Applied Microbiology and Biotechnology (2024) 108:493
https://doi.org/10.1007/s00253-024-13333-w

APPLIED GENETICS AND MOLECULAR BIOTECHNOLOGY

Development of a xylose-inducible and glucose-insensitive expression system for *Parageobacillus thermoglucosidasius*

Junyang Wang[1,2,3] · Weishan Wang[1,2] · Yihua Chen[1,2] · Zihe Liu[3] · Xu Ji[3] · Guohui Pan[1,2] · Zilong Li[1] · Keqiang Fan[1]

Received: 25 January 2024 / Revised: 6 September 2024 / Accepted: 15 October 2024 / Published online: 23 October 2024

Abstract

Inducible expression systems are pivotal for governing gene expression in strain engineering and synthetic biotechnological applications. Therefore, a critical need persists for the development of versatile and efficient inducible expression mechanisms. In this study, the xylose-responsive promoter *xylA5*p and its transcriptional regulator XylR were identified in *Parageobacillus thermoglucosidasius* DSM 2542. By combining promoter *xylA5*p with its regulator XylR, fine-tuning the expression strength of XylR, and reducing the glucose catabolite repression on xylose uptake, we successfully devised a xylose-inducible and glucose-insensitive expression system, denoted as IExyl*. This system exhibited diverse promoter strengths upon induction with xylose at varying concentrations and remained unhindered in the presence of glucose. Moreover, we showed the applicability of IExyl* in *P. thermoglucosidasius* by redirecting metabolic flux towards riboflavin biosynthesis, culminating in a 2.8-fold increase in riboflavin production compared to that of the starting strain. This glucose-insensitive and xylose-responsive expression system provides valuable tools for designing optimized biosynthetic pathways for high-value products and facilitates future synthetic biology investigations in *Parageobacillus*.

Key points

- *A xylose-inducible and glucose-insensitive expression system IExyl* was developed.*
- *IExyl* was applied to enhance the riboflavin production in P. thermoglucosidasius*
- *A tool for metabolic engineering and synthetic biology research in Parageobacillus strains.*

Keywords Inducible promoter · Xylose · Xylose transporter · Carbon catabolite repression · *Parageobacillus thermoglucosidasius* · Riboflavin

Introduction

Research in synthetic biology relies heavily on well-characterized elements and tools for introducing new functionalities into organisms (Canton et al. 2008). However, the availability and ease of use of these elements often constrain the selection of microbial chassis. As a result, significant research efforts have been dedicated to developing synthetic biological elements tailored for diverse microorganisms. While the majority of research has centered around model organisms such as *Escherichia coli* and *Saccharomyces cerevisiae*, recent initiatives have been extended to include non-model microorganisms, especially mesophiles like *Pseudomonas* spp., *Synechococcus* spp., *Bacteroides thetaiotaomicron*, *Pichia pastoris*, and *Rhodosporidium toruloides* (Guo and Shi 2023; Markley et al. 2015; Mimee

✉ Zilong Li
lizl@im.ac.cn

✉ Keqiang Fan
fankq@im.ac.cn

1 State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China

2 University of Chinese Academy of Sciences, Beijing 100049, China

3 College of Life Science and Technology, Beijing Advanced Innovation Center for Soft Matter Science and Engineering, Beijing University of Chemical Technology, Beijing 100029, China

et al. 2015; Nikel et al. 2014; Schultz et al. 2019), as well as thermophiles, including strains from the *Parageobacillus* genus (Cripps et al. 2009). These investigations seek to expand the landscape of synthetic biology by diversifying the array of organisms that can be engineered and manipulated for various applications.

Parageobacillus spp. are facultatively anaerobic bacterium classified among thermophiles. These microorganisms exhibit robust growth across a wide temperature range, particularly thriving between 50 and 70 °C, with optimal growth observed at temperatures between 60 and 65 °C (Cripps et al. 2009; Liang et al. 2022a, b). *Parageobacillus thermoglucosidasius* DSM 2542 (previously named as *Geobacillus thermoglucosidasius* DSM 2542) is a model strain of *Parageobacillus* genus, which can efficiently utilize oligosaccharides, cellobiose, pentose and hexose sugars from various feedstocks including lignocellulosic biomass (Cripps et al. 2009; Liang et al. 2022a, b). Due to these distinctive traits of *P. thermoglucosidasius*, it is used as an appealing chassis for industrial biotechnology applications such as the bioproduction of ethanol, lactate, formate, succinate, and riboflavin (Cripps et al. 2009; Liang et al. 2022a, b; Lin et al. 2014; Wang et al. 2022a; Yang et al. 2021). Furthermore, the risk of contamination during large-scale cultivation is minimized in fermentation of thermophiles like *P. thermoglucosidasius* at higher temperatures, as common contaminants are mesophiles unable to thrive under these conditions. Additionally, high-temperature fermentations can reduce the need for additional cooling or heating steps in post-feedstock pretreatment or product recovery (Liang et al. 2022b; Taylor et al. 2009).

