





CRISPR-assisted rational flux-tuning and arrayed CRISPRi screening of an L-proline exporter for L-proline hyperproduction

Jiao Liu^{1,2,4}, Moshi Liu^{1,2,3,4}, Tuo Shi^{1,2}, Guannan Sun^{1,2,3}, Ning Gao^{1,2,3}, Xiaojia Zhao^{1,2,3}, Xuan Guo^{1,2}, Xiaomeng Ni^{1,2}, Qianqian Yuan ^{1,2}, Jinhui Feng^{1,2}, Zhemin Liu^{1,2}, Yanmei Guo^{1,2}, Jiuzhou Chen^{1,2}, Yu Wang ^{1,2,3}✉, Ping Zheng ^{1,2,3}✉ & Jibin Sun ^{1,2,3}

Development of hyperproducing strains is important for biomanufacturing of biochemicals and biofuels but requires extensive efforts to engineer cellular metabolism and discover functional components. Herein, we optimize and use the CRISPR-assisted editing and CRISPRi screening methods to convert a wild-type *Corynebacterium glutamicum* to a hyperproducer of L-proline, an amino acid with medicine, feed, and food applications. To facilitate L-proline production, feedback-deregulated variants of key biosynthetic enzyme γ -glutamyl kinase are screened using CRISPR-assisted single-stranded DNA recombineering. To increase the carbon flux towards L-proline biosynthesis, flux-control genes predicted by in silico analysis are fine-tuned using tailored promoter libraries. Finally, an arrayed CRISPRi library targeting all 397 transporters is constructed to discover an L-proline exporter Cgl2622. The final plasmid-, antibiotic-, and inducer-free strain produces L-proline at the level of 142.4 g/L, 2.90 g/L/h, and 0.31 g/g. The CRISPR-assisted strain development strategy can be used for engineering industrial-strength strains for efficient biomanufacturing.

¹Key Laboratory of Systems Microbial Biotechnology, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin 300308, China. ²National Technology Innovation Center of Synthetic Biology, Tianjin 300308, China. ³University of Chinese Academy of Sciences, Beijing 100049, China. ⁴These authors contributed equally: Jiao Liu, Moshi Liu. ✉email: wang_y@tib.cas.cn; zheng_p@tib.cas.cn

Industrial biomanufacturing that converts biomass-derived carbon sources into target chemicals holds promise for addressing global concerns over limited fossil resources and environmental problems. Microbial strains are considered as the key biocatalysts of biomanufacturing¹. However, it is still challenging and time-consuming to develop microbial strains that meet the requirements of industrialization and commercialization, such as high production level (titer, yield, and productivity), free of plasmid, inducer, and antibiotic, and stable performance in scale-up fermentation^{2–4}. Random mutagenesis and screening strategies are widely used in the initial stage of strain development, especially in the breeding of amino acid producers^{3,5,6}. Stepwise rational metabolic engineering at a systems level can further improve strain performance but also requires efficient genome editing tools and components with specific functions like catalysis, regulation, and transport^{7–9}.

The CRISPR (Clustered regularly interspaced short palindromic repeats) system has recently been explored as a leading-edge tool for microbial genome editing. Besides genome editing applications such as gene deletion and insertion¹⁰, the CRISPR system offers enormous potential for genotype-phenotype mapping at a genome scale using arrayed or pooled guide RNA (gRNA) libraries¹¹. Genome-wide pooled CRISPR interference (CRISPRi) screening has been performed in several microorganisms including *Saccharomyces cerevisiae*¹², *Escherichia coli*¹³, and *Synechocystis*¹⁴ for identifying crucial genes for viability and tolerance. Conversely, arrayed CRISPRi libraries for microorganisms have been rarely reported, possibly due to the higher cost and demand for library construction and screening methods. Only arrayed CRISPRi libraries targeting ~300 essential genes were constructed for *Bacillus subtilis*¹⁵ and *Streptococcus pneumoniae*¹⁶. Its potential in mining functional components for industrial strain development has not been fully exploited.

Corynebacterium glutamicum is a major workhorse in industrial biomanufacturing¹⁷. It is now used for the industrial production of over six million tons of amino acids (such as L-lysine and L-glutamate) per year. Additionally, this microorganism has been engineered to produce many more compounds ranging from alcohols, organic acids, and plant secondary metabolites¹⁸. L-Proline, the only proteinogenic amino acid with a secondary amine, is a high-value amino acid with applications in the medicine and food industry and has large potential for use as a feed additive^{19,20}. Compared with the fermentative production of amino acids like L-lysine and L-glutamate with a high yield of ~0.70 g product/g glucose^{21,22}, the L-proline production has a relatively low yield of ~0.20 g/g²³ (Supplementary Table 1). Plasmid-based inducible expression systems requiring antibiotics and expensive chemical inducers are still used in L-proline producing strains^{23,24}, which preclude their use in industrial-scale fermentation. In the process of transforming *C. glutamicum* into efficient producers of molecules of interest, transport engineering has played a prominent role besides metabolic engineering^{25,26}. Traditional methods for discovering carrier proteins for molecule transport include phenotype screening of random mutation or genomic DNA libraries^{27,28}, transcriptome analysis²⁹, and sequence similarity search³⁰. Because of the limitations of these methods in comprehensiveness and universality, exporters for many compounds, including L-proline, are yet to be discovered^{25,26}.

In this study, to de novo develop an industrial-strength L-proline producing strain, we first optimize the CRISPR system in *C. glutamicum* for decreasing the cytotoxicity and increasing editing efficiency. CRISPR-assisted chromosomal editing and arrayed CRISPRi screening are then employed for engineering crucial enzymes, fine-tuning metabolic fluxes, and discovering L-proline exporters. Fed-batch fermentations of the final plasmid-

antibiotic-, and inducer-free strain result in the highly efficient production of L-proline. This study demonstrates how CRISPR facilitates de novo development of high-performance strains by enabling efficient chromosomal engineering and the discovery of functional components. The arrayed CRISPRi library for all *C. glutamicum* transporters will be useful for investigating the transport of other molecules of interest.

Results

Optimization of CRISPR/Cas9-assisted genome editing methods for *C. glutamicum*. Efficient methods for genetic manipulation are required for strain development. Genome editing tools based on both CRISPR/Cas9^{31–34} and CRISPR/Cas12a^{35–38} have been developed for *C. glutamicum*. Comparing these two systems for application in GC-rich *C. glutamicum* (53.8% GC-content for type strain ATCC 13032), the CRISPR/Cas9 system recognizing G-rich protospacer adjacent motif (PAM) sequences should have a wider genome targeting scope than the CRISPR/Cas12a system recognizing T-rich PAMs³⁹. Besides, the CRISPR/Cas9 system can distinguish a single nucleotide change in the seed sequence, which is preferred for precise editing⁴⁰. However, the CRISPR/Cas9 system is cytotoxic to *C. glutamicum* and previous methods usually suffer from few transformants and low accuracy, especially for manipulation of large DNA fragments^{32,34}.

The double-stranded DNA break (DSB) generated by the CRISPR/Cas9 can be a two-edged sword, which is lethal to most bacterial cells but can be used for counter-selection of edited cells⁴¹. Therefore, the unwanted and untimely leaky expression of Cas9 proteins should be minimized. To optimize the CRISPR/Cas9 system for *C. glutamicum*, we first constructed a temperature-sensitive plasmid that expressed Cas9 and gRNA and harbored homologous recombination (HR) arms. With a very low dosage of isopropyl- β -D-thiogalactopyranoside (IPTG) (0.01 mM) to induce Cas9-facilitated counter-selection, a few dozen transformants and a moderate editing efficiency of 33.3% were obtained for deletion of a 1.7 kb DNA fragment (*cgl0620-cgl0622*) with two 1.0 kb HR arms. Interestingly, with the increase in IPTG dosage, the transformant number and editing efficiency gradually decreased (Fig. 1a and Supplementary Fig. 1), again demonstrating the importance of controllable Cas9 expression in genome editing.

Next, the expression of Cas9 was regulated by using a perfectly symmetric *lac* operator (LacO) that binds the *lac* repressor (LacI) very tightly⁴² and a weaker ribosome binding site (RBS). By using a green fluorescent protein (GFP) as a reporter, it was found that these modifications almost eliminated leaky expression. Gradually enhanced gene expression was obtained by increasing the IPTG dosage (Supplementary Fig. 2a). When used for *cas9* expression, these modifications on LacO and RBS also reduced the cytotoxicity of *cas9* expressing plasmid and produced more transformants (Supplementary Fig. 2b). Then, the optimized system was tested for Cas9-mediated gene deletion. The editing efficiency was largely improved up to 91.3% with 0.05 mM IPTG dosage (Fig. 1b and Supplementary Fig. 1). Hundreds of transformants could be obtained because of the strict control of Cas9 expression. More tests were then conducted to evaluate the optimized system. When shorter HR arms (0.5 kb) were used for deletion of a longer DNA fragment (20 kb fragment in the prophage CGP3), an editing efficiency of 46.4% was obtained, which was over the previously reported efficiency of 26.9% for the same test³⁴. The editing efficiency remained 34.8% for deletion of a 219 kb DNA fragment (complete prophage CGP3), which has not been achieved by previous methods³⁴. For insertion of a 4 kb DNA fragment (an artificial *proBAC* operon), an efficiency of 43.5% was obtained (Fig. 1c and Supplementary Fig. 3).

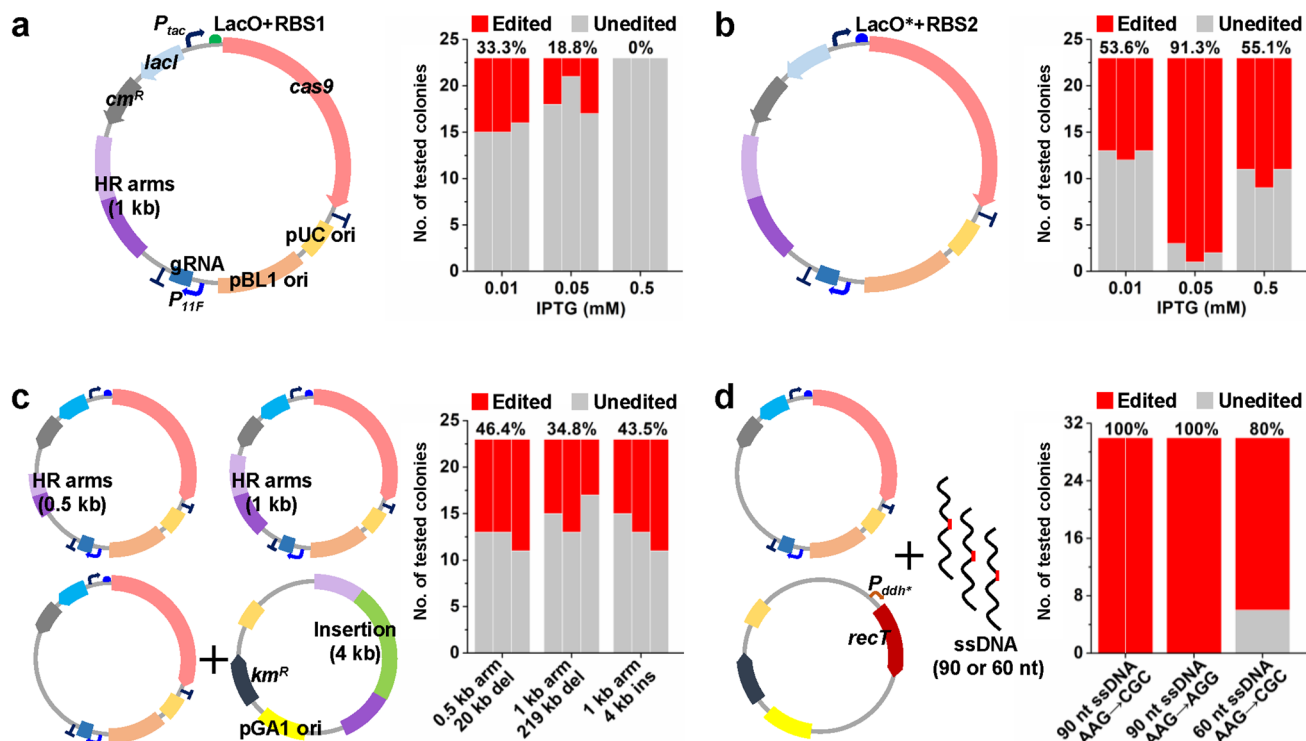


Fig. 1 Gene deletion, insertion, and ssDNA recombining in *C. glutamicum* using the optimized CRISPR/Cas9 system. **a** Deletion of a 1.7 kb DNA fragment (*cgI0620-cgI0622*) using the all-in-one plasmid. Different concentrations of IPTG (0.01, 0.05, and 0.5 mM) were used for counter-selection. **b** Deletion of the same 1.7 kb DNA fragment (*cgI0620-cgI0622*) using the optimized all-in-one plasmid with a modified *lac* operator that binds LacI very tightly (LacO^{*}) and a weak RBS (RBS2). Different concentrations of IPTG (0.01, 0.05, and 0.5 mM) were used for counter-selection. **c** Deletion and insertion of large DNA fragments. Three tests, deletion of a 20 kb DNA fragment (part of prophage CGP3) with 0.5 kb HR arms, deletion of a 219 kb DNA fragment (complete prophage CGP3) with 1.0 kb HR arms, and insertion of a 4 kb DNA fragment (an artificial *proBAC* operon inserted to *putA*) with 1.0 kb HR arms were conducted. For gene insertion, the inserted DNA fragment and HR arms were provided using a second plasmid. IPTG (0.05 mM) was used for counter-selection. For **a**, **b**, and **c**, twenty-three colonies were randomly selected and verified by PCR for each test. Results of three independent replicates are shown. Red and grey fractions of bars represent edited and unedited colonies, respectively. The average editing efficiency (%) from three independent replicates is shown above the bars. **d** Introduction of triple and single nucleotide changes using CRISPR/Cas9-assisted ssDNA recombining. Three tests, triple nucleotide changes with 90 nt ssDNA (two independent replicates), single nucleotide change with 90 nt ssDNA, and triple nucleotide changes with 60 nt ssDNA, were conducted to introduce *rpsL*^{K43R} mutation that produced streptomycin resistance phenotype. IPTG (0.05 mM) was used for counter-selection. Thirty colonies were randomly selected and verified by streptomycin resistance phenotype test. Red and grey fractions of bars represent edited and unedited colonies, respectively. The editing efficiency (%) is shown above the bars. Genetic elements on the plasmids: *P*_{tac}, IPTG-inducible *tac* promoter; *P*_{11F}, 11F constitutive promoter³²; *P*_{ddh}^{*}, a variant of the promoter of *ddh* gene (Supplementary Data 6); LacO, wild-type *lac* operator; LacO^{*}, a modified *lac* operator that binds LacI tightly⁴²; RBS1, a strong RBS AAAGGAGTTGAGA; RBS2, a weak RBS AAAGGCACCCGAT; pUC ori, pUC origin of replication; pBL1 ori, pBL1 origin of replication; pGA1 ori, pGA1 origin of replication; *cas9*, *S. pyogenes* Cas9 gene; gRNA, guide RNA expression cassette; HR arms, homologous recombination arms; *cmr*^R, chloramphenicol resistance gene; *kmr*^R, kanamycin resistance gene; *recT*, gene encoding the recombinase from the Rac prophage of *E. coli*; Insertion, the 4 kb DNA fragment for insertion test. Images for DNA gel electrophoresis and streptomycin resistance phenotype test are provided in Supplementary Figs. 1 and 3. Source data underlying Fig. 1 are provided as a Source Data file.

