCGCTTCAATGTATTCC	
Chapter 4 - Studies about Cra-regulation		
tetR-F	GAGCGCAACGCAATTAATGTGAGtttacggctagct cagtcctag	Clone aTc induced glycerol production with or without Cra-binding site
tetR-R	CTCCTCTTTCTCTAGTAGTGCTCAGTATC	
B34GPD1-F	CTGAGCACTACTAGAGAAAGAGGAGAAAtacta gatgTCTGCT	
regulator-R	CTCACATTAATTGCGTTGCGCTC	
tetR_Cra_R	GTAGACTGAAACGCTTCAGCTGTGCTCAGTA TCTCTATCACTGATAGG	
+Cra-F	ACAGCTGAAGCGTTTCAGTCTACTAGAG	
Seq-R40-F	CCGCCATTATTACGACAAGC	Sequencing
Chapter 5 - Burden monitor		
GFP production driven by different native promoters (high-copy number)		
regulator-F	TACTAGAGAAAGAGGAGAAAtactagatgTCTGC TG	Amplification of pBAD-HC plasmid as a backbone
regulator-R	CTCACATTAATTGCGTTGCGCTC	
Pacs-F	GCGCAACGCAATTAATGTGAGGACATTGCTC GCCCCTATGTG	Different native promoters amplified from *E. coli*
Pacs-R	GCAGAcatctagtaTTTCTCCTCTTTCTCTAGTAA TGTTAATAATATGTGGCATAAGCGTTAAATGT AG	
Pcrp-F	GCGCAACGCAATTAATGTGAGTCCTGACAGA GTACGCGTACTAACCAAATC	

Pcrp-R	GCAGAcatctagtaTTTCTCCTCTTTCTCTAGTAACCTGGGAAGGGGCTATCAACTG	
PcydA-F	GCGCAACGCAATTAATGTGAGGCTATAAATTGATCACCGTCGAAAAATG	
PcydA-R	gcagacatctagtaTTTCTCCTCTTTCTCTAGTAGAACACTTTCATCATACCATGTGAATGC	
Pepd-WT-BW-F	GCAGAcatctagtaTTTCTCCTCTTTCTCTAGTAGAAGCCATTTATCGCTACGCGTAC	
Pepd-WT-BW-R	GCGCAACGCAATTAATGTGAGGAGATGACTGAAGAAGGCCGGTATC	
Pepd-non-F	GCGCAACGCAATTAATGTGAGGGCTGGACAAACATTCCTTTTATTC	
Pepd-non-R	GCAGAcatctagtaTTTCTCCTCTTTCTCTAGTAGAAGCCATTTATCGCTACGCGTAC	
Pepd-scram-F	GCGCAACGCAATTAATGTGAGGGATCCGAGATGACTGAAGAAGG	
Pepd-scram-R	GCAGAcatctagtaTTTCTCCTCTTTCTCTAGTACTCGAGGAAGCCATTTATCGCTAC	
Pepd-icd-F	GCGCAACGCAATTAATGTGAGTTTCATGACGGCAAACAATAGGGTAG	
Pepd-icd-R	GCAGAcatctagtaTTTCTCCTCTTTCTCTAGTAGAAGCCATTTATCGCTACGCGTAC	
PgltB-F	GCGCAACGCAATTAATGTGAGTTTAGCTCAATTCGGCGATGAC	
PgltB-R	GCAGAcatctagtaTTTCTCCTCTTTCTCTAGTAGGATTTCCAACTTATCGGGAAAG	
Picd-F	GCGCAACGCAATTAATGTGAGGGTTTACGCCTGGTAGAACGTTG	
Picd-R	GCAGAcatctagtaTTTCTCCTCTTTCTCTAGTAGTCATGAAAGATGCGCGTTATTAG	
Ppck-F	GAGCGCAACGCAATTAATGTGAGAAAGTTAGCGTGGTGAATCGATAC	
Ppck-R	CAGCAGAcatctagtaTTTCTCCTCTTTCTCTAGTACGTGATTCCTGTCACGAAAC	
ppsA-F	GGTTAAATATGCAAAGATAAATGCGCAGAAATGTGTTTCTCAAACTACTAGAGAAAGAGGAGAAAtactagatgtctgctg	
ppsA-R	GGATGAAAAAAACGGTGAATCGTTCAAGCAAATATATTTTTTTACCTCACATTAATTGCGTTGCGCTC	
PsoxS-F	AGGAATTATACTCCCCAACAGATGAATTAACGAACTGAACTACTAGAGAAAGAGGAGAAAtactagatgTCTGCTG	
PsoxS-R	CAAGTTAACTTGAGGTAAAGCGATTTatggaaaagaaattaCTCACATTAATTGCGTTGCGCTC	
galEp1-F	tgttatgctatggttatttcAtaccataagcctaatggagcTACTAGAGAAAGAGGAGAAAtactagatgTCTGCTG	
galEp1-R	aagatgcgaaaagtgtgacatggaataaattagtggaatcgtCTCACATTAATTGCGTTGCGCTC	
LacZp1-F	cactttatgcttccggctcgtatgttgtgtggAattgtgagcggataacaTACTAGAGAAAGAGGAGAAAtactagatgTCTGCTG	
PfbaB-F	GCGCAACGCAATTAATGTGAGGACAATACCACTTAATAATACCTTTTAAATACCTTTGC	
PfbaB-R	GCAGAcatctagtaTTTCTCCTCTTTCTCTAGTAGTTACGGGCGAGTTTTTCCATCT	
PgapA-F	GCGCAACGCAATTAATGTGAGTGACTGATTCTAACAAAACATTAACACCAAC	