Despite numerous studies developing synthetic biology elements for *P. thermoglucosidasius*, such as origins of replication, selectable markers, promoter libraries, ribosome binding site libraries, fluorescent reporter proteins, plasmid vectors, genome editing tools, and riboswitch-mediated gene regulatory control tools, a notable gap remains in the availability of inducible promoters (Frenzel et al. 2018; Lau et al. 2021; Madika et al. 2022; Marcano-Velazquez et al. 2019; Reeve et al. 2016). Inducible promoters are crucial for finely tuning the expression of target genes at specific times. Particularly, inducible promoters responsive to sugar inducers offer a cost-effective means of strictly regulating gene expression. These promoters have been extensively developed in various species, including xylose-regulation systems like *xyl*p (in *Bacillus subtilis*, *Clostridium perfringens*, *Brevibacillus choshinensis*) and *xylT*p (in *Lactococcus lactis*) (Bhavsar et al. 2001; D'Urzo et al. 2013; Guzman et al. 1995; Nariya et al. 2011).

The well-established xylose isomerase pathway for xylose metabolism is conserved across a broad spectrum of bacteria, including *Parageobacillus* spp. This pathway encompasses the uptake of xylose into cells, its isomerization to xylulose, and subsequent phosphorylation to form xylulose-5-phosphate (Kim and Woo 2018; You et al. 2022). In Firmicutes, xylose uptake is facilitated by two types of transporters: an ATP-binding cassette (ABC) xylose transporter encoded by *xylFGH* and a major facilitator superfamily (MFS) transporter encoded by *xylT* (Gu et al. 2010). The components of the xylose ABC transporter include *xylF*, *xylG*, and *xylH*, which encode the periplasmic xylose-binding protein, ATPase, and permease, respectively (Gu et al. 2010). The conversion of xylose to xylulose and then to xylulose-5-phosphate is catalyzed by the enzymes D-xylose isomerase (XylA) and xylulose kinase (XylB), respectively. Notably, these two encoding genes, *xylA* and *xylB*, are organized into an operon (Gu et al. 2010). The transcription of genes involved in xylose transport and metabolism is typically under negative regulation by an ROK (Repressor, Open reading frame, Kinase) family regulator known as XylR, which binds to its operator sites in the promoter regions of these genes in the absence of xylose (Gu et al. 2010; Rodionov et al. 2001; Scheler and Hillen 1994).

Promoters of these genes together with the regulator XylR were used in the construction of xylose-induced expression systems in various bacteria and fungi. For instance, *Brevibacillus choshinensis* harnessed the xylose-responsive promoter of the *xylA* gene to achieve robust intracellular expression of heterologous proteins (D'Urzo et al. 2013). In *Paenibacillus polymyxa* ATCC 842, a xylose-responsive operon was identified, demonstrating an impressive 186-fold increase in the transcription of relative operon, and a more than four-fold enhancement in the relative fluorescence intensity of enhanced green fluorescent protein (eGFP) upon xylose induction (Wang et al. 2022b). In *Streptomyces avermitilis* MA-4680, a system for protein expression triggered by xylose was devised, utilizing XylR and the promoter of *xylA*. This resulted in a notable 24-fold increase in chitobiase activity under xylose-induced conditions compared to non-induced circumstances (Noguchi et al. 2018). To refine transcriptional control in *Saccharomyces cerevisiae*, a hybrid strategy involving XylR from *Caulobacter crescentus* was deployed. This approach effectively inhibited the transcription activities of synthetic promoters derived from the constitutive promoter TEF from *Ashbya gossypii*. Notably, the resulting promoters exhibited repression when no xylose was added, and showcased up to a 25-fold induction with the addition of xylose (Hector and Mertens 2017).

