






# Titration bacterial growth and chemical biosynthesis for efficient *N*-acetylglucosamine and *N*-acetylneuraminic acid bioproduction

Rongzhen Tian <sup>1,2</sup>, Yanfeng Liu <sup>1,2</sup>✉, Yanting Cao<sup>1,2</sup>, Zhongjie Zhang<sup>1,2</sup>, Jianghua Li<sup>1,2</sup>, Long Liu <sup>1,2</sup>, Guocheng Du <sup>1,2</sup> & Jian Chen <sup>2,3</sup>✉

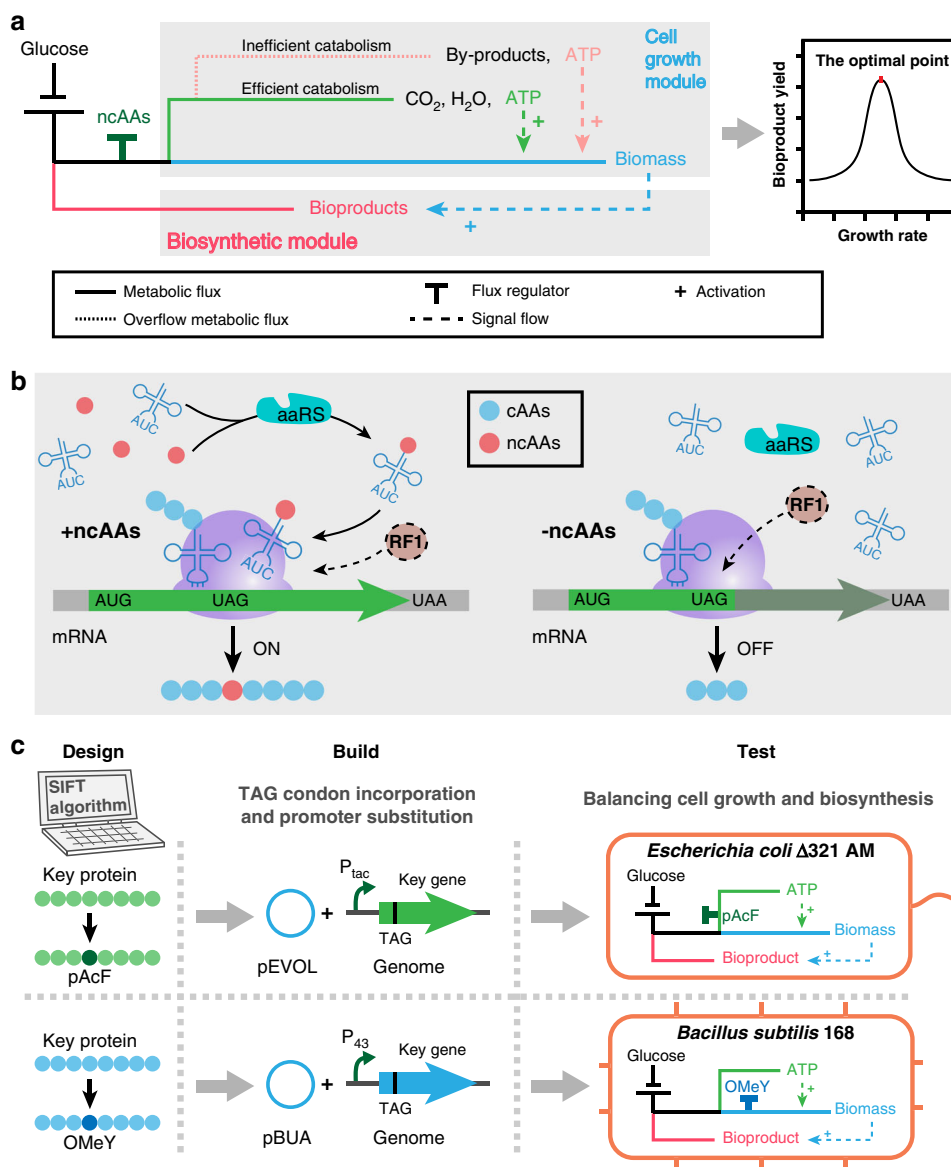
Metabolic engineering facilitates chemical biosynthesis by rewiring cellular resources to produce target compounds. However, an imbalance between cell growth and bioproduction often reduces production efficiency. Genetic code expansion (GCE)-based orthogonal translation systems incorporating non-canonical amino acids (ncAAs) into proteins by reassigning non-canonical codons to ncAAs qualify for balancing cellular metabolism. Here, GCE-based cell growth and biosynthesis balance engineering (GCE-CGBBE) is developed, which is based on titrating expression of cell growth and metabolic flux determinant genes by constructing ncAA-dependent expression patterns. We demonstrate GCE-CGBBE in genome-recoded *Escherichia coli* Δ321AM by precisely balancing glycolysis and *N*-acetylglucosamine production, resulting in a 4.54-fold increase in titer. GCE-CGBBE is further expanded to non-genome-recoded *Bacillus subtilis* to balance growth and *N*-acetylneuraminic acid bioproduction by titrating essential gene expression, yielding a 2.34-fold increase in titer. Moreover, the development of ncAA-dependent essential gene expression regulation shows efficient biocontainment of engineered *B. subtilis* to avoid unintended proliferation in nature.

<sup>1</sup>Key Laboratory of Carbohydrate Chemistry and Biotechnology, Ministry of Education, Jiangnan University, Wuxi 214122, China. <sup>2</sup>Science Center for Future Foods, Jiangnan University, Wuxi 214122, China. <sup>3</sup>National Engineering Laboratory for Cereal Fermentation Technology, Jiangnan University, Wuxi 214122, China. ✉email: [yanfengliu@jiangnan.edu.cn](mailto:yanfengliu@jiangnan.edu.cn); [jchen@jiangnan.edu.cn](mailto:jchen@jiangnan.edu.cn)

Metabolic engineering is an enabling technology for developing sustainable biomanufacturing processes for chemicals, fuels, and nutraceuticals using renewable feedstock by repurposing cellular metabolism<sup>1–4</sup>. High-level bio-production of 1,3-propanediol and 1,4-butanediol demonstrates the success of bio-based chemical production<sup>5–7</sup>. However, metabolic fluxes in target chemical biosynthetic pathways need to be shifted from cellular endogenous metabolism for cell growth to bioproduction, which conflicts with a cell's need for ensuring survival<sup>3</sup>. When redirection of target metabolic pathways is insufficient, abundant substrate is mainly used for cell growth, which is often

accompanied by by-product synthesis, resulting in low production yield and titer. In contrast, excessive shunting affects cell growth, and subsequently reduces productivity (Fig. 1a)<sup>8</sup>. Therefore, balancing cell growth with biosynthesis is one of the core issues in constructing highly efficient metabolically engineered strains<sup>9,10</sup>.

Efforts have been made to solve these issues through various metabolic engineering strategies, such as de-coupling cell growth with biosynthesis based on genetic circuits, co-coupling cell growth with biosynthesis by coupling production to overall energy and biomass formation, and parallel metabolic pathway engineering via separating bioproduction from cell growth by implementing parallel



**Fig. 1** Balancing cell growth and product synthesis based on genetic code expansion. **a** Metabolic flux distribution during cell growth and biosynthesis. Theoretically, the balance between cell growth and product synthesis can be achieved by regulating the distribution of metabolic fluxes to obtain optimal yield. **b** Utilization of the non-canonical amino acid (ncAA) incorporation tool to build a molecular switch. Translation can be interrupted when an amber stop codon (TAG) is inserted into the gene sequence of interest (green arrow) in the absence of ncAAs. However, after ncAAs are added, aminoacyl-tRNA synthetase (aaRS) can catalyze ncAA-loading of tRNAs (blue clover-leaf structure), which are able to restore gene expression. While peptide chain release factor RF1 was depleted in genetically recoded *E. coli*  $\Delta 321$ AM, competition between RF1 and aminoacyl-tRNA molecules may have reduced the efficiency of the ncAA system in non-genome-recoded *B. subtilis* 168 (the dotted circle). **c** Workflow of the GCE-CGBBE strategy in *E. coli* and *B. subtilis*. First, amber stop codon (TAG) substitution sites of key genes that control cell growth were designed based on the SIFT algorithm. Next, strains harboring ncAA-incorporation tools were constructed to achieve regulatory expression of genes with amber stop codon insertion. Finally, cell growth and product balance were tested in two of the best-studied Gram-negative and Gram-positive bacteria, i.e., genetically recoded *E. coli*  $\Delta 321$ AM and non-genome-recoded *B. subtilis* 168. cAAs, canonical amino acids.

dual carbon sources pathways<sup>11–13</sup>. De-coupling cell growth with biosynthesis commonly uses toggles, autonomous and quorum-sensing system-based genetic circuits, or optogenetic circuits to shift growth and production phases<sup>11,14–17</sup>. However, a potential interference of intracellular metabolites or the culture environment with genetic circuits and the requirement of large-scale fermentation equipment redesign for optogenetic circuit implementation impact either orthogonality or large-scale applicability<sup>18,19</sup>. Moreover, incorporating synthetic regulatory circuits into native pathways requires extensive genome editing, the designing and building of which is challenging<sup>20</sup>. Co-coupling cell growth with biosynthesis is suitable for obtaining sufficient titer, but difficult to use for improving yield, due to difficulties in regulating cell growth, especially for biosensor-based cell growth regulation that often shows atypical physiological properties<sup>21,22</sup>. Although parallel metabolic pathway engineering improves production yields, the lack of precise cell growth regulation causes overflowing metabolism, resulting in by-product (acetate) formation<sup>13</sup>. Moreover, parallel metabolic pathway engineering could be limited by precursor availability depending on multiple endogenous pathways, owing to an increased complexity of pathway designs, and the metabolic burden of introducing multiple substrate utilization pathways<sup>23,24</sup>. Designing a generic and robust cell growth and biosynthesis balancing strategy is challenging.

Metabolic connections between cell growth and target product biosynthesis increase the complexity of balancing cellular metabolic pathways. Therefore, an orthogonal gene expression system should regulate target metabolic pathways to avoid interference by other pathways. For example, the genetic code expansion (GCE)-based orthogonal protein translation system introduces orthogonal aminoacyl-tRNA synthetase (aaRS)/tRNA pairs into a host to incorporate non-canonical amino acids (ncAAs) rather than the 20 proteinogenic amino acids into proteins by reassigning non-canonical codons to ncAAs<sup>25</sup>. More than 200 ncAAs can be site-specifically incorporated into proteins, thereby expanding the genetic code of microorganisms, which lays the foundation for expanding functional diversity of proteins and the construction of orthogonal protein translation systems (Fig. 1b)<sup>25–29</sup>. The amber stop codon (TAG) is least frequently used in bacteria, and it has only one corresponding release factor; therefore, it is usually used in orthogonal tRNA design to minimize direct impacts on the expression levels of non-target genes. Recent progress in genome engineering and writing has enabled the construction of genetically recoded *Escherichia coli*, including *E. coli* with all amber stop codons replaced by ochre stop codons (TAA), or with a 61-codon genome (serine codons (TCG and TCA) and amber stop codons were replaced by their synonyms AGC, AGT, and TAA, respectively), which provides a favorable chassis cell for applying GCE-based orthogonal translation systems by reassigning free codons to ncAAs<sup>30,31</sup>. The expression of genes with amber stop codon insertions requires ncAAs, and their expression levels are presumed ncAA-dependent. At the same time, tunable translation control tools using site-specific ncAA incorporation have been studied<sup>29,32–34</sup>. Therefore, by inserting amber stop codons into genes that are important for controlling cell growth, such as genes involved in central carbon metabolic pathways, should be regulated by ncAA abundances. Further titration of ncAAs to determine optimum concentrations to precisely control cell growth rates can potentially be used for balancing cell growth and bioproduction (Fig. 1a, c).

