Shih-I Tan ORCID iD: 0000-0002-7632-9991

I-Son Ng ORCID iD: 0000-0003-1659-5814

CRISPRi-mediated Programming Essential Gene *can* as a Direct Enzymatic Performance Evaluation & Determination (DEPEND) System

Shih-I Tan, Peng-Jui Yu, I-Son Ng*

Department of Chemical Engineering, National Cheng Kung University, Tainan 701, Taiwan

*Corresponding author: Prof. I-Son Ng

E-mail: yswu@mail.ncku.edu.tw

ORCID:

I-Son Ng: 0000-0003-1659-5814

Shih-I Tan: 0000-0002-7632-9991

Tel: +866-62757575; Fax: +86-62344496

Abstract

Harnessing enzyme expression for production of target chemicals is a critical and multifarious process, where screening of different genes by inspection of

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/bit.27443.

enzymatic activity plays an imperative role. Here, we conceived an idea to improve the time-consuming and labor-intensive process of enzyme screening. Controlling cell growth was achieved by the CRISPRi system with different sgRNA targeting the essential gene *can* (CRISPRi::CA) that encodes a carbonic anhydrase for CO_2 uptake. CRISPRi::CA comprises a whole-cell biosensor to monitor CO_2 concentration, ranging from 1% to 5%. Based on CRISPRi::CA, an effective and simple Direct Enzymatic Performance Evaluation & Determination (DEPEND) system was developed by a single step of plasmid transformation for targeted enzymes. As a result, the activity of different carbonic anhydrases corresponded to the colony forming units (CFU). Furthermore, the enzymatic performance of 5-aminolevulinic acid synthetase (ALAS), which converts glycine and succinate-CoA to release a molecule of CO_2, has also been distinguished, and the effect of the chaperone GroELS on ALAS enzyme folding was successfully identified in the DEPEND system. We provide a highly feasible, time-saving, and flexible technology for the screening and inspection of high-performance enzymes, which may accelerate protein engineering in the future.

Keywords: CRISPRi, essential gene, carbonic anhydrase, aminolevulinic acid synthetase, performance, whole-cell biosensor

1. Introduction

In metabolic engineering, reconstruction of new cellular pathways depends on the meticulous perfection of enzymatic performance, including codon optimization (Grote et al., 2005; Webster et al., 2017), 5'-untranslated regions (5'-UTR) (Cheong et al., 2015; Seo et al., 2014), fusion tag (Tan et al., 2018), and

enzyme selection (Jeske et al., 2019). With all due considerations, Cambray and his co-workers constructed a library of 24400 synthetic green fluorescent protein DNA sequences to evaluate the translational level and revealed a primitive design principle (Cambray et al., 2018). Except for optimization of native protein expression, enzyme evolution is a crucial strategy to increase enzymatic performance, because mutations change the codon usage and amino acid sequence; thus, the intrinsic quality (i.e. catalytic constant), quantity (i.e., protein expression level), and extrinsic quality (solubility ratio) would also change (Fig. 1). Recently, Wang and her colleagues could maintain a high solubility of mutants in an evolution process called “soluble expression phage-assisted continuous evolution” (SE-PACE), which was used to revolutionize the maltose-binding protein and simultaneously improve its solubility and activity (Wang et al., 2018). SE-PACE successfully included the factor of extrinsic quality to obtain higher enzymatic performance. However, it required a complex design with a split T7 RNAP reporter for soluble expression and an AND gate, as well as individual verification of enzymatic activity. Therefore, to establish a more expedient system or platform for characterization of enzymatic performance is urgent and meaningful for metabolic engineering.

Essential genes play a vital role for survival, and 248 of the 4288 genes are essential in *E. coli* (Goodall et al., 2018). Manipulating an essential gene is a splendid approach to achieve the goals of synthetic biology. Previous studies have shown that knocking out some essential genes is a successful strategy to establish an auxotrophic system for enzyme evolution (Atkinson et al., 2019), a biocontainment system (Lee et al., 2018), and plasmid stabilization (Kang et al.,

2018). Notably, in essential gene-mediated plasmid stabilization, achieved by knocking out *inf*A from the chromosome and introducing it under different promoter strength in the plasmid, the cell regulated the plasmid copy number to appropriately complement the defect of *inf*A, showing that increasing *inf*A promoter strength in plasmid resulted in lower plasmid copy numbers, and vice versa (Kang et al., 2018). Furthermore, regulation of essential genes by replacing the promoter region with an inducible promoter enables researchers to control the cell growth phase (Min et al., 2012). Brockman and Prather demonstrated a two-fold increase in myo-inositol concertation via dynamic control of the essential *pfk*A gene by adding a protein degradation tag (Brockman and Prather, 2015).

Can is an essential gene that encodes a continuously expressed β-class carbonic anhydrase (CA) in *E. coli* (Merlin et al., 2003). CA performs the enzymatic interconversion between the CO_2 and bicarbonate (Jo et al., 2019). *Can*-deficient *E. coli* only survived in a 5% CO_2 environment (Hashimoto and Kato, 2003). Besides, *E. coli* demand a supply of bicarbonate/CO_2 as a metabolic substrate during normal growth for biosynthesis of various small molecules, including arginine, pyrimidines, and purines, and fatty acid biosynthesis in central metabolism (Charles and Roberts, 1968).

In recent years, Cluster Regularly Interspaced Short Palindromic Repeat (CRISPR) and CRISPR interference have been explored to programming the cell (Jinek et al., 2012); control the gene expression levels (Qi et al., 2013), tuning multi-complexed metabolic flux in *E. coli* (Gao et al., 2018) and even capture the image of genome editing and transcription as a powerful tool (Wang et al., 2019). However, this is the first attempt of the use of CRISPR interference targeting the

can gene (CRISPRi::CA) to control cell growth. The different targeted sites on the *can* gene offer a different cell inhibition efficiency. CRISPRi::CA is utilized as a Direct Enzymatic Performance Evaluation & Determination (DEPEND) system to characterize the enzymatic performance of the CO_2-converting enzymes CA and 5-aminolevulinic acid synthetase (ALAS).

2. Materials and Methods

2.1 Strains, Plasmids, Primers, and Media

The *E. coli* strains and plasmids used in this study have been listed in Table 1, and the primers have been listed in Table S1. *E. coli* DH5α was used in regular cloning. CRISPRi-mediated cell inhibition, and the CO_2 whole-cell biosensor used DH5α. For the screening of enzymatic performance, BL21(DE3) was used. All cultures were in a 37°C incubator shaking at 200 rpm, except when CRISPR::CA was applied as a CO_2 biosensor, where the cells were cultured at 25°C at 250 rpm stirring. For regular culture, we used the Luria–Bertani (LB) medium (1.5% tryptone, 1.5% NaCl, and 0.5% yeast extract), supplemented with 100 mg/L Ampicillin (Ap), 50 mg/L Kanamycin (Km), and 25 mg/L Chlorophenol (Cm) for a single plasmid; 50 mg/L Ap or 12.5 mg/L Cm for dual plasmids; and 25 mg/L Ap, 12.5 mg/L Km, and 8.25 mg/L Cm for 3 plasmids. The cells were first cultured in LB overnight and then inoculated (2 % *v/v*) into the MM9G medium (12.8 g/L Na_2HPO_4, 3 g/L KH_2PO_4, 0.5 g/L NaCl, 1 g/L NH_4Cl, 0.24 g/L $MgSO_4$, 0.011 g/L $CaCl_2$, and 4 g/L glucose) with proper antibiotics. For characterization, LB or LB supplemented with 3 g/L glycine, 1 g/L succinate, and 50 μM pyridoxal

5’-phosphate (PLP) was used. If not specially noted, the IPTG concentration was 0.1 mM.