In our previous study, we characterized a set of xylose-inducible promoters in *P. thermoglucosidasius* DSM 2542, and used them for constructing strains capable of dynamically controlling riboflavin production (Wang et al. 2022a). These promoters exhibited a wide range of strengths and demonstrated dose-dependent induction in response to xylose. While we initially proposed these promoters as promising starting points for the development of xylose-inducible

systems in *P. thermoglucosidasius*, it is worth noting that their inducible properties were significantly inhibited in the presence of glucose as a fermentation substrate, severely hampering their practical application (Wang et al. 2022a). Therefore, further enhancements are imperative to improve the performance of these xylose-inducible promoters, informed by a comprehensive understanding of xylose metabolism and its regulation in *P. thermoglucosidasius*.

In this study, we identified several genes involved in the xylose metabolism and regulation, including the ROK family regulator XylR (*AOT13_16590*) and a set of sugar ABC transporter (*AOT13_11435*, *AOT13_11440*, and *AOT13_11445*) in *P. thermoglucosidasius* DSM 2542. We introduced a highly adaptable xylose-inducible expression system, denoted as IExyl*, developed based on the *xylA* promoter and its associated regulator XylR. IExyl* exhibited minimal leaky expression in the absence of xylose and demonstrated a broad spectrum of promoter strengths upon induction with xylose at varying concentrations in *P. thermoglucosidasius*. A significant feature of this system is its operational effectiveness even in the presence of glucose, achieved through successful mitigation of carbon catabolite repression (CCR) on xylose metabolism. Moreover, we leveraged IExyl* to augment riboflavin production by optimizing the metabolic flux between cell growth and riboflavin biosynthesis in *P. thermoglucosidasius*. This xylose-inducible and glucose-insensitive expression system provides important tools for the systematic design of optimized biosynthetic pathways for valuable products and streamlines synthetic biology studies in *Parageobacillus*.

Materials and methods

Strains, media, and growth conditions

The bacterial strains employed in this study are detailed in Table S1. *E. coli* strains were cultivated in lysogeny broth (LB) medium. For genetic manipulation and fermentation of *P. thermoglucosidasius*, USYE media were utilized following previously established protocols (Wang et al. 2022a). The USYE medium was amended with D-glucose and/or D-xylose at varying concentrations, as specified. The mutant strains of *P. thermoglucosidasius* were cultivated in 250 mL shake flasks. The seed culture contained 50 mL of USYE medium supplemented with 40 mM each (final concentrations) of Bis-Tris, PIPES, and HEPES, adjusted to pH 7.0. Subsequently, 0.5 mL of the seed culture was inoculated into 50 mL of USYE medium (with glucose and/or xylose added as specified) and subjected to fermentation at 60 °C and 250 rpm under designated conditions, including the designated time, concentration of glucose and/or xylose. When required, antibiotics were added to the media at the following concentrations: 100 μg/mL of ampicillin for *E. coli* and 12.5 μg/mL kanamycin, 10 mg/L spectinomycin for *P. thermoglucosidasius*.

Construction of plasmids and engineered strains

All the engineered strains and plasmids used in this study are listed in Table S1 and S2 in detail, while the sequences of all primers and oligonucleotides are listed in Table S3. Genomic DNA of *P. thermoglucosidasius* DSM 2542 was extracted using the Ultraclean® Microbial DNA Isolation kit (Mo Bio Laboratories, Inc., Cambio Ltd., Cambridge, UK) following the manufacturer's protocol. Plasmids extraction were performed using the NucleoSpin Plasmid EasyPure kit (Macherey-Nagel). PCR screening for correct transformants was performed using Taq 2× Master Mix (TsingKe, China).

To obtain the smallest fragment of the xylose-inducible promoter, a set of truncated promoters of *xyl* operon was PCR-amplified from the genomic DNA of *P. thermoglucosidasius* DSM 2542 using primer pairs *xylA500*p-F/R, *xylA400*p-F/R, *xylA300*p-F/R, *xylA200*p-F/R, *xylA100*p-F/R, *xylA1*p-F/R, *xylA2*p-F/R, *xylA3*p-F/R, *xylA4*p-F/R, *xylA5*p-F/R, and *xylA6*p-F/R. Subsequently, these fragments were assembled through Gibson assembly with the corresponding plasmid backbone, which was PCR-amplified from pUCG18-*sfgfp* using primer pairs pUCG18-*sfgfp*-GJ-F/R. This procedure yielded 11 plasmids designated as pUCG18-n-*sfgfp* (where n represents the name of the 11 promoters, as outlined in Table S2).