With exogenous recombinases, recombination with single-stranded DNA (ssDNA) can be more efficient than recombination with plasmid-borne double-stranded DNA (dsDNA)^{32,43}. To test genome editing with ssDNAs as templates, the RecT recombinase from the Rac prophage of *E. coli* was expressed using a low-copy and easily curable plasmid under the control of a strong constitutive promoter *P*_{ddh}^{*}. By providing 90 nt ssDNA templates targeting the *rpsL*^{K43R} mutation, triple or single nucleotide changes were efficiently introduced to the target site with 100% efficiencies and over 1000 transformants for each test. The editing efficiency remained 80% when the length of ssDNA was reduced to 60 nt (Fig. 1d and Supplementary Fig. 3). Compared with the CRISPR/Cas9 systems previously developed for *C. glutamicum*^{31–34}, the present system shows overall improved editing efficiency. A routine procedure for genome editing and plasmid curing takes as short as four days

(Supplementary Fig. 4). The sizable quantity of transformants also make this system easy to operate.

Screening γ -glutamyl kinase variants to facilitate L-proline production. With the optimized genome editing method, we conducted metabolic engineering of *C. glutamicum* ATCC 13032 for L-proline production. L-Proline is synthesized from the tricarboxylic acid cycle (TCA cycle) intermediate 2-oxoglutarate by L-glutamate dehydrogenase (Gdh), γ -glutamyl kinase (ProB), γ -glutamyl phosphate reductase (ProA), spontaneous cyclization, and pyrroline-5-carboxylate reductase (ProC) in *C. glutamicum*. L-Proline biosynthesis is tightly regulated by feedback inhibition of ProB by the end-product L-proline (Fig. 2)¹⁹. Therefore, releasing the feedback inhibition of ProB by L-proline is the initial step for the development of L-proline producers⁴⁴. The ~100%

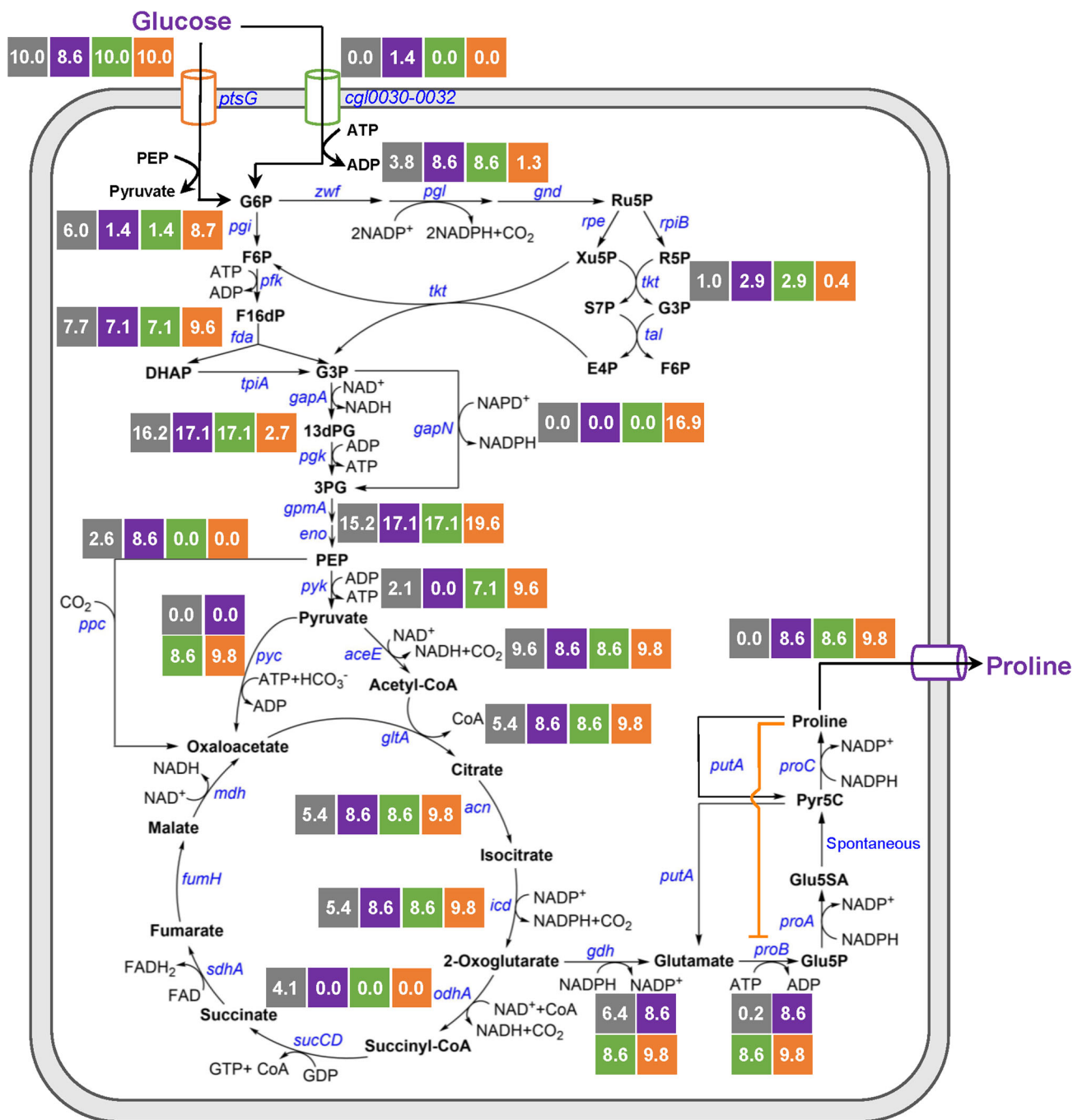


Fig. 2 *In silico* simulation of the L-proline biosynthetic pathway from glucose. Glucose input was set at 10 mmol/gCDW-h. The numbers in squares represent metabolic fluxes of reactions. Grey squares represent the metabolic fluxes for maximum biomass formation. Purple squares represent the metabolic fluxes for maximum L-proline biosynthesis using the *ppc*-based C4 anaplerotic pathway. Green squares represent the metabolic fluxes for maximum L-proline biosynthesis using the *pyc*-based C4 anaplerotic pathway. Orange squares represent the metabolic fluxes for maximum L-proline biosynthesis using the *gapN*-based G3P oxidation pathway and *pyc*-based C4 anaplerotic pathway. The feedback inhibition of Pro by L-proline is indicated with orange line. Source data for *in silico* simulation are provided in Supplementary Data 3-5.

single nucleotide editing efficiency of CRISPR/Cas9-assisted ssDNA recombineering makes it possible to perform codon saturation mutagenesis in *C. glutamicum* chromosome without introducing adjacent synonymous mutations, which do not alter the encoded protein but can influence gene expression⁴⁵. To identify the target amino acid residues for mutagenesis, 946 ProB sequences from UniProt database were aligned and analyzed (Fig. 3a and Supplementary Data 1). It has been reported that the substrate L-glutamate and the feedback inhibitor L-proline bind at

overlapping sites⁴⁶. The predicted binding sites of ProB from *C. glutamicum* ATCC 13032 (CgProB) are D154 and N155³⁵, which are conserved in ProB enzymes from different sources. Interestingly, the adjacent region is less conserved, especially positions 148 and 150. The sequence alignment suggests that T148, G149, V150, and N151 of CgProB are all low-probability amino acid residues for these four sites (probability of 4.55, 0.95, 3.07, and 1.37%, respectively) (Fig. 3a). The 3D protein structure modelling of CgProB indicates that the less conserved region forms a flexible

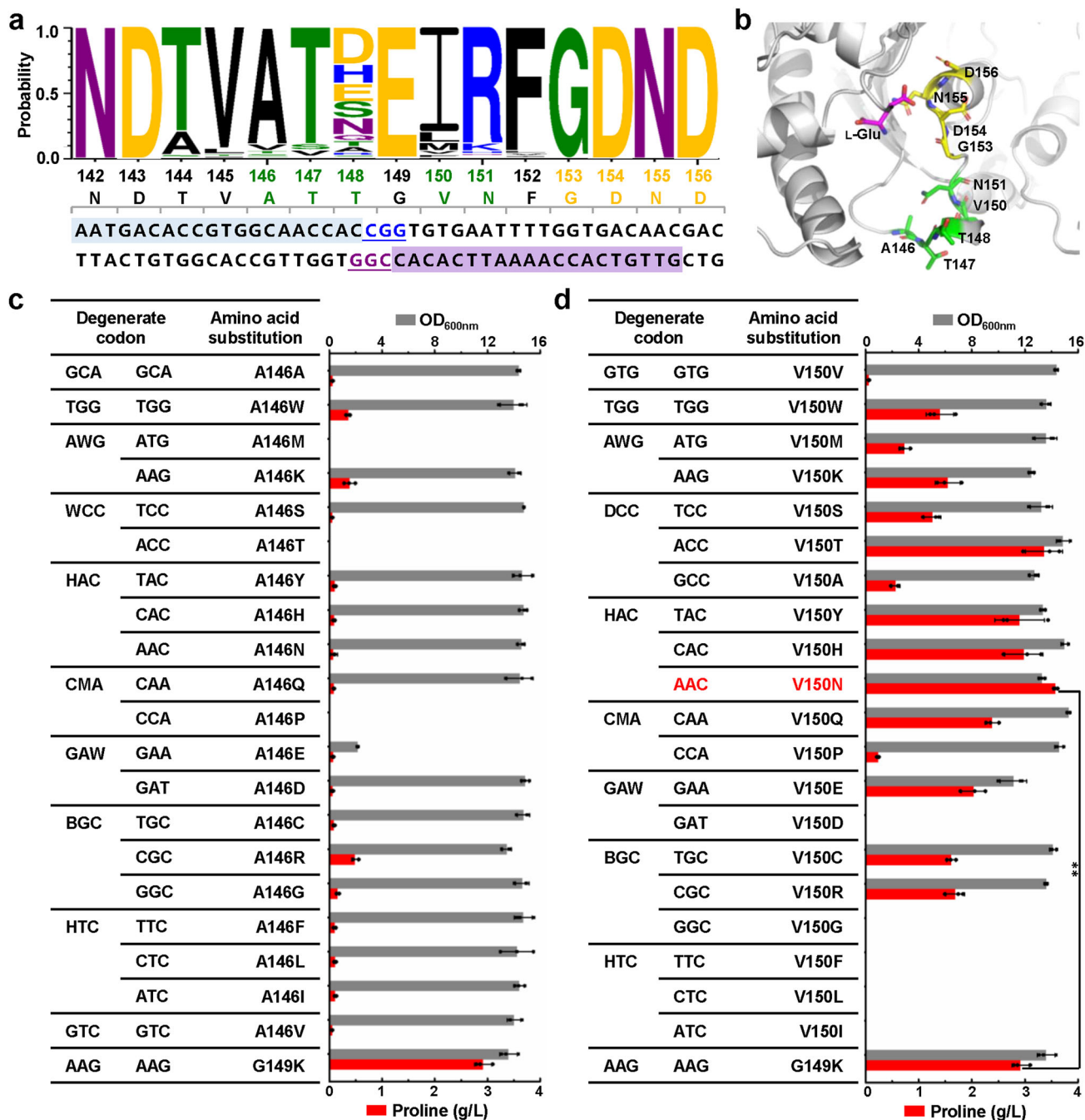


Fig. 3 Codon saturation mutagenesis of γ -glutamyl kinase of *C. glutamicum* (CgProB) for improved L-proline production via CRISPR/Cas9-assisted ssDNA recombineering. **a** Conservation analysis of γ -glutamyl kinases. The conserved region involved in L-proline binding is indicated in yellow. The A146, T147, T148, V150, and N151 residues mutated in this study are highlighted in green. The two gRNAs used and their corresponding PAM sequences (underlined) are highlighted in blue and purple, respectively. γ -Glutamyl kinases sequences used for conservation analysis are provided in Supplementary Data 1. **b** The model structure of CgProB. The model structure was constructed with the crystal structure of γ -glutamyl kinase from *E. coli* (PDB ID: 2J5T)⁸⁰ as a template (93% coverage and 38% sequence identity with CgProB) using Discovery Studio 2018 software. The backbone of CgProB is shown in a ribbon model on which several key residues are shown in a stick model. The substrate L-glutamate is indicated in magenta. The conserved amino acid residues G153, D154, N155, and D156 are indicated in yellow. The engineered amino acid residues A146, T147, T148, V150, and N151 are indicated in green. The full structure is shown in Supplementary Fig. 5. **c** Growth and L-proline production by *C. glutamicum* strains with amino acid substitution at codon 146 of CgProB. The strain harboring reported CgProB^{G149K} mutation³⁵ was used as a control. L-Proline production was conducted in 96-deep-well plates. Data are presented as mean values \pm SD ($n = 3$ independent experiments). **d**, Growth and L-proline production by *C. glutamicum* strains with amino acid substitution at codon 150 of CgProB. The strain harboring reported CgProB^{G149K} mutation³⁵ was used as a control. The V150N mutation used in strain PRO-1 is indicated in red. L-Proline production was conducted in 96-deep-well plates. Data are presented as mean values \pm SD ($n = 3$ independent experiments). ** $P = 0.0024$, Student's two-tailed t -test. Source data underlying Fig. 3c, d are provided as a Source Data file.

loop and maps away from the binding sites (Fig. 3b and Supplementary Fig. 5a). Analysis of the receptor-ligand interactions suggests that these amino acid residuals unlikely interact with the substrate L-glutamate directly (Supplementary Fig. 5b). However, previous studies on ProB of *E. coli* show that this flexible loop modulates L-proline inhibition⁴⁷. Therefore, we hypothesize that mutations in this region may release the feedback inhibition of CgProB by L-proline without affecting the substrate L-glutamate binding. Previous research on G149 mutagenesis provides support for our hypothesis³⁵. Therefore, codon saturation mutagenesis at A146, T147, T148, V150, and N151 were performed for CgProB with CRISPR/Cas9-assisted ssDNA recombineering.