PgapA-R	GCAGAcatctagtaTTTCTCCTCTTTCTCTAGTAT GCCTGTAAAATTACAAAAACCTTACGC	
PgpmM-F	GCGCAACGCAATTAATGTGAGCAGCCAGCGC GTTAACTGGAATG	
PgpmM-R	GCAGAcatctagtaTTTCTCCTCTTTCTCTAGTAA CCTCATACTCAAGAGTCAAAATTTGCGTAATT TTAC	
PpfkA-F	GCGCAACGCAATTAATGTGAGCAGACCCGCA TTTTGTGTATAAAATACC	
PpfkA-R	GCAGAcatctagtaTTTCTCCTCTTTCTCTAGTAT GAACTTTGGAATGCAAAATGAAATCTG	
Pppc-F	GCGCAACGCAATTAATGTGAGGCATCTTATC CGACCTACACCTTTGGTG	
Pppc-R	GCAGAcatctagtaTTTCTCCTCTTTCTCTAGTAT CGGGCTTTGCTTTTCGTCGTCT	
PpykF-F	GCGCAACGCAATTAATGTGAGATATTTTTTGA AACGCTGTTTTTGTTTTCCTTTTG	
PpykF-R	GCAGAcatctagtaTTTCTCCTCTTTCTCTAGTAG AAACTTGCTTTCTGGGCGCACTG	
GFP production driven by different native promoters (low-copy number)		
CopyN-NotI-F	ATAAGAATGCGGCCGCGAGCGCAACGCAATT AATGTGAG	Change copy number of any GFP plasmid in this study
CopyN-PstI-R	AAAACTGCAGACGCAAAAAGGCCATCCGTC	
Change glycerol production plasmid with different native promoters		
TF-F	CTTCTAGAGCAACGCAATTAATGTGAG	Amplification of native promoters from constructed GFP plasmids
TF-R	ctttacgcatctagtaTTTCTCCTCTTTCTCTAGTA	
IF-backbone-F	TACTAGAGAAAGAGGAGAAAtactagatgcgtaaag gagaagaacttttcactggag	Amplification of pBAD-HC plasmid as a backbone
IF-backbone -R	CTCACATTAATTGCGTTGCTCTAGAAGCGGC CGCGAATTC	
Construction of a genome-based GFP system as a burden monitor		
SS9-sgRNA-F	TCTGGCGCAGTTGATATGTAgttttagagctagaaat agcaag	Construction of sgRNA plasmid from pKDsgRNA-ack
NO-SCAT-sgRNA-R	gtgctcagtatctctatcactga	
2-SS9-left arm-J100-F	GATGACGTAATTAAGCATTGATAATTGAGATC CCTCTCCCTGACAGGttgacggctagctcagtcctag	Homologous arms
2-SS9-Right arm-J100-R	GCTTCCTGAGTAATAACTTCCTGAGTGAATAT TTAACCTGAGCTTGATCCgatggtagtgtggggtctcc	
Seq-SS9-F	GTTTGGGTGAAAACGAAAATTCC	For sequencing
Seq-SS9-R	CGTTTATTATGCCACAGAGAATCG	
Chapter 6 – CRISPR interference		
J05-dCas9-RBS-lib-F	DDRRGGAW TACTAGATGGATAAGAAATACTCAATAGGCTT AG	Clone different strength of dCas9 by CPEC cloning
J05-dCas9-RBS-lib-R	CTAGTAGCTAGCATAGTACCTAGGACTGAGC	
dCas9-RBS-lib-F	NVDGGRGG TAACAT ATGGATAAGAAATACTCAATAGGCTTAGCTAT CG	
dCas9-RBS-lib-R	TTGTAT GCTAGCATAGTACCTAGGACTGAGCTAGCC	
F-105	actatgctagcaaagaggagaaatactagATGGATAAGA AATACTCAATAGGC	J23105_dCas9

R-105	acctaggactgagctagccgtaaaCTGCAGTCTAGAGACGTCG	
F-114	acaatgctagcaaagaggagaaatactagATGGATAAGAAATACTCAATAGGC	J23114_dCas9
R-114	acctaggactgagctagccataaaCTGCAGTCTAGAGACGTCG	
pBAD-back-F	GATTTAGGGtaaTACTAGAGccaggcatcaaataaaacgaaag	Clone backbone from plasmid pBAD-LC
pBAD-back-R	GTAATAAGAAAACACctagtaTTTCTCCTCTTTCTCTAGTA	
agrA-H15Y-F	GGTCGAGGGATTCCGCCATTgttttagagctagaaatagcaag	Mutation of *argA* gene by No-SCAR method
agrA-H15Y-R	ATATCAATACCCACCGGG	
argA-seq	ATCCTGACATGCCTCTCCCGAG	Sequencing *argA* gene in the genome