To evaluate the promoter strength of the xylose-inducible promoters and available strong constitutive promoters, pLdh was PCR-amplified from the genomic DNA of *Geobacillus thermodenitrificans* NG80-2 using primer pairs pUCG18-pLdh-*sfgfp*-F/R. Additionally, pRplsWT was synthesized by TsingKe. The two fragments containing promoters pLdh and pRplsWT were assembled with the corresponding plasmid backbones, which were PCR-amplified from pUCG18-*sfgfp* using primer pairs pUCG18-pLdh-*sfgfp*-GJ-F/R and pUCG18-pRp-*sfgfp*-GJ-F/R, respectively. This process resulted in the generation of plasmids pUCG18-pLdh-*sfgfp* and pUCG18-pRplsWT-*sfgfp*.

To assess the fluorescence intensities of the reporter genes *sfgfp* and *mCherry* at different temperatures, plasmid pUB31-pLdh-*sfgfp* was constructed by Gibson assembly of five fragments. These fragments included Ldh-up, amplified from the genomic DNA of *P. thermoglucosidasius* using primer pairs Ldh-up-F/R; pUB31-pLdh, amplified from the genomic DNA of *G. thermodenitrificans* NG80-2 using primer pairs pUB31-pLdh-F/R; UB31-*sfgfp*, amplified from pUCG18-pLdh-*sfgfp* using primer pairs pUB31-*sfgfp*-F/R; Ldh-down, amplified from the genomic DNA of *P. thermoglucosidasius* using primer pairs Ldh-down-F/R; and pUB31-GJ, amplified from pUB31 using primer

pairs pUB31-GJ-F/R. Additionally, plasmid pUB31-pLdh-*mCherry* was constructed by Gibson assembly of two fragments: pUB31-*mCherry*, amplified from pUC19-*mCherry* using primer pairs pUB31-*mCherry*-F/R, and pUB31-*mCherry*-GJ, amplified from pUB31 using primer pairs pUB31-*mCherry*-GJ-F/R.

To enhance the expression of the mCherry protein, various combinations of promoters and ribosome binding site (RBS) were characterized. P43-*mCherry1*, containing RBS1, was assembled using three fragments: RBS1 amplified from puBS4.4-p43 with primer pairs P43-*mCherry1*-F/R, *mCherry1* amplified from pUB31-pLdh-*mCherry* with primer pairs *mCherry1*-F/R, and pUCG18-*mCherry1*-GJ amplified from pUCG18-*mCherry* with primer pairs pUCG18-*mCherry1*-GJ-F/R. These fragments were Gibson assembled to produce plasmid pUCG18-p43-*mCherry*-RBS1. For fragments pLdh-*mCherry1*, pRplsWT-*mCherry1*, and *xylR*p-*mCherry1*, amplification was carried out from pUCG18-pLdh-*sfgfp*, pUCG18-pRplsWT-*sfgfp*, and the genomic DNA of *P. thermoglucosidasius*, respectively. The primer pairs used were pLdh-*mCherry1*-F/R, pRplsWT-*mCherry1*-F/R, and *xylR*p-*mCherry1*-F/R, respectively. Additionally, psyn-*mCherry1* was synthesized by TsingKe. These fragments were then Gibson assembled with pUCG18-pL-mCh1-GJ, pUCG18-pR-mCh1-GJ, pUCG18-px-mCh1-GJ, and pUCG18-ps-mCh1-GJ (amplified from pUCG18-*mCherry* using primer pairs pUCG18-pL-mCh1-GJ-F/R, pUCG18-pR-mCh1-GJ-F/R, pUCG18-pX-mCh1-GJ-F/R, and pUCG18-pS-mCh1-GJ-F/R, respectively). This resulted in the plasmids pUCG18-pLdh-*mCherry*-RBS1, pUCG18-pRplsWT-*mCherry*-RBS1, pUCG18-*syn*p-*mCherry*-RBS1, and pUCG18-*xylR*p-*mCherry*-RBS1, respectively.