In this study, a GCE-based cell growth and biosynthesis balancing engineering (GCE-CGBBE) strategy for efficient bioproduction is developed, which is based on titrating expression of cell growth and metabolic flux determinant genes by constructing ncAA-dependent expression patterns (Fig. 1c). A translation system for incorporating the ncAAs into proteins via amber stop

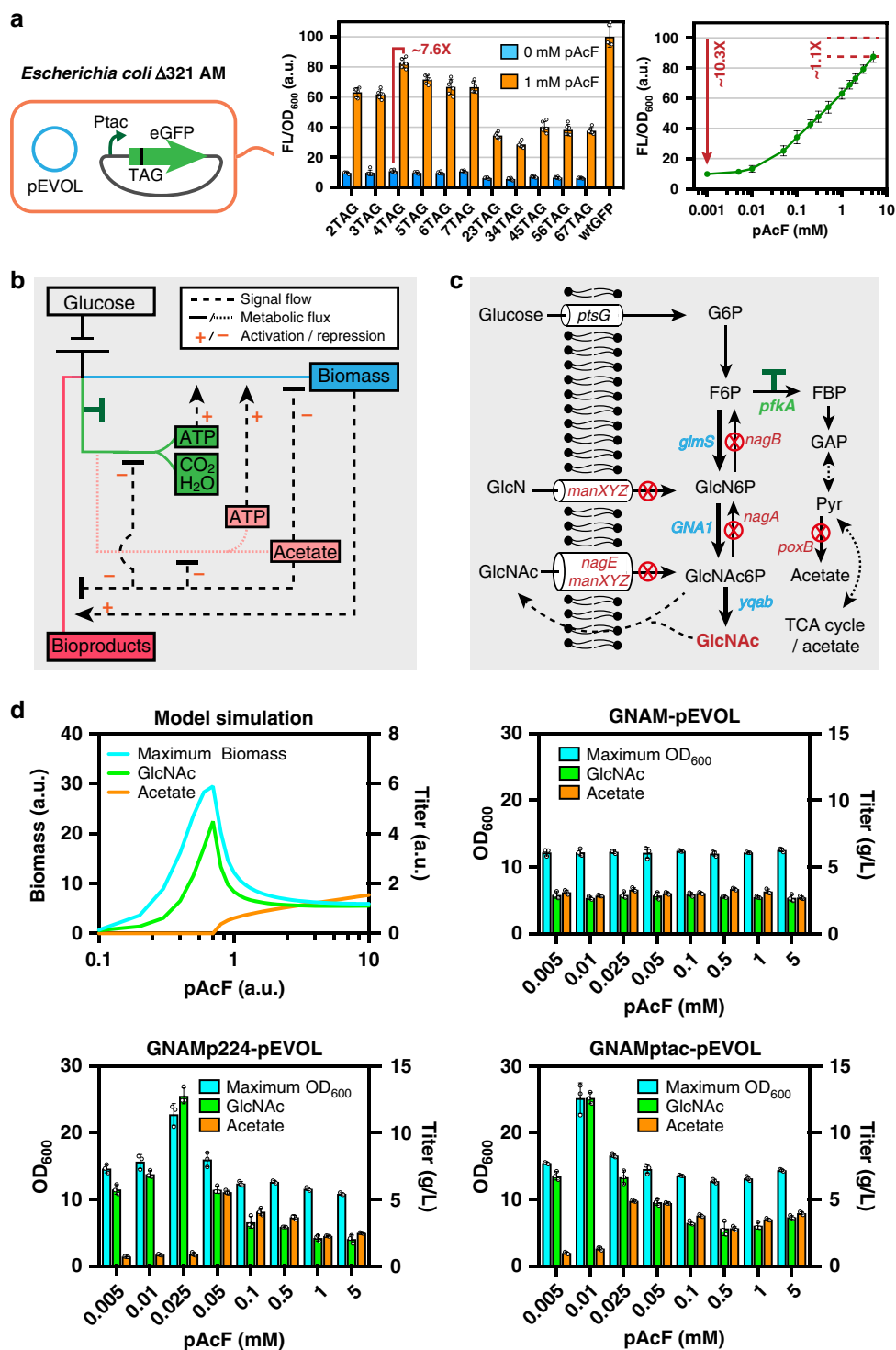
codon insertion is optimized for genetically recoded *E. coli* and non-genome-recoded *Bacillus subtilis*, two of the best-studied Gram-negative and Gram-positive bacteria, respectively. We demonstrate GCE-CGBBE in genome-recoded *E. coli* by precisely balancing glycolysis and *N*-acetylglucosamine (GlcNAc) production. We further expand GCE-CGBBE to non-genome-recoded *B. subtilis* to improving *N*-acetylneuraminic acid (NeuAc) production. In addition, the development of ncAA-dependent essential gene expression regulation shows efficient biocontainment of engineered *B. subtilis*.

## Results

**Testing the ncAA-incorporation system in *E. coli* Δ321AM.** The orthogonal translation system for ncAA incorporation is composed of mutant aaRS/tRNA pairs for translating non-canonical codons, which has been well established in *E. coli*<sup>35</sup>. Therefore, the widely used pEVOL-based ncAA/ *p*-acetyl-L-phenylalanine (pAcF) incorporation system was selected, which is specific for translating amber stop codons and have only limited activity for other 20 proteinogenic amino acids<sup>35</sup>. In our pAcF incorporation system, we selected the *E. coli* Δ321AM strain, which has been genetically recoded by (1) exchanging all 321 amber stop codons in the genome with ochre stop codons for translation termination; (2) knockout of peptide chain release factor RF1 to avoid competition between RF1 and the orthogonal tRNA for efficient translation; (3) recovering almost the same cell growth rate as wild type *E. coli* via adaptive evolution<sup>30,36</sup>. Therefore, *E. coli* Δ321AM was suitable for specifically using the amber stop codon for incorporating ncAAs.

Next, the effects of replacing amber stop codons at different positions in the N-terminal coding region on the dynamic range of gene expression in the presence or absence of pAcF were tested using green fluorescent protein (GFP) under the control of the *P*<sub>tac</sub> promoter. By replacing one or two original codons from second to seventh position with amber stop codons (six mutants with a single amber stop codon replacement (2TAG to 7TAG), and five mutants with double adjacent amber stop codons replacement (23TAG to 67TAG)), it was found that the pEVOL-pAcF system can enhance gene expression by 6.48-fold and 7.65-fold in *E. coli* Δ321AM when the second (2TAG) and fourth (4TAG) codon of the GFP coding sequence are replaced, respectively. Furthermore, we selected the gene of 2TAG GFP to test the regulation range of gene expression by adding different concentrations of pAcF (0.001–5 mM). With increasing pAcF concentration, the fluorescence intensity increased logarithmically, achieving a 9.01-fold enhancement of gene expression (Fig. 2a). In addition, the cytotoxicity of pAcF for *E. coli* Δ321AM harboring pEVOL was also tested, showing that the addition of different concentrations of pAcF had no significant effect on cell growth (Supplementary Fig. 1). Tests on the ncAA-incorporation tool indicated that the gene expression level was dependent on pAcF concentration, which laid the foundation for its further application in titrating gene expression in *E. coli* Δ321AM.

**Balancing *E. coli* growth and GlcNAc biosynthesis.** In bioproduction systems, cells are essential biocatalysts for improving production efficiency. However, excess or improper control of cell growth often causes decreased production yield or by-product accumulation. When *E. coli* is provided with glucose in excess, high growth rates cause an overflow of the glycolytic pathway, accompanied by acetate synthesis and ATP production<sup>37</sup>. Although ATP produced in the acetate synthesis pathway may be beneficial for intracellular thermodynamically unfavorable reactions, thereby improving cell growth, acetate accumulation seriously impairs cell growth and substrate to target product conversion (Fig. 2b).



**Fig. 2** Non-canonical amino acid (ncAA) incorporation tool-based pathway regulation in genetically recoded *Escherichia coli* Δ321AM. **a** Testing the efficiency of ncAA tool-pEVOL using the reporter protein GFP (left panel). The histogram (middle panel) represents the degree of GFP expression recovery after replacing one or two original codons by amber stop codons in the sequence from the second to the seventh codon (2TAG to 7TAG and 23TAG to 67TAG). The data are expressed as the mean ± SD from six ( $n = 6$ ) biologically independent replicates. The line graph (right panel) shows the use of the 2TAG GFP gene to test GFP expression levels in the presence of different concentrations of *p*-acetyl-L-phenylalanine (pAcF). The data are expressed as the mean ± SD from three ( $n = 3$ ) biologically independent replicates. **b** Intracellular metabolic feedback loop of engineered *E. coli*. 'T' marker represents the regulation of metabolic flow. **c** Construction of the *N*-acetylglucosamine (GlcNAc) biosynthetic *E. coli* Δ321AM strain. Red circular symbols represent gene knockouts, the 'T' marker represents the regulation of metabolic flux, and bold arrows represent an increase in metabolic flux. **d** Model simulated (half-saturation constant  $K_S = 15$ ) and actual fermentation results of the three GlcNAc biosynthetic strains measured in the presence of different pAcF concentrations. The data are expressed as the mean ± SD from three ( $n = 3$ ) biologically independent replicates. Source data underlying (a), (d) are provided as a Source data file.

Therefore, balancing cell growth and target compound biosynthesis is a win-win strategy, which cannot only balance the competition between cell growth and target compound synthesis, but also serves to avoid the formation of overflow by-products.

Consequently, biosynthesis of GlcNAc, a widely used nutraceutical in cartilage protection and arthritis treatment, was chosen as an example, which competes for both carbon flux in the glycolytic pathway and precursors of cell wall synthesis (Fig. 2c). Using *E. coli*  $\Delta$ 321AM as the initial host, we first constructed the GlcNAc-producing strain GNAM by knocking out seven genes (GlcNAc degradation-related genes *nagABE* and *manXYZ*, and pyruvate oxidase *poxB*), and by strengthening the GlcNAc biosynthetic pathway through integrated expression and promoter optimization of three pathway genes, i.e., *glmS*, *GNAI*, and *yqaB*, which originated from *E. coli* K-12 *glmS* mutant 2123-72 with a moderately strong promoter ( $P_{566}$ ), *Saccharomyces cerevisiae* S288C *GNAI* with a strong promoter ( $P_{tac}$ ), and *E. coli* K-12 *yqaB* with a strong promoter ( $P_{tac}$ ), respectively (Supplementary Figs. 2, 3)<sup>38,39</sup>. The GNAM strain produced 2.81 g/L of GlcNAc and 3.13 g/L of acetate (Supplementary Fig. 3). To aid the design of a balancing strategy, we constructed a kinetic model of GlcNAc production and cell growth of the engineered strains. In this model, the pAcF concentration represented a controller of the glycolytic flux, which controlled the expression of phosphofructokinase by inserting an amber stop codon into the corresponding *pfkA* gene. When ncAA was added in excess in silico, excess glycolytic flux enabled rapid cell growth, and activated the synthesis of acetate and ATP; consequently, with an accumulation of acetate, bacterial growth was rapidly inhibited. As pAcF was gradually decreased in silico, the glycolytic flux decreased, which slowed acetate accumulation, and allowed for increased cell division. Therefore, more GlcNAc was produced with increasing numbers of cells as biocatalysts. As the concentration of pAcF continued to decrease in silico, the carbon flux of the glycolytic pathway gradually decreased to limited cell growth and further reduced production (Fig. 2d; Supplementary Fig. 4). Through model simulations, we found that there should be a balance between cell growth and product synthesis, and GlcNAc can be efficiently biosynthesized within a narrow pAcF concentration range around the balance point.