2.2 Plasmid Construction

All the sequences of the genetic elements used in this study have been listed in Table S2. The SynP promoter was obtained by primer annealing of XbaI-SynP-F and BamHI-SynP-R. The fragment was then inserted into pSUI-PlacI-sfGFP for promoter replacement by digestion with *Xba*I/*Bam*HI. For pSUI-PlacI-sfGFP, the *lacI* gene was amplified from *E. coli* BL21(DE3), and the construction was completed with *Bgl*II/*Pst*I digestion. Primer PstI-SpeI-Scaffold-R was designed to acquire sgRNA for dCas9 targeting on multiple sites on the *can* gene with corresponding primers. The sgRNA fragments were further implemented into pSUI-SynP-sfGFP with *Bam*HI/*Pst*I restriction sites and removal of sfGFP. The sequence of hCAII, SyCA, RshemA, and RchemA, which was synthesized by IDT, were amplified by PCR with *Eco*RI/*Hind*III restriction sites and constructed into the pSIT plasmid. The chaperone *gro*ELS was acquired from the *E. coli* genome and inserted down-stream of RchemA in pSIT-RchemA with *Nde*I/*Xho*I to obtain pSIT-RchemA-GroELS.

2.3 Optical Density Measurement of Biomass and Fluorescence Measurement

A 200 μL of the broth was added into a transparent 96-well plate. Optical density in terms of biomass and fluorescence was measured at 600 nm wavelength, 480 nm excitation, and 510 nm emission using the SpectraMax M2 Microplate Readers (Molecular Devices, Sunnyvale, CA).

2.4 Dose-Response of IPTG for Synthetic and PlacI Promoter

The DH5α harboring the pSUI-SynP-sfGFP or pSUI-PlacI-sfGFP plasmid was precultured at 37°C shaking at 200 rpm for 12 h and then inoculated at 2% ratio into 10 mL of the LB medium with different concentrations of IPTG for culturing. The cells were cultured for 12 h, followed by measuring the optical density and fluorescence intensity. The data was then simulated by the Hill's equation, as follows:

$$y = y_0 + \frac{ax^b}{c^b + x^b}$$

In the Hill's equation, y indicates the specific fluorescence corresponding to the concentration of IPTG at the level x; y_o denotes the leakage expression of the inducible promoter using the same units as y; a, b, and c indicate the dynamic range, Hill coefficient, and IPTG concentration at the point achieving half of the dynamic range, respectively. All parameters can be calculated by fitting the experimental data via the SigmaPlot 12.1 software (Systat Software, San Jose, CA) according to the manufacturer's instructions. Non-linear or linear least-squares regression was used to minimize errors between the fitted and experimental data.

2.5 Determination of CA Activity

The cells expressing the CA enzyme were collected, washed 3 times with deionized water, and finally adjusted to OD of 5 at 600 nm for further use. The CA activity assay was conducted as in our previous publication (Tan et al., 2018). Briefly, 9 mL of chilled Tris-HCl (20 mM, pH 8.3) buffer and 0.2 mL of the

enzyme were mixed and put in a 20-mL sample bottle at 0°C while stirring. Then, 6 mL of chilled CO_2-saturated solution was added into the sample bottle immediately. The time (in seconds) for the pH value to drop from 8.3 to 6.3 was recorded. CA activity was calculated by a Wilbur–Anderson unit (WAU) per milliliter of sample. The definition for WAU is $(T_0-T)/T$, where T and T_0 are the time from pH 8.3 to 6.3 with and without CA.

2.6 Determination of CFU and Cell Recovery in the DEPEND System

The competent BL21(DE3) cells were incubated with 400 ng of each plasmid, pSM-dCas9::Ap, pSUI-SynP-CAPsg, and pSIT-GOI in an ice bath for 30 min, and then the mixture was transferred to 42°C for 90 seconds, followed by immediate incubation in an ice bath for 10 min. After heat shock treatment, the cells were mixed with 0.4 mL of the LB medium and cultured in a 37°C incubator, shaking at 200 rpm for a one-hour recovery. The cells were spread on the LB plate with appropriate antibiotics and cultured in a 37°C incubator for 12 h. The CFU was calculated. If a colony could not form, the process was changed to use 0.4 mL of the LB medium supplemented with 3 g/L glycine, 1 g/L succinate, and 50 μM PLP, followed by inoculation in a 96-well plate for monitoring cell recovery every 10 min for 5 h.

3. Results and Discussion

3.1 Controlled Cell Growth by CRISPRi of the Essential Gene *can*

In order to program cell growth by CRISPRi, we designed a synthetic promoter (Psyn) which contained a -35 (5’-TTGACG) and -10 (5’-TACAGT) sequences

copy from the constitutive J23100 promoter, a truncated *lac*O' of 15 bp inserted between both binding sites, and *lac*O of 23 bp was before -35 binding sequences (Fig. 2a). The Psyn was then cloned in front of sfGFP, and the resulting plasmid was used to verify the promoter strength (Fig. 2b). The dose-response curve showed that the Psyn responded to a different IPTG induction (Fig. 2c) with a dynamic range (a) of 2163 a.u. and a basal expression level (y_0) of 1192 a.u., showing a 3.82-fold change (Table 2). As a comparison, the strong PlacI promoter, was selected and characterized by a dose-response curve (Figure S1). The fitting parameters displaying a basal expression (y_0) of 3660 a.u., a dynamic range (a) of 33960 a.u., and a half-dynamic range (c) of 0.009 mM were much higher than those of the Psyn promoter. Conversely, both promoters showed identical Hill's coefficient of 1.4. Using the provided methodology of *in-silico* analysis (Salis et al., 2009), we found the initial translation rate for the Psyn and PlacI promoter of 33.5 a.u. and 538.3 a.u., respectively (Fig. S2). We speculated that the extra sequences (GGAUCC) in front of the B0034 ribosome binding site (RBS) of the PlacI promoter assisted ribosome binding, accounting for the higher translation rate.

The molecular ratio of the regulator and operon and promoter strength are two important factors for developing a tight inducible system. Recently, Schuller et al. discovered that *lac*I and *lac*O must be tuned to a well-balanced status to develop an inducible genome-integrated recombinant protein expression system (Sculler et al., 2020). Therefore, we considered that our designed Psyn was imbalanced, so it served as a constitutive promoter for RNA synthesis, but it also behaved as a weak

inducible promoter for protein expression with a low basal expression level, due to the RBS design.

After characterization of Psyn, we constructed dual plasmids in a CRISPRi system called “CRISPRi::CA”, where one plasmid was responsible for sgRNA targeting of the *can* gene driven by Psyn, and the other plasmid encoded the dCas9 protein (Fig. 2d). Four targeting sites on the *can* gene were assessed for its capability to control cell growth (Fig. 2e). Individual targeting sites conferred cell growth rates in the order of wild type > S2 > S1 > SNT > SP (Fig. 2f and Table 3). Targeting the template strain showed less efficacy for the inhibition of cell growth (i.e. 18.8% and 33.8% for S2 and S1, respectively), but targeting the non-template strain achieved an inhibition efficiency of 41.3%. Furthermore, targeting the promoter region markedly inhibited cell growth with an efficiency of 65%. The results were consistent with a previous study (Cao et al., 2017), which showed that sgRNA targeting of the promoter region and non-template strand of the GFP gene conferred a strong inhibition of GFP expression. On the other hand, it is worth mentioning that the final OD is similar for different targeting sites (Table 3), indicating that the *can* gene is essential but the essentiality must be further studied.