IV. References

1. Berla, B.M. et al. Synthetic biology of cyanobacteria: unique challenges and opportunities. *Frontiers in Microbiology* **4** (2013).
2. Nielsen, A.A., Segall-Shapiro, T.H. & Voigt, C.A. Advances in genetic circuit design: novel biochemistries, deep part mining, and precision gene expression. *Curr Opin Chem Biol* **17**, 878-892 (2013).
3. Straathof, A.J. Transformation of biomass into commodity chemicals using enzymes or cells. *Chem Rev* (2013).
4. Wang, Y.H., Wei, K.Y. & Smolke, C.D. Synthetic biology: advancing the design of diverse genetic systems. *Annu Rev Chem Biomol Eng* **4**, 69-102 (2013).
5. Kent, R. & Dixon, N. Contemporary tools for regulating gene expression in bacteria. *Trends Biotechnol* **38**, 316-333 (2020).
6. Rugbjerg, P. & Sommer, M.O.A. Overcoming genetic heterogeneity in industrial fermentations. *Nat Biotechnol* **37**, 869-876 (2019).
7. Garcia, S. & Trinh, C.T. Modular design: Implementing proven engineering principles in biotechnology. *Biotechnol Adv* **37**, 107403 (2019).
8. Burg, J.M. et al. Large-scale bioprocess competitiveness: the potential of dynamic metabolic control in two-stage fermentations. *Current Opinion in Chemical Engineering* **14**, 121-136 (2016).
9. Soma, Y., Fujiwara, Y., Nakagawa, T., Tsuruno, K. & Hanai, T. Reconstruction of a metabolic regulatory network in *Escherichia coli* for purposeful switching from cell growth mode to production mode in direct GABA fermentation from glucose. *Metab Eng* **43**, 54-63 (2017).
10. Wiktor, J., Lesterlin, C., Sherratt, D.J. & Dekker, C. CRISPR-mediated control of the bacterial initiation of replication. *Nucleic Acids Research* (2016).
11. Li, S. et al. Enhanced protein and biochemical production using CRISPRi-based growth switches. *Metab Eng* **38**, 274-284 (2016).
12. Izard, J. et al. A synthetic growth switch based on controlled expression of RNA polymerase. *Molecular Systems Biology* **11** (2015).
13. Lo, T.-M., Chng, Si H., Teo, Wei S., Cho, H.-S. & Chang, Matthew W. A two-layer gene circuit for decoupling cell growth from metabolite production. *Cell Systems* **3**, 133-143 (2016).
14. Tokuyama, K. et al. Magnesium starvation improves production of malonyl-CoA-derived metabolites in *Escherichia coli*. *Metab Eng* **52**, 215-223 (2019).
15. Zeng, W., Guo, L., Xu, S., Chen, J. & Zhou, J. High-throughput ccreening technology in industrial biotechnology. *Trends Biotechnol* (2020).
16. Klamt, S., Mahadevan, R. & Hadicke, O. When do two-stage processes outperform one-stage processes? *Biotechnol J* **13** (2018).
17. von Kamp, A. & Klamt, S. Growth-coupled overproduction is feasible for almost all metabolites in five major production organisms. *Nat Commun* **8**, 15956 (2017).
18. Voigt, C.A. Genetic parts to program bacteria. *Curr Opin Biotechnol* **17**, 548-557 (2006).
19. Liebermeister, W. Flux cost functions and the choice of metabolic fluxes. *arXiv preprint arXiv:1801.05742* (2018).
20. Venayak, N., Anesiadis, N., Cluett, W.R. & Mahadevan, R. Engineering metabolism through dynamic control. *Curr Opin Biotechnol* **34**, 142-152 (2015).

21. Sander, T., Wang, C.Y., Glatter, T. & Link, H. CRISPRi-Based Downregulation of Transcriptional Feedback Improves Growth and Metabolism of Arginine Overproducing E. coli. *ACS synthetic biology* **8**, 1983-1990 (2019).
22. Ginesy, M., Belotserkovsky, J., Enman, J., Isaksson, L. & Rova, U. Metabolic engineering of *Escherichia coli* for enhanced arginine biosynthesis. *Microb Cell Fact* **14**, 29 (2015).
23. Liu, D., Mannan, A.A., Han, Y., Oyarzun, D.A. & Zhang, F. Dynamic metabolic control: towards precision engineering of metabolism. *J Ind Microbiol Biotechnol* **45**, 535-543 (2018).
24. Ramseier, T.M. Cra and the control of carbon flux via metabolic pathways. *Research in Microbiology* **147**, 489-493 (1996).
25. Nakano, M., Ogasawara, H., Shimada, T., Yamamoto, K. & Ishihama, A. Involvement of cAMP-CRP in transcription activation and repression of the pck gene encoding PEP carboxykinase, the key enzyme of gluconeogenesis. *FEMS Microbiol Lett* **355**, 93-99 (2014).
26. Pisithkul, T., Patel, N.M. & Amador-Noguez, D. Post-translational modifications as key regulators of bacterial metabolic fluxes. *Curr Opin Microbiol* **24**, 29-37 (2015).
27. Deutscher, J., Francke, C. & Postma, P.W. How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiol Mol Biol Rev* **70**, 939-1031 (2006).
28. Gupta, A., Reizman, I.M., Reisch, C.R. & Prather, K.L. Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. *Nat Biotechnol* **35**, 273-279 (2017).
29. Shi, S., Ang, E.L. & Zhao, H. In vivo biosensors: mechanisms, development, and applications. *J Ind Microbiol Biotechnol* **45**, 491-516 (2018).
30. Richter, F. et al. Engineering of temperature- and light-switchable Cas9 variants. *Nucleic Acids Res* (2016).
31. Repoila, F. & Gottesman, S. Temperature Sensing by the dsrA Promoter. *Journal of Bacteriology* **185**, 6609-6614 (2003).
32. Maury, J. et al. Glucose-dependent promoters for dynamic regulation of metabolic pathways. *Front Bioeng Biotechnol* **6**, 63 (2018).
33. Ceroni, F. et al. Burden-driven feedback control of gene expression. *Nat Methods* (2018).
34. Rogers, J.K. & Church, G.M. Genetically encoded sensors enable real-time observation of metabolite production. *Proceedings of the National Academy of Sciences* **113**, 2388-2393 (2016).
35. Siedler, S. et al. SoxR as a single-cell biosensor for NADPH-consuming enzymes in *Escherichia coli*. *ACS Synth Biol* **3**, 41-47 (2014).
36. Moser, F. et al. Dynamic control of endogenous metabolism with combinatorial logic circuits. *Mol Syst Biol* **14**, e8605 (2018).
37. Meyer, A.J., Segall-Shapiro, T.H. & Voigt, C.A. Marionette: *E. coli* containing 12 highly-optimized small molecule sensors (2018).
38. Shin, J., Zhang, S., Der, B.S., Nielsen, A.A. & Voigt, C.A. Programming *Escherichia coli* to function as a digital display. *Mol Syst Biol* **16**, e9401 (2020).
39. De Paepe, B., Peters, G., Coussement, P., Maertens, J. & De Mey, M. Tailor-made transcriptional biosensors for optimizing microbial cell factories. *J Ind Microbiol Biotechnol* (2016).
40. Lalwani, M.A., Zhao, E.M. & Avalos, J.L. Current and future modalities of dynamic control in metabolic engineering. *Curr Opin Biotechnol* **52**, 56-65 (2018).