In order to verify the above hypothesis, the *pfkA* gene of *E. coli* was selected for modulation, because its expression determines the carbon flux of both glycolytic and GlcNAc biosynthetic pathways. First, pEVOL was transformed into GNAM to obtain GNAM-pEVOL. Next, an amber stop codon was inserted after the translation start codon (ATG) of the *pfkA* coding sequence in the genome of the engineered strain. At the same time, in order to increase transcription and subsequently reduce the working concentration of pAcF, we replaced the promoter of *pfkA* with strong  $P_{224}$  and  $P_{tac}$  promoters, yielding GNAMP<sub>224</sub>-pEVOL and GNAMP<sub>tac</sub>-pEVOL strains, respectively. Therefore, *pfkA* expression could be regulated by titrating pAcF in the medium, further fine-tuning the carbon flux between glycolytic and GlcNAc biosynthetic pathways.

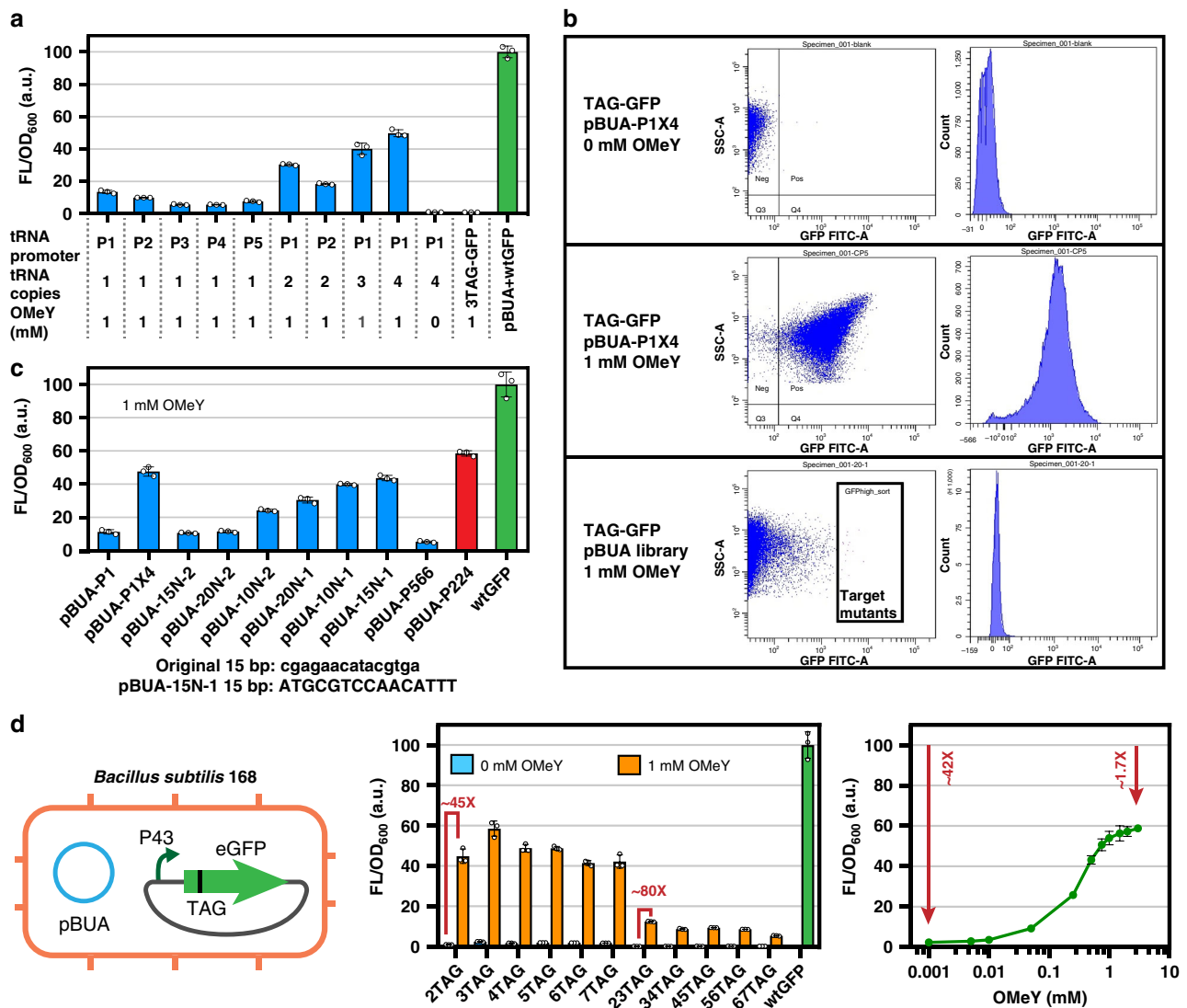
Indeed, different concentrations of pAcF influenced both cell growth and GlcNAc biosynthesis. Compared with engineered *E. coli* without balanced growth and bioproduction, both cell growth and GlcNAc biosynthesis were greatly improved with decreased acetate. The GNAMP<sub>tac</sub>-pEVOL strain grown with 0.01 mM pAcF resulted in a 2.06-fold increase in maximal OD<sub>600</sub> (25.21), 83.75% decreased acetate levels (5.54 vs. 0.90 g/L), and a 4.54-fold increased GlcNAc titer (12.77 g/L) with 0.48 g<sub>GlcNAc</sub>/g<sub>glucose</sub>, reaching 88.52% of the theoretical pathway yield (Fig. 2d, Supplementary Fig. 5). In order to ensure that the strategy is not only effective at high concentrations of glucose, fermentation experiments were also performed at a gradient glucose

concentration ranging from 5 to 30 g/L. Strain GNAMP<sub>tac</sub>-pEVOL showed higher GlcNAc titer and lower acetate titer regardless of the glucose concentration, indicating the effectiveness of the strategy under different culture conditions (Supplementary Fig. 6). Besides, the expression level of PfkA was also tested by monitoring fluorescence of GFP fused with PfkA, which increased logarithmically with an increase of pAcF concentration as expected (Supplementary Fig. 7).

### B. subtilis expression system for incorporating O-methyl-L-tyrosine.

In order to further verify the universality of the GCE-CGBBE strategy for non-genetically recoded bacteria, the strategy was further expanded to the non-genome-recoded Gram-positive model bacterium *B. subtilis* to balance growth and bioproduction. Since ncAA-incorporation tools are lacking for *B. subtilis*, an ncAA O-methyl-L-tyrosine (OMeY) incorporation tool for *B. subtilis*, pBUA, was first constructed and tested. We selected an aaRS/tRNA mutant pair derived from *Methanococcus jannaschii* that can incorporate OMeY into amber stop codons<sup>25</sup>. The coding gene of the mutant aaRS/tRNA pair for OMeY incorporation was inserted into pHT01 plasmid for expression. Accordingly, successful expression of genes with amber stop codon insertion should be observed in the presence of OMeY. Therefore, efficiency of the OMeY incorporation tool (pBUA) was tested by comparing fluorescence intensities of engineered *B. subtilis* harboring a GFP gene with or without amber stop codon substitution. The higher the percentage of GFP gene expression with TAG substitution compared to that without TAG substitution, the more efficient the OMeY incorporation tool. The first version of pBUA tools (pBUA-P1 to pBUA-P5) was constructed based on the pHT01 plasmid by integrating aaRS and mutant tRNA gene copies under the control of  $P_{43}$  or one of five original *B. subtilis* tRNA promoters (P1 to P5), respectively (Fig. 3a). When one copy of the tRNA mutant was under the control of a P1 or P2 promoter, the expression level of GFP with TAG substitution (at the third codon in the N-terminal region) reached 13.63%, and only 10.04% of GFP expression without TAG substitution, which showed that P1 and P2 were more efficient than the other original *B. subtilis* tRNA promoters. To further enhance OMeY incorporation efficiency, copy number and promoter selection were optimized for this tRNA mutant. As four copies of the tRNA mutant were under the control of the P1 promoter, an effective OMeY incorporation tool, pBUA-P1X4, was constructed to restore ~49.78% of GFP expression with only 2.21% leakage (Fig. 3a). For promoter optimization of tRNA genes, two strategies were used. The first strategy was to construct random mutation libraries to screen promoters for efficiently driving tRNA expression. We constructed three random mutation libraries (pBUA-10N, pBUA-20N, and pBUA-15N) based on pBUA-P1 by random A, T, G, and C insertions 10 bp upstream of the tRNA coding sequence, 20 bp upstream of the tRNA coding sequence, and 15 bp between the -35 and -10 regions of promoter P1 via degenerate primer design and PCR. After screening 6,000,000 cells using flow cytometry, and further characterization of 1000 strains obtained from each library, the most efficient pBUA mutant in all libraries was finally obtained (Fig. 3b). The efficiency of the pBUA system with a single tRNA mutant copy obtained by this strategy was 2.87-fold higher than that of pBUA-P1, which restored the expression level of GFP to 43.65%. The second strategy was to rationally replace promoter P1 with strong synthetic promoters to enhance tRNA mutant expression. The strong promoters  $P_{224}$  and  $P_{566}$  of *B. subtilis* were selected to express mutant tRNAs using the pBUA system. The pBUA-P<sub>224</sub> construct further restored GFP levels to 58.50%, which was 1.23-fold higher than that of the four tRNA copies with the P1





**Fig. 3 Construction and optimization of *Bacillus subtilis* non-canonical amino acid (ncAA) incorporation tools. a** Optimization of tRNA promoter and copy number. **b** Screening of three random mutant libraries for efficient pBUA mutants using flow cytometry. SSC-A: granularity, FITC-A: GFP fluorescence. **c** Degree of GFP expression level recovery by efficient pBUA mutants obtained from three random mutant libraries, harboring pBUA constructs with synthetic promoters. **d** Efficiency verification of pBUA in *B. subtilis*. The histogram (left panel) represents the degree of GFP expression recovery after replacing one or two original codons with amber stop codons in the sequence from the second to the seventh codon (2TAG to 7TAG and 23TAG to 67TAG). The line graph (right panel) shows the use of the 3TAG GFP gene to test GFP expression levels in the presence of different concentrations of *p*-acetyl-L-phenylalanine (pACF). All the data are expressed as the mean  $\pm$  SD from three ( $n = 3$ ) biologically independent replicates. Source data underlying (a), (c), (d) are provided as a Source data file.

promoter system that was selected for OMeY incorporation in *B. subtilis* (Fig. 3c).

We then tested the effects of different amber stop codon insertion sites on the dynamic range of expression of pBUA-P<sub>224</sub> in the presence or absence of OMeY. By replacing one or two codons from the second to seventh position with amber stop codons (six mutants with a single amber stop codon replacement (2TAG to 7TAG), and five mutants with double adjacent amber stop codon replacements (23TAG to 67TAG)), it was found that pBUA-P<sub>224</sub> can activate up to 77.75-fold gene expression when the second and third codons of the GFP coding sequence were replaced (23TAG), while GFP expression was difficult to detect in the absence of OMeY, demonstrating that the corresponding tRNA was charged exclusively with the reassigned amino acid by an aARS that did not interact with any other tRNA. In parallel,

when the third codon was replaced with TAG, the expression level of GFP could be maximally rescued, reaching 58.50% of GFP without TAG codon insertion. Furthermore, we selected the gene of 3TAG GFP and tested the regulation range of gene expression by adding different concentrations of OMeY (0.001–3 mM). Fluorescence intensity increased with increasing OMeY concentration, achieving 24.82-fold activation of gene expression (Fig. 3d; Supplementary Fig. 8). In addition, the cytotoxicity of the OMeY incorporation system was tested by culturing *B. subtilis* harboring pBUA-P<sub>224</sub> at different concentrations of OMeY. This showed that the OMeY incorporation system had no significant effect on cell growth in the OMeY concentration range tested (Supplementary Fig. 9). In summary, a pBUA-P<sub>224</sub>-based OMeY incorporation system was constructed and optimized for *B. subtilis*, which can efficiently incorporate OMeY into proteins at

amber stop codons, reaching 58.50% of the gene expression level of non-engineered protein, with up to 24.82-fold activation under different OMeY concentrations.