To reduce the library size and avoid rare codons, 8–9 ssDNAs containing degenerate bases were designed and synthesized to cover all the 19 amino acid substitutions (Fig. 3c, d). Sixty transformants were randomly picked and tested for editing at each site. Site-directed mutagenesis targeting A146 and V150 produced several mutants with extracellular L-proline accumulation using glucose as a carbon source. Sequencing of *proB* in these mutants showed that sixteen different amino acid substitutions were obtained for A146 mutation and fourteen different amino acid substitutions were obtained for V150 mutation. All the sequenced *proB* variants had one to three nucleotide changes, demonstrating a 100% editing efficiency. Compared with A146 substitutions leading to low-level L-proline production (A146W, A146K, and A146R) (Fig. 3c), several V150 substitutions facilitated up to 3.6 g/L L-proline production (V150N), which was even 24.1% higher than that produced by the strain harboring previously reported G149K substitution (2.9 g/L) (Fig. 3d). The L-proline producing mutants overall showed decreased biomass compared with the wild-type strain, suggesting a redirection of metabolic flux from the production of biomass to the target metabolite. The strain expressing the *proB*^{V150N} mutant was designated as strain PRO-1.

Metabolic flux analysis for optimal L-proline biosynthesis.

Next, to predict the flux-control reactions leading to the optimal L-proline biosynthesis, flux balance analysis (FBA) was performed using a genome-scale metabolic model of *C. glutamicum*, *iCW773*²⁴. FBA is a mathematical approach for analyzing the fluxes of metabolites through a metabolic network and has been widely used to predict the growth rate of an organism or the theoretical yield of producing a metabolite of interest⁴⁸. By comparing the metabolic flux distribution between maximum biomass formation and maximum L-proline production, the flux-control reactions can be predicted. The *in silico* simulation suggests that when the substrate glucose is used for maximum biomass formation, the glycolysis and TCA cycle are active and few metabolic fluxes are channeled to L-proline biosynthesis. When maximum L-proline biosynthesis is the objective of simulation, 2-oxoglutarate flux is drained from the TCA cycle to enter the L-proline biosynthetic pathway, resulting in a theoretical L-proline yield from glucose of 0.86 mol/mol (0.55 g/g) (Fig. 2 and Supplementary Data 2 and 3). Since the TCA cycle is highly active in *C. glutamicum*⁴⁹, a strong L-proline biosynthetic pathway (Gdh, ProB, ProA, and ProC) is expected to channel 2-oxoglutarate flux from the TCA cycle²³.

According to the simulation, an efficient C4 anaplerotic pathway is needed for the optimal L-proline biosynthesis. The phosphoenolpyruvate carboxylase (Ppc)-based C4 anaplerotic pathway converts phosphoenolpyruvate (PEP) and CO₂ to oxaloacetate⁵⁰. To save PEP for oxaloacetate biosynthesis, glucose needs to be partially transported into cells via the nonphosphotransferase system (non-PTS) (Fig. 2), which is inefficient in *C. glutamicum*⁵¹. The pyruvate carboxylase (Pyc)-based C4

anaplerotic pathway offers a promising alternative⁵⁰. First, Pyc converts pyruvate and CO₂ to oxaloacetate. As a result, sufficient PEP can be used as the phosphoryl donor for phosphotransferase system (PTS)-mediated glucose uptake, which allows complete glucose uptake via the efficient PTS without the involvement of inefficient non-PTS. Second, the simulation suggests that using the Pyc-based C4 anaplerotic pathway leads to excess ATP production from glucose (2.04 excess ATP mol/mol glucose), which can support cell growth. In the simulation, the excess ATP is consumed by meaningless cyclic reactions, e.g., ATP-consuming L-cysteine transport reactions shown in Supplementary Data 4.

Since significant amount of NADPH is needed for L-proline biosynthesis, almost half carbon flux is channeled into the oxidative pentose phosphate pathway (PPP) for NADPH regeneration at the cost of CO₂ release, which consequently lowers the L-proline yield. The nonphosphorylating NADP-dependent glyceraldehyde-3-phosphate dehydrogenase (GapN) can be used to replace the native NAD-dependent isoenzyme for NADPH regeneration with improved carbon conservation^{52–54}. However, GapN catalyzes the conversion of glyceraldehyde-3-phosphate (G3P) to glycerate-3-phosphate with an NADPH formation but no ATP formation. Therefore, the GapN-based pathway produces less ATP than the native GapA-Pgk (NAD-dependent glyceraldehyde-3-phosphate dehydrogenase and phosphoglycerate kinase) pathway that produces an NADH and an ATP (Fig. 2). By combining the GapN-based G3P oxidation pathway and the previously described Pyc-based C4 anaplerotic pathway, it is possible to provide sufficient ATP and NADPH for cell growth and L-proline biosynthesis. After adding the GapN-catalyzed reaction to the model, a simulation suggests the L-proline yield from glucose increases from 0.86 mol/mol to 0.98 mol/mol (0.63 g/g) (Fig. 2 and Supplementary Data 5). The metabolic flux through the oxidative PPP decreases by 84.9%. The G3P flux is divided into the GapN and GapA-Pgk pathways by a ratio of 86.2% to 13.8%. Based on these analyses, L-proline biosynthetic genes (*gdh*, *proB*, *proA*, and *proC*), *pyc*, and *gapN* are flux-control genes and need to be finely tuned for enhanced L-proline production.

Chromosomally transcriptional tuning of flux-control genes using tailored promoter libraries.

Promoter engineering is widely used to enhance gene expression at the transcription level. A number of homologous and heterologous promoters have been developed for constitutive or inducible gene expression in *C. glutamicum*⁵⁵. However, the choice of strong and constitutive natural promoters is still limited¹⁷, and *P_{sod}* and *P_{tuf}* promoters are the mostly used ones for strain development^{54,56}. Moreover, because the 5' region of coding sequence (CDS) significantly affects gene expression^{57,58}, a promoter usually shows different strengths for a reporter gene (e.g., a fluorescent protein-encoding gene) and the target gene, making it difficult to predict the promoter performance for metabolic tuning. To overcome this problem, we fused the first 180 bp of target gene with a flexible linker (GGGGG)₃ and a red fluorescent protein (RFP) gene to construct a tailored reporter system. The native promoter of the target gene was used as a template to build promoter libraries for good adaptability to its naturally controlled gene (Fig. 4a). Regarding the screening method, fluorescence-activated cell sorting (FACS) is a commonly used high-throughput method for the first round of screening⁵⁹. However, exponential reproduction of microbial cells during the cultivation of transformants may lower the proportion of variants that are overburdened by high-level protein expression⁶⁰. To maximize the efficiency of the first round of screening, we used a fluorescence imaging system to

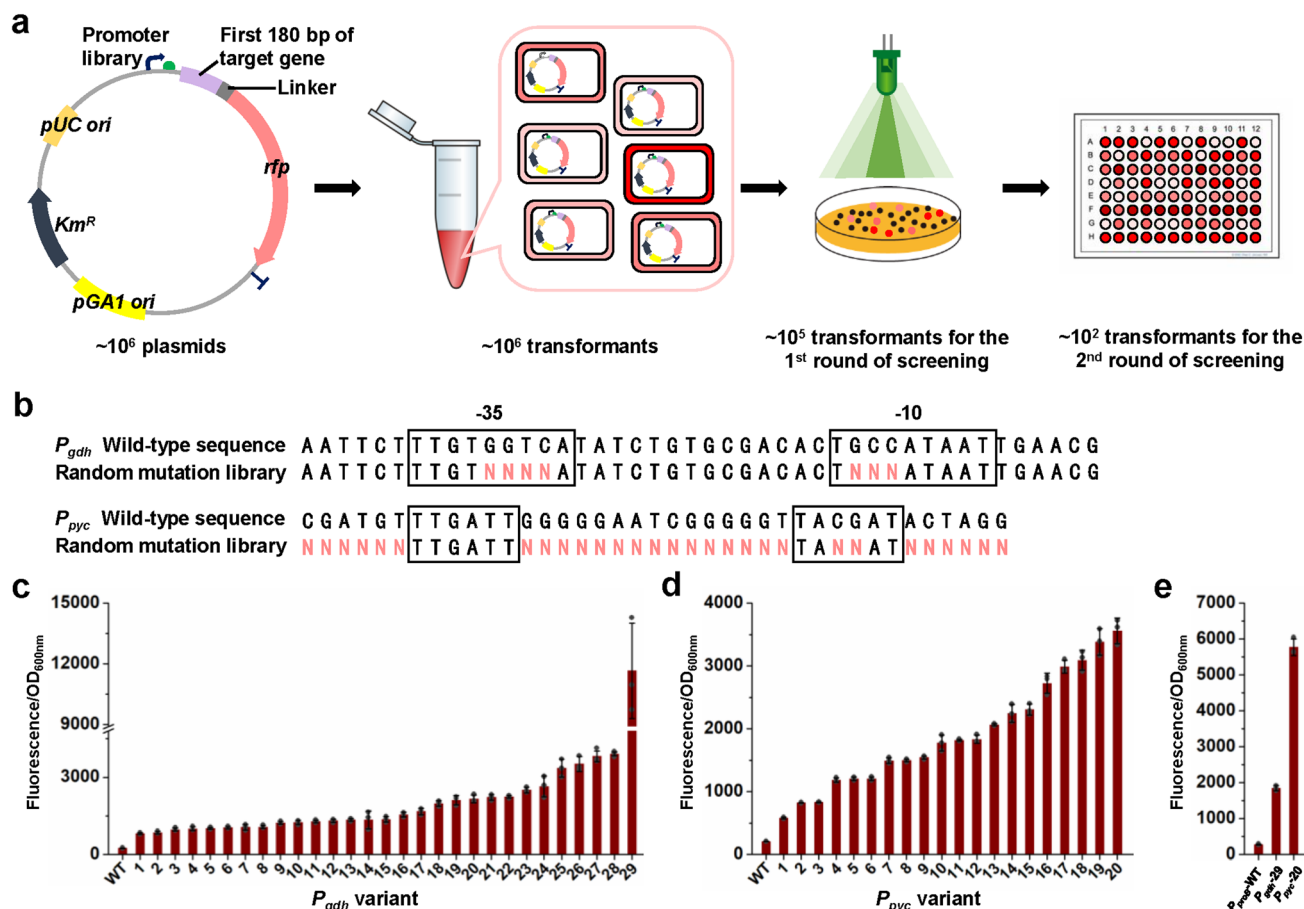


Fig. 4 Construction and screening of tailored promoter libraries for target genes. **a** Workflow of construction and screening of tailored promoter libraries. The reporter gene was constructed by fusing the first 180 bp of the target gene (*gdh*, *pyc*, or *proB*) with a flexible linker (GGGGS)₃ and an *rfp* gene. Random mutation libraries of native promoters were constructed by PCR with primers containing several degenerate bases. The plasmid libraries ($\sim 10^6$) were transformed into *C. glutamicum*. Approximately 1/10 transformants were first screened by colony fluorescence imaging and those with increased RFP fluorescence were cultivated in 96-deep-well plates for a second round of screening. **b** Representative *gdh* and *pyc* promoter libraries tested in this study. The sequences of wild-type promoter and random mutation library are shown. The degenerate bases introduced during PCR primer synthesis are highlighted in red. The predicted -35 and -10 regions are highlighted in black boxes. Other tested promoter libraries for *gdh*, *pyc*, and *proB* are shown in Supplementary Fig. 7. **c** Strength analysis of the *P_{gdh}* variants using the *gdh-rfp* fusion gene as a reporter. WT represents the wild-type *P_{gdh}* control. The number represents the serial number of *P_{gdh}* variants. Data are presented as mean values \pm SD ($n = 3$ independent experiments). **d** Strength analysis of the *P_{pyc}* variants using the *pyc-rfp* fusion gene as a reporter. WT represents the wild-type *P_{pyc}* control. The number represents the serial number of *P_{pyc}* variants. Data are presented as mean values \pm SD ($n = 3$ independent experiments). **e** Strength analysis of *P_{gdh}-29* and *P_{pyc}-20* using the *proB-rfp* fusion gene as a reporter. All the *P_{proB}* libraries failed to generate good diversity and the strongest *P_{gdh}-29* and *P_{pyc}-20* variants were used for expression of the *proB-rfp* fusion gene. The wild-type *P_{proB}* was used as a control. Data are presented as mean values \pm SD ($n = 3$ independent experiments). Cells of the stationary growth phase were used to detect their fluorescence outputs using a microplate reader (λ excitation = 560 nm, λ emission = 607 nm). Source data underlying Fig. 4c–e are provided as a Source Data file.

directly assay the RFP fluorescence of 1/10 transformants on agar plates ($\sim 10^5$ colonies) (Fig. 4a and Supplementary Fig. 6). Cultivation in deep-well plates was used for the second round of screening of $\sim 10^2$ colonies with enhanced RFP fluorescence.

Several libraries for native promoters of *gdh*, *proB*, and *pyc* were constructed by installing degenerate bases via chemical synthesis (Fig. 4b and Supplementary Fig. 7). The *P_{gdh}* library produced variants with 2.3- to 46.4-fold increases in the expression level of the tailored fluorescent reporter compared with the wild-type promoter. The *P_{pyc}* library covered 1.8- to 16.1-fold fluorescence increases compared with the original *P_{pyc}* (Fig. 4c and Supplementary Data 6). However, the construction of *P_{proB}* libraries failed to generate variants with stronger transcription initiation capabilities. Therefore, we tested the strongest promoter variants from the *P_{gdh}* and *P_{pyc}* libraries for the expression of the *proB* reporter. *P_{gdh}-29* and *P_{pyc}-20* led to

6.7- and 20.8-fold increases in the expression level of the *proB* reporter, respectively. Although the strengths of *P_{gdh}-29* and *P_{pyc}-20* for the *proB* reporter were different from those for their corresponding reporters, these two promoters can be used to enhance the expression of *proB* (Fig. 4c).