41. Zhang, F., Carothers, J.M. & Keasling, J.D. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nat Biotechnol* **30**, 354-359 (2012).
42. Becker, J. & Wittmann, C. Systems metabolic engineering of *Escherichia coli* for the heterologous production of high value molecules-a veteran at new shores. *Curr Opin Biotechnol* **42**, 178-188 (2016).
43. Lin, G.-M., Warden-Rothman, R. & Voigt, C.A. Retrosynthetic design of metabolic pathways to chemicals not found in nature. *Current Opinion in Systems Biology* (2019).
44. Baumgärtner, F., Seitz, L., Sprenger, G.A. & Albermann, C. Construction of *Escherichia coli* strains with chromosomally integrated expression cassettes for the synthesis of 2´-fucosyllactose. *Microbial Cell Factories* **12**, 1-13 (2013).
45. van Heerden, J.H., Bruggeman, F.J. & Teusink, B. Multi-tasking of biosynthetic and energetic functions of glycolysis explained by supply and demand logic. *Bioessays* **37**, 34-45 (2015).
46. Ye, J. & Medzhitov, R. Control strategies in systemic metabolism. *Nature Metabolism* **1**, 947-957 (2019).
47. Li, M. et al. Recent advances of metabolic engineering strategies in natural isoprenoid production using cell factories. *Nat Prod Rep* (2019).
48. Lee, S.Y. et al. A comprehensive metabolic map for production of bio-based chemicals. *Nature Catalysis* **2**, 18-33 (2019).
49. Lee, J.W. et al. Systems metabolic engineering of microorganisms for natural and non-natural chemicals. *Nat Chem Biol* **8**, 536-546 (2012).
50. Yang, D., Park, S.Y., Park, Y.S., Eun, H. & Lee, S.Y. Metabolic engineering of *Escherichia coli* for natural product biosynthesis. *Trends Biotechnol* **38**, 745-765 (2020).
51. Kochanowski, K. et al. Functioning of a metabolic flux sensor in *Escherichia coli*. *Proceedings of the National Academy of Sciences* **110**, 1130-1135 (2013).
52. Lehning, C.E., Siedler, S., Ellabaan, M.M.H. & Sommer, M.O.A. Assessing glycolytic flux alterations resulting from genetic perturbations in *E. coli* using a biosensor. *Metab Eng* (2017).
53. Ramseier, T.M., Bledig, S., Michotey, V., Feghali, R. & Jr Saier, M.H. The global regulatory protein FruR modulates the direction of carbon flow in *Escherichia coli*. *Mol. Microbiol.* **16**, 1157-1169 (1995).
54. Chavarria, M. & de Lorenzo, V. The imbroglio of the physiological Cra effector clarified at last. *Mol Microbiol* (2018).
55. Singh, D. et al. Protein-protein interactions with fructose-1-kinase alter function of the central *Escherichia coli* transcription regulator, Cra. *bioRxiv*, 201277 (2017).
56. Okano, H., Hermsen, R., Kochanowski, K. & Hwa, T. Regulation underlying hierarchical and simultaneous utilization of carbon substrates by flux sensors in Escherichia coli. *Nat Microbiol* **5**, 206-215 (2020).
57. Shimizu, K. & Matsuoka, Y. Regulation of glycolytic flux and overflow metabolism depending on the source of energy generation for energy demand. *Biotechnol Adv* (2018).
58. Kochanowski, K. et al. Few regulatory metabolites coordinate expression of central metabolic genes in *Escherichia coli*. *Mol Syst Biol* **13**, 903 (2017).
59. Peeters, K. et al. Fructose-1,6-bisphosphate couples glycolytic flux to activation of Ras. *Nat Commun* **8**, 922 (2017).