**Balancing cell growth and NeuAc synthesis in *B. subtilis*.** *B. subtilis* has proven important for the bioproduction of vitamins, functional sugars, and industrial enzymes; however, cell growth often competes with target chemical biosynthesis for energy and precursors. Therefore, a robust cell growth and precise bio-production regulation method is needed for *B. subtilis*. NeuAc biosynthesis involves not only competition for precursors of glycolysis and cell wall synthesis pathways, but also for phosphoenolpyruvate, a precursor of the TCA cycle. Therefore, NeuAc biosynthesis was selected as an example for *B. subtilis*, and the NeuAc biosynthetic strain with pBUA-P<sub>224</sub>, NAB, was constructed. Precisely controlling the expression level of essential genes by titrating OMeY should achieve precise regulation of cell growth, thereby fine-tuning the balance between cell growth and product synthesis. However, the regulation of the expression level of essential genes may affect endogenous metabolism of cells, and thus the generality of the GCE-CGBBE strategy. In order to minimize the impact on endogenous metabolism and identify metabolic pathway-independent target genes for cell growth control, ten genes related to cell wall synthesis and structural integrity were selected as candidates for the construction of OMeY-dependent *B. subtilis* growth. In order to minimize damage to protein function, the SIFT algorithm was used to predict TAG substitution sites<sup>40</sup>. To minimize the dosage of OMeY, the strong promoter P<sub>43</sub> was used to replace endogenous promoters. In the end, three OMeY-dependent NeuAc producing strains were successfully constructed: NABd4, NABd8, and NABd9 by controlling the expression of *murB*, *walR*, and *cdsA*, respectively. The three genes encode UDP-*N*-acetylenolpyruvoylglucosamine reductase (peptidoglycan precursor biosynthesis), two-component response regulator (cell wall metabolism control) and phosphatidate cytidyltransferase (phospholipids biosynthesis), respectively. Precise regulation of the expression levels of these three genes can theoretically slow cell division and further control cell growth rate. To test the effect of OMeY on growth regulation, the growth curves of three engineered and control strains were first determined at different concentrations of OMeY, and specific growth rates were calculated (Fig. 4). The growth of OMeY-dependent strains was able to respond to changes in OMeY concentration: (1) with an increase in OMeY concentration, the maximum specific growth rate of NABd4 increased logarithmically; (2) the maximum specific growth rate of NABd8 first increased linearly and then remained basically unchanged; (3) the maximum specific growth rate of NABd9 oscillated (Fig. 4). And after fusion of *murB*, *walR*, and *cdsA* with *gfp* gene, respectively, the expression levels of these three essential genes were also tested by monitoring fluorescence dynamics, which all increased first and then flattened with increasing OMeY concentration (Supplementary Fig. 10). Because GlcN6P is the precursor of cell wall peptidoglycan synthesis, the control of MurB expression in peptidoglycan synthesis pathway may affect metabolic flux to NeuAc biosynthetic pathway. Bio-production and cell growth are difficult to be completely separated, therefore, the effects of both controlling cell growth and improving precursor GlcN6P supply may contribute to improved NeuAc production.

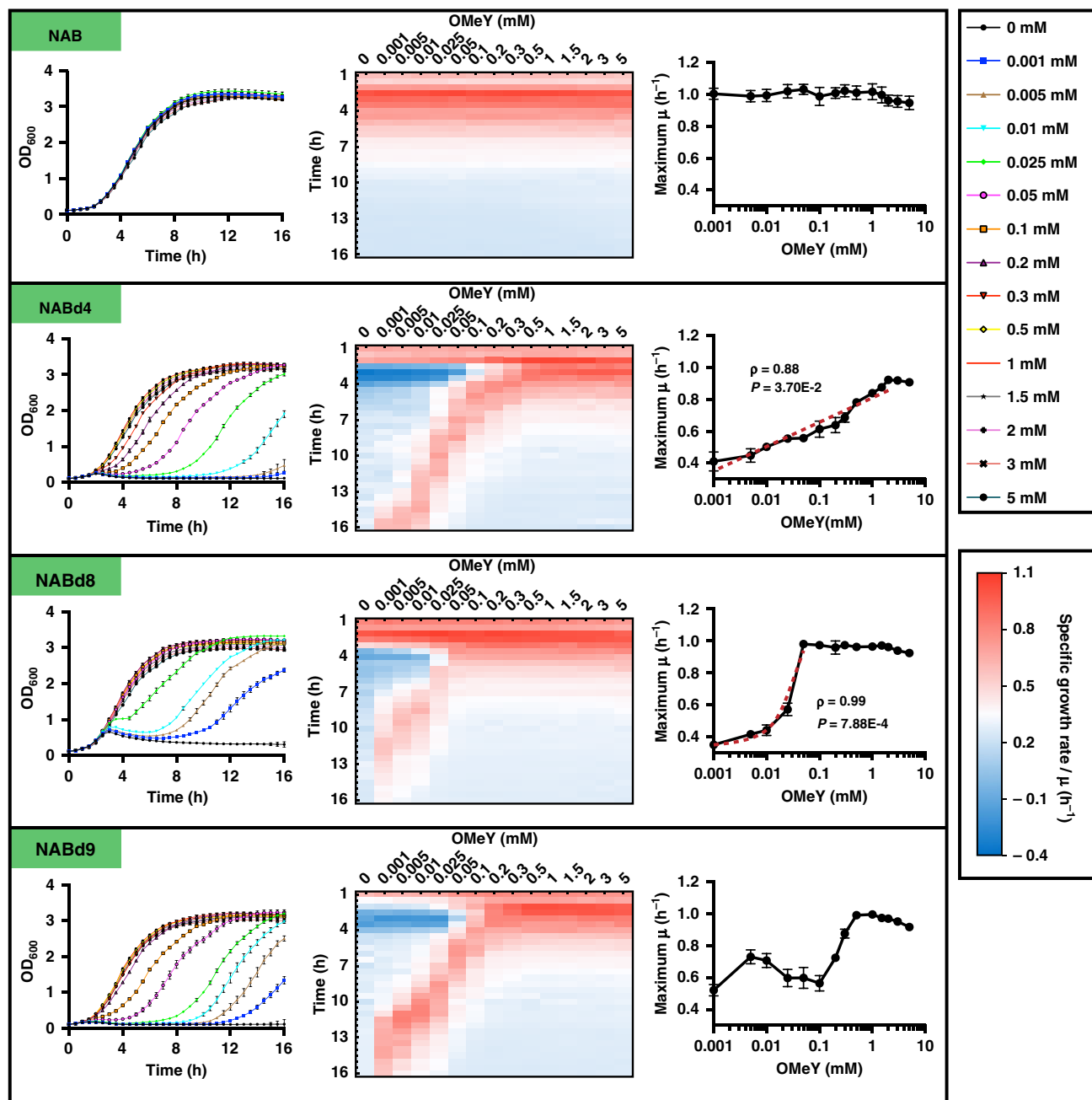
Fermentation tests on 24-well plates were carried out to verify a balance between cell growth and biosynthesis. By controlling essential genes using different concentrations of OMeY, the NeuAc titer of NABd8 increased to 2.34-fold, from 2.02 to 4.72 g/L, and from 0.13 to 0.34 g/L/OD<sub>600</sub> (Fig. 5a, b; Supplementary Figs. 11, 12) when adding 0.1 mM OMeY. Besides, the molar ratio of

NeuAc to by-product acetoin increased up to 5.12-fold from 0.034 to 0.174 mol/mol after growth control. When the titer of NeuAc reached its maximum level, the molar ratio of NeuAc to acetoin increased to 3.65-fold, reaching 0.124 mol/mol (Fig. 5b).

In order to test the effectiveness and robustness of the system at larger scale, we first constructed a single-plasmid strain by grafting gene *age* and *neuB* from plasmid pP43NMK to pBUA-P<sub>224</sub> (Fig. 5c)<sup>41</sup>. At the same time, seven different promoters were tested to express gene *age* and *neuB* on pBUA-P<sub>224</sub>. Finally, a single-plasmid strain named NAB-SpP566 was obtained, which produced 1.87 g/L NeuAc (the strain harboring two plasmids, NAB, produced 2.02 g/L NeuAc) in 24-well plate (Fig. 5d). Then the strain NABd8-SpP566 was constructed by P<sub>43</sub> promoter substitution and TAG codon incorporation for essential gene *walR* on the basis of NAB-SpP566. Subsequently, fed-batch fermentation was performed in a 3-L fermenter for strain NABd8-SpP566 with NAB-SpP566 as control. The NeuAc titer of NABd8-SpP566 reached 9.71 g/L, which was 2.97-fold of NAB-SpP566 (3.27 g/L). Moreover, NeuAc production of NABd8-SpP566 on cell reached 0.155 g/L/OD<sub>600</sub>, which was 4.43-fold of NAB-SpP566 (0.035 g/L/OD<sub>600</sub>, Fig. 5e). And the production of acetoin by-product decreased 56.42%, from 30.47 to 13.28 g/L. The above results preliminarily indicate the universality and stability of the GCE-CGBBE strategy in different strains using ncAA-incorporation tools.

**Biocontainment of OMeY-dependent *B. subtilis* strains.** Unintended proliferation of genetically modified organisms (GMOs) is one of the major concerns for the biosafety of metabolic engineering research and industrial applications, which may pose potential threats to the survival of environmental microorganisms, plant growth, human and animal health, and may cause gene leakage by horizontal gene transfer<sup>42–44</sup>. Therefore, it is necessary to carry out biological containment of engineered microorganisms. For genetically recoded *E. coli* Δ321AM possessing an altered genetic code, ncAA-dependent strains have been constructed by inserting an amber stop codon to computationally redesigned essential genes to achieve biological containment<sup>45,46</sup>. From our design perspective, construction of the above OMeY-dependent *B. subtilis* cannot only achieve precise regulation of cell growth, but theoretically achieve the goal of biocontainment. Therefore, we tested the escape frequencies of NABd4, NABd8, and NABd9. The different strains were first cultured to mid-log phase in presence of 1 mM OMeY, washed, and subsequently incubated without OMeY to remove excess OMeY inside and outside the cells. By plating on OMeY-free plates, we found that the 15-day escape frequencies of all OMeY-dependent *B. subtilis* were lower than 3.67 × 10<sup>-10</sup> escapees/c.f.u., meeting the requirement of 10<sup>-8</sup> recommended by the National Institute of Health (Fig. 6a, b)<sup>47</sup>. These results demonstrate the effectiveness of ncAA-incorporation tools in biocontainment of non-genetically recoded industrial strains.

Due to the regulation of cell wall synthesis genes, cell morphology may change and further lead to a decline of production performance during large-scale fermentation. In order to test whether OMeY-dependent *B. subtilis* cell morphology was normal and whether the cell wall was damaged, cellular morphology was examined. In the presence of OMeY, the cell morphology was essentially the same as that of the control strain, NAB, except that NABd8 cells were longer than the control strain. In the absence of OMeY, the cell wall was ruptured, accompanied by content leakage (Fig. 6c). This provides evidence for the industrial application potential of ncAA-dependent strains, while at the same time highlighting their applicability in artificially controlled environments only.