CRISPRi is an emerging technology that has been applied to control cell morphology, map phenotypes, localize gene loci, redirect carbon flux, and develop biosensors (Elhadi et al., 2016; Wang et al., 2018; Pickar-Oliver and Gersbach, 2019; Li et al., 2019). Recently, a novel application of CRISPRi was conceived to develop a cell phase self-switchable system, where the cell would automatically and reversibly switch its cell phase between the replication and chemical

production phases (Zheng et al., 2019). Here, we investigated a new design of the *can* gene to control the cell growth rate by using CRISPRi::CA.

3.2 CRISPRi::CA Serves as a CO_2 Whole-Cell Biosensor

In a previous report, the *can*-deficient strain could only survive strictly in an environment with 5% CO_2 (Hashimoto and Kato, 2003). However, using the CRISPRi system is more feasible, because the strain expressing sgRNA targeting the promoter and dCas9 protein (i.e., strain SP) survived under air conditions but had a lower cell growth rate. Therefore, we hypothesized that CRISPRi::CA could be developed as a CO_2 whole-cell biosensor. We verified the cell growth of the SP strain in the CO_2-bubbling device with different CO_2 concentrations (Fig. 3a). The results showed that CO_2 supplementation at 5% could reasonably increase the cell growth, while supplying 10% CO_2 led to severe cell growth inhibition. The cell growth rate was calculated and applied in the dose-response curve (Fig. 3b). The data were fitted by the Hill's equation, showing high sensitivity to CO_2 concentration ranging from 1% to 5%. Decreased cell growth was observed in a 10% CO_2 environment, which resembled anaerobic conditions (Wang et al., 2019). As *E. coli* cultured in minimal medium with 20% CO_2 showed extremely slow growth and adopted the mutation, indicating a strong stress response to increased CO_2 supply (Chless and Rober, 1968).

Another essential gene, *inf*A is directly related to cell viability which the lower expression of *inf*A has guaranteed to control plasmid with minimized cell-to-cell variations (Kang et al., 2018). In our design, the sgRNA binding to the promoter of CA is the most effective one, thus the biomass can be controlled very well even

up to 48 h (Fig. 2f). Therefore, the CO_2 whole-cell biosensor is designed by this sgRNA and shows the high sensitive and stability to CO_2 concentration (Fig. 3a). Actually, the stability of plasmids can be maintained by sufficient antibiotics, while using the effective sgRNA targeted on promoter is also important.

The next generation of valuable renewable resources includes the one-carbon compounds (C1), such as methane, methanol, CO, CO_2, and formate (Dürre and Eikmanns, 2015). A whole-cell biosensor to detect the concentration of C1 compounds is a promising technology to monitor their consumption *de novo*. For liquid C1 compounds (i.e. formate and methanol), sensing modules originate from the acetogens and methanotrophs, and they have been developed as a whole-cell biosensor (Lee et al., 2019), but there remains limited literature regarding gas-phase C1 compound detection through synthetic biology. Now, we have successfully established a CO_2 whole-cell biosensor using essential gene features, coupled with the CRISPRi system.

3.3 Screening CA Performance by CRISPRi::CA

In addition to the exterior effect achieved by CO_2 supplementation (Fig. 3b), interior introduction of heterologous CA expression also restored cell growth. Therefore, we selected two different carbonic anhydrases from *Sulfurihydrogenibium yellowstonense* (SyCA) and human cells (hCAII) and cloned them into the plasmid with RSF ori to achieve replication compatibility and high copy numbers in *E. coli* (Novick, 1987). First, the individual CA expression level and activity were examined, as shown in Fig. 4a and 4b. The two different CAs were successfully produced with considerable expression levels, as

the arrows indicated in Fig. 4a. Inspection of activity was performed by the Wilbur-Anderson assay, where SyCA and hCAII activity under IPTG induction was 1503 and 382 WAU, respectively (Fig. 4b). On the other hand, we considered leakage expression due to the extremely high number of copies by RSF ori, even though there was no obvious expression shown by SDS-PAGE (Fig. 4a). The results showed that the cell continued to show CA activity without induction, but the enzymatic performance (EP) was only 176.7 WAU for SyCA and 86.45 WAU for hCAII, respectively (Fig. 4b). Indeed, CA would continue to be slightly expressed without IPTG induction, because *E. coli* also express the CA gene with certain activity (Tan and Ng, 2020). Furthermore, we have demonstrated that overexpression of CA in a non-deficient strain arrested cell growth and suggested that the CA-deficient strain could rehabilitate its growth (Hsu et al., 2019).

To develop CRISPRi::CA as a DEPEND platform, we co-transformed three plasmids, pSM-dCas9::Ap, pSUI-SynP-SPCAsg, and pSIT-CA, for expression of dCas9, CAPsg, and CA in BL21(DE3). The results showed that the CFU corresponded to CA activity. SyCA showed a higher CFU of 235 compared to 30 for hCAII (Fig. 4c and Fig. S3), indicating that this platform is available for a direct screening of recombinant CA performance. Alvizo and colleagues performed random mutagenesis to generate a CA mutant library and examined a highly active CA mutant from the library via a series of cell treatments (i.e. cell culturing, induction, cell disruption) and an enzymatic assay (Alvizo et al., 2014). The DEPEND system facilitated the process by a single step of comparing the CFU without further procedures. Additionally, we also examined cell growth and showed that the cell in grew faster non-induction conditions, which was consistent

with our previous report (Hsu et al., 2019). Finally, the I-SyCA strain had a higher cell growth rate than that the I-hCAII strain, which was related to its activity and CFU analysis (Fig. 4d). As enzyme evolution by error-prone PCR or DNA shuffling is a ready-to-use technology (Hossain et al., 2016), DEPEND is used for the screening of enzymatic activity after evolution. Herein, we can apply DEPEND system to isolate the best enzyme, which is involving of CO_2 uptake and release.

3.4 Characterization of CO_2-evolving Enzymatic Reaction by CRISPRi::CA

The DEPEND system screens recombinant CA enzymatic performance by the principle of CA gene supplementation. Next, as previous data showed that addition of external CO_2 could rescue cell growth (Fig. 3b), we speculated whether interior CO_2 supply from an enzymatic reaction could also work and further applied the DEPEND system to other kinds of enzyme. Therefore, we selected 5-aminolevulinic acid synthetase (ALAS) from *Rhodobacter spheroides* (RsHemA) and *Rhodobacter capsulatus* (RcHemA), which reacted with glycine and succinyl-CoA to produce 5-aminolevulinic acid and CO_2. Three plasmids, used for dCas9, CAPsg, and ALAS, were co-transformed into BL21(DE3) by induction, but failed to form a colony, indicating that the CO_2 supply from the enzymatic reaction was limited. Consequently, after switching to monitor the recovery process in heat-shock transformation, we found that the addition of RcHemA or RsHemA could not support cell growth in non-induction conditions, while RcHemA with induction could solely support cell growth (Fig. 5a). Furthermore, the cell growth rate was calculated as shown in Fig. 5b. Even though the cell growth rates of RsHemA and RcHemA are both negative, the death rate

(i.e. the negative value of specific growth rate) implies the enzymatic activity of RcHemA is higher than that of RsHemA. Recently, we also discovered that RsHemA achieved 100 mg/L of ALA production and RcHemA achieved 2 g/L (Tan et al., 2020; Yu et al., 2019). Moreover, the plasmid-driven T7 system equipped with RsHemA was a powerful platform for portable and universe expression that achieved 300 mg/L of ALA production (Tan and Ng, 2020). Moreover, Lou et al. also showed that specific activity was higher in RcHemA than that in RsHemA (Jou et al., 2014).