Next, the flux-control reactions were tuned using the screened promoter variants. All the genetic modifications were conducted in the *C. glutamicum* chromosome using the optimized CRISPR/Cas9-assisted genome editing method. First, based on strain PRO-1 expressing a chromosomal copy of *proB*^{V150N}, an artificial *proB*^{V150N}AC operon controlled by *P_{pyc}-20* was inserted into the *putA* locus for simultaneous enhancement of L-proline biosynthesis and disruption of L-proline degradation, resulting in strain PRO-2 (Supplementary Fig. 8). *putA* encodes the multifunctional L-proline utilization A flavoenzyme that catalyzes the oxidation of L-proline to L-glutamate in two reaction steps⁶¹ (Fig. 2). Deletion

of *putA* has shown a beneficial effect for L-proline production²⁴. Insertion of *proB*^{V150N}AC operon into the *putA* locus resulted in a 75.8% increase in L-proline production (from 3.3 g/L to 5.8 g/L) in 24-deep-well plate cultivation (Fig. 5a). Meanwhile, biomass production and glucose consumption largely decreased, leading to a 120% increase in the conversion yield (from 0.05 g/g to 0.11 g/g). Second, *gdh* responsible for competition with TCA cycle for 2-oxoglutarate flux was overexpressed in strain PRO-2 by in-situ replacing the original *P_{gdh}* by its variants with gradually increased strengths (*P_{gdh}-1*, *P_{gdh}-16*, *P_{gdh}-23*, *P_{gdh}-26*, and *P_{gdh}-29*, with 2.3-, 5.3-, 9.2-, 13.3-, and 46.4-fold strength increases, respectively). Strain PRO-6 in which *gdh* was overexpressed using *P_{gdh}-26* showed the largest improvement in L-proline production compared with strain PRO-2 (34.5% and 18.5% improvement in titer and yield, respectively) (Fig. 5a). The use of the strongest promoter *P_{gdh}-29* slightly decreased L-proline titer, which was possibly due to the trade-offs between the resources used for maintaining cellular metabolism and synthesizing L-proline. Third, the C4 anaplerotic pathway was enhanced by in-situ overexpressing *pyc* using *P_{pyc}* variants with gradually increased strengths (*P_{pyc}-1*, *P_{pyc}-9*, *P_{pyc}-13*, *P_{pyc}-16*, and *P_{pyc}-20*, with 1.8-, 6.4-, 8.9-, 12.1-, and 16.1-fold strength increases, respectively). Moderate improvement in L-proline production was observed. Strain PRO-11 overexpressing *pyc* with the *P_{pyc}-16* variant produced 9.7% more L-proline with 5.8% higher yield compared with its parent strain PRO-6 (Fig. 5b). Finally, the NADPH-generating glycolysis branch was introduced to strain PRO-11 by expressing *gapN* from *Streptococcus mutans* using five *P_{pyc}* variants (*P_{pyc}-1*, *P_{pyc}-9*, *P_{pyc}-13*, *P_{pyc}-16*, and *P_{pyc}-20*). By using a tailored reporter consisting of the first 180 bp of *gapN* and an RFP gene, it was observed that the five *P_{pyc}* variants led to gradual expression levels of *gapN* (Supplementary Fig. 9). The strength of *P_{pyc}* variants for expressing *gapN* was different from that for expressing its native *pyc* gene, which was reasonable considering the effects of the 5' region of CDS as discussed previously^{57,58}. Only expression of *gapN* with *P_{pyc}-1* (strain PRO-13) improved L-proline production. The titer and yield of strain PRO-13 reached 9.0 g/L and 0.15 g/g, which were 13.9% and 15.9% higher than its parent strain PRO-11, respectively. Use of stronger *P_{pyc}* variants caused growth retardation, suggesting very high-level expression of *gapN* negatively affected cellular metabolism (Fig. 5c).

Identification of an L-proline exporter by arrayed CRISPRi screening. Chromosomally transcriptional tuning of flux-control genes improved L-proline production by nearly 3-fold. Besides metabolic pathways, transport systems are also pivotal for the hyperproduction of biochemicals²⁶. To test whether L-proline excretion is a limiting step for its production, the intracellular L-proline concentration of strain PRO-13 during batch fermentation in shake flasks was measured and found to be 300–400 mM (Supplementary Fig. 10), which is almost 10-fold higher than the wild-type nonproducing strain⁶². Therefore, strengthening L-proline excretion is required for further strain improvement. As a compatible solute essential for protection against hyperosmotic shock, uptake systems of L-proline have been well characterized in microorganisms⁶³, whereas L-proline exporter that transports the intracellularly synthesized L-proline molecules to the culture medium has not been discovered yet²⁵. By using TransportDB 2.0⁶⁴, *C. glutamicum* was predicted to possess nearly 400 membrane transporters (Supplementary Data 7). For a comprehensive analysis of *C. glutamicum* transporters and identification of L-proline exporters, we adopted CRISPRi technology for *C. glutamicum* and constructed an arrayed CRISPRi library consisting of 397 plasmids, each of which was designed to repress the transcription of a membrane transporter gene. The plasmid library

was then transformed into an L-proline producing *C. glutamicum* chassis PRO-CRISPRi (*C. glutamicum* expressing a deregulated *ProB*^{G149D} variant²⁴) to evaluate the effect of gene repression on L-proline biosynthesis. Transformation of five plasmids failed to produce any transformants possibly because repression of these genes seriously inhibited cell growth. Because impaired amino acid excretion could significantly decrease amino acid production³⁰, the transporters whose repression brought significant decrease in L-proline production were considered as potential L-proline exporters and further characterized (Fig. 6a).

Two rounds of screening were conducted for identifying potential L-proline exporters. The first round of screening was performed by cultivating all the 392 strains with transporter-targeting CRISPRi systems and a negative control with a nontargeting CRISPRi system. Repression of twenty-one transporters were found to negatively affect L-proline production (Fig. 6b), which were subjected to a second round of screening by cultivation in 24-deep-well plates. After two rounds of screening, repression of six transporters (*Cgl2622*, *Cgl1436*, *Cgl1933*, *Cgl2348*, *Cgl1360*, and *Cgl1937*) decreased L-proline production (Fig. 6c). *Cgl1933* (PtsI), *Cgl1360* (PtsG), and *Cgl1937* (PtsH) belong to the PTS for sugar uptake⁶⁵, repression of which would hinder glucose uptake and consequently decrease L-proline production. Therefore, these three genes were not further investigated here. *cgl2622*, *cgl1436*, and *cgl2348* were individually deleted in strain PRO-CRISPRi by CRISPR/Cas9 to verify their functions. While CRISPRi of *cgl1436* slightly decreased L-proline production, its gene deletion showed no significant affect. Interestingly, the strain with repression of *cgl2348* only produced ~20% L-proline compared with the control without CRISPRi. However, deletion of *cgl2348* did not decrease L-proline production significantly. *Cgl2348* was annotated as a DNA uptake protein or related DNA-binding protein. According to a previous transcriptomic analysis, *cgl2348* belongs to an operon containing *proB* and *proA* involved in L-proline biosynthesis⁶⁶. CRISPRi of *cgl2348* may also affect the expression of its neighboring genes including *proB* and *proA*, leading to decreased L-proline production. Partial deletion of the CDS of *cgl2348* should not affect the expression of its neighboring genes, which could explain why CRISPRi and deletion of *cgl2348* showed different influences on L-proline production. Only deletion of *cgl2622* significantly decreased L-proline production, making *Cgl2622* a candidate for L-proline exporter (Fig. 6d).

To verify the function of *Cgl2622* as an L-proline exporter, two experiments were conducted. First, *cgl2622* was deleted, complemented, and overexpressed in the L-proline producing strain PRO-CRISPRi to test the effects on intracellular and extracellular accumulation of L-proline. The resultant strains were cultivated in shake flasks to obtain enough cells for the measurement. In strain PRO-CRISPRi, the intracellular L-proline accumulated to very high levels of 150–300 mM (Fig. 6e). Deletion of *cgl2622* further increased the intracellular L-proline concentration especially in the early phase of fermentation and largely decreased extracellular accumulation of L-proline. By complementing or overexpressing *cgl2622* in a plasmid with IPTG-inducible promoter *P_{trc}*, intracellular L-proline concentration was maintained at relatively low levels between 50 mM to 100 mM, while extracellular L-proline production was largely increased (Fig. 6e). The similar results for complementation and overexpression of *cgl2622* are possibly because of the high-level expression of *cgl2622* using multi-copy plasmid and strong IPTG-inducible promoter. The results suggest that *Cgl2622* plays an important role in maintaining the intracellular L-proline homeostasis and transporting L-proline outside cells.

Second, peptide uptake and amino acid export assay was conducted. *cgl2622* was deleted and complemented in *C.*

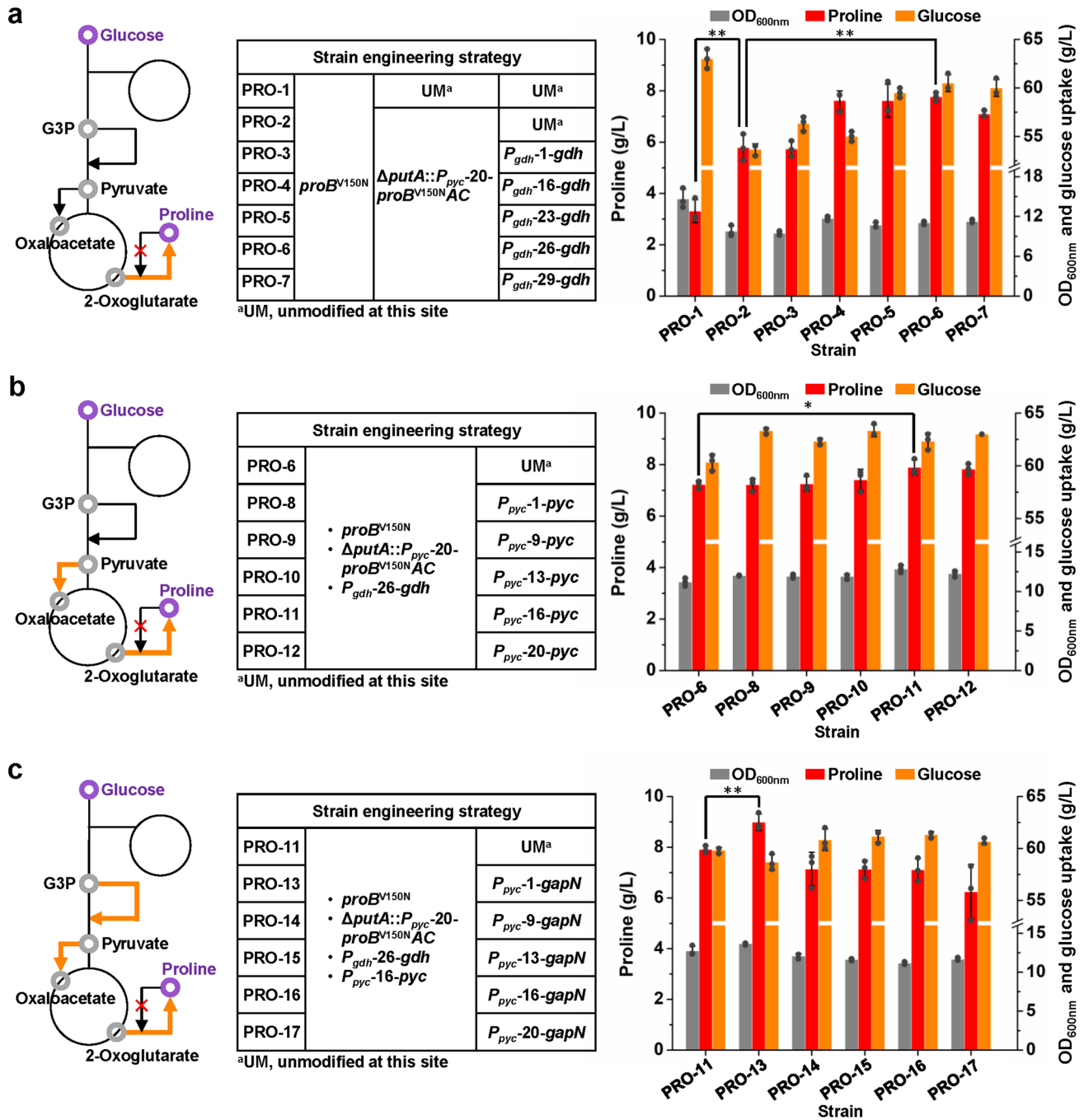


Fig. 5 Enhancing L-proline production by tuning the expression of flux-control genes. **a** Channeling metabolic flux into L-proline biosynthesis by overexpressing an artificial *proB*^{V150N}*AC* operon and *gdh*. Strain PRO-1 was constructed by introducing a V150N mutation to the *proB* gene in the chromosome of wild-type *C. glutamicum* ATCC 13032. Strain PRO-2 was constructed by inserting an artificial *proB*^{V150N}*AC* operon in to the *putA* gene of strain PRO-1. Strains PRO-3 to PRO-7 were constructed by replacing the wild-type *P_{gdh}* promoter with different *P_{gdh}* promoter variants for *gdh* overexpression in strain PRO-2. L-Proline production was conducted in 24-deep-well plates. Data are presented as mean values \pm SD ($n = 3$ independent experiments). $**P = 0.0032$ for PRO-1 vs. PRO-2, $**P = 0.0031$ for PRO-2 vs. PRO-6, Student's two-tailed *t*-test. **b** Strengthening C4 anaplerotic pathway by overexpressing *pyc*. Strains PRO-8 to PRO-12 were constructed by replacing the wild-type *P_{pyc}* promoter with different *P_{pyc}* promoter variants for *pyc* overexpression in strain PRO-6. L-Proline production was conducted in 24-deep-well plates. Data are presented as mean values \pm SD ($n = 3$ independent experiments). $*P = 0.0260$, Student's two-tailed *t*-test. **c** Switching glycolysis cofactor from NADH to NADPH by overexpressing *gapN* from *S. mutans*. Strains PRO-13 to PRO-17 were constructed by introducing *gapN* downstream the *cgl2334* gene under the control of different *P_{pyc}* promoter variants in strain PRO-11. L-Proline production was conducted in 24-deep-well plates. Data are presented as mean values \pm SD ($n = 3$ independent experiments). $**P = 0.0059$, Student's two-tailed *t*-test. Source data underlying Fig. 5 are provided as a Source Data file.