**Fig. 4** Effects of different O-methyl-L-tyrosine (OMeY) concentrations on the growth of OMeY-dependent, engineered *Bacillus subtilis* strains. The growth curves of the four engineered strains (left panels). Specific growth rates and maximum specific growth rates of the engineered strains at different OMeY concentrations are shown in the middle and right panels, respectively.  $\rho$ , Pearson correlations;  $P$ ,  $p$ -value. All the data displayed as the mean  $\pm$  SD ( $n = 4$ ) of four biologically independent replicates. The statistical analysis was performed by one-sided  $t$ -test. Source data are provided as a Source data file.

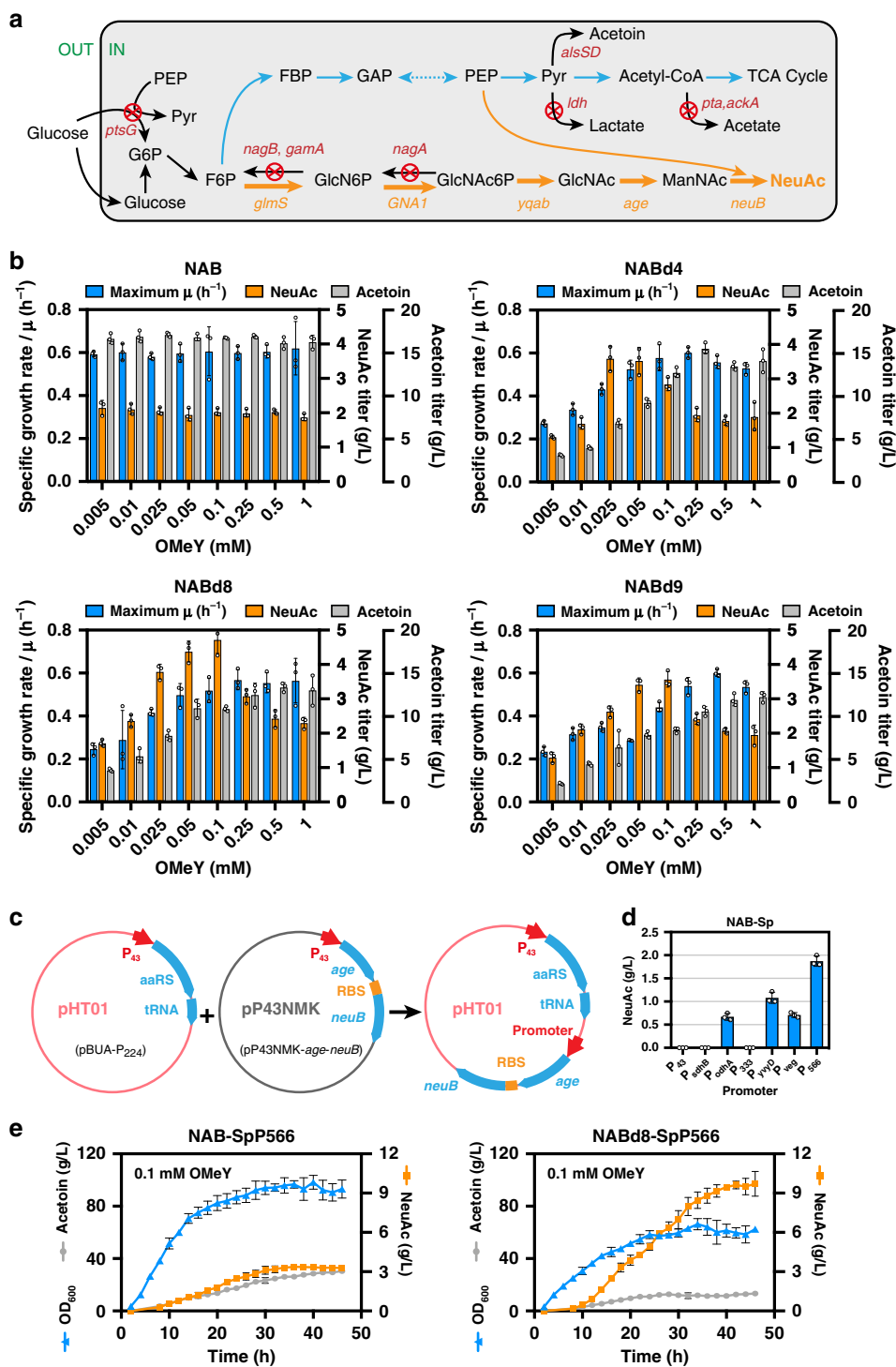
## Discussion

Fine-tuning cell growth and biosynthesis are important for efficient bioproduction. To precisely balance these processes, a GCE-based orthogonal protein translation system was introduced into genetically recoded *E. coli* and non-genome-recoded *B. subtilis* to incorporate ncAA and further regulate metabolic flux and cell growth determinant genes by amber stop codon insertion. Proof-of-concept applications of metabolic engineering were verified by increasing GlcNAc and NeuAc in *E. coli* and *B. subtilis*, respectively. Thus, we demonstrated the potential of the combination of GCE and genetically recoded bacteria for precise regulation of cellular metabolism. Moreover, our highly efficient ncAA OMeY

incorporation tool (pBUA), the first ever developed for *B. subtilis*, expanded ncAA-incorporated protein production to *B. subtilis*, which will potentially prove useful due to its strong protein secretion ability<sup>48</sup>. Given that the GCE-CGBBE strategy is pathway-independent and inactive without ncAAs that do not naturally occur in natural growth environments, it should hold generic and robust features for other important biochemical production in engineered bacteria.

Because only trace amounts of ncAAs need to be added to balance cell growth and bioproduction, GCE-CGBBE should be economically feasible. We calculated the costs of two selected ncAAs for large-scale fermentation. For pAcF, it would cost ~\$14

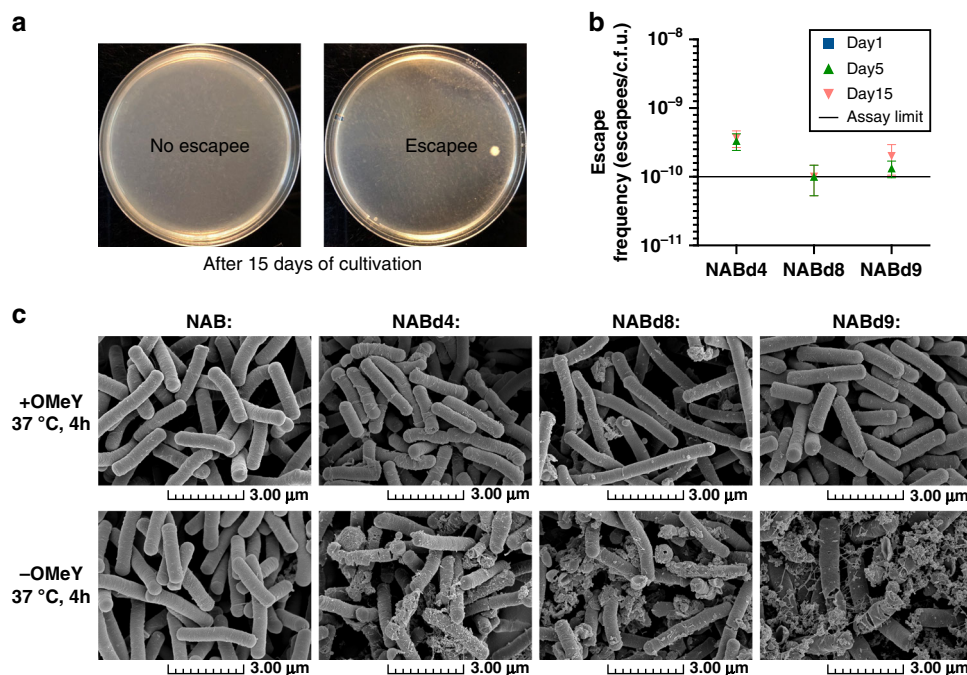




**Fig. 5** Balancing cell growth with biosynthesis of *N*-acetylneuraminic acid (NeuAc). **a** NeuAc biosynthetic pathways in *B. subtilis*. G6P, Glucose-6P; F6P, Fructose-6P; FBP, Fructose-1,6-2P; GAP, glyceraldehyde-3-P; PEP, phosphoenolpyruvate; GlcNAc, N-acetylglucosamine; ManNAc, N-acetylmannose. **b** Fermentation efficiencies of engineered *B. subtilis* strains at different *O*-methyl-L-tyrosine (OMeY) concentrations. **c** Amplifying gene *age* and *neuB* from plasmid pP43NMK and inserting into pBUA-P<sub>224</sub> under the control of different promoters. **d** NeuAc production of seven single-plasmid NeuAc biosynthesis strains. **e** Fermentation efficiencies of NAB-SpP566 and NABd8-SpP566 in a 3-L fermenter. All the data are expressed as the mean  $\pm$  SD from three ( $n = 3$ ) biologically independent replicates. Source data underlying **(b)**, **(d)**, **(e)** are provided as a Source data file.

per ton of fermentation broth at an optimal initial concentration of 0.01 mM. In parallel, for OMeY, it would cost ~\$27 per ton of fermentation broth with an optimal initial concentration of 0.1 mM (Supplementary Table 1). These costs should be acceptable for bioproducing high value-added compounds. Furthermore, addition of ncAAs abolishes the need for monitoring glucose

concentration, thereby simplifying the fermentation process, and reducing human labor. Moreover, based on the explosive growth of DNA sequencing throughput, and the availability of powerful bioinformatics tools, biosynthetic pathways for ncAAs from cheap precursors have been explored to produce at least 700 ncAAs, which promisingly further reduces ncAA cost<sup>49</sup>.



**Fig. 6** Determination of escape frequencies of O-methyl-L-tyrosine (OMeY)-dependent, engineered *Bacillus subtilis*. **a** After 15 days of cultivation, an escapee appeared on OMeY-free plates. **b** Escape frequencies of three OMeY-dependent *B. subtilis* strains were determined by plate coating. The escape efficiencies of all three strains on day 1 were under the assay limit, which is not shown in the plot. The data are expressed as the mean  $\pm$  SEM from three ( $n = 3$ ) biologically independent replicates. Source data are provided as a Source data file. **c** Scanning electron microscopic morphological identification of four engineered strains in the presence or absence of OMeY. Scanning electron microscopy experiments were performed three times independently with similar results. Source data underlying **b** are provided as a Source data file.

Potential applications of GCE-based ncAA-dependent strains are optimization processes for improving ratios of modular co-cultures in bioproduction, which is a promising strategy for overcoming unavailability of machineries and resources in individual cells by distributing precursor biosynthesis modules to multiple strains<sup>50,51</sup>. Currently, the control of the ratio of certain strains during modular co-culture relies on inoculation optimization and quorum-sensing system-based growth-regulation circuits<sup>52,53</sup>. A strategy for controlling this can be theoretically achieved in co-culture fermentation by titrating different ncAAs by constructing multiple different ncAA-dependent sub-populations.

Overall, the GCE-CGBBE strategy was developed as a generic approach for balancing bacterial growth and chemical biosynthesis for efficient bioproduction. Proof-of-concept applications for metabolic engineering revealed increased GlcNAc and NeuAc production in the best-studied Gram-negative (*E. coli*) and Gram-positive (*B. subtilis*) bacteria, yielding 4.54-fold and 2.34-fold increased titers, respectively. In addition, the developed ncAA-dependent essential gene expression regulation shows efficient biocontainment of engineered *B. subtilis*, which can prevent its unintended proliferation during metabolic research and industrial application. Because the GCE-CGBBE strategy is pathway-independent and does not interfere with ncAAs that do not naturally occur in any medium or growth environment, it should hold generic and robust features for additional important biochemical production in engineered bacteria.