Next, we examined the DEPEND system with a chaperone effect on RcHemA, as addition of the chaperone GroELS could enhance the solubility of RcHemA and further increase the ALA production from 2 g/L to 8 g/L (Yu et al., 2019). The cell survived longer (from 70 min to 100 min) with a larger biomass, but the specific growth rate was similar (Fig. 5a and 5b). Ultimately, the DEPEND system had the potential to screen the activity of various enzymes.

Since the DEPEND system was based on CRISPRi for selection pressure, it is a more flexible and feasible strategy to examine enzymatic performance. Linder et al. demonstrated that a NADPH-auxotrophic *E. coli* (i.e. *zwf, gnd, maeB, icd,* and *pnt*AB were deleted) provided a platform for high-throughput screening of enzyme evolving NADPH regeneration by examination of cell growth (Lindner et al., 2018), which could also be achieved by the DEPEND system without tedious work for the deletion of multiple genes.

4. Conclusion

Biotechnological tools have contributed to advances in synthetic biology, protein engineering, and metabolic engineering, where the CRISPRi tool is one of the most robust approaches. CRISPRi::CA provides a new concept for regulating cell growth by sgRNA targeting of an essential gene. Moreover, CRISPRi::CA has been further explored as a CO_2 biosensor, which is the first development such a biosensor. On the other hand, CRISPRi::CA could also be applied as an enzyme evaluating platform "DEPEND", which only required the examination of cell growth to distinguish the enzymatic performance of proteins that interact with CO_2. By switching the sgRNA design or using multiple target sites, the DEPEND system may be a flexible technology to screen various proteins and even accelerate the process of direct enzyme evolution due to the advantages of faster and simpler inspection of enzymatic performance.

Author Contributions

SIT and ISN conceived the study. SIT and PJY performed all the experiments. SIT sketch the original draft investigation. ISN did methodology validation, supervised the experiments, prepared, review and edited the manuscript.

Declaration of Competing Interest

The authors declare no competing financial interest.

Acknowledgements

The authors are grateful to the financial support for this study provided by the Ministry of Science and Technology (MOST 108-2621-M-006-015, MOST 108-2221-E-006-004-MY3 and MOST 108-2218-E-006-006) in Taiwan.

Reference

1. Alvizo, O., Nguyen, L. J., Savile, C. K., Bresson, J. A., Lakhapatri, S. L., Solis, E. O.,... & Novick, S. J. (2014). Directed evolution of an ultrastable carbonic anhydrase for highly efficient carbon capture from flue gas. *Proceedings of the National Academy of Sciences*, 111(46), 16436-16441. doi: 10.1073/pnas.1411461111

2. Atkinson, J. T., Campbell, I. J., Thomas, E. E., Bonitatibus, S. C., Elliott, S. J., Bennett, G. N., & Silberg, J. J. (2019). Metalloprotein switches that display chemical-dependent electron transfer in cells. *Nature Chemical Biology*, *15*(2), 189. doi: 10.1038/s41589-018-0192-3

3. Brockman, I. M., & Prather, K. L. (2015). Dynamic knockdown of *E. coli* central metabolism for redirecting fluxes of primary metabolites. *Metabolic Engineering*, *28*, 104-113. doi:10.1016/j.ymben.2014.12.005

4. Cambray, G., Guimaraes, J. C., & Arkin, A. P. (2018). Evaluation of 244,000 synthetic sequences reveals design principles to optimize translation in *Escherichia coli*. *Nature Biotechnology*, *36*(10), 1005. doi: 10.1038/nbt.4238

5. Cao, Y., Li, X., Li, F., & Song, H. (2017). CRISPRi–sRNA: transcriptional–translational regulation of extracellular electron transfer in *Shewanella oneidensis*. *ACS Synthetic Biology*, *6*(9), 1679-1690. doi: 10.1021/acssynbio.6b00374

6. Charles, H. P., & Roberts, G. A. (1968). Carbon dioxide as a growth factor for mutants of *Escherichia coli*. *Microbiology*, *51*(2), 211-224. doi: 10.1099/00221287-51-2-211

7. Cheong, D. E., Ko, K. C., Han, Y., Jeon, H. G., Sung, B. H., Kim, G. J.,... & Song, J. J. (2015). Enhancing functional expression of heterologous proteins through random substitution of genetic codes in the 5' coding region. *Biotechnology and Bioengineering*, *112*(4), 822-826. doi: 10.1002/bit.25478

8. Dürre, P., & Eikmanns, B. J. (2015). C1-carbon sources for chemical and fuel production by microbial gas fermentation. *Current Opinion in Biotechnology*, *35*, 63-72. doi: 10.1016/j.copbio.2015.03.008

9. Elhadi, D., Lv, L., Jiang, X. R., Wu, H., & Chen, G. Q. (2016). CRISPRi engineering E. coli for morphology diversification. *Metabolic Engineering*, *38*, 358-369. doi: 10.1016/j.ymben.2016.09.001

10. Gao, C., Wang, S., Hu, G., Guo, L., Chen, X., Xu, P., & Liu, L. (2018). Engineering *Escherichia coli* for malate production by integrating modular pathway characterization with CRISPRi-guided multiplexed metabolic tuning. *Biotechnology and Bioengineering*, 115(3), 661-672. doi: 10.1002/bit.26486

11. Goodall, E. C., Robinson, A., Johnston, I. G., Jabbari, S., Turner, K. A., Cunningham, A. F.,... & Henderson, I. R. (2018). The essential genome of *Escherichia coli* K-12. *MBio*, *9*(1), e02096-17. doi: 10.1128/mBio.02096-17

12. Grote, A., Hiller, K., Scheer, M., Münch, R., Nörtemann, B., Hempel, D. C., & Jahn, D. (2005). JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Research*, *33*(suppl_2), W526-W531. doi: 10.1093/nar/gki376

13. Hashimoto, M., & Kato, J. I. (2003). Indispensability of the *Escherichia coli* carbonic anhydrases YadF and CynT in cell proliferation at a low CO_2 partial pressure. *Bioscience, Biotechnology, and Biochemistry*, *67*(4), 919-922. doi: 10.1271/bbb.67.919

14. Hossain, S. G., Shin, H. D., Li, J., Wang, M., Du, G., Liu, L., & Chen J. (2016) Integrating error-prone PCR and DNA shuffling as an effective molecular evolution strategy for the production of α-ketoglutaric acid by L-amino acid deaminase. *RSC Advances*, 6(52), 46149-46158. doi: 10.1039/C6RA02940J

15. Hsu, K. P., Tan, S. I., Chiu, C. Y., Chang, Y. K., & Ng, I. S. (2019). ARduino-pH Tracker and screening platform for characterization of recombinant carbonic anhydrase in *Escherichia coli*. *Biotechnology Progress*, 35(5), e2834. doi: 10.1002/btpr.2834

16. Jeske, L., Placzek, S., Schomburg, I., Chang, A., & Schomburg, D. (2018). BRENDA in 2019: a European ELIXIR core data resource. *Nucleic Acids Research*, *47*(D1), D542-D549. doi: 10.1093/nar/gky1048

17. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2019). A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. *Science*, 337(6096), 816-821. doi: 10.1126/science.1225829

18. Jo, B. H., Moon, H., & Cha, H. J. (2019). Engineering the genetic components of a whole-cell catalyst for improved enzymatic CO_2 capture and utilization. *Biotechnology and Bioengineering*. doi: 10.1002/bit.27175