60. Kim, D. et al. Systems assessment of transcriptional regulation on central carbon metabolism by Cra and CRP. *Nucleic Acids Res* (2018).
61. Saier Jr, M.H. & Ramseier, T.M. The catabolite repressor/activator (Cra) protein of enteric bacteria. *Journal of bacteriology* **178**, 3411 (1996).
62. Bley Folly, B. et al. Assessment of the interaction between the flux-signaling metabolite fructose-1,6-bisphosphate and the bacterial transcription factors CggR and Cra. *Mol Microbiol* (2018).
63. Larson, M.H. et al. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat Protoc* **8**, 2180-2196 (2013).
64. Deltcheva, E. et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471**, 602-607 (2011).
65. Karvelis, T. et al. crRNA and tracrRNA guide Cas9-mediated DNA interference in *Streptococcus thermophilus*. *RNA Biol* **10**, 841-851 (2013).
66. Dominguez, A.A., Lim, W.A. & Qi, L.S. Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nat Rev Mol Cell Biol* **17**, 5-15 (2016).
67. Zalatan, J.G. et al. Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell* **160**, 339-350 (2015).
68. Qi, L.S. et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173-1183 (2013).
69. Jakociunas, T., Jensen, M.K. & Keasling, J.D. CRISPR/Cas9 advances engineering of microbial cell factories. *Metab Eng* **34**, 44-59 (2015).
70. La Russa, M.F. & Qi, L.S. The new state of the art: Cas9 for gene activation and repression. *Mol Cell Biol* **35**, 3800-3809 (2015).
71. Fontana, J. et al. Effective CRISPRa-mediated control of gene expression in bacteria must overcome strict target site requirements. (2019).
72. Fontana, J. et al. Effective CRISPRa-mediated control of gene expression in bacteria must overcome strict target site requirements. *Nat Commun* **11**, 1618 (2020).
73. Wang, H.H. et al. Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**, 894-898 (2009).
74. Beuter, D. et al. Selective enrichment of slow-growing bacteria in a metabolism-wide CRISPRi library with a TIMER protein. *ACS Synth Biol* **7**, 2775-2782 (2018).
75. Shi, H., Colavin, A., Lee, T.K. & Huang, K.C. Strain library imaging protocol for high-throughput, automated single-cell microscopy of large bacterial collections arrayed on multiwell plates. *Nat Protoc* **12**, 429-438 (2017).
76. Bassalo, M.C. et al. Rapid and efficient one-step metabolic pathway integration in *E. coli*. *ACS Synth Biol* (2016).
77. Reisch, C.R. & Prather, K.L. The no-SCAR (Scarless Cas9 Assisted Recombineering) system for genome editing in *Escherichia coli*. *Sci Rep* **5**, 15096 (2015).
78. Sharan, S.K., Thomason, L.C., Kuznetsov, S.G. & Court, D.L. Recombineering: a homologous recombination-based method of genetic engineering. *Nat Protoc* **4**, 206-223 (2009).
79. Datsenko, K.A. & Wanner, B.L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences* **97**, 6640-6645 (2000).
80. Li, X.T., Thomason, L.C., Sawitzke, J.A., Costantino, N. & Court, D.L. Positive and negative selection using the tetA-sacB cassette: recombineering and P1 transduction in *Escherichia coli*. *Nucleic Acids Res* **41**, e204 (2013).

81. Salles, I.M., Forchhammer, N., Croux, C., Girbal, L. & Soucaille, P. Evolution of a *Saccharomyces cerevisiae* metabolic pathway in *Escherichia coli*. *Metabolic engineering* **9**, 152-159 (2007).
82. Hermsen, R., Okano, H., You, C., Werner, N. & Hwa, T. A growth-rate composition formula for the growth of *E. coli* on co-utilized carbon substrates. Molecular Systems Biology **11** (2015).
83. Baba, T. et al. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2**, 2006 0008 (2006).
84. Grenier, F., Matteau, D., Baby, V. & Rodrigue, S. Complete genome sequence of *Escherichia coli* BW25113. *Genome Announc* **2** (2014).
85. Sastry, A.V. et al. The *Escherichia coli* transcriptome mostly consists of independently regulated modules. *Nat Commun* **10**, 5536 (2019).
86. Rugbjerg, P., Sarup-Lytzen, K., Nagy, M. & Sommer, M.O.A. Synthetic addiction extends the productive life time of engineered Escherichia coli populations. *Proc Natl Acad Sci U S A* **115**, 2347-2352 (2018).
87. Rosano, G.L. & Ceccarelli, E.A. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front Microbiol* **5**, 172 (2014).
88. Keren, L. et al. Promoters maintain their relative activity levels under different growth conditions. *Mol Syst Biol* **9**, 701 (2013).
89. Rinschen, M.M., Ivanisevic, J., Giera, M. & Siuzdak, G. Identification of bioactive metabolites using activity metabolomics. *Nat Rev Mol Cell Biol* **20**, 353-367 (2019).
90. Piazza, I. et al. A map of protein-metabolite interactions reveals principles of chemical communication. *Cell* **172**, 358-372 e323 (2018).
91. Monteiro, F. et al. Measuring glycolytic flux in single yeast cells with an orthogonal synthetic biosensor. *bioRxiv*, 682302 (2019).
92. Kotte, O., Volkmer, B., Radzikowski, J.L. & Heinemann, M. Phenotypic bistability in *Escherichia coli*'s central carbon metabolism. *Mol Syst Biol* **10**, 736 (2014).
93. Gerosa, L. et al. Pseudo-transition analysis identifies the key regulators of dynamic metabolic adaptations from steady-state data. *Cell Syst* **1**, 270-282 (2015).
94. Gerosa, L., Kochanowski, K., Heinemann, M. & Sauer, U. Dissecting specific and global transcriptional regulation of bacterial gene expression. *Mol Syst Biol* **9**, 658 (2013).
95. Ow, D.S. et al. Inactivating FruR global regulator in plasmid-bearing *Escherichia coli* alters metabolic gene expression and improves growth rate. *J Biotechnol* **131**, 261-269 (2007).
96. Shimada, T., Fujita, N., Yamamoto, K. & Ishihama, A. Novel roles of cAMP receptor protein (CRP) in regulation of transport and metabolism of carbon sources. *PLoS One* **6**, e20081 (2011).
97. Plumbridge, J. Regulation of gene expression in the PTS in *Escherichia coli*: the role and interactions of Mlc. *Current Opinion in Microbiology* **5**, 187-193 (2002).
98. Kang, Y., Weber, K.D., Qiu, Y., Kiley, P.J. & Blattner, F.R. Genome-wide expression analysis indicates that FNR of *Escherichia coli* K-12 regulates a large number of genes of unknown function. *J Bacteriol* **187**, 1135-1160 (2005).
99. Bradley, M.D., Beach, M.B., de Koning, A.P., Pratt, T.S. & Osuna, R. Effects of Fis on *Escherichia coli* gene expression during different growth stages. *Microbiology* **153**, 2922-2940 (2007).
100. Fillat, M.F. The FUR (ferric uptake regulator) superfamily: diversity and versatility of key transcriptional regulators. *Arch Biochem Biophys* **546**, 41-52 (2014).