## Methods

**Plasmids and strains.** Gibson assembly was used to construct all plasmids used in this study with pET28a, pHT01, and p43NMK as backbones<sup>54–57</sup>. For genome editing of *E. coli*  $\Delta$ 321AM,  $\lambda$ -Red recombination was used, and pCP20 served to eliminate resistance selection markers on the *E. coli* genome. A homologous recombination box with 1000 bp homology arms was used to perform genome editing in *B. subtilis*, and the Cre-lox system was used to eliminate resistance selection markers on the genome.

Construction of the three pBUA libraries was as follows: based on pBUA-P1, we first selected 10- and 20-base sequences upstream of the mutant tRNA sequence, and replaced them with sequences of the same length composed of A, T, C, and G in equal probabilities to construct two random mutation libraries, named pBUA-10N and pBUA-20N, via degenerate primer design and PCR. Next, we mutated the 15-base DNA sequence between the  $-35$  and  $-10$  regions into a 15-base random sequence to construct the pBUA-15N library. For each mutant library, more than 6,000,000 cells were screened using flow cytometry, and  $\sim$ 1000 transformants with the highest single-cell fluorescence intensities were selected.

Construction of OMeY-dependent *B. subtilis*: First, a homologous recombination cassette with 1000 bp left homology arm, spectinomycin resistance cassette, P<sub>43</sub> promoter and 1000 bp right homology arm (coding sequence of essential genes incorporated with TAG) was constructed by fusion PCR (for example, Seq S9-S11). Then the homologous recombination cassette was transformed into a strain harboring pBUA-P<sub>224</sub> and the bacterial suspension was plated on a plate containing 1 mM OMeY and 100 mg/L spectinomycin. After 16–24 h plate incubation, all colonies were further screened with OMeY-free LB medium, in which OMeY-dependent *B. subtilis* could not grow successfully. And if there are no colonies on the plate or no positive transformants are obtained after three experiments, it is considered that the essential genes cannot be used to construct OMeY-dependent strains.

All plasmids and strains used in this study are shown in Supplementary Data 1. Snapgene 4.3.6 was used for gene sequence analysis and primer design, and the list of all primers is Supplementary Data 2. All sequences of genes, constructed plasmids and gene expression cassettes used in this study are provided as Supplementary Data 3.

**Strains cultivation.** For cell culture, Luria-Bertani (LB) medium (10 g/L tryptone, 5 g/L yeast extract, and 10 g/L NaCl) was used unless noted otherwise. To produce GlcNAc by fermentation in *E. coli*, TMM (TB-added minimal medium: 30 g/L glucose, 2.4 g/L tryptone, 4.8 g/L yeast extract, 1 g/L glycerol, 5.8 g/L KH<sub>2</sub>PO<sub>4</sub>, 3.3 g/L K<sub>2</sub>HPO<sub>4</sub>·3H<sub>2</sub>O, 3.55 g/L citric acid·H<sub>2</sub>O, 2.5 g/L MgSO<sub>4</sub>·7H<sub>2</sub>O, 5 g/L (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 6 g/L urea, 0.025 g/L CaCl<sub>2</sub>·H<sub>2</sub>O, 0.1 mg/L CoCl<sub>2</sub>·H<sub>2</sub>O, 0.1 mg/L CuSO<sub>4</sub>·5H<sub>2</sub>O, 5 mg/L FeSO<sub>4</sub>·7H<sub>2</sub>O, 0.33 mg/L MnSO<sub>4</sub>·H<sub>2</sub>O, 3.8 mg/L ZnSO<sub>4</sub>·7H<sub>2</sub>O, pH 7.0, adjusted using NH<sub>4</sub>OH). To produce NeuAc by fermentation in *B. subtilis*, BFM (*B. subtilis* fermentation medium: 12 g/L yeast extract, 6 g/L tryptone, 6 g/L (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 12.5 g/L K<sub>2</sub>HPO<sub>4</sub>·4H<sub>2</sub>O, 2.5 g/L KH<sub>2</sub>PO<sub>4</sub>, 3 g/L MgSO<sub>4</sub>·7H<sub>2</sub>O, and 60 g/L glucose, pH 7.0, adjusted using NH<sub>4</sub>OH).

Cell culture using 96-well plates: first, 96-well deep-well plate with 300  $\mu$ L fresh LB medium per well, which were then cultured for 10 h at 37 °C with shaking at 750 rpm. Next, 10  $\mu$ L of the culture were inoculated in a 96-well microtiter plate

containing 190  $\mu$ L fresh LB medium with different concentrations of ncAA (pAcF or OMeY) and incubated at 37 °C with shaking at 750 rpm until OD<sub>600</sub> reached 1. Subsequently, 10  $\mu$ L of the culture were similarly inoculated into 96-well black plates (Corning 3603) containing 190  $\mu$ L fresh LB medium, cultured at 37 °C with shaking at 750 rpm, and fluorescence intensities and cell growth curves were determined.

For determining the growth curves of OMeY-dependent *B. subtilis* at different OMeY concentrations, a 96-well plate was used to reduce the effect of the sampling process on cell growth. OMeY-dependent *B. subtilis* were first cultured in fresh LB medium supplemented with 1 mM OMeY for 10 h at 37 °C and 220 rpm. Before inoculation with different concentrations of OMeY, the bacterial suspension was washed to reduce residual OMeY inside and outside the cells. Then the bacterial suspension was first centrifuged at 5000  $\times$  g for 3 min and then resuspended in fresh OMeY-free LB medium. After performing this step twice, the bacterial suspension was incubated for 1 h at 37 °C with shaking at 220 rpm. Next, the bacterial suspension was centrifuged again at 5000  $\times$  g for 3 min, and resuspended in fresh LB medium. Other operations and detection methods were the same as those used for culturing cells in a 96-well plate.

Fermentation experiments performed in 24-well deep-well plates: For *E. coli*, the colonies on the plates were first picked into 3 mL fresh LB medium and cultured for 9 h at 37 °C and 220 rpm. Then, 15  $\mu$ L of bacterial suspension was inoculated into a 24-well deep-well plate containing 1.5 mL of TMM medium per well, and fermentation was performed at 37 °C and 220 rpm for 33 h. For OMeY-dependent *B. subtilis*, the colonies on the plates were first picked into 1.5 mL fresh BFM medium with 1 mM OMeY. Before inoculation, the bacterial suspension was washed to remove additional OMeY; the culture was centrifuged at 5000  $\times$  g for 3 min and resuspended in 3 mL fresh BFM medium, which was repeated twice. Thereafter, 15  $\mu$ L of bacterial suspension was inoculated into a 24-well deep-well plate containing 1.5 mL of fermentation medium per well, and fermentation was performed at 37 °C and 220 rpm for 60 h. A breathable sterile film was placed on the 24-well plate to ensure sufficient oxygen supply. At 30 h of fermentation, 1.5 mL of the culture was concentrated to ~1.3 mL, and at 60 h of fermentation, 1.5 mL of the culture was concentrated to ~1.1 mL.

Fed-batch fermentation in 3-L fermenter: The fermentation medium consists of 24 g/L yeast extract, 24 g/L tryptone, 8 g/L (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 13 g/L K<sub>2</sub>HPO<sub>4</sub>·4H<sub>2</sub>O, 3 g/L KH<sub>2</sub>PO<sub>4</sub>, 3.5 g/L MgSO<sub>4</sub>·7H<sub>2</sub>O, 2 g/L NaCl, 12 g/L urea, and 30 g/L initial glucose. First, colonies were picked from the plates and inoculated into baffled 500-mL shake flasks containing 50 mL of BFM medium, which was then cultured at 37 °C with shaking at 220 rpm until OD<sub>600</sub> reached 15. Next, 100 mL of the seed was inoculated into a 3-L fermenter (T&J-Atype; T&J Bio-engineering Co., Ltd, Shanghai, China) with an initial 1.8 L of fermentation medium. Keep pH at 7.1 using 15% NH<sub>3</sub> and maintain glucose concentration within 10–30 g/L by feeding with 700 g/L glucose throughout the fermentation period. The temperature was maintained at 37 °C, the aeration rate was maintained at 1.5 v.v.m. and the agitation speed was maintained at 800 rpm. Sampling every 2 h after inoculation to determine OD<sub>600</sub>, NeuAc titer, by-product acetoin titer and glucose concentration.

**Essential gene selection and TAG codon incorporation.** Selection of essential genes for *B. subtilis* was based on three databases<sup>38–60</sup>. In order to avoid an influence of the essential gene expression change on target metabolites, ten essential genes related to cell wall synthesis were selected: *tagO*, *ftsW*, *ftsZ*, *murB*, *murC*, *groes*, *rodA*, *walR*, *cdsA*, and *yerQ*.

For the selection of TAG integration sites, the SIFT algorithm was used to predict residue tolerance of the proteins (<http://sift.jcvi.org>)<sup>40</sup>. The SIFT algorithm can predict amino acid substitutions that have little effect on protein function by using sequence homology. Of all the tolerable residue sites, we selected those sites closest to the translation start codon (ATG) to facilitate genome editing. However, if the distance between all predicted tolerable residue sites and the translation start codon was longer than 45 bases, and the active site of the protein was not at the N-terminus, the amber stop codon was inserted next to the translation start codon, which is the simplest method to prevent the production of truncated proteins<sup>32,45</sup>. The ten selected essential genes in the genome underwent P<sub>43</sub> promoter (strong) substitution and TAG codon incorporation, respectively, by homologous recombination. Finally, three OMeY-dependent strains (based on three essential genes, *murB*, *walR*, and *cdsA*) were successfully constructed and sequenced: NABd4, NABd8, and NABd9 (Supplementary Data 1).

**Analytical methods.** Cell mass was determined by measuring OD<sub>600</sub>. Fluorescence intensities were measured using a Cytation 3 Multi-Mode Reader (BIOTEK) at excitation and emission wavelengths of 488 and 523 nm, respectively. Gen5 CHS 2.06 was used to collect the biomass and fluorescence intensity data. While measuring the fluorescence intensities (FLS) and OD<sub>600</sub> (ODS) of the samples, fluorescence intensity (FLB) and OD<sub>600</sub> (ODB) of the blank medium were measured for fluorescence intensity correction. The final fluorescence intensity per unit biomass (FL/OD<sub>600</sub>) was calculated as (FLS-FLB)/(ODS-ODB).

High-performance liquid chromatography (HPLC, Agilent 1200 Series, USA) with an ultraviolet absorption detector (210 nm) was used to detect the concentrations of GlcNAc, NeuAc, and acetate in each sample. An Aminex HPX-87H column (300  $\times$  7.8 mm, Bio-Rad, Hercules, CA, USA) was utilized, and the

mobile phase was 10 mM H<sub>2</sub>SO<sub>4</sub> with a flow rate of 0.5 mL/min at 40 °C. Agilent OpenLAB Control Panel was used to collect the HPLC data.

The glucose concentration during fermentation was measured using a glucose-lactate analyzer (MI100, Shenzhen Sieman Technology Co., Ltd, Shenzhen, China). Microsoft Excel 2019 16.37 was used to analyze the biomass, yield and fluorescence intensity data, and carry out statistical calculation. Fluorescent-activated cell sorting (FACS) technology was used to screen single cells with high fluorescence intensities. Cell analyzer (BD FACSAria III, BD, Franklin Lakes, NJ, USA) with FlowJo\_V10 (FlowJo, LLC) software was used for cell sorting. For each random mutation library, 6,000,000 transformants were determined and screened to obtain ~1000 potential target transformants.