19. Kang, C. W., Lim, H. G., Yang, J., Noh, M. H., Seo, S. W., & Jung, G. Y. (2018). Synthetic auxotrophs for stable and tunable maintenance of plasmid copy number. *Metabolic Engineering*, *48*, 121-128. doi: 10.1016/j.ymben.2018.05.020

20. Lee, J. W., Chan, C. T., Slomovic, S., & Collins, J. J. (2018). Next-generation biocontainment systems for engineered organisms. *Nature Chemical Biology*, *14*(6), 530-537. doi: 10.1038/s41589-018-0056-x

21. Lee, J. Y., Sung, B. H., Oh, S. H., Kwon, K. K., Lee, H., Kim, H.,... & Lee, S. G. (2019). C1 compound biosensors: design, functional study, and applications. *International Journal of Molecular Sciences*, 20(9), 2253. doi: 10.3390/ijms20092253

22. Li, Y., Li, S., Wang, J., & Liu, G. (2019). CRISPR/Cas systems towards next-generation biosensing. *Trends in Biotechnology*, *37*(7), 730-743. doi: 10.1016/j.tibtech.2018.12.005

23. Lindner, S. N., Ramirez, L. C., Krüsemann, J. L., Yishai, O., Belkhelfa, S., He, H.,... & Bar-Even, A. (2018). NADPH-auxotrophic *E. coli:* a sensor strain for testing *in vivo* regeneration of NADPH. *ACS Synthetic Biology*, *7*(12), 2742-2749. doi: 10.1021/acssynbio.8b00313

24. Lou, J. W., Zhu, L., Wu, M. B., Yang, L. R., Lin, J. P., & Cen, P. L. (2014). High-level soluble expression of the *hem*A gene from *Rhodobacter capsulatus* and comparative study of its enzymatic properties. *Journal of Zhejiang University Science B*, 15(5), 491-499. doi: 10.1631/jzus.B1300283

25. Merlin, C., Masters, M., McAteer, S., & Coulson, A. (2003). Why is carbonic anhydrase essential to *Escherichia coli*? *Journal of Bacteriology, 185*(21), 6415-6424. doi: 10.1128/JB.185.21.6415-6424.2003

26. Min, B. E., Seo, S. W., & Jung, G. Y. (2012). Switching control of an essential gene for reprogramming of cellular phenotypes in *Escherichia coli*. *Biotechnology and Bioengineering*, *109*(7), 1875-1880. doi: 10.1002/bit.24468

27. Novick, R. P. (1987). Plasmid incompatibility. *Microbiological Reviews*, *51*(4), 381. doi: 10.1016/0147-619X(78)90001-X

28. Pickar-Oliver, A., & Gersbach, C. A. (2019). The next generation of CRISPR–Cas technologies and applications. *Nature Reviews Molecular Cell Biology*, *20*(8), 490-507. doi: 10.1038/s41580-019-0131-5

29. Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., & Lim, W. A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell, 152*(5), 1173-1183. doi: 10.1016/j.cell.2013.02.022

30. Salis, H. M., Mirsky, E. A., & Voigt, C. A. (2009). Automated design of synthetic ribosome binding sites to control protein expression. *Nature Biotechnology*, *27*(10), 946. doi: 10.1038/nbt.1568

31. Schuller, A., Cserjan-Puschmann, M., Tauer, C., Jarmer, J., Wagenknecht, M., Reinisch, D.,... & Striedner, G. (2020). *Escherichia coli* σ 70 promoters allow expression rate control at the cellular level in genome-integrated expression systems. *Microbial Cell Factories*, *19*(1), 1-11. doi: 10.1186/s12934-014-0186-0

32. Seo, S. W., Yang, J. S., Cho, H. S., Yang, J., Kim, S. C., Park, J. M.,... & Jung, G. Y. (2014). Predictive combinatorial design of mRNA translation initiation regions for systematic optimization of gene expression levels. *Scientific Reports*, *4*, 4515. doi: 10.1038/srep04515

33. Tan, S. I., Han, Y. L., Yu, Y. J., Chiu, C. Y., Chang, Y. K., Ouyang, S.,... & Ng, I. S. (2018). Efficient carbon dioxide sequestration by using recombinant carbonic anhydrase. *Process Biochemistry*, *73*, 38-46. doi: 10.1016/j.procbio.2018.08.017

34. Tan, S. I., Ng, I. S., & Yu, Y. J. (2017). Heterologous expression of an acidophilic multicopper oxidase in *Escherichia coli* and its applications in biorecovery of gold. *Bioresources and Bioprocessing*, *4*(1), 20. doi: 10.1186/s40643-017-0150-z

35. Tan, S. I., You, S. C., Shih, I. T., & Ng, I. S. (2020). Quantification, regulation and production of 5-aminolevulinic acid by green fluorescent protein in

recombinant *Escherichia coli*. *Journal of Bioscience and Bioengineering*, *129*(4), 387-394. doi: 10.1016/j.jbiosc.2019.10.005

36. Tan, S. I., & Ng, I. S. (2020). New insight into plasmid-driven T7 RNA polymerase (PDT7) in *Escherichia coli* and used as genetic amplifier for biosensor. *ACS Synthetic Biology*. *20*(3), 613-622 doi.org/10.1021/acssynbio.9b00466

37. Wang, B., Zhang, X., Yu, X., Cui, Z., Wang, Z., Chen, T., & Zhao, X. (2019). Evolutionary engineering of *Escherichia coli* for improved anaerobic growth in minimal medium accelerated lactate production. *Applied Microbiology and Biotechnology*, *103*(5), 2155-2170. doi: 10.1007/s00253-018-09588-9

38. Wang, H., Nakamura, M., Abbott, T. R., Zhao, D., Luo, K., Yu, C.,..., & Qi, L. S. (2019). CRISPR-mediated live imaging of genome editing and transcription. *Science, 365*(6459):1301-1305. doi: 10.1126/science.aax7852

39. Wang, T., Badran, A. H., Huang, T. P., & Liu, D. R. (2018). Continuous directed evolution of proteins with improved soluble expression. *Nature Chemical Biology*, *14*(10), 972. doi: 10.1038/s41589-018-0121-5

40. Wang, T., Guan, C., Guo, J., Liu, B., Wu, Y., Xie, Z.,... & Xing, X. H. (2018). Pooled CRISPR interference screening enables genome-scale functional genomics study in bacteria with superior performance. *Nature Communications*, *9*(1), 1-15. doi: 10.1038/s41467-018-04899-x

41. Webster, G. R., Teh, A. Y. H., & Ma, J. K. C. (2017). Synthetic gene design—The rationale for codon optimization and implications for molecular

pharming in plants. *Biotechnology and Bioengineering*, *114*(3), 492-502. doi: 10.1002/bit.26183

42. Yu, T. H., Yi, Y. C., Shih, I. T., & Ng, I. S. (2019). Enhanced 5-aminolevulinic acid production by co-expression of codon-optimized *hem*A gene with chaperone in genetic engineered *Escherichia coli*. *Applied Biochemistry and Biotechnology*, 1-14. doi: 10.1007/s12010-019-03178-9

43. Zheng, Y., Meng, F., Zhu, Z., Wei, W., Sun, Z., Chen, J.,... & Chen, G. Q. (2019). A tight cold-inducible switch built by coupling thermosensitive transcriptional and proteolytic regulatory parts. *Nucleic Acids Research*, *47*(21), e137-e137. doi: 10.1093/nar/gky830

Figures

Figure 1 Illustration of the recombinant protein quality. The quality of recombinant proteins includes three factors: intrinsic quality (left), quantity (middle), and extrinsic quality (right). Intrinsic quality represents the different protein amino acid sequences with different catalytic constants in the Michaelis–Menten model. Quantity and extrinsic quality describe recombinant protein expression. The DNA or amino acid sequence affects the protein expression level (from white triangle to white rectangle), and inclusion body (closed symbols) to soluble form (white symbols). Total enzymatic performance was calculated as the product of three factors.