101. Zhang, C., Meng, X., Gu, H., Ma, Z. & Lu, L. Predicted glycerol 3-phosphate dehydrogenase homologs and the glycerol kinase GlcA coordinately adapt to various carbon sources and osmotic stress in *Aspergillus fumigatus*. *G3 (Bethesda)* **8**, 2291-2299 (2018).
102. Szenk, M., Dill, K.A. & de Graff, A.M.R. Why Do Fast-Growing Bacteria Enter Overflow Metabolism? Testing the Membrane Real Estate Hypothesis. *Cell Syst* **5**, 95-104 (2017).
103. Basan, M. et al. Overflow metabolism in *Escherichia coli* results from efficient proteome allocation. *Nature* **528**, 99-104 (2015).
104. Zampieri, M., Horl, M., Hotz, F., Muller, N.F. & Sauer, U. Regulatory mechanisms underlying coordination of amino acid and glucose catabolism in *Escherichia coli*. *Nat Commun* **10**, 3354 (2019).
105. Rahman, M. & Shimizu, K. Altered acetate metabolism and biomass production in several *Escherichia coli* mutants lacking rpoS-dependent metabolic pathway genes. *Mol Biosyst* **4**, 160-169 (2008).
106. Battesti, A., Majdalani, N. & Gottesman, S. The RpoS-mediated general stress response in *Escherichia coli*. *Annu Rev Microbiol* **65**, 189-213 (2011).
107. Dubey, A.K. et al. CsrA regulates translation of the *Escherichia coli* carbon starvation gene, cstA, by blocking ribosome access to the cstA transcript. *Journal of Bacteriology* **185**, 4450-4460 (2003).
108. Lee, M.E., Aswani, A., Han, A.S., Tomlin, C.J. & Dueber, J.E. Expression-level optimization of a multi-enzyme pathway in the absence of a high-throughput assay. *Nucleic Acids Res* **41**, 10668-10678 (2013).
109. Zelcbuch, L. et al. Spanning high-dimensional expression space using ribosome-binding site combinatorics. *Nucleic Acids Res* **41**, e98 (2013).
110. Levin-Karp, A. et al. Quantifying translational coupling in *E. coli* synthetic operons using RBS modulation and fluorescent reporters. *ACS Synth Biol* **2**, 327-336 (2013).
111. Tripathi, N.K. & Shrivastava, A. Recent developments in bioprocessing of recombinant proteins: expression hosts and process development. *Front Bioeng Biotechnol* **7**, 420 (2019).
112. Rohe, P., Venkanna, D., Kleine, B., Freudl, R. & Oldiges, M. An automated workflow for enhancing microbial bioprocess optimization on a novel microbioreactor platform. *Microbial cell factories* **11**, 144-144 (2012).
113. Gasmi, N. et al. Development of a cultivation process for the enhancement of human interferon alpha 2b production in the oleaginous yeast, *Yarrowia lipolytica*. *Microbial cell factories* **10**, 90-90 (2011).
114. Yang, J. & Guo, L. Biosynthesis of β-carotene in engineered *E. coli* using the MEP and MVA pathways. *Microbial cell factories* **13**, 160 (2014).
115. Albrecht, M., Misawa, N. & Sandmann, G. Metabolic engineering of the terpenoid biosynthetic pathway of *Escherichia coli* for production of the carotenoids β-carotene and zeaxanthin. *Biotechnology letters* **21**, 791-795 (1999).
116. Harker, M. & Bramley, P. Expression of prokaryotic 1-deoxy-d-xylulose-5-phosphatases in *Escherichia coli* increases carotenoid and ubiquinone biosynthesis. *FEBS letters* **448**, 115-119 (1999).
117. Studier, F.W. Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification* **41**, 207-234 (2005).

118. Wessels, J.M., Ford, W.E., Szymczak, W. & Schneider, S. The complexation of tetracycline and anhydrotetracycline with Mg2+ and Ca2+: a spectroscopic study. *The Journal of Physical Chemistry B* **102**, 9323-9331 (1998).
119. Ammar, E.M., Wang, X. & Rao, C.V. Regulation of metabolism in *Escherichia coli* during growth on mixtures of the non-glucose sugars: arabinose, lactose, and xylose. *Sci Rep* **8**, 609 (2018).
120. Vinuselvi, P., Kim, M.K., Lee, S.K. & Ghim, C.M. Rewiring carbon catabolite repression for microbial cell factory. *BMB Rep* **45**, 59-70 (2012).
121. Fabich, A.J. et al. Comparison of carbon nutrition for pathogenic and commensal *Escherichia coli* strains in the mouse intestine. *Infect Immun* **76**, 1143-1152 (2008).
122. Sarkar, D. & Shimizu, K. Effect of cra gene knockout together with other genes knockouts on the improvement of substrate consumption rate in *Escherichia coli* under microaerobic condition. *Biochemical Engineering Journal* **42**, 224-228 (2008).
123. Yao, R., Kurata, H. & Shimizu, K. Effect of cra gene mutation on the metabolism of *Escherichia coli* for a mixture of multiple carbon sources. (2013).
124. Son, Y.-J., Phue, J.-N., Trinh, L.B., Lee, S. & Shiloach, J. The role of Cra in regulating acetate excretion and osmotic tolerance in *E. coli* K-12 and *E. coli* B at high density growth. *Microbial Cell Factories* **10**, 52 (2011).
125. Zhu, L.W. & Tang, Y.J. Current advances of succinate biosynthesis in metabolically engineered *Escherichia coli*. *Biotechnol Adv* **35**, 1040-1048 (2017).
126. Zhu, L.W. et al. Collaborative regulation of CO_2 transport and fixation during succinate production in *Escherichia coli*. *Sci Rep* **5**, 17321 (2015).
127. Pontrelli, S. et al. Metabolic repair through emergence of new pathways in *Escherichia coli*. *Nat Chem Biol* **14**, 1005-1009 (2018).
128. Gleizer, S. et al. Conversion of *Escherichia coli* to generate all biomass carbon from CO_2. *Cell* **179**, 1255-1263 e1212 (2019).
129. Sun, Z.Z., Yeung, E., Hayes, C.A., Noireaux, V. & Murray, R.M. Linear DNA for rapid prototyping of synthetic biological circuits in an *Escherichia coli* based TX-TL cell-free system. *ACS Synth Biol* **3**, 387-397 (2014).
130. Zucca, S., Pasotti, L., Politi, N., Cusella De Angelis, M.G. & Magni, P. A standard vector for the chromosomal integration and characterization of BioBrick™ parts in *Escherichia coli*. *Journal of Biological Engineering* **7**, 12 (2013).
131. Kelwick, R. et al. A forward-design approach to increase the production of poly-3-hydroxybutyrate in genetically engineered *Escherichia coli*. *PLoS One* **10**, e0117202 (2015).
132. Qi, L., Haurwitz, R.E., Shao, W., Doudna, J.A. & Arkin, A.P. RNA processing enables predictable programming of gene expression. *Nat Biotechnol* **30**, 1002-1006 (2012).
133. Chappell, J., Jensen, K. & Freemont, P.S. Validation of an entirely in vitro approach for rapid prototyping of DNA regulatory elements for synthetic biology. *Nucleic Acids Res* **41**, 3471-3481 (2013).
134. Lewis, D.D., Villarreal, F.D., Wu, F. & Tan, C. Synthetic biology outside the cell: linking computational tools to cell-free systems. *Front Bioeng Biotechnol* **2**, 66 (2014).
135. Jones, K.L., Kim, S.-W. & Keasling, J. Low-copy plasmids can perform as well as or better than high-copy plasmids for metabolic engineering of bacteria. *Metabolic engineering* **2**, 328-338 (2000).