**Model simulation.** MATLAB R2019a was used for kinetic modeling of cell growth, as well as for target metabolite (GlcNAc) and by-product (acetate) synthesis<sup>61–63</sup>. The development of the kinetic model was mainly based on four kinetics, including the growth kinetic model of acetic acid-producing *E. coli* developed by Luli et al., Monod equation describing nutrient limitation, acetate synthesis and specific growth rate equations determined by Basan et al., and Michaelis–Menten equation describing the biosynthesis pathway of GlcNAc<sup>61,64,65</sup>. This model mainly captures three characteristics of engineered *E. coli* in fermentation experiments: (1) addition of pAcF can change the specific growth rate by regulating the flux of the glycolytic pathway; (2) accumulation of acetate can inhibit both cell growth and biosynthesis of target metabolites and acetate; and (3) cell growth is also limited by the concentration of nutrients (especially nitrogen sources) in the medium. Besides, concentrations of metabolites and all parameters in the model are expressed in arbitrary units (a.u.) which is sufficient to qualitatively simulate the dynamic characteristics of metabolites and biomass. Based on these points, model construction and parameter selection were described as stated in the following sections.

First, according to the experimental results regarding the regulation of pAcF concentration on gene expression, the flux rate of glycolytic pathway ( $v_{EMP}$ ) per unit biomass per unit time was calculated as a function of the pAcF concentration ( $x_{pAcF}$ ):

$$v_{EMP} = A \cdot \ln(x_{pAcF}) + B \quad (1)$$

where  $x_{pAcF}$  ranged from 0.1 to 10, and  $v_{EMP}$  ranged from 0.3 to 1 according to the experimental results of Hollinshead et al. regarding the culture test of *p/kA*-deficient *E. coli*<sup>66</sup>. Therefore, the parameters in the kinetic model were calculated as  $A = 0.152$  and  $B = 0.65$ . According to the experimental results of Basan et al., as the glycolytic pathway flux rate ( $v_{EMP}$ ) gradually increases, the specific growth rate ( $\mu$ ) subsequently increases. Further, as the glycolytic pathway flux rate ( $v_{EMP}$ ) continues to increase, ‘overflow metabolism’ activates the acetate synthesis pathway, and  $\mu$  increases linearly with the increase of acetate synthesis rate ( $v_{ace}$ )<sup>65</sup>. Therefore, the promotion of the glycolytic pathway flux rate ( $v_{EMP}$ ) to  $\mu$  is divided into two phases, including non-acetate-producing ( $v_{EMP} = 0.3–0.6$ ) and acetate-producing ( $v_{EMP} = 0.6–1$ ) phases. The  $\mu$  of the intersection of the two phases ( $\mu_0$  was set to 0.76 according to the experimental data of Basan et al., 0.76 h<sup>-1</sup>, to match the further parameters of cell growth) is taken as the rate of growth in the absence of all limitations. Therefore, in the model described, ‘ $I_{EMP}$ ’ is used as a coefficient to calculate the effect of glycolytic pathway flux rate on cell growth rate. During the non-acetate-producing phase,  $I_{EMP}$  (0.3–1) increases linearly with increasing  $v_{EMP}$ . During the acetate-producing stage, the synthesis rate of acetate ( $v_{ace}$  was ranged from 0 to 0.24 according to the experimental data of Basan et al., 0–0.24 g/OD<sub>600</sub>/h) increases linearly with increasing  $v_{EMP}$ , while  $I_{EMP}$  ( $I_{EMP}$  was ranged from 1 to 1.6 to achieve an up to 1.6 times’ increase on  $\mu$  according to the results of Basan et al.) increases linearly with increasing  $v_{ace}$ <sup>65,67</sup>. Therefore, the relationship between  $I_{EMP}$  and  $x_{pAcF}$  could be modeled by the following equation with parameters calculated as  $C = 7/3$ ,  $D = -0.4$ ,  $E = 1.5$ , and  $F = 0.1$ .

$$I_{EMP} = \begin{cases} C \cdot v_{EMP} + D & \text{for } 0.3 \leq v_{EMP} \leq 0.6 \\ E \cdot v_{EMP} + F & \text{for } 0.6 < v_{EMP} \leq 1 \end{cases} \quad (2)$$

The acetate biosynthesis rate ( $v_{ace}$ ) could be modeled by the following equation:  $G = 0.6$ ,  $H = -0.36$ .

$$v_{ace} = \begin{cases} 0 & \text{for } 0.3 \leq v_{EMP} \leq 0.6 \\ G \cdot v_{EMP} + H & \text{for } 0.6 < v_{EMP} \leq 1 \end{cases} \quad (3)$$

And the inhibition factor  $I_{ace}$  of acetate for cell growth was calculated according to the results of Luli et al.<sup>64</sup>:

$$I_{ace} = \frac{k_1}{\mu} \quad (4)$$

where  $k_1$  was calculated as a function of  $x_{ace}$  according to the results of Pinhal et al., and  $x_{ace}$  has the following dynamic:

$$k_1 = 0.76 \cdot 0.0534^{x_{ace}} + 0.1 \quad (5)$$

$$\frac{dx_{ace}}{dt} = I_{ace} \cdot v_{ace} \cdot x_{cell} \quad (6)$$

In addition to the effects of glycolytic flux rate and acetic acid accumulation, cell growth rate is also limited by nitrogen source ( $x_N$ ) under conditions of sufficient carbon source (glucose) in fermentation experiment. Therefore, we also set the



coefficient ' $I_N$ ' to calculate the effects of nutrients on cell growth rate according to Monod equation:

$$I_N = \frac{x_{N_0}}{x_{N_0} + K_S} \quad (7)$$

where  $x_N$  is the residual nitrogen source in the fermentation system,  $x_{N_0}$  is the initial nitrogen source, and  $K_S$  is the half-saturation constant ( $K_S = 5, 10, 15, 20$ ) of Monod equation (Supplementary Fig. 4). In the simulated fermentation system, the nitrogen source is used for conversion into biomass and GlcNAc, the theoretical conversion rates of which are  $Y_{X/N} = 0.7$  and  $Y_{GNA/N} = 0.54$ , respectively. Therefore,  $x_{N_0}$  could be modeled by the following equation:

$$x_N = x_{N_0} - \frac{x_{\text{cell}}}{Y_{X/N}} - \frac{x_{\text{GNA}}}{Y_{GNA/N}} \quad (8)$$

After calculating the three parameters,  $I_{\text{EMP}}$ ,  $I_{\text{ace}}$ , and  $I_N$ , which affect cell growth rate, the biomass ( $x_{\text{cell}}$ ) has the following dynamics:

$$\mu = I_{\text{EMP}} \cdot I_{\text{ace}} \cdot I_N \cdot \mu_0 \quad (9)$$

$$\frac{dx_{\text{cell}}}{dt} = x_{\text{cell}} \cdot \mu \quad (10)$$

In addition, Michaelis–Menten equation was used for qualitative simulation of GlcNAc biosynthesis. All  $K_m$  values in the metabolic pathway were equal to 1, and all maximal reaction rates ( $v_{\text{max}}$ ) were 0.24, in order to match the synthesis rate of acetate. At the same time, the biosynthesis rate ( $v_{\text{GNA}}$ ) of GlcNAc was multiplied by  $I_{\text{ace}}$  to inhibit accumulation. Therefore, the GlcNAc concentration ( $x_{\text{GNA}}$ ) has the following dynamics:

$$v_{\text{GNA}} = I_{\text{ace}} \cdot \frac{v_{\text{GNAmax}} \cdot x_{\text{GlcNAc6P}}}{x_{\text{GlcNAc6P}} + K_{m6}} \quad (11)$$

$$\frac{dx_{\text{GNA}}}{dt} = x_{\text{cell}} \cdot v_{\text{GNA}} \quad (12)$$

where  $v_{\text{max}}$  of Yqab per unit biomass per unit time ( $v_{\text{GNAmax}}$ ) was 0.24, and  $K_m$  of Yqab ( $K_{m6}$ ) was 1. The initial parameters of the model are initial nitrogen source  $x_{N_0} = 50$  and initial biomass  $x_{\text{cell}0} = 0.1$ .

**Solid-media escape assays.** Escape frequency was determined according to methods already reported in previous studies<sup>46</sup>. Briefly, the OMeY-dependent strain was first cultured in 50 mL of fresh LB medium containing 1 mM OMeY for 10 h at 37 °C and 220 rpm. Cells were then washed by centrifugation and resuspended in fresh LB medium. After the cells were washed twice, OD<sub>600</sub> was measured. Three technical replicates were performed for each strain. For each sample, ~10<sup>10</sup> cells (it is estimated that the bacterial suspension with an OD<sub>600</sub> of 1 contains ~20,000,000 cells per mL) were plated onto OMeY-free, solid LB plates. Next, the numbers of escapees growing on the plates were counted after 1, 5, and 15 days of culture. At the same time, the bacterial suspension was diluted and plated on an LB plate with 1 mM OMeY to count the total number of cells. Escape frequencies for each OMeY-dependent strain were calculated as mean of escapees per colony forming unit (c.f.u.), plated on three OMeY-free LB solid plates. Standard error of the mean (SEM) was calculated as described by Mandell et al.<sup>46</sup>.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

The authors declare that all data supporting the findings of this study are available within the paper and its Supplementary Information files. A reporting summary for this article is available as a Supplementary Information file. The datasets generated and analyzed during the current study are also available from the corresponding author upon request. Source data are provided with this paper.

## Code availability

The MATLAB code used in this study can be found in GitHub [[https://github.com/Tian-RZ/Eco\\_kinetic\\_model.git](https://github.com/Tian-RZ/Eco_kinetic_model.git)].

Received: 10 May 2020; Accepted: 21 September 2020;