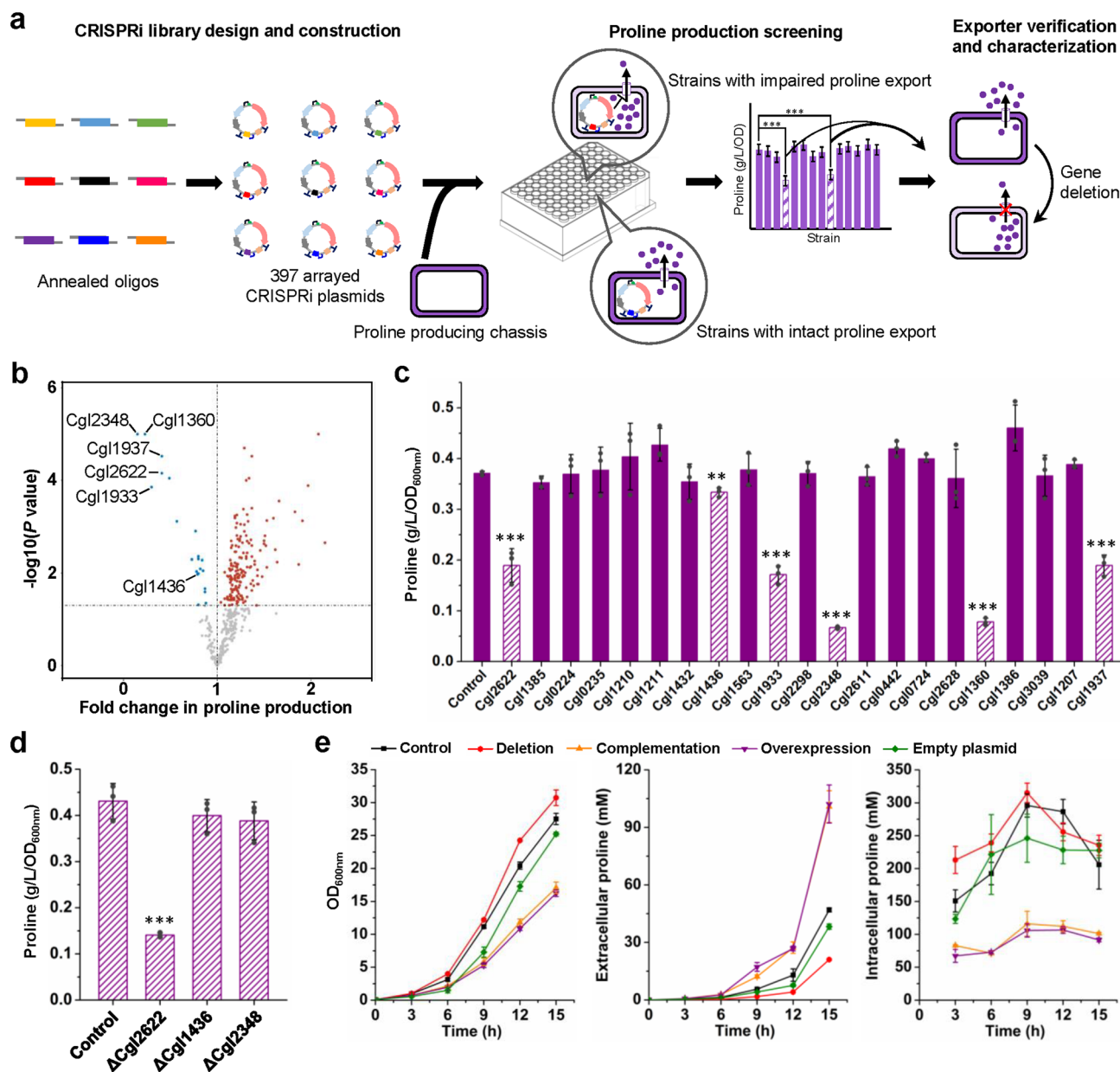


Fig. 6 CRISPRi screening and characterization of L-proline exporter in *C. glutamicum*. **a** Workflow of screening L-proline exporter by constructing an arrayed CRISPRi library targeting the potential transporter genes. A pair of oligos were synthesized and annealed to generate dsDNAs harboring a spacer sequence. Totally 397 CRISPRi plasmids were constructed using the dsDNAs and pdCas9gRNA-*ccdB*. Each CRISPRi plasmid targeted a potential transporter gene of *C. glutamicum*. The CRISPRi plasmids were individually transformed into an L-proline producing *C. glutamicum* PRO-CRISPRi expressing a deregulated ProB^{G149D} variant²⁴. The resultant 392 strains were cultivated in 96-deep-well plates for the first round of screening. Strains with significantly reduced L-proline production were selected for a second round of screening by cultivation in 24-deep-well plates. The transporter genes whose repression significantly reduced L-proline production were considered as potential L-proline exporters for characterization. **b** Volcano plot of differential L-proline production levels caused by CRISPRi repression of transporter genes. The results of the first round of screening are analyzed and shown. Three parallel L-proline production experiments were conducted in 96-deep-well plates for all the 392 strains. Strain PRO-CRISPRi harboring a nontargeting CRISPRi system was used as a control. Gene repressions causing significant increases and decreases in L-proline production are indicated in red and blue dots, respectively ($P < 0.05$, Student's two-tailed *t*-test). Grey dots represent those with non-significant changes in L-proline production level. Twenty-one exporters (blue dots) were selected for a second round of screening. **c** The second round of screening of the 21 L-proline exporter candidates by cultivation in 24-deep-well plates. Data are presented as mean values \pm SD ($n = 3$ independent experiments). All Student's two-tailed *t*-tests compare the L-proline production levels of strains expressing a gene-targeting CRISPRi system with the control strain expressing a nontargeting CRISPRi system (*** $P = 0.0007$ for Cgl2622, ** $P = 0.0027$ for Cgl1436, *** $P = 0.00005$ for Cgl1933, *** $P = 3 \times 10^{-8}$ for Cgl2348, *** $P = 4 \times 10^{-7}$ for Cgl1360, *** $P = 0.0001$ for Cgl1937). **d** Effects of deletion of *cgl2622*, *cgl1436*, and *cgl2348* on L-proline production. Strains were cultivated in 24-deep-well plates for L-proline production. Data are presented as mean values \pm SD ($n = 3$ independent experiments). *** $P = 0.00003$, Student's two-tailed *t*-test. **e** Effects of deletion, complementation, and overexpression of *cgl2622* on growth and extracellular and intracellular accumulation of L-proline. Complementation was conducted by expressing *cgl2622* via a plasmid under the control of IPTG-inducible P_{trc} in the *cgl2622*-deleted PRO-CRISPRi strain. Strains were cultivated in shake flasks to obtain enough cells for the measurement of intracellular L-proline. Data are presented as mean values \pm SD ($n = 3$ independent experiments). Source data underlying Fig. 6b–e are provided as a Source Data file.

glutamicum wild-type strain. Upon the addition of Thr-Pro peptide, the *cgl2622*-deleted mutant showed higher intracellular L-proline level and dramatically decreased L-proline export rate (Supplementary Fig. 11). Conversely, complementation of *cgl2622* in a plasmid with IPTG-inducible promoter P_{trc} largely decreased intracellular L-proline concentration but accelerated L-proline export (Supplementary Fig. 11). These results further verify the function of Cgl2622 as an L-proline exporter.

Transport engineering to enhance L-proline production. To overcome the limitation of L-proline excretion in engineered producing strains, overexpression of Cgl2622 is required. To quantify the expression level of Cgl2622, we used the developed strategy of building a tailored fluorescent reporter by fusing the first 180 bp of *cgl2622* with the flexible linker (GGGGS)₃ and fluorescent protein-encoding gene. Both RFP and GFP were tested, whereas the fusion protein failed to produce fluorescence, which was possibly due to the membrane topology of Cgl2622. Therefore, we directly added a second copy of *cgl2622* controlled by the derepressed P_{trc} promoter in the chromosome of strain PRO-13 for overexpression. The resultant strain PRO-18 produced 15.7 g/L L-proline with 0.25 g/g yield, which are 177.4% and 170.5% of the production level of its parent strain PRO-13 (Fig. 7a). Overexpression of Cgl2622 also led to ~100% decrease in intracellular L-proline concentration (Supplementary Fig. 12), which was considered as a major cause of improved L-proline production.

Fed-batch fermentation in a 5 L bioreactor was then conducted to evaluate the L-proline production by strain PRO-18. After 49 h cultivation, 87.2 g/L L-proline was produced with a productivity of 1.78 g/L/h and a yield of 0.19 g/g (Fig. 7b). The biomass reached a OD_{600nm} value over 180. It seemed that too much carbon flux was directed to the production of biomass rather than the target molecule L-proline. Since *C. glutamicum* is a biotin-auxotrophic bacterium⁶⁷, biotin limitation can be used to balance

the biosynthesis of biomass and target molecules. Therefore, a biotin-limited fed-batch fermentation was conducted. Although biomass formation was reduced, L-proline production was only moderately improved to 97.8 g/L with a productivity of 2.00 g/L/h and a yield of 0.23 g/g. Product analysis showed the high-level accumulation of extracellular L-glutamate up to 77.5 g/L (Fig. 7c).

Improving L-proline production by eliminating biotin limitation-induced L-glutamate excretion.

L-Glutamate production by *C. glutamicum* can be induced by treatments such as biotin limitation, Tween 40 addition, and penicillin addition. Biotin limitation decreases the activity of 2-oxoglutarate dehydrogenase complex (ODHC) that converts 2-oxoglutarate to succinyl-CoA, causing accumulation of intracellular 2-oxoglutarate. Meanwhile, the cell membrane tension is altered and mechanosensitive ion channels responsible for L-glutamate excretion are activated^{30,67}. To reduce L-glutamate production, the major L-glutamate exporter MscCG encoded by *cgl1270* was deleted in strain PRO-18 by CRISPR/Cas9. The resultant strain PRO-19 was then used for L-proline production in biotin-limited fed-batch fermentation. After 49 h cultivation, 142.4 g/L L-proline was produced with a productivity of 2.90 g/L/h and a yield of 0.31 g/g (Fig. 7d). The titer, productivity, and yield of strain PRO-19 were improved by 44.9, 45.0, and 34.8%, respectively, compared with strain PRO-18 under the same cultivation condition. The byproduct L-glutamate was largely decreased to 12.3 g/L, which was only 15.6% of that produced by strain PRO-18. The amino acid profile of the broth from the end of fermentation was further determined using an amino acid analyzer. Besides L-glutamate, 7.9 g/L of L-alanine and 3.2 g/L of L-valine were also detected as byproducts (Supplementary Fig. 13). Because the L-glutamate exporter MscCG encoding gene *cgl1270* was deleted in strain PRO-19, L-glutamate may accumulate inside the cells. L-Glutamate is an important amine donor for the biosynthesis of many amino acids⁶⁸. Its accumulation may enhance the transamination and biosynthesis of other amino acids, which

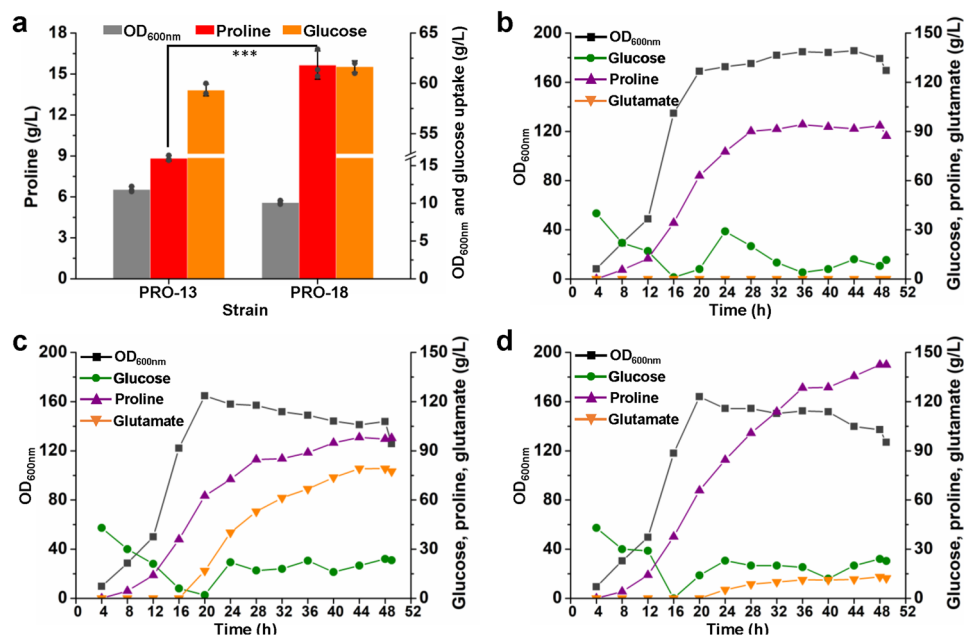


Fig. 7 L-Proline production using strains with transport engineering. **a** Effects of *cgl2622* overexpression on L-proline production in 24-deep-well plates. Strain PRO-18 was constructed by adding a second copy of *cgl2622* controlled by the derepressed P_{trc} promoter in the chromosome of strain PRO-13. Data are presented as mean values \pm SD ($n = 3$ independent experiments). *** $P = 0.0004$, Student's two-tailed t -test. **b** Fed-batch fermentation of strain PRO-18 in a 5 L bioreactor with 1 mg/mL biotin for a biotin-rich condition. **c** Fed-batch fermentation of strain PRO-18 in a 5 L bioreactor with 45 μ g/mL biotin for a biotin-limited condition. **d** Fed-batch fermentation of strain PRO-19 (strain PRO-18 with *cgl1270* (*mscCG*) deleted) in a 5 L bioreactor. Biotin was used at 45 μ g/mL for a biotin-limited condition. Source data underlying Fig. 7 are provided as a Source Data file.

may explain the formation of L-alanine and L-valine as byproducts. Therefore, increasing the expression of genes involved in L-proline biosynthesis (such as *gdh* and *proB*) and deleting related amino-transferase encoding genes may further improve L-proline production and decrease byproduct formation.

Discussion

To develop industrial strains for biomanufacturing, efficient genetic manipulation tools and multifarious components with catalytic, regulatory, or transport functions are needed. Here, CRISPR/Cas9-assisted rational flux-tuning and CRISPRi-mediated comprehensive transporter discovery were applied to transfer the wild-type *C. glutamicum* strain to an industrial-strength L-proline producer. Taking advantage of the discovered L-proline exporter Cgl2622, the titer, yield, and productivity of the final strain PRO-19 are superior to those of most speciality amino acids (Supplementary Table 1). No plasmid, antibiotic marker, or chemically inducible promoter is contained in the final strain, ensuring the stable production performance in scale-up fermentations.