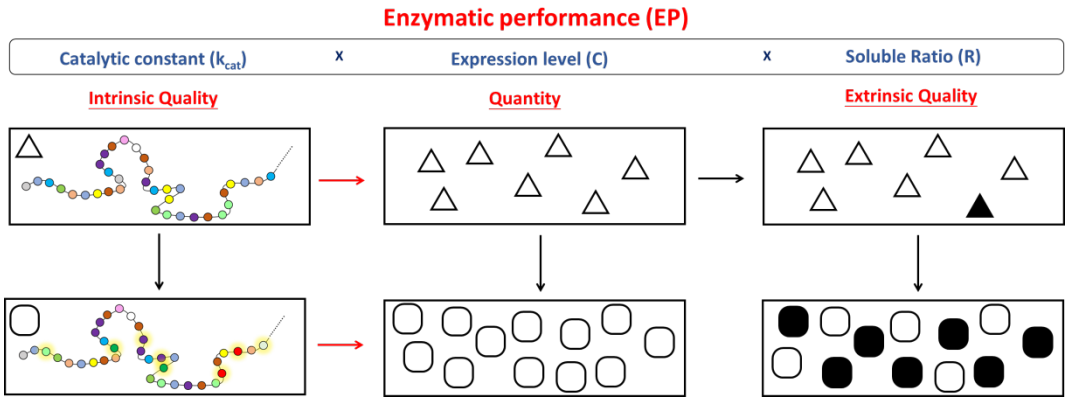

Figure 2. Establishment of the CRISPRi system targeting the essential gene *can* in *E. coli*. (a) Design of synthetic promoter, (b) map of pSUI-SynP-sfGFP, (c) dose-response curve of DH5α harboring pSUI-SynP-sfGFP. The red line indicates the fitting curve based on Hill's equation, and the black dots indicate the data points. (d) Plasmid design by CRISPRi targeting the *can* essential gene (CRISPRi::CA) included one plasmid driving a strongly expressed sgRNA and another plasmid with weak constitutively expressed dCas9. (e) Scheme of the selected targeting site on the *can* essential gene. (f) Growth curve of strains with different sgRNA. The orange, black, green, pink, and blue colors represent the cell growth curve of DH5α WT, S1, S2, SNT, and SP strains, respectively.

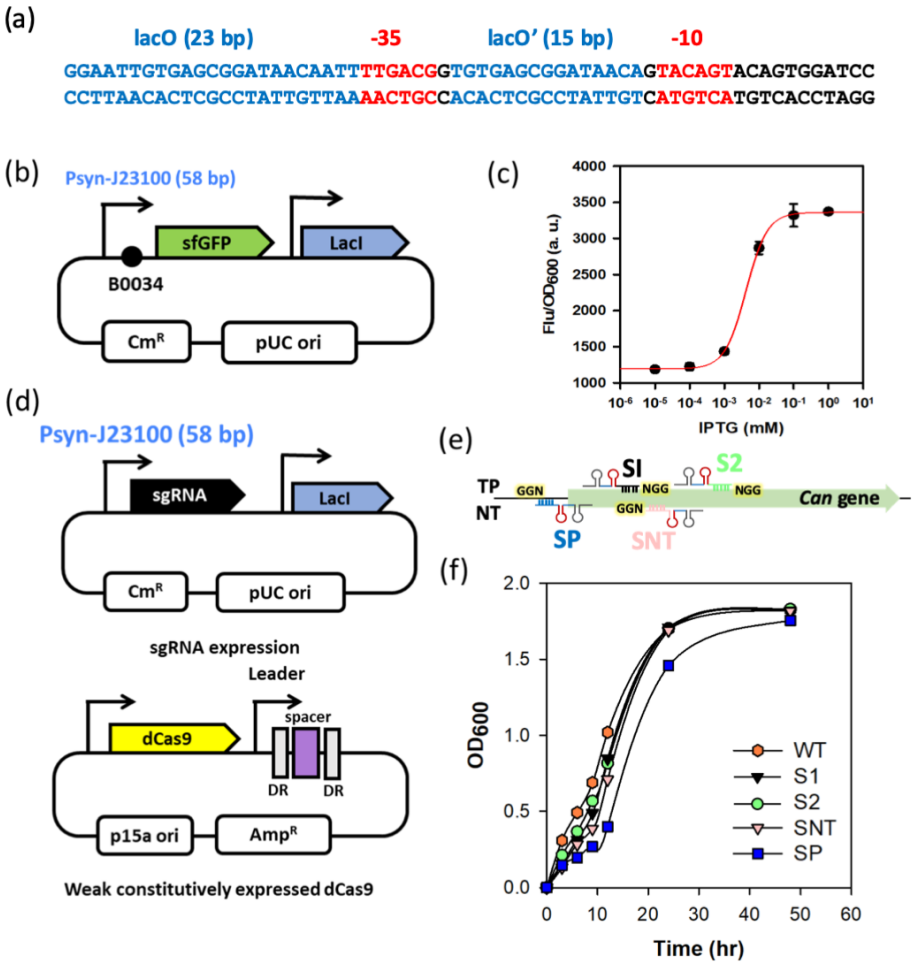

Figure 3 CRISPRi::CA system used as a CO_2 biosensor. (a) Growth curve of SP strain under bubbling of CO_2 concentration from 1% to 10%. The white circle, black circle, white triangle, and black triangle represent the growth curve of SP under air bubbling, 2%, 5%, and 10% CO_2, respectively. (b) Dose-response curve of the CO_2 biosensor. The red dots represent the calculated specific growth rate, and the black line represent the fitted curve based on the Hill's equation, except for the black dot.