Published online: 08 October 2020

## References

- Lu, H., Villada, J. C. & Lee, P. K. H. Modular metabolic engineering for biobased chemical production. *Trends Biotechnol.* **37**, 152–166 (2019).
- Choi, K. R. et al. Systems metabolic engineering strategies: integrating systems and synthetic biology with metabolic engineering. *Trends Biotechnol.* **37**, 817–837 (2019).
- Nielsen, J. & Keasling, J. D. Engineering cellular metabolism. *Cell* **164**, 1185–1197 (2016).
- Woolston, B. M., Edgar, S. & Stephanopoulos, G. Metabolic engineering: past and future. *Annu. Rev. Chem. Biomol.* **4**, 259–288 (2013).
- Yim, H. et al. Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. *Nat. Chem. Biol.* **7**, 445–452 (2011).
- Przystałowska, H. et al. 1,3-Propanediol production by new recombinant *Escherichia coli* containing genes from pathogenic bacteria. *Microbiol. Res.* **171**, 1–7 (2015).
- Liu, H.-J. et al. Microbial production of 1,3-propanediol from glycerol by *Klebsiella pneumoniae* under micro-aerobic conditions up to a pilot scale. *Biotechnol. Lett.* **29**, 1281–1285 (2007).
- Falls, K. C., Williams, A. L., Bryksin, A. V. & Matsumura, I. *Escherichia coli* deletion mutants illuminate trade-offs between growth rate and flux through a foreign anabolic pathway. *PLoS ONE* **9**, e88159 (2014).
- Xu, P., Li, L., Zhang, F., Stephanopoulos, G. & Koffas, M. Improving fatty acids production by engineering dynamic pathway regulation and metabolic control. *PNAS* **111**, 11299 (2014).
- Kochanowski, K. et al. Functioning of a metabolic flux sensor in *Escherichia coli*. *PNAS* **110**, 1130 (2013).
- Lo, T.-M., Chng, S. H., Teo, W. S., Cho, H.-S. & Chang, M. W. A two-layer gene circuit for decoupling cell growth from metabolite production. *Cell Syst.* **3**, 133–143 (2016).
- Rugbjerg, P., Sarup-Lytzen, K., Nagy, M. & Sommer, M. O. A. Synthetic addition extends the productive life time of engineered *Escherichia coli* populations. *PNAS* **115**, 2347–2352 (2018).
- Fujiwara, R., Noda, S., Tanaka, T. & Kondoh, A. Metabolic engineering of *Escherichia coli* for shikimate pathway derivative production from glucose-xylose co-substrate. *Nat. Commun.* **11**, 279 (2020).
- Cui, S. et al. Engineering a bifunctional Phr60-Rap60-Spo0A quorum-sensing molecular switch for dynamic fine-tuning of menaquinone-7 synthesis in *Bacillus subtilis*. *ACS Synth. Biol.* **8**, 1826–1837 (2019).
- Gupta, A., Reizman, I. M. B., Reisch, C. R. & Prather, K. L. J. Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. *Nat. Biotechnol.* **35**, 273–279 (2017).
- Zhao, E. M. et al. Optogenetic regulation of engineered cellular metabolism for microbial chemical production. *Nature* **555**, 683–687 (2018).
- Bothfeld, W., Kapov, G. & Tyo, K. E. J. A glucose-sensing toggle switch for autonomous, high productivity genetic control. *ACS Synth. Biol.* **6**, 1296–1304 (2017).
- Moser, F. et al. Genetic circuit performance under conditions relevant for industrial bioreactors. *ACS Synth. Biol.* **1**, 555–564 (2012).
- Rugbjerg, P., Myling-Petersen, N., Porse, A., Sarup-Lytzen, K. & Sommer, M. O. A. Diverse genetic error modes constrain large-scale bio-based production. *Nat. Commun.* **9**, 787 (2018).
- Xia, P.-F., Ling, H., Foo, J. L. & Chang, M. W. Synthetic genetic circuits for programmable biological functionalities. *Biotechnol. Adv.* **37**, 107393 (2019).
- Buerger, J., Gronenberg, L. S., Genee, H. J. & Sommer, M. O. Wiring cell growth to product formation. *Curr. Opin. Biotech.* **59**, 85–92 (2019).
- Sandberg, T. E., Salazar, M. J., Weng, L. L., Palsson, B. O. & Feist, A. M. The emergence of adaptive laboratory evolution as an efficient tool for biological discovery and industrial biotechnology. *Metab. Eng.* **56**, 1–16 (2019).
- Liu, N., Santala, S. & Stephanopoulos, G. Mixed carbon substrates: a necessary nuisance or a missed opportunity? *Curr. Opin. Biotech.* **62**, 15–21 (2020).
- Rugbjerg, P. & Sommer, M. O. A. Overcoming genetic heterogeneity in industrial fermentations. *Nat. Biotechnol.* **37**, 869–876 (2019).
- Wang, L., Brock, A., Herberich, B. & Schultz, P. G. Expanding the genetic code of *Escherichia coli*. *Science* **292**, 498 (2001).
- Arranz-Gibert, P., Vanderschuren, K. & Isaacs, F. J. Next-generation genetic code expansion. *Curr. Opin. Chem. Biol.* **46**, 203–211 (2018).
- Liu, C. C. & Schultz, P. G. Recombinant expression of selectively sulfated proteins in *Escherichia coli*. *Nat. Biotechnol.* **24**, 1436–1440 (2006).
- Pearson, A. D. et al. Trapping a transition state in a computationally designed protein bottle. *Science* **347**, 863–867 (2015).
- Kato, Y. Tight translational control using site-specific unnatural amino acid incorporation with positive feedback gene circuits. *ACS Synth. Biol.* **7**, 1956–1963 (2018).
- Lajoie, M. J. et al. Genomically recoded organisms expand biological functions. *Science* **342**, 357–360 (2013).
- Fredens, J. et al. Total synthesis of *Escherichia coli* with a recoded genome. *Nature* **569**, 514–518 (2019).
- Minaba, M. & Kato, Y. High-yield, zero-leakage expression system with a translational switch using site-specific unnatural amino acid incorporation. *Appl. Environ. Microbiol.* **80**, 1718–1725 (2014).
- Kato, Y. Tunable translational control using site-specific unnatural amino acid incorporation in *Escherichia coli*. *PeerJ* **3**, e904–e904 (2015).
- Volkwein, W., Maier, C., Krafczyk, R., Jung, K. & Lassak, J. A versatile toolbox for the control of protein levels using Nε-acetyl-L-lysine dependent amber suppression. *ACS Synth. Biol.* **6**, 1892–1902 (2017).



35. Young, T. S., Ahmad, I., Yin, J. A. & Schultz, P. G. An enhanced system for unnatural amino acid mutagenesis in *E. coli*. *J. Mol. Biol.* **395**, 361–374 (2010).
36. Wannier, T. M. et al. Adaptive evolution of genomically recoded *Escherichia coli*. *PNAS* **115**, 3090–3095 (2018).
37. Pfeiffer, T., Schuster, S. & Bonhoeffer, S. Cooperation and competition in the evolution of ATP-producing pathways. *Science* **292**, 504 (2001).
38. Deng, M.-D. et al. Directed evolution and characterization of *Escherichia coli* glucosamine synthase. *Biochimie* **88**, 419–429 (2006).
39. Liu, Y. et al. Pathway engineering of *Bacillus subtilis* for microbial production of N-acetylglucosamine. *Metab. Eng.* **19**, 107–115 (2013).
40. Kumar, P., Henikoff, S. & Ng, P. C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* **4**, 1073–1081 (2009).
41. Shukal, S., Chen, X. & Zhang, C. Systematic engineering for high-yield production of viridiflorol and amorphadiene in auxotrophic *Escherichia coli*. *Metab. Eng.* **55**, 170–178 (2019).
42. Lee, J. W., Chan, C. T. Y., Slomovic, S. & Collins, J. J. Next-generation biocontainment systems for engineered organisms. *Nat. Chem. Biol.* **14**, 530–537 (2018).
43. Torres, L., Krüger, A., Csibra, E., Gianni, E. & Pinheiro, V. B. Synthetic biology approaches to biological containment: Pre-emptively tackling potential risks. *Essays Biochem.* **60**, 393–410 (2016).
44. Dana, G. V., Kuiken, T., Rejeski, D. & Snow, A. A. Four steps to avoid a synthetic-biology disaster. *Nature* **483**, 29 (2012).
45. Rovner, A. J. et al. Recoded organisms engineered to depend on synthetic amino acids. *Nature* **518**, 89–93 (2015).
46. Mandell, D. J. et al. Biocontainment of genetically modified organisms by synthetic protein design. *Nature* **518**, 55–60 (2015).
47. Wilson, D. J. NIH guidelines for research involving recombinant DNA molecules. *Acc. Res.* **3**, 177–185 (1993).
48. van Dijl, J. M. & Hecker, M. *Bacillus subtilis*: From soil bacterium to super-secreting cell factory. *Microb. Cell Fact.* **12**, 3 (2013).
49. Völler, J.-S. & Budisa, N. Coupling genetic code expansion and metabolic engineering for synthetic cells. *Curr. Opin. Biotech.* **48**, 1–7 (2017).
50. Zhou, K., Qiao, K., Edgar, S. & Stephanopoulos, G. Distributing a metabolic pathway among a microbial consortium enhances production of natural products. *Nat. Biotechnol.* **33**, 377–383 (2015).
51. Zhang, H. & Wang, X. Modular co-culture engineering, a new approach for metabolic engineering. *Metab. Eng.* **37**, 114–121 (2016).
52. Dinh, C. V., Chen, X. & Prather, K. L. J. Development of a quorum-sensing based circuit for control of coculture population composition in a naringenin production system. *ACS Synth. Biol.* **9**, 590–597 (2020).
53. Wang, L., York, S. W., Ingram, L. O. & Shanmugam, K. T. Simultaneous fermentation of biomass-derived sugars to ethanol by a co-culture of an engineered *Escherichia coli* and *Saccharomyces cerevisiae*. *Bioresour. Technol.* **273**, 269–276 (2019).
54. Gibson, D. G. Enzymatic assembly of overlapping DNA fragments. *Meth. Enzymol.* **498**, 349–361 (2011).
55. Zhang, X.-Z., Cui, Z.-L., Hong, Q. & Li, S.-P. High-level expression and secretion of methyl parathion hydrolase in *Bacillus subtilis* WB800. *Appl. Environ. Microbiol.* **71**, 4101–4103 (2005).
56. Nguyen, H. D., Phan, T. T. P. & Schumann, W. Expression vectors for the rapid purification of recombinant proteins in *Bacillus subtilis*. *Curr. Microbiol.* **55**, 89–93 (2007).
57. Zhang, X. et al. Modular pathway engineering of key carbon-precursor supply-pathways for improved N-acetylneuraminic acid production in *Bacillus subtilis*. *Biotechnol. Bioeng.* **115**, 2217–2231 (2018).
58. Koo, B.-M. et al. Construction and analysis of two genome-scale deletion libraries for *Bacillus subtilis*. *Cell Syst.* **4**, 291–305.e7 (2017).
59. Peters, J. M. et al. A comprehensive, CRISPR-based functional analysis of essential genes in bacteria. *Cell* **165**, 1493–1506 (2016).
60. Kobayashi, K. et al. Essential *Bacillus subtilis* genes. *PNAS* **100**, 4678–4683 (2003).
61. Liu, Y. et al. A dynamic pathway analysis approach reveals a limiting futile cycle in N-acetylglucosamine overproducing *Bacillus subtilis*. *Nat. Commun.* **7**, 11933 (2016).
62. Tian, R. et al. Synthetic N-terminal coding sequences for fine-tuning gene expression and metabolic engineering in *Bacillus subtilis*. *Metab. Eng.* **55**, 131–141 (2019).
63. Ramos, J. R. C., Rath, A. G., Genzel, Y., Sandig, V. & Reichl, U. A dynamic model linking cell growth to intracellular metabolism and extracellular by-product accumulation. *Biotechnol. Bioeng.* **117**, 1533–1553 (2020).
64. Luli, G. W. & Strohl, W. R. Comparison of growth, acetate production, and acetate inhibition of *Escherichia coli* strains in batch and fed-batch fermentations. *Appl. Environ. Microbiol.* **56**, 1004–1011 (1990).
65. Basan, M. et al. Overflow metabolism in *Escherichia coli* results from efficient proteome allocation. *Nature* **528**, 99–104 (2015).
66. Hollinshead, W. D. et al. Examining *Escherichia coli* glycolytic pathways, catabolite repression, and metabolite channeling using  $\Delta$ pfk mutants. *Biotechnol. Biofuels.* **9**, 212 (2016).
67. Renilla, S. et al. Acetate scavenging activity in *Escherichia coli*: interplay of acetyl-CoA synthetase and the PEP-glyoxylate cycle in chemostat cultures. *Appl. Microbiol. Biotechnol.* **93**, 2109–2124 (2012).

## Acknowledgements

We are grateful to Dr. Elad Noor from ETH Zurich for constructive advice on the manuscript. This study is financially supported by the National Key Research and Development Program of China (2018YFA0900300), National Natural Science Foundation of China (31972854, 21676119), Key Research and Development Program of Jiangsu Province (BE2019628), Fundamental Research Funds for the Central Universities (JUSRP52020A), National First-class Discipline Program of Light Industry Technology and Engineering (LITE2018-16), Postgraduate Research & Practice Innovation Program of Jiangsu Province (JNKY19\_015).

## Author contributions

Y.L. and R.T. designed the experiments. R.T., Y.C., and Z.Z. performed the biochemical experiments and analyzed the data. Y.L. and R.T. designed and performed the modeling. Y.L., J.L., L.L., G.D., and J.C. conceived the project and supervised the research. R.T., Y.L., Y.C., Z.Z., J.L., L.L., G.D., and J.C. wrote the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-020-18960-1>.

Correspondence and requests for materials should be addressed to Y.L. or J.C.

Peer review information *Nature Communications* thanks Yusuke Kato, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020