136. Ryall, B., Eydallin, G. & Ferenci, T. Culture history and population heterogeneity as determinants of bacterial adaptation: the adaptomics of a single environmental transition. *Microbiol Mol Biol Rev* **76**, 597-625 (2012).
137. Ferenci, T. & Maharjan, R. Mutational heterogeneity: a key ingredient of bet-hedging and evolutionary divergence?: The broad spectrum of mutations and their flexible frequency in populations provides a source of risk avoidance and alternative evolutionary strategies. *Bioessays* **37**, 123-130 (2015).
138. Mahillon, J. & Chandler, M. Insertion sequences. *Microbiology and molecular biology reviews* **62**, 725-774 (1998).
139. Schleif, R. Regulation of the l-arabinose operon of *Escherichia coli*. *Trends in Genetics* **16**, 559-565 (2000).
140. Schleif, R. AraC protein, regulation of the l-arabinose operon in *Escherichia coli*, and the light switch mechanism of AraC action. *FEMS Microbiology Reviews* **34**, 779-796 (2010).
141. Rogers, J.K. et al. Synthetic biosensors for precise gene control and real-time monitoring of metabolites. *Nucleic Acids Res* **43**, 7648-7660 (2015).
142. Kafri, M., Metzl-Raz, E., Jona, G. & Barkai, N. The cost of protein production. *Cell Rep* **14**, 22-31 (2016).
143. Lee, T.S. et al. BglBrick vectors and datasheets: A synthetic biology platform for gene expression. *J Biol Eng* **5**, 12 (2011).
144. Zucca, S., Pasotti, L., Mazzini, G., Cusella De Angelis, M.G. & Magni, P. Characterization of an inducible promoter in different DNA copy number conditions. *BMC Bioinformatics* **13**, S11 (2012).
145. Tegel, H., Ottosson, J. & Hober, S. Enhancing the protein production levels in *Escherichia coli* with a strong promoter. *FEBS J* **278**, 729-739 (2011).
146. Cheng, J.K. & Alper, H.S. Transcriptomics-guided design of synthetic promoters for a mammalian system. *ACS Synth Biol* **5**, 1455-1465 (2016).
147. Schlabach, M.R., Hu, J.K., Li, M. & Elledge, S.J. Synthetic design of strong promoters. *Proc Natl Acad Sci U S A* **107**, 2538-2543 (2010).
148. Jeschek, M., Gerngross, D. & Panke, S. Rationally reduced libraries for combinatorial pathway optimization minimizing experimental effort. *Nat Commun* **7** (2016).
149. Mori, M., Marinari, E. & De Martino, A. A yield-cost tradeoff governs *Escherichia coli*'s decision between fermentation and respiration in carbon-limited growth. *NPJ Syst Biol Appl* **5**, 16 (2019).
150. Paul, L., Mishra, P.K., Blumenthal, R.M. & Matthews, R.G. Integration of regulatory signals through involvement of multiple global regulators: control of the *Escherichia coli* gltBDF operon by Lrp, IHF, Crp, and ArgR. *BMC Microbiol* **7**, 2 (2007).
151. Sander, T. et al. Allosteric feedback inhibition enables robust amino acid biosynthesis in *E. coli* by enforcing enzyme overabundance. *Cell Systems* (2019).
152. Didovyk, A., Borek, B., Hasty, J. & Tsimring, L. Orthogonal modular gene repression in *Escherichia coli* using engineered CRISPR/Cas9. *ACS Synth Biol* **5**, 81-88 (2016).
153. Westbrook, A.W., Moo-Young, M. & Chou, C.P. Development of a CRISPR-Cas9 toolkit for comprehensive engineering of *Bacillus subtilis*. *Appl Environ Microbiol* (2016).
154. Camsund, D. et al. Time-resolved imaging-based CRISPRi screening. *Nat Methods* **17**, 86-92 (2020).