All assembly segments for plasmids pUCG18-n-*mCherry*-RBS2, pUCG18-n-*mCherry*-RBS3, and pUCG18-n-*mCherry*-RBS4 were amplified as described above, using their corresponding primers in Table S3, and employing pUCG18-p43-*mCherry*-RBS1, pUCG18-pLdh-*mCherry*-RBS1, pUCG18-pRplsWT-*mCherry*-RBS1, pUCG18-*syn*p-*mCherry*-RBS1, and pUCG18-*xylR*p-*mCherry*-RBS1 as templates, respectively.

To generate the *xylR* knockout strain DSM 2542 Δ*xylR*, plasmid pUB31-*xylR*-*mCherry*-*sfgfp* was constructed by assembling several segments, each amplified using the primers in Table S3. The templates for amplification included pUCG18-pLdh-*sfgfp*, pUCG18-pRplsWT-*mCherry*-RBS4, and the genomic DNA of *P. thermoglucosidasius*, respectively.

To optimize the expression levels of *xylR* gene in *P. thermoglucosidasius*, plasmids pUCG18-p43-*xylR*-*xylA5*p-*sfgfp*, pUCG18-pLdh-*xylR*-*xylA5*p-*sfgfp*, pUCG18-pRplsWT-*xylR*-*xylA5*p-*sfgfp*, and pUCG18-*xylR*p-*xylR*-*xylA5*p-*sfgfp* were constructed. These plasmids contained the *xylR* gene controlled by different promoters and were assembled using the corresponding primer pairs in Table S3. The templates for amplification were pUCG18-p43-*mCherry*-RBS1, pUCG18-pLdh-*mCherry*-RBS1, pUCG18-pRplsWT-*mCherry*-RBS1, pUCG18-*xylR*p-*mCherry*-RBS1, and pUCG18-*xylA5*p-*sfgfp*, respectively.

To assess the contribution of *AOT13_11435*, *AOT13_11440*, and *AOT13_11445* genes to the uptake of xylose, the knockout plasmid pUB31-transporter-*mCherry*-*sfgfp* was constructed. This was achieved by assembly of the vector pUB31 backbone with three segments, including pLdh-*sfgfp*, amplified from pUCG18-pLdh-*sfgfp*, pRplsWT-*mCherry*, amplified from pUCG18-pRplsWT-*mCherry*-RBS4, and the transporter knock-out cassette, amplified from the genomic DNA of *P. thermoglucosidasius* using the primers in Table S3.

The plasmid IExyl3 was constructed by assembly of three segments including *xylR*p-*xylR*, *xylA5*p-*sfgfp*, and pLdh-transporter, which were amplified using the corresponding primer pairs detailed in Table S3, and pUCG18-*xylR*p-*xylR*-*xylA5*p-*sfgfp*, pUCG18-*xylA5*p-*sfgfp*, and pUB31-transporter-*mCherry*-*sfgfp* as templates, respectively.

To introduce mutations into the *cre* sequence in *AOT13_11435*, the mutated sequences *creM1*, *creM2*, and *creM3* were synthesized and cloned into the vector IExyl3 using Gibson assembly. This resulted in the generation of plasmids IExyl3-*creM1*, IExyl3-*creM2*, and IExyl3-*creM3*, respectively.

To enhance the expression level of the *zwf* and *ribA* genes, fragments of these genes were PCR-amplified from the genomic DNA of *P. thermoglucosidasius* and cloned into the vector IExyl3-*creM3* (also referred to as IExyl*) using Gibson assembly. This process yielded two plasmids: IExyl*-*zwf* and IExyl*-*ribA*.

The resulting plasmids were then introduced into *P. thermoglucosidasius* DSM 2542 using the transformation protocol previously described by Cripps (Cripps et al. 2009), generating the corresponding engineered strains of *P. thermoglucosidasius* (see Table S1).

Double reporter gene-based markerless gene knockout method for *P. thermoglucosidasius*

To assess the performance of fluorescent protein markers superfolder Green Fluorescent Protein (sfGFP) and mCherry, the *sfgfp* and *mCherry* genes, driven by promoter pLdh, were integrated into the *P. thermoglucosidasius* genome at the *AOT13_05975* locus. Fluorescence of sfGFP and mCherry was evaluated at various temperatures on agar plates (48, 49, 50, 51, 52, and 53 °C for mCherry; 60, 62, 64, 66, and 68 °C for sfGFP). To enhance the expression of the mCherry protein, the combination of different promoters and RBS was systematically optimized. A highly convenient and efficient reporter gene-based method for DNA replacement,

enabling rapid chromosomal editing in *P. thermoglucosidasius* was developed.