Both the CRISPR/Cas9^{31–34} and CRISPR/Cas12a^{35–38} systems have been adopted for genetic manipulation of *C. glutamicum*. Considering the capability of accepting G-rich PAMs and distinguishing a single nucleotide mismatch⁴⁰, the CRISPR/Cas9 system was selected and optimized for genome editing of *C. glutamicum* in this study. By combating the cytotoxicity of CRISPR/Cas9 through strict control of *cas9* expression, the optimized method showed satisfactory performance in codon saturation mutagenesis, biosynthetic operon integration, and in-situ promoter replacement, which lays the technical groundwork for the development of high-performance strains for the production of L-proline and other biobased chemicals. Previous studies suggest that the CRISPR/Cas12a system is less toxic and facilitates efficient genome editing in *C. glutamicum*^{35–38}. Cas12a possesses both DNase and RNase activities and thus it can process a CRISPR RNA array into multiple mature gRNAs, providing unique benefits for multiplex editing⁶⁹. Since the CRISPR/Cas12a system recognizes T-rich PAMs that are different from the CRISPR/Cas9 system³⁹, they can be complementary tools for the engineering of *C. glutamicum*.

Since microorganisms are commonly evolved for converting carbon sources to biomass, metabolic flux redirection is essential for hyperproduction of target chemicals⁷⁰. We performed in silico FBA using a genome-scale metabolic model of *C. glutamicum*²⁴ to predict the flux-control reactions in L-proline biosynthesis. The L-proline biosynthetic pathway, C4 anaplerotic pathway, and NADPH-generating glycolytic pathway were suggested as the key targets for engineering. Overexpression of deregulated *proB* and *gdh* is obviously important for L-proline overproduction²³. Fixation of CO₂ via the C4 anaplerotic pathway has been proven a useful strategy for improving the production yield of several chemicals, especially the L-aspartate family amino acids such as L-lysine^{50,71}. Here, we demonstrate that the biosynthesis of L-proline also benefits from the enhanced C4 anaplerotic pathway. In previous studies, increasing the supply of NADPH for L-proline biosynthesis was commonly accomplished by strengthening the PPP²³. To decrease CO₂ release and increase L-proline yield, the in silico simulation suggests a PPP-independent alternative for NADPH regeneration based on the nonphosphorylating NADP-dependent G3P dehydrogenase (GapN). GapN was initially used for improving L-lysine biosynthesis that requires 4 mol NADPH/mol L-lysine. It was found that cell growth on glucose was significantly retarded by the complete replacement of the native NADH-generating glycolytic pathway with the GapN-based NADPH-generating glycolytic pathway⁵². Since the in silico

simulation also suggests the collaboration between these two glycolytic pathways, GapN was expressed using different promoters in this study to achieve the balance between them. These efforts for flux-tuning led to a nearly 3-fold improvement in L-proline titer compared with the strain only harboring a feedback-insensitive γ -glutamyl kinase variant. Besides the catalytic reactions inside the cell, transport reactions that happen on the membrane also influence flux distribution by providing a strong driving force for continuous biosynthesis^{72,73}. Because exporters for L-proline have not been discovered yet, transport engineering was not applied for L-proline production in previous studies. We constructed an arrayed CRISPRi library targeting all the 397 potential membrane transporter genes for a comprehensive screening in *C. glutamicum*, and identified Cgl2622 as an L-proline exporter. Its overexpression largely enhanced L-proline production. During the CRISPRi screening, it was observed that repression of several transporter genes increased L-proline production, which provided interesting targets for further strain improvement (Fig. 6b). This arrayed CRISPRi library can be readily used to screen undiscovered transporters for other metabolites of interest, such as the unavailable exporters for glycine, L-glutamine, and L-asparagine^{25,26}. The results demonstrate that arrayed CRISPRi screening is a useful strategy to identify undiscovered functional components such as transporters.

Cgl2622 (previously known as ThrE) was previously characterized as an L-threonine exporter, which also accepted L-serine as a substrate but with a lower efflux rate²⁸. Cgl2622 belongs to a ubiquitous ThrE family of putative transmembrane amino acid efflux that exists in both Gram-negative and Gram-positive bacteria, as well as in archaea and eukaryotes⁷⁴. However, upon Cgl2622 overexpression and further metabolic engineering, very limited improvement in L-threonine production was observed⁷⁵. Therefore, it is suggested that Cgl2622 is possibly not natively designed to export L-threonine but an ancient genomic relict using a still unknown compound as substrate²⁵. Based on the observation that overexpression of Cgl2622 improved L-proline production by ~1.7-fold regarding both titer and yield, we hypothesize that L-proline is possibly the native export substrate for Cgl2622. It is no surprise that a transporter has more than one substrate. For instance, BrnFE of *C. glutamicum* exports branched-chain amino acids (L-leucine, L-isoleucine, and L-valine) and L-methionine^{29,76}. Considering L-proline is one of the compatible solutes used by *C. glutamicum*⁶³, Cgl2622 may cooperate with mechanosensitive ion channels for osmotic adaptation^{30,77,78}. Studies on the expression regulation of Cgl2622 possibly provides clues to the biological functions of this ubiquitous transporter. Besides, the discovery of Cgl2622 as an L-proline exporter should be helpful for identifying L-proline exporters in other species.

In conclusion, we developed a highly efficient L-proline producing *C. glutamicum* strain using CRISPR-assisted metabolic flux-tuning and arrayed CRISPRi screening of the L-proline exporter Cgl2622. The optimized CRISPR/Cas9 system for genome editing, arrayed CRISPRi library for transporter genes, and interdisciplinary strategy for strain development will provide a pipeline for developing high-performance strains for biomanufacturing.

Methods

Strains and culture conditions. Strains used in this study are listed in Supplementary Data 8. *E. coli* strains Trans1-T1 and DB3.1 (TransGen Biotech, Beijing, China) used for plasmid cloning were cultivated aerobically at 37 °C in Luria-Bertani (LB) broth. Kanamycin (50 μ g/mL) or chloramphenicol (20 μ g/mL) was added to the medium as required. *C. glutamicum* ATCC 13032, ATCC 13869, and the derivatives were cultivated aerobically at 30 °C in TSB medium containing 5 g/L glucose, 9 g/L soya peptone, 5 g/L yeast extract, 1 g/L K₂HPO₄·3H₂O, 0.1 g/L MgSO₄·7H₂O, 3 g/L urea, 0.5 g/L succinic acid, 10 μ g/L biotin, 100 μ g/L vitamin B1,

and 20 g/L MOPS (pH 7.2). Kanamycin (25 µg/mL), chloramphenicol (5 µg/mL), or IPTG (0.01 to 0.5 mM) was added as required.

Plasmid construction. Plasmids used in this study are listed in Supplementary Data 8. Plasmids were constructed via recombination or Golden Gate assembly. Recombination was conducted using the ClonExpress MultiS One Step Cloning Kit (Vazyme, Nanjing, China). Restriction endonucleases and T4 DNA ligase used for Golden Gate assembly were purchased from New England Biolabs (Beijing) (Beijing, China). DNA polymerase and reagents used for PCR were purchased from TransGen Biotech (Beijing, China). Services of primer and gene synthesis and DNA sequencing were provided by GENEWIZ Inc. (Suzhou, China). Primers and details for constructing plasmids are described in Supplementary Data 9.

Genetic manipulation of *C. glutamicum*. For DNA deletion in *C. glutamicum*, 1 µg plasmid (for example pCas9gRNA-1 or pCas9gRNA-2) expressing Cas9, a targeting gRNA, and two HR arms was transformed into *C. glutamicum* via electroporation. Transformants were plated on TSB solid medium supplemented with 0.05 mM IPTG and 5 µg/mL chloramphenicol for counter-selection. After 24 h of cultivation at 30 °C, colonies were verified by PCR. For DNA insertion in *C. glutamicum*, 1 µg plasmid (for example pCas9gRNA-5) expressing Cas9 and a targeting gRNA and 1 µg plasmid (for example, pEC-1) harboring the inserted fragment and two HR arms were co-transformed into *C. glutamicum* via electroporation. Transformants were plated on TSB solid medium supplemented with 0.05 mM IPTG, 5 µg/mL chloramphenicol, and 25 µg/mL kanamycin for counter-selection. After 48 h of cultivation at 30 °C, colonies were verified by PCR. For ssDNA recombineering for codon saturation mutagenesis of CgProB, 1 µg plasmid pRecT expressing RecT was first transformed into *C. glutamicum* via electroporation. Transformants were plated on TSB solid medium supplemented with 25 µg/mL kanamycin. After 24 h of cultivation at 30 °C, colonies were verified by PCR. A correct transformant was used for preparation of competent cells and transformation of 1 µg plasmid (for example pCas9gRNA-6) expressing Cas9 and a targeting gRNA and 5 µg synthetic ssDNA (Supplementary Data 10). Transformants were plated on TSB solid medium supplemented with 0.05 mM IPTG, 5 µg/mL chloramphenicol, and 25 µg/mL kanamycin for counter-selection. After 18–24 h of cultivation at 30 °C, colonies were picked and cured for the plasmids. For plasmid curing, transformants were cultivated in TSB medium without antibiotics at 37 °C overnight and spread on TSB plates without antibiotics. Colonies were confirmed as cured by determining their sensitivity to antibiotics. Transformants with plasmid cured were used for the next round of genetic manipulation or other tests. An operation scheme of CRISPR/Cas9-assisted gene deletion, insertion, and ssDNA recombineering is shown in Supplementary Fig. 4. For gene over-expression in the plasmid, the target gene was amplified and inserted into the pEC-XK99E plasmid under the control of IPTG-inducible promoter P_{trc} . The recombinant plasmid was transformed into *C. glutamicum* via electroporation. Gene overexpression in the transformant was induced by the addition of 0.05 mM IPTG. All the plasmids and gRNAs used for genetic manipulation of *C. glutamicum* are shown in Supplementary Data 8 and Supplementary Data 11, respectively.

Electro-competent cells for *C. glutamicum* wild-type strains and low-level L-proline producing strains PRO-1 and PRO-CRISPRi were prepared using a hypertonic medium containing glycine and DL-threonine for weakening cell walls and Tween 80 and isonicotinic acid hydrazide for increasing cytoplasmic membrane fluidity⁷⁹. Electro-competent cells for high-level L-proline producing strain PRO-2 and its derivatives were prepared using a modified procedure because of the low growth rate caused by L-proline overproduction. Cells were cultivated overnight in TSB medium and transferred into 100 mL fresh TSB medium supplemented with 3% glycine and 0.1% Tween 80 to an OD_{600nm} value of 0.3 for aerobic cultivation at 30 °C. When the OD_{600nm} of culture reached ~0.8, cells were collected by centrifugation at 6,000 × g and 4 °C for 10 min. After washed with ice-cold deionized distilled water for four times, cells were resuspended in 0.5 mL 10.0% (v/v) glycerol and 100 µL aliquots of competent cells were obtained. DNA (less than 10 µL) was added to the competence cells and transferred to a 2 mm electroporation cuvette (Bio-Rad Laboratories (Shanghai), China). Electroporation was performed with the Eppendorf Electroporator 2510 (Eppendorf China, Shanghai, China) with parameters set at 2,500 V and 5 ms. After electroporation, 1 mL TSB medium was added immediately and the suspension was quickly incubated for 6 min at 46 °C. Cells were incubated for 3 h at 30 °C, and spread on TSB plates supplemented with antibiotics and IPTG as required. The plates were incubated at 30 °C until colonies appeared.

Homologous modeling of CgProB. The model structure of CgProB was constructed with the crystal structure of γ-glutamyl kinase from *E. coli* (PDB ID: 2J5T)⁸⁰ as a template (93% coverage and 38% sequence identity with CgProB) using Discovery Studio 2018 software (Accelrys, USA). Molecular docking with L-glutamate and selection of the optimal conformation was performed using AutoDock Tools 1.5.6⁸¹. Receptor-ligand interaction analysis and figure rendering were performed using Discovery Studio 2018 software (Accelrys, USA).

In silico analysis of the optimal L-proline biosynthetic pathway. The genome-scale metabolic model iCW773²⁴ was used to predict the optimal L-proline

biosynthetic pathway in *C. glutamicum* by performing FBA⁴⁸. Simulations were performed using the COBRApy toolbox (v 0.22.1)⁸². Uptake rate of glucose was set as 10 mmol/gCDW-h. To simulate the Ppc-based C4 anaplerotic pathway for L-proline biosynthesis, no modification to the model was needed. To simulate the Pyc-based C4 anaplerotic pathway for L-proline biosynthesis, the Ppc-catalyzed reaction was manually closed and oxaloacetate was synthesized from pyruvate via Pyc-catalyzed reaction. To simulate the GapN-Pyc-based pathway for L-proline biosynthesis, the GapN-catalyzed reaction was added to the model.

Promoter library construction and screening. Native promoters of *gdh*, *proB*, and *pyc* were amplified from the genomic DNA of *C. glutamicum* ATCC 13032. The promoter was then ligated with the corresponding tailored RFP reporter gene and pEC-XK99E plasmid to construct a basic plasmid for library construction. The basic plasmid was used as a template for library construction via PCR with primers containing degenerate bases as shown in Fig. 4b and Supplementary Fig. 7. The PCR product was self-ligated and transformed into *E. coli* Trans1-T1. All the transformants were collected and used for plasmid extraction. The plasmid library was transformed into *C. glutamicum* via electroporation and cells were plated on TSB agar plates. After 24 h cultivation, the RFP fluorescences of 1/10 transformants on agar plates (~10⁵ colonies) were directly assayed using a Tanon 5200 Multi Fluorescence Imaging system (Tanon Technology, Shanghai, China). Deep-well plates were used for the second round of screening of ~10² colonies with enhanced RFP fluorescence. Cells of the stationary growth phase were used to detect their fluorescence outputs using a microplate reader (SpectraMax M5, Molecular Devices, USA, λ excitation = 560 nm, λ emission = 607 nm).