155. Kim, D. et al. Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nat Biotechnol* **34**, 863-868 (2016).
156. McCarty, N.S., Graham, A.E., Studena, L. & Ledesma-Amaro, R. Multiplexed CRISPR technologies for gene editing and transcriptional regulation. *Nat Commun* **11**, 1281 (2020).
157. Bin Moon, S. et al. Highly efficient genome editing by CRISPR-Cpf1 using CRISPR RNA with a uridinylate-rich 3'-overhang. *Nat Commun* **9**, 3651 (2018).
158. Liu, R.M. et al. Synthetic chimeric nucleases function for efficient genome editing. *Nat Commun* **10**, 5524 (2019).
159. Gill, R.T., Garst, A. & Lipscomb, T.E.W. (Google Patents, 2018).
160. Schmid-Burgk, J.L. et al. Highly parallel profiling of Cas9 variant specificity. *Mol Cell* **78**, 794-800 e798 (2020).
161. Hanna, R.E. & Doench, J.G. Design and analysis of CRISPR–Cas experiments. *Nat. Biotechnol.* (2020).
162. Legut, M. et al. High-throughput screens of PAM-flexible Cas9 variants for gene knockout and transcriptional modulation. *Cell Rep* **30**, 2859-2868 e2855 (2020).
163. Tanner, L.B. et al. Four key steps control glycolytic fux in mammalian cells. *Cell Syst* **7**, 49-62 e48 (2018).
164. Monteiro, F. et al. Measuring glycolytic flux in single yeast cells with an orthogonal synthetic biosensor. *Mol Syst Biol* **15**, e9071 (2019).
165. Zhang, L. et al. The catabolite repressor/activator, Cra, bridges a connection between carbon metabolism and host colonization in the plant drought resistance-promoting bacterium *Pantoea alhagi* LTYR-11Z. *Appl Environ Microbiol* (2018).
166. Rutledge, A.C. & Adeli, K. Fructose and the metabolic syndrome: pathophysiology and molecular mechanisms. *Nutrition Reviews* **65**, 13-23 (2007).
167. Zhao, S. et al. Dietary fructose feeds hepatic lipogenesis via microbiota-derived acetate. *Nature* **579**, 586-591 (2020).
168. Zhang, L., Feng, G. & Declerck, S. Signal beyond nutrient, fructose, exuded by an arbuscular mycorrhizal fungus triggers phytate mineralization by a phosphate solubilizing bacterium. *ISME J* **12**, 2339-2351 (2018).
169. Bernal, V., Castano-Cerezo, S. & Canovas, M. Acetate metabolism regulation in *Escherichia coli:* carbon overflow, pathogenicity, and beyond. *Appl Microbiol Biotechnol* **100**, 8985-9001 (2016).
170. Sasnow, S.S., Wei, H. & Aristilde, L. Bypasses in intracellular glucose metabolism in iron-limited *Pseudomonas putida*. *Microbiologyopen* **5**, 3-20 (2016).
171. Quan, J. & Tian, J. Circular polymerase extension cloning for high-throughput cloning of complex and combinatorial DNA libraries. *Nat Protoc* **6**, 242-251 (2011).
172. Craft, N.E. & Soares, J.H. Relative solubility, stability, and absorptivity of lutein and. beta.-carotene in organic solvents. *Journal of Agricultural and Food Chemistry* **40**, 431-434 (1992).
173. Sielaff, M. et al. Evaluation of FASP, SP3, and iST protocols for proteomic sample preparation in the low microgram range. *J Proteome Res* **16**, 4060-4072 (2017).
174. Moggridge, S., Sorensen, P.H., Morin, G.B. & Hughes, C.S. Extending the compatibility of the SP3 paramagnetic bead processing approach for proteomics. *J Proteome Res* **17**, 1730-1740 (2018).
175. Consortium, G.O. The gene ontology resource: 20 years and still GOing strong. *Nucleic acids research* **47**, D330-D338 (2019).
176. Lin, M.T. et al. Novel utilization of terminators in the design of biologically adjustable synthetic filters. *ACS Synth Biol* **5**, 365-374 (2016).

177. Ceroni, F., Algar, R., Stan, G.B. & Ellis, T. Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat Methods* **12**, 415-418 (2015).
178. Ishihama, A., Shimada, T. & Yamazaki, Y. Transcription profile of *Escherichia coli*: genomic SELEX search for regulatory targets of transcription factors. *Nucleic Acids Res* **44**, 2058-2074 (2016).
179. Santos-Zavaleta, A. et al. RegulonDB v 10.5: tackling challenges to unify classic and high throughput knowledge of gene regulation in *E. coli* K-12. *Nucleic Acids Res* **47**, D212-D220 (2019).
180. Myers, K.S. et al. Genome-scale analysis of *Escherichia Coli* FNR reveals complex features of transcription factor binding. *PLoS Genet* **9**, e1003565 (2013).
181. Santos-Zavaleta, A. et al. A unified resource for transcriptional regulation in *Escherichia coli* K-12 incorporating high-throughput-generated binding data into RegulonDB version 10.0. *BMC Biol* **16**, 91 (2018).
182. Shimada, T., Yamamoto, K. & Ishihama, A. Novel members of the Cra regulon involved in carbon metabolism in *Escherichia coli*. *J Bacteriol* **193**, 649-659 (2011).

V. Statement of authorship (Eigenständigkeitserklärung)

Ich versichere, dass ich meine Dissertation:

”Transcriptional regulation of *Escherichia coli* metabolism and engineered metabolic pathways”

selbstständig und ohne unerlaubte Hilfe angefertigt und ich mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen und Hilfsmittel bedient habe.

Diese Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

I hereby declare that I am the sole author of this thesis

”Transcriptional regulation of *Escherichia coli* metabolism and engineered metabolic pathways”

and that I have not used any sources other than those listed in the references. I further declare that I have not submitted this thesis for the award of any other degree or diploma in any other tertiary institution.

_____	_____
Ort, Datum	Chun-Ying Wang