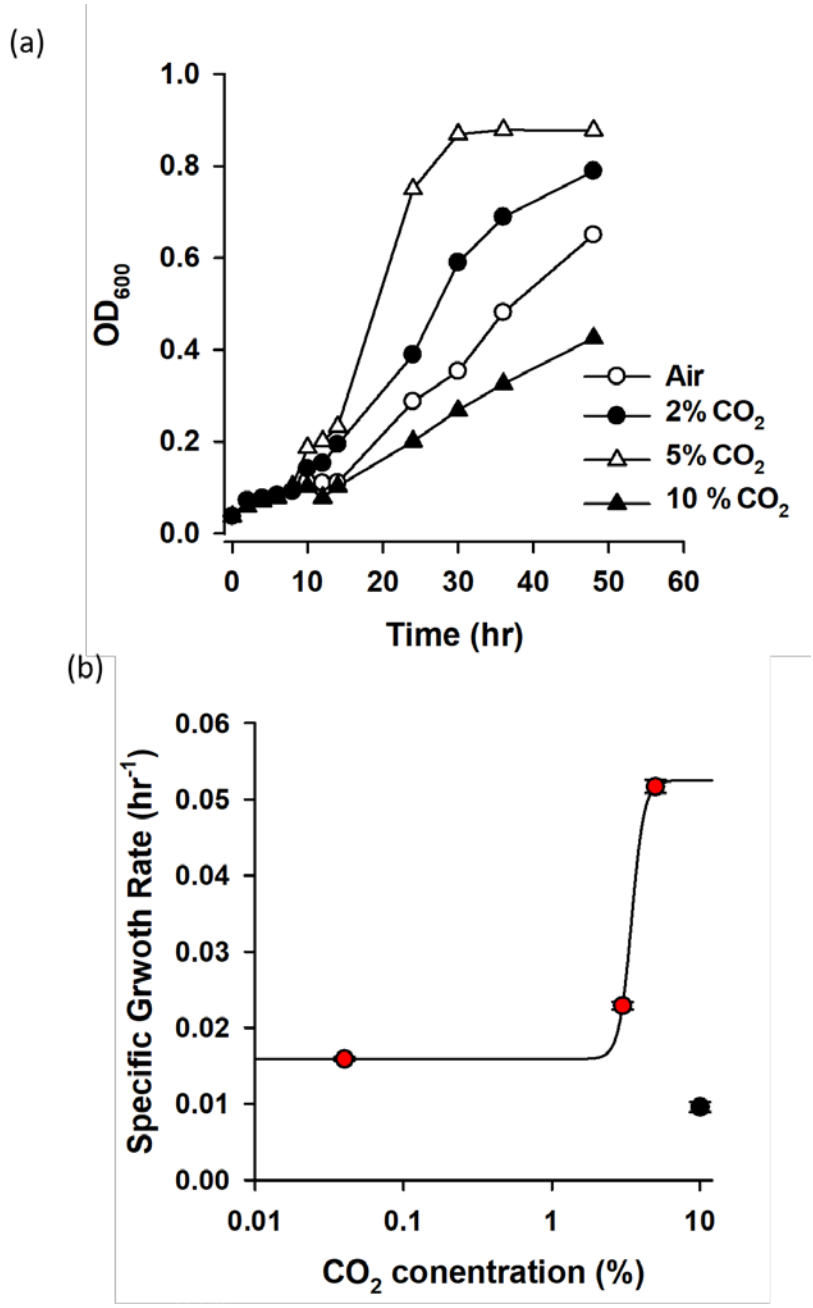

Figure 4 The DEPEND system applied for the determination of high-performance recombinant carbonic anhydrase. (a) SDS-PAGE analysis of SyCA and hCAII. (b) WAU results of SyCA and hCAII. (c) CFU analysis of the DEPEND system. (d) Growth curve of I-SyCA and I-hCAII with or without IPTG induction. The circle and rectangle represent the I-SyCA and hCAII strains, respectively. The closed and opened symbol represent the conditions with and without IPTG induction.

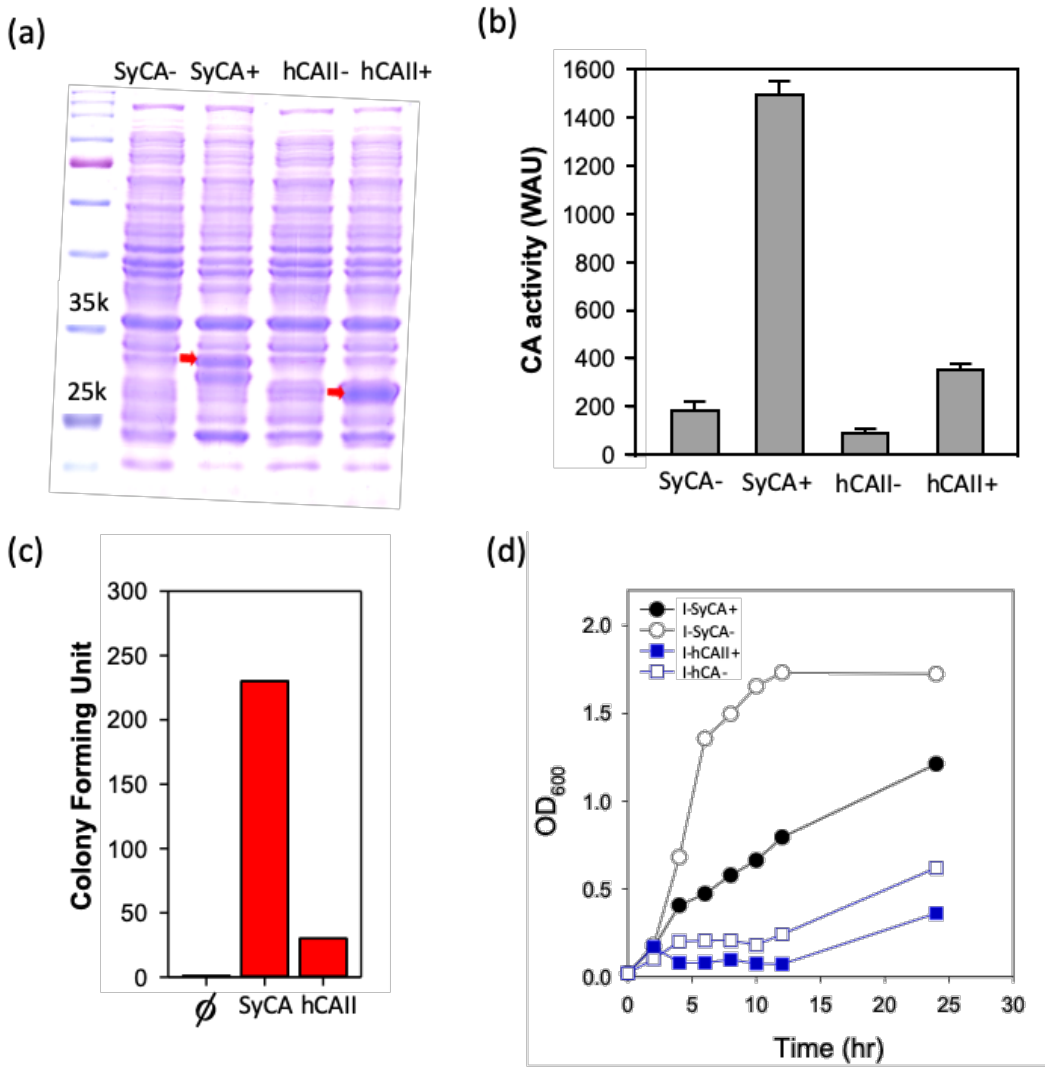

Figure 5 The DEPEND system evaluating the performance of recombinant ALAS. (a) Growth curve of different conditions in cell recovery. The circle, triangle, and rectangle represent the recovery curve with RsHemA (Rs), RcHemA (Rc), and RcHemA coupled with EcGroELS (RcG). The colored and white symbols represent the condition with (+) and without (-) IPTG induction, respectively. (b) Specific growth rate and maximum OD_{600} difference. The pink bars and black dots indicate the specific growth rate and maximum OD_{600} difference of individual conditions.