L-Proline production by cultivation in 24-deep-well plates. The seed medium (TSC medium) contains 5 g/L glucose, 3 g/L soya peptone, 1 g/L yeast extract, 1 g/L K₂HPO₄·3H₂O, 0.1 g/L MgSO₄·7H₂O, 3 g/L urea, 0.5 g/L succinic acid, 10 µg/L biotin, 100 µg/L vitamin B1, and 20 g/L MOPS (pH 7.2). The fermentation medium (TSD medium) contains 80 g/L glucose, 1 g/L soya peptone, 1 g/L yeast extract, 1 g/L NaCl, 1 g/L (NH₄)₂SO₄, 1 g/L K₂HPO₄·3H₂O, 0.45 g/L MgSO₄·7H₂O, 0.05 g/L FeSO₄·7H₂O, 6 g/L urea, 400 µg/L biotin, 100 µg/L vitamin B1, and 40 g/L MOPS (pH 7.2). For seed preparation, strains were cultivated in 24-deep-well plates containing 800 µL TSC medium in each well at 30 °C, 800 rpm, and 90% moisture for 8 h. The seed culture was then used to inoculate 24-deep-well plates containing 800 µL TSD medium in each well to an initial OD_{600nm} of 0.06. The 24-deep-well plates were cultivated at 30 °C, 800 rpm, and 90% moisture for 18 h and cultivation was stopped for glucose and L-proline assay.

L-Proline production by cultivation in shake flasks. TSC medium was used for seed preparation. The fermentation medium is TSD medium with urea changed from 6 g/L to 10 g/L for providing more nitrogen source for cell growth and L-proline production. For seed preparation, strains were cultivated in 250 mL shake flasks containing 25 mL TSC medium at 30 °C and 220 rpm for 8 h. The seed culture was then used to inoculate 250 mL shake flasks containing 25 mL TSD medium (with 10 g/L urea) to an initial OD_{600nm} of 0.1. The shake flasks were cultivated at 30 °C and 220 rpm for 15 h and cultivation was stopped for glucose and L-proline assay.

L-Proline production by cultivation in 5 L bioreactors. TSB medium was used for seed preparation. The fermentation medium contains 50 g/L glucose, 10 g/L corn steep powder, 10 g/L soya peptone, 10 g/L bovine bone peptone, 30 g/L (NH₄)₂SO₄, 6 g/L KH₂PO₄, 0.5 g/L MgSO₄·7H₂O, 0.05 g/L FeSO₄·7H₂O, 0.03 g/L MnSO₄·H₂O, 45 or 1000 µg/L biotin, 45 µg/L vitamin B2, and 0.5 g/L antifoam. Fed-batch fermentations were conducted in 5 L bioreactors (Baoning Bio., Shanghai, China). For seed preparation, strains were cultivated overnight in 500 mL shake flasks containing 75 mL TSB medium at 30 °C and 220 rpm. The seed culture was then used to inoculate 5 L bioreactors containing 1.5 L fermentation medium with an inoculation volume of 10%. Cultivation temperature was maintained at 30 °C, and pH was automatically controlled at 7.0 by the addition of NH₄OH. The dissolved oxygen was maintained at 30% by automatic adjustment of agitation and aeration rate. Glucose solution (800 g/L) was fed into the bioreactor at appropriate rates to maintain the glucose concentration in the range of 5–30 g/L. Samples were withdrawn from the bioreactor every 4 h for measuring cell biomass, glucose, and products.

CRISPRi library construction and screening. gRNAs targeting transporter genes of *C. glutamicum* were identified and analyzed for potential off-target sites using the sgRNACas9 bioinformatics tool⁸³. For each gRNA, a pair of 24-nt oligos were synthesized and annealed to generate dsDNAs harboring a spacer sequence. Golden Gate assembly strategy was applied to construct CRISPRi plasmid harboring gRNA targeting a specific gene using plasmid pCas9gRNA-*ccdB* and the annealed dsDNA⁸⁴. The predicted transporter genes and the corresponding gRNAs are shown in Supplementary Data 7. Totally 397 CRISPRi plasmids were constructed and individually transformed into an L-proline producing *C. glutamicum* with CgProB^{G149K} (strain PRO-CRISPRi). Plasmid and strain construction was conducted using the BioFoundry infrastructure at Tianjin Institute of Industrial

Biotechnology and the developed automated Golden Gate assembly protocol⁸⁴. The transformation of five plasmids failed to produce any colonies possibly because repression of these genes seriously inhibited cell growth. Two rounds of screening were conducted to identify the L-proline exporter. For the first round of screening, the resultant 392 strains with CRISPRi system targeting 392 potential transporter genes and a control strain with a nontargeting CRISPRi system were cultivated in 96-deep-well plates for L-proline production. Three replicates were conducted. For seed preparation, strains were cultivated in 96-deep-well plates containing 200 μ L TSC medium in each well at 30 °C, 800 rpm, and 90% moisture for 8 h. Because determining and adjusting the OD_{600nm} values of seed cultures for all the tested strains is laborious, 3 μ L of each seed culture was used to inoculate 96-deep-well plates containing 200 μ L TSD medium in each well. Several wells were sampled to determine the initial OD_{600nm}, which was approximately 0.06. The inoculated 96-deep-well plates were cultivated at 30 °C, 800 rpm, and 90% moisture for 21 h and cultivation was stopped for glucose and L-proline assay. IPTG (0.03 mM) was added to induce dCas9 expression. The strains that showed significant decrease in L-proline production were selected for a second round of screening in 24-deep-well plates. Three replicates were conducted. For seed preparation, strains were cultivated in 24-deep-well plates containing 800 μ L TSC medium in each well at 30 °C, 800 rpm, and 90% moisture for 8 h. The OD_{600nm} values of seed cultures were detected and adjusted to the same value of 2.4 by dilution with TSC medium. To start the cultivation, 20 μ L of the diluted seed culture was used to inoculate 24-deep-well plates containing 800 μ L TSD medium in each well to an initial OD_{600nm} of 0.06. The 24-deep-well plates were cultivated at 30 °C, 800 rpm, and 90% moisture for 18 h and cultivation was stopped for glucose and L-proline assay. The target genes whose repression caused significant decrease in L-proline production were further analyzed.

Peptide uptake and amino acid export assay. To determine the peptide uptake and amino acid export, *C. glutamicum* wild-type strain, *cgl2622*-deleted strain, and *cgl2622* complemented strain were cultivated overnight. Cells were used to inoculate CGXII medium⁸⁵ containing 20 g/L glucose and 1 mM Thr-Pro peptide to an initial OD_{600nm} of 2.0. After incubation at 30 °C and 220 rpm for 2 h, cells were harvested by centrifugation at 5,000 \times g for 10 min and resuspended in pre-warmed CGXII medium (30 °C) containing 20 g/L glucose and 3 mM Thr-Pro peptide to an initial OD_{600nm} of 8.0. Cells were incubated at 30 °C and 220 rpm and samples were taken every 20 min. Intracellular and extracellular L-proline concentrations were quantified using HPLC as described below.

Analytical methods. Glucose in the medium was quantified using an SBA-40D biosensor analyzer (Institute of Biology of Shandong Province Academy of Sciences, Jinan, China). Extracellular L-proline concentrations of cultivation in deep-well plates and shake flasks were quantified according to the method based on the acid-ninhydrin reaction⁸⁶. Extracellular amino acids of cultivation in 5 L bioreactors were quantified using a HPLC method and an L-8900 Amino Acid Analyzer (Hitachi, Japan). Cell cultures were centrifuged at 13,000 \times g for 5 min, and the supernatant was used for detection after appropriate dilution. The HPLC system consists of a Prominence UFLC (Shimadzu, Japan) equipped with a Zorbax Eclipse AAA column (4.6 mm \times 150 mm, 5 μ m, Agilent Technologies, USA) and a UV detector³⁰. A gradient of 50 mM sodium acetate buffer at pH 6.4 with a gradient solution containing acetonitrile-water (50%, v/v) was used as the eluent. Amino acids were detected as their 2,4-dinitrofluorobenzene derivatives at 360 nm by following the precolumn derivation method. The L-8900 Amino Acid Analyzer (Hitachi, Japan) was used to analyze the amino acid profile of the fermentation broth of strain PRO-19 using the amino acids mixture standard solutions (Type AN-2 and Type B, FUJIFILM Wako Pure Chemical Corporation, Japan) as standards. Analysis was performed according to the manufacturer's instructions. Intracellular amino acids were extracted using a centrifugal separation procedure involving centrifugation of cells through a layer of silicon oil⁸⁷ and then quantified using the abovementioned HPLC method. The intracellular volume used to calculate the internal amino acid concentration was 1.7 μ L/mg DCW⁸⁸.

Statistics. Error bars indicate standard deviations from three parallel experiments. All *P* values were generated from two-tailed *t*-tests using the Microsoft Excel 2016 (Microsoft Corporation, USA).

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

Data availability

The data supporting the findings of this work are available within the paper and the Supplementary Information files. A reporting summary for this article is available as a Supplementary Information file. The crystal structure of γ -glutamyl kinase from *E. coli* (PDB ID: 2J5T, www.rcsb.org/structure/2J5T) was used as a template for modeling of CgProB. Source data are provided with this paper.

Received: 17 September 2021; Accepted: 24 January 2022;

Published online: 16 February 2022

References

- Lee, S. Y. et al. A comprehensive metabolic map for production of bio-based chemicals. *Nat. Catal.* **2**, 18–33 (2019).
- Lee, S. Y. & Kim, H. U. Systems strategies for developing industrial microbial strains. *Nat. Biotechnol.* **33**, 1061–1072 (2015).
- Park, S. H. et al. Metabolic engineering of *Corynebacterium glutamicum* for L-arginine production. *Nat. Commun.* **5**, 4618 (2014).
- Meadows, A. L. et al. Rewriting yeast central carbon metabolism for industrial isoprenoid production. *Nature* **537**, 694–697 (2016).
- Zheng, B. et al. Utilization of rare codon-rich markers for screening amino acid overproducers. *Nat. Commun.* **9**, 3616 (2018).
- Sun, X. et al. Isoleucyl-tRNA synthetase mutant based whole-cell biosensor for high-throughput selection of isoleucine overproducers. *Biosens. Bioelectron.* **172**, 112783 (2021).
- Carqueijeiro, I. et al. Beyond the semi-synthetic artemisinin: metabolic engineering of plant-derived anti-cancer drugs. *Curr. Opin. Biotechnol.* **65**, 17–24 (2020).
- Choi, K. R. et al. Systems metabolic engineering strategies: integrating systems and synthetic biology with metabolic engineering. *Trends Biotechnol.* **37**, 817–837 (2019).
- Lee, J. W. et al. Systems metabolic engineering of microorganisms for natural and non-natural chemicals. *Nat. Chem. Biol.* **8**, 536–546 (2012).
- Nishida, K. & Kondo, A. CRISPR-derived genome editing technologies for metabolic engineering. *Metab. Eng.* **63**, 141–147 (2021).
- Canver, M. C. et al. Integrated design, execution, and analysis of arrayed and pooled CRISPR genome-editing experiments. *Nat. Protoc.* **13**, 946–986 (2018).
- Lian, J., Schultz, C., Cao, M., Hamedirad, M. & Zhao, H. Multi-functional genome-wide CRISPR system for high throughput genotype-phenotype mapping. *Nat. Commun.* **10**, 5794 (2019).
- Wang, T. et al. Pooled CRISPR interference screening enables genome-scale functional genomics study in bacteria with superior performance. *Nat. Commun.* **9**, 2475 (2018).
- Yao, L. et al. Pooled CRISPRi screening of the cyanobacterium *Synechocystis* sp PCC 6803 for enhanced industrial phenotypes. *Nat. Commun.* **11**, 1666 (2020).
- Peters, J. M. et al. A comprehensive, CRISPR-based functional analysis of essential genes in bacteria. *Cell* **165**, 1493–1506 (2016).
- Liu, X. et al. High-throughput CRISPRi phenotyping identifies new essential genes in *Streptococcus pneumoniae*. *Mol. Syst. Biol.* **13**, 931 (2017).
- Becker, J., Rohles, C. M. & Wittmann, C. Metabolically engineered *Corynebacterium glutamicum* for bio-based production of chemicals, fuels, materials, and healthcare products. *Metab. Eng.* **50**, 122–141 (2018).
- Wolf, S. et al. Advances in metabolic engineering of *Corynebacterium glutamicum* to produce high-value active ingredients for food, feed, human health, and well-being. *Essays Biochem.* **65**, 197–212 (2021).
- Wendisch, V. F. Microbial production of amino acids and derived chemicals: synthetic biology approaches to strain development. *Curr. Opin. Biotechnol.* **30**, 51–58 (2014).
- Zhang, Q. et al. Amino acids in swine nutrition and production in Amino acids in nutrition and health: Amino acids in the nutrition of companion, zoo and farm animals. (ed. Wu, G.) 81–107 https://doi.org/10.1007/978-3-030-54462-1_6 (Springer International Publishing, Cham; 2021).
- Xu, J. Z., Ruan, H. Z., Yu, H. B., Liu, L. M. & Zhang, W. Metabolic engineering of carbohydrate metabolism systems in *Corynebacterium glutamicum* for improving the efficiency of L-lysine production from mixed sugar. *Microb. Cell Fact.* **19**, 39 (2020).
- Ikeda, M. & Takeno, S. Recent advances in amino acid production in *Corynebacterium glutamicum*: Biology and Biotechnology. (eds. M. Inui & K. Toyoda) 175–226 https://doi.org/10.1007/978-3-030-39267-3_7 (Springer International Publishing, Cham; 2020).
- Zhang, J. et al. De novo engineering of *Corynebacterium glutamicum* for L-proline production. *ACS Synth. Biol.* **9**, 1897–1906 (2020).
- Zhang, Y. et al. A new genome-scale metabolic model of *Corynebacterium glutamicum* and its application. *Biotechnol. Biofuels* **10**, 169 (2017).
- Eggeling, L. Exporters for production of amino acids and other small molecules. *Adv. Biochem. Eng. Biotechnol.* **159**, 199–225 (2017).
- Perez-García, F. & Wendisch, V. F. Transport and metabolic engineering of the cell factory *Corynebacterium glutamicum*. *FEMS Microbiol. Lett.* **365**, fny166 (2018).
- Vrljic, M., Sahm, H. & Eggeling, L. A new type of transporter with a new type of cellular function: L-lysine export from *Corynebacterium glutamicum*. *Mol. Microbiol.* **22**, 815–826 (1996).