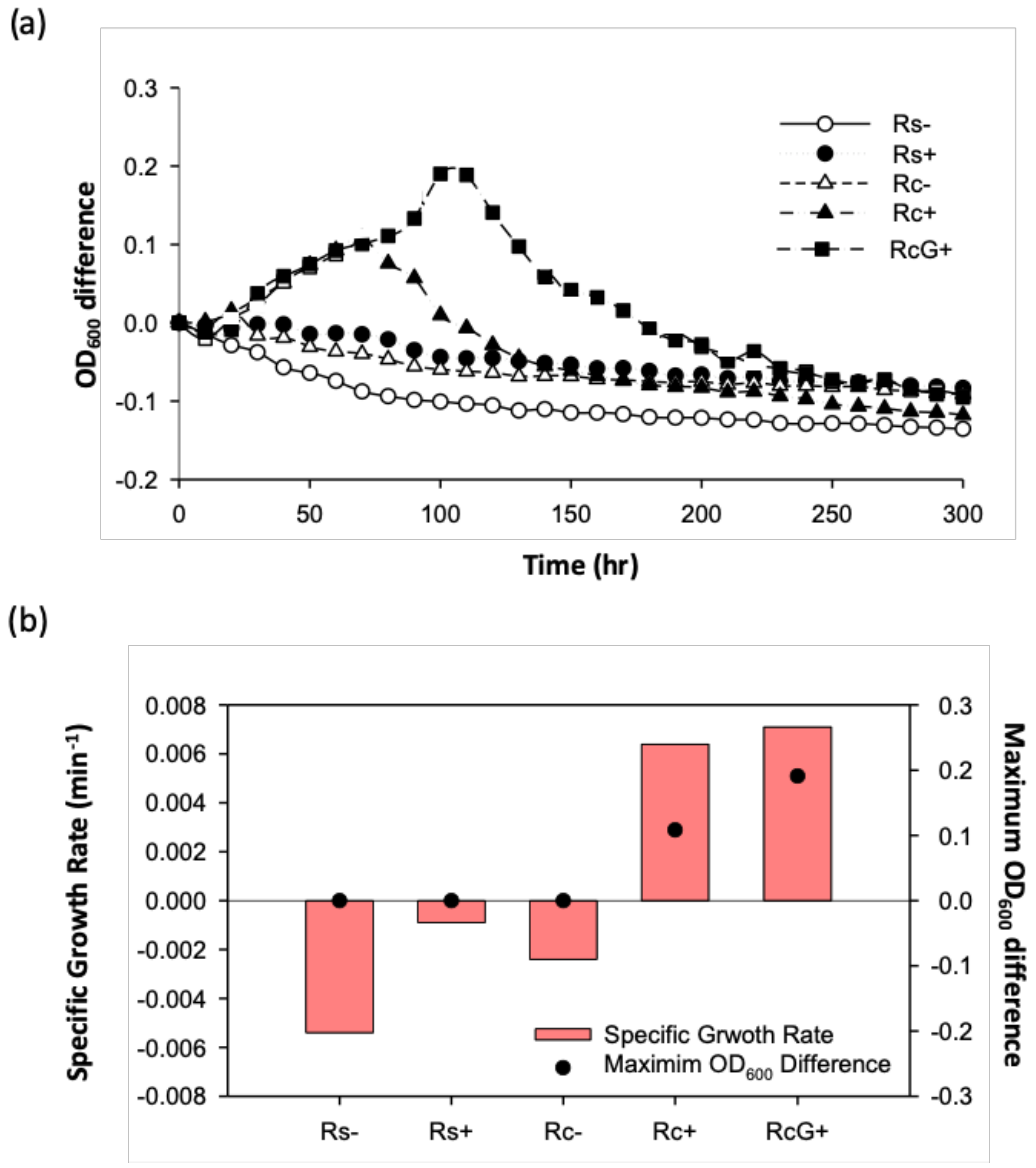

Table 1 Strains and plasmids used in this study

Name	Description	Remark
Strains		
E. coli DH5α	F-, Δ(argF-lac)169, φ80dlacZ58(M15), ΔphoA8, glnX44(AS), λ-, deoR481, rfbC1, gyrA96(NalR), recA1, endA1, thiE1, hsdR17	Molecular cloning and characterization
E. coli BL21(DE3)	F^- *ompT gal dcm lon hsdS*$_B$($r_B^-m_B^-$) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K\text{-}12}$(λ^S)	T7-based expression
S1	DH5α harboring pSM-dCas9::Ap and pSUI-SynP-CAS1sg	This study
S2	DH5α harboring pSM-dCas9::Ap and pSUI-SynP-CAS2sg	This study
SNT	DH5α harboring pSM-dCas9::Ap and pSUI-SynP-CANTsg	This study
SP	DH5α harboring pSM-dCas9::Ap and pSUI-SynP-CAPsg	This study
I-SyCA	BL21(DE3) harboring pSUI-SynP-CAPsg-PI-anti-dCas9sg-PDT7, pSM-dCas9::Ap and pSIT-SyCA	This study

I-hCAII	BL21(DE3) harboring pSUI-SynP-CAPsg-PI-anti-dCas9sg-PDT7, pSM-dCas9::Ap and pSIT-hCAII	This study
Plasmids		
pSU-PlacI-sfGFP	3099 bp, Cm^R, pUC ori, PlacI promoter, B0034 RBS, sfGFP	Tan and Ng 2020
pSUI-PlacI-sfGFP	2922 bp, Cm^R, pUC ori, PlacI promoter, B0034 RBS, sfGFP	This study
pSUI-SynP-sfGFP	2935 bp, Cm^R, pUC ori, PI promoter, B0034 RBS, sfGFP	This study
pSM-dCas9::Ap	9406 bp, Ap^R, p15a ori, dCas9, mobilization cluster	This study
pSUI-SynP-CAS1sg	3578 bp, Cm^R, pUC ori, Synthetic promoter (SynP), CAS1-sgRNA (CAS1sg), LacI	This study
pSUI-SynP-CAS2sg	3578 bp, Cm^R, pUC ori, Synthetic promoter (SynP), CAS2-sgRNA (CAS2sg), LacI	This study
pSUI-SynP-CANTsg	3578 bp, Cm^R, pUC ori, Synthetic promoter (SynP), CANT-sgRNA (CANTsg), LacI	This study
pSUI-SynP-CAPsg	3578 bp, Cm^R, pUC ori, Synthetic promoter (SynP), CAP-sgRNA	This study

	(CAPsg), LacI	
pSIT	3802 bp, Km^R, RSF ori, T7 promoter, T7 terminator, LacI	Tan and Ng 2020
pSIT-hCAII	4531 bp, Km^R, RSF ori, T7 promoter, hCAII, LacI	This study
pSIT-SyCA	4293 bp, Km^R, RSF ori, T7 promoter, SyCA, LacI	This study
pSIT-RsHemA	5005 bp, Km^R, RSF ori, T7 promoter, RsHemA, LacI	This study
pSIT-RcHemA	5005 bp, Km^R, RSF ori, T7 promoter, RcHemA, LacI	This study
pSIT-RcHemA-EcGroELS	6941 bp, Km^R, RSF ori, T7 promoter, RcHemA, EcGroELS, LacI	This study

Table 2 Fitting parameters of the IPTG dose-response curve by Hill's equation for DH5α harboring pSUI-PlacI-sfGFP and pSUI-Psyn-sfGFP, respectively.

Plasmid	y_o (a.u.)	a (a.u.)	b	c (mM)	Fold change
pSUI-PlacI-sfGFP	3663	33960	1.4	0.009	3.82
pSUI-Psyn-sfGFP	1192	2163	1.4	0.004	10.98

* Hill's equation was followed the formula: $y_0 = \frac{ax^b}{c^b + x^b}$ where the *y*, y_0, *x*, *a*, *b* and *c* denotes the specific fluorescence at any level of *x*, the leakage expression of specific fluorescence, IPTG concentration, dynamic range, Hill's coefficient and the IPTG concentration that achieve half the dynamic range. The fold change was calculated as $\frac{a}{y_0} + 1$.

Table 3 Average growth rate at exponential phase and final biomass of the CRISPRi::CA mediated strains and wild type (WT).

Strains	Average growth rate at 9 h (OD/h)	Inhibition efficiency (%)*	Biomass at 48 h (OD)
WT	0.080	0	1.812
S1	0.053	33.8	1.813
S2	0.065	18.8	1.810
SNT	0.047	41.3	1.813
SP	0.028	65.0	1.745

*Inhibition efficacy was calculated as

$$\left(1 - \frac{average\ growth\ rate\ at\ 9\ h\ of\ each\ strain}{average\ growth\ rate\ at\ 9h\ of\ WT}\right) \times 100\%$$