28. Simic, P., Sahm, H. & Eggeling, L. L-Threonine export: use of peptides to identify a new translocator from *Corynebacterium glutamicum*. *J. Bacteriol.* **183**, 5317–5324 (2001).
29. Trotschel, C., Deutenberg, D., Bathe, B., Burkovski, A. & Kramer, R. Characterization of methionine export in *Corynebacterium glutamicum*. *J. Bacteriol.* **187**, 3786–3794 (2005).
30. Wang, Y. et al. A novel *Corynebacterium glutamicum* L-glutamate exporter. *Appl. Environ. Microbiol.* **84**, e02691–17 (2018).
31. Cho, J. S. et al. CRISPR/Cas9-coupled recombineering for metabolic engineering of *Corynebacterium glutamicum*. *Metab. Eng.* **42**, 157–167 (2017).
32. Liu, J. et al. Development of a CRISPR/Cas9 genome editing toolbox for *Corynebacterium glutamicum*. *Microb. Cell Fact.* **16**, 205 (2017).
33. Peng, F. et al. Efficient gene editing in *Corynebacterium glutamicum* using the CRISPR/Cas9 system. *Microb. Cell Fact.* **16**, 201 (2017).
34. Wang, B. et al. A RecET-assisted CRISPR–Cas9 genome editing in *Corynebacterium glutamicum*. *Microb. Cell Fact.* **17**, 63 (2018).
35. Jiang, Y. et al. CRISPR–Cpf1 assisted genome editing of *Corynebacterium glutamicum*. *Nat. Commun.* **8**, 15179 (2017).
36. Krumbach, K., Sonntag, C. K., Eggeling, L. & Marienhagen, J. CRISPR/Cas12a mediated genome editing to introduce amino acid substitutions into the mechanosensitive channel MscCG of *Corynebacterium glutamicum*. *ACS Synth. Biol.* **8**, 2726–2734 (2019).
37. Zhang, J., Yang, F., Yang, Y., Jiang, Y. & Huo, Y. X. Optimizing a CRISPR–Cpf1-based genome engineering system for *Corynebacterium glutamicum*. *Microb. Cell Fact.* **18**, 60 (2019).
38. Zhao, N. et al. Multiplex gene editing and large DNA fragment deletion by the CRISPR/Cpf1–RecE/T system in *Corynebacterium glutamicum*. *J. Ind. Microbiol. Biotechnol.* **47**, 599–608 (2020).
39. Leenay, R. T. et al. Identifying and visualizing functional PAM diversity across CRISPR–Cas systems. *Mol. Cell* **62**, 137–147 (2016).
40. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
41. Cui, L. & Bikard, D. Consequences of Cas9 cleavage in the chromosome of *Escherichia coli*. *Nucleic Acids Res.* **44**, 4243–4251 (2016).
42. Sadler, J. R., Sasnor, H. & Betz, J. L. A perfectly symmetric *lac* operator binds the *lac* repressor very tightly. *Proc. Natl Acad. Sci. U. S. A.* **80**, 6785–6789 (1983).
43. Binder, S., Siedler, S., Marienhagen, J., Bott, M. & Eggeling, L. Recombineering in *Corynebacterium glutamicum* combined with optical nanosensors: a general strategy for fast producer strain generation. *Nucleic Acids Res.* **41**, 6360–6369 (2013).
44. Long, M. et al. Significantly enhancing production of *trans*-4-hydroxy-L-proline by integrated system engineering in *Escherichia coli*. *Sci. Adv.* **6**, eaba2383 (2020).
45. Kudla, G., Murray, A. W., Tollervey, D. & Plotkin, J. B. Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* **324**, 255–258 (2009).
46. Pérez-Arellano, I., Rubio, V. & Cervera, J. Mapping active site residues in glutamate-5-kinase. The substrate glutamate and the feed-back inhibitor proline bind at overlapping sites. *FEBS Lett.* **580**, 6247–6253 (2006).
47. Pérez-Arellano, I., Carmona-Álvarez, F., Gallego, J. & Cervera, J. Molecular mechanisms modulating glutamate kinase activity. Identification of the proline feedback inhibitor binding site. *J. Mol. Biol.* **404**, 890–901 (2010).
48. Orth, J. D., Thiele, I. & Palsson, B. Ø. What is flux balance analysis? *Nat. Biotechnol.* **28**, 245–248 (2010).
49. Bott, M. & Eikmanns, B. J. TCA cycle and glyoxylate shunt of *Corynebacterium glutamicum* in *Corynebacterium glutamicum*: Biology and Biotechnology. (eds. Yukawa, H. & Inui, M.) 281–313 https://doi.org/10.1007/978-3-642-29857-8_10 (Springer Berlin Heidelberg, Berlin, Heidelberg; 2013).
50. Sauer, U. & Eikmanns, B. J. The PEP–pyruvate–oxaloacetate node as the switch point for carbon flux distribution in bacteria. *FEMS Microbiol. Rev.* **29**, 765–794 (2005).
51. Chen, Z. et al. Metabolic engineering of *Corynebacterium glutamicum* for the production of 3-hydroxypropionic acid from glucose and xylose. *Metab. Eng.* **39**, 151–158 (2017).
52. Takeno, S., Murata, R., Kobayashi, R., Mitsuhashi, S. & Ikeda, M. Engineering of *Corynebacterium glutamicum* with an NADPH-generating glycolytic pathway for L-lysine production. *Appl. Environ. Microbiol.* **76**, 7154–7160 (2010).
53. Komati Reddy, G., Lindner, S. N. & Wendisch, V. F. Metabolic engineering of an ATP-neutral Embden–Meyerhof–Parnas pathway in *Corynebacterium glutamicum*: growth restoration by an adaptive point mutation in NADH dehydrogenase. *Appl. Environ. Microbiol.* **81**, 1996–2005 (2015).
54. Hoffmann, S. L. et al. Cascaded valorization of brown seaweed to produce L-lysine and value-added products using *Corynebacterium glutamicum* streamlined by systems metabolic engineering. *Metab. Eng.* **67**, 293–307 (2021).
55. Lee, M. J. & Kim, P. Recombinant protein expression system in *Corynebacterium glutamicum* and its application. *Front. Microbiol.* **9**, 2523 (2018).
56. Walter, T. et al. Fermentative *N*-methylanthranilate production by engineered *Corynebacterium glutamicum*. *Microorganisms* **8**, 866 (2020).
57. Mutalik, V. K. et al. Precise and reliable gene expression via standard transcription and translation initiation elements. *Nat. Methods* **10**, 354–360 (2013).
58. 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).
59. Rohlhill, J., Sandoval, N. R. & Papoutsakis, E. T. Sort-seq approach to engineering a formaldehyde-inducible promoter for dynamically regulated *Escherichia coli* growth on methanol. *ACS Synth. Biol.* **6**, 1584–1595 (2017).
60. 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).
61. Moxley, M. A. & Becker, D. F. Rapid reaction kinetics of proline dehydrogenase in the multifunctional proline utilization A protein. *Biochemistry* **51**, 511–520 (2012).
62. Skjerdal, O. T. et al. Changes in intracellular composition in response to hyperosmotic stress of NaCl, sucrose or glutamic acid in *Brevibacterium lactofermentum* and *Corynebacterium glutamicum*. *Appl. Microbiol. Biotechnol.* **44**, 635–642 (1996).
63. Peter, H., Weil, B., Burkovski, A., Kramer, R. & Morbach, S. *Corynebacterium glutamicum* is equipped with four secondary carriers for compatible solutes: identification, sequencing, and characterization of the proline/ectoine uptake system, ProP, and the ectoine/proline/glycine betaine carrier, EctP. *J. Bacteriol.* **180**, 6005–6012 (1998).
64. Elbourne, L. D., Tetu, S. G., Hassan, K. A. & Paulsen, I. T. TransportDB 2.0: a database for exploring membrane transporters in sequenced genomes from all domains of life. *Nucleic Acids Res.* **45**, D320–D324 (2017).
65. Ikeda, M. Sugar transport systems in *Corynebacterium glutamicum*: features and applications to strain development. *Appl. Microbiol. Biotechnol.* **96**, 1191–1200 (2012).
66. Pfeifer-Sancar, K., Mentz, A., Rückert, C. & Kalinowski, J. Comprehensive analysis of the *Corynebacterium glutamicum* transcriptome using an improved RNAseq technique. *BMC Genomics* **14**, 888 (2013).
67. Hirasawa, T. & Wachi, M. Glutamate fermentation-2: Mechanism of L-glutamate overproduction in *Corynebacterium glutamicum*. *Adv. Biochem. Eng. Biotechnol.* **159**, 57–72 (2017).
68. Walker, M. C. & van der Donk, W. A. The many roles of glutamate in metabolism. *J. Ind. Microbiol. Biotechnol.* **43**, 419–430 (2016).
69. Li, M. et al. Efficient multiplex gene repression by CRISPR–dCpf1 in *Corynebacterium glutamicum*. *Front. Bioeng. Biotechnol.* **8**, 357 (2020).
70. Wang, Y. et al. Switch of metabolic status: redirecting metabolic flux for acetoin production from glycerol by activating a silent glycerol catabolism pathway. *Metab. Eng.* **39**, 90–101 (2017).
71. Li, Y. et al. Current status on metabolic engineering for the production of L-aspartate family amino acids and derivatives. *Bioresour. Technol.* **245**, 1588–1602 (2017).
72. Widhalm, J. R. et al. Identification of a plastidial phenylalanine exporter that influences flux distribution through the phenylalanine biosynthetic network. *Nat. Commun.* **6**, 8142 (2015).
73. Han, T., Kim, G. B. & Lee, S. Y. Glutaric acid production by systems metabolic engineering of an L-lysine-overproducing *Corynebacterium glutamicum*. *Proc. Natl Acad. Sci. U. S. A.* **117**, 30328–30334 (2020).
74. Yen, M. R. et al. The ubiquitous ThrE family of putative transmembrane amino acid efflux transporters. *Res. Microbiol.* **153**, 19–25 (2002).
75. Simic, P., Willuhn, J., Sahm, H. & Eggeling, L. Identification of *glyA* (encoding serine hydroxymethyltransferase) and its use together with the exporter ThrE to increase L-threonine accumulation by *Corynebacterium glutamicum*. *Appl. Environ. Microbiol.* **68**, 3321–3327 (2002).
76. Kennerknecht, N. et al. Export of L-isoleucine from *Corynebacterium glutamicum*: a two-gene-encoded member of a new translocator family. *J. Bacteriol.* **184**, 3947–3956 (2002).
77. Borngen, K. et al. The properties and contribution of the *Corynebacterium glutamicum* MscS variant to fine-tuning of osmotic adaptation. *Biochim. Biophys. Acta* **1798**, 2141–2149 (2010).
78. Hashimoto, K. et al. The protein encoded by NCgl1221 in *Corynebacterium glutamicum* functions as a mechanosensitive channel. *Biosci. Biotechnol. Biochem.* **74**, 2546–2549 (2010).
79. Ruan, Y., Zhu, L. & Li, Q. Improving the electro-transformation efficiency of *Corynebacterium glutamicum* by weakening its cell wall and increasing the cytoplasmic membrane fluidity. *Biotechnol. Lett.* **37**, 2445–2452 (2015).
80. Marco-Marín, C. et al. A novel two-domain architecture within the amino acid kinase enzyme family revealed by the crystal structure of *Escherichia coli* glutamate 5-kinase. *J. Mol. Biol.* **367**, 1431–1446 (2007).

81. Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461 (2010).
82. Ebrahim, A., Lerman, J. A., Palsson, B. O. & Hyduke, D. R. COBRApy: COncstraints-Based Reconstruction and Analysis for Python. *BMC Syst. Biol.* **7**, 74 (2013).
83. Xie, S., Shen, B., Zhang, C., Huang, X. & Zhang, Y. sgRNAs9: a software package for designing CRISPR sgRNA and evaluating potential off-target cleavage sites. *PLoS One* **9**, e100448 (2014).
84. Wang, Y. et al. MACBETH: multiplex automated *Corynebacterium glutamicum* base editing method. *Metab. Eng.* **47**, 200–210 (2018).
85. Keilhauer, C., Eggeling, L. & Sahm, H. Isoleucine synthesis in *Corynebacterium glutamicum*: molecular analysis of the *ilvB-ilvN-ilvC* operon. *J. Bacteriol.* **175**, 5595–5603 (1993).
86. Bates, L. S., Waldren, R. P. & Teare, I. D. Rapid determination of free proline for water-stress studies. *Plant Soil* **39**, 205–207 (1973).
87. Forget, R. S., Martin, J. E. & Cote, R. H. A centrifugal separation procedure detects moderate affinity cGMP binding sites in membrane-associated proteins and permeabilized cells. *Anal. Biochem.* **215**, 159–161 (1993).
88. Ruffert, S., Lambert, C., Peter, H., Wendisch, V. F. & Krämer, R. Efflux of compatible solutes in *Corynebacterium glutamicum* mediated by osmoregulated channel activity. *Eur. J. Biochem.* **247**, 572–580 (1997).

Acknowledgements

This work was supported by the National Key R&D Program of China (2018YFA0901400 [JS]), National Natural Science Foundation of China (31800038 [JL], 31870044 [PZ], and 21908239 [QY]), Key R&D Program of Shandong Province (2021CXGC010602 [PZ]), International Partnership Program of Chinese Academy of Sciences (153D31KYSB20170121 [JS]), Youth Innovation Promotion Association of Chinese Academy of Sciences (2021177 [YW]), Tianjin “Project+Team” Key Training Program (XC202038 [YW]), and Tianjin Synthetic Biotechnology Innovation Capacity Improvement Project (TSBICIP-KJGG-005 [PZ]).

Author contributions

Conceptualization: J.L., Y.W., P.Z., J.S. Methodology: J.L., M.L., T.S., X.N., Q.Y., J.C., Y.W., P.Z. Investigation: J.L., M.L., T.S., G.S., N.G., X.Z., X.G., X.N., Q.Y., J.F., Z.L., Y.G., J.C., Y.W. Visualization: J.L., Y.W. Supervision: Y.W., P.Z., J.S. Writing-original draft: Y.W. Writing-review & editing: J.L., Y.W., P.Z., J.S.

Competing interests

The engineered enzymes, promoters, and proline producing strains described in this paper are covered by patents CN202010696984.7, CN202010696983.2, CN202010838604.9, CN202011316277.7, CN202111045175.4, and CN202110952846.5. J.L., M.L., T.S., G.S., X.G., J.C., Y.W., P.Z., and J.S. are listed as co-inventors of the patents. The rest of authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-28501-7>.

Correspondence and requests for materials should be addressed to Yu Wang or Ping Zheng.

Peer review information *Nature Communications* thanks Reinhard Krämer, Zhiming Rao and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

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) 2022