Since the utilized vector lacked replication capability at temperatures above 65 °C, we employed a three-step selection process. First, colonies containing the complete vectors integrated into the genome through homologous recombination were selected based on sfGFP fluorescence and kanamycin resistance at 68 °C. Subsequently, a single sequential subculture at 50 °C without kanamycin allowed for the selection of potential double crossover clones using the mCherry fluorescence phenotype on agar plates. Subsequently, the identification of desired mutants was achieved through the observation of a loss in mCherry fluorescence phenotype on agar plates.

Green fluorescent protein (GFP) fluorescence assay

The protein sfGFP was employed as a reporter to evaluate the activities of promoters (Reeve et al. 2016). All strains of *P. thermoglucosidasius* carrying the GFP reporter system were cultured at 60 °C for 24 h in USYE medium with different concentrations of glucose or xylose. The GFP fluorescence intensities of cells were quantified using a Becton-Dickinson FACSCalibur flow cytometer, as previously described (Wang et al. 2022a).

Analysis of gene expression by reverse transcription - quantitative real-time PCR (RT-qPCR)

The engineered strains of *P. thermoglucosidasius* were cultured in liquid USYE medium, and harvested after fermentation for 24 h by fast filtration. Subsequently, the harvested cells were flash frozen in liquid nitrogen and ground into powder. Total RNA extraction was then carried out using TRNzol (Tiangen, China). The quantity and integrity of the isolated RNA were assessed as previously described (Wang et al. 2022a). The first-strand synthesis of cDNA was performed with 1 µg total RNA, using a PrimeScriptTM RT Reagent Kit with gDNA Eraser (TaKaRa, Japan) according to the manufacturer's instructions. All oligonucleotides used for RT-qPCR are listed in Table S3. The RT-qPCR experiments were performed using previously described process (Wang et al. 2022a). The gene expression levels were calculated based on the 2DDCt method (Livak and Schmittgen 2001), while the housekeeping gene *gap* was used as a reference (Song et al. 2016).

Determination of transcriptional start site (TSS)

To map the transcriptional start points of *xylA5*p, 5'-RACE (Rapid Amplification of cDNA Ends) were performed using the RLM-RACE kit (Ambion) according to the manufacturer's instructions. The total RNA was extracted from *P. thermoglucosidasius* as previously described. P1 and P2 was used as special outer and inner primers, respectively (Table S3).

Analytical methods

The quantitative determination of intracellular ribulose-5-phosphate, ribose-5-phosphate, and phosphoribosylpyrophosphate (PRPP) was performed using the previously described method (You et al. 2022). High performance liquid chromatograph was used to quantify the concentration of glucose or xylose. The chromatographic conditions were as follows: Chromatographic column: Xtiamte® Suger-H, (7.8 mm×300 mm, 5.0 µm, Weich Co., LTD.), column temperature chamber: 40 °C, mobile phase: 5 mM H_2SO_4, flow rate: 0.6 mL/min, sample size: 5 µL. The specific glucose uptake rate was calculated according to the glucose concentration (quantified by HPLC) at different fermentation time.

Statistical analysis

All statistical analysis and plots were performed with Origin (OriginPro 2021, OriginLab) or Prism7 (version 7.0, GraphPad Software Inc.). All quantitative data were presented as mean ± standard deviation.

Bioinformatic analysis

Protein homologs related to xylose utilization were searched in genomes of *P. thermoglucosidasius* DSM 2542 (CP012712.1) and NCIMB 11955 (CP016622.1) using standalone BLASTP program 2.13.0+ (Camacho et al. 2009). The protein sequences of reported xylose utilization related genes from *Bacillus subtilis* 168, *Clostridium berjerinckii* NCIMB 8052, and *Geobacillus kaustophilus* HTA426 were used as query sequences (Gu et al. 2010). Sequence comparison of protein homologs was performed using parasail_aligner with its nw_stats_striped_16 algorithm with the following parameters: gap open: 10, gap extend: 0.5, exact match length cutoff: 3 (Daily 2016). Multiple sequence alignment and construction of phylogenetic trees were preformed using MEGA 11 with its default parameters (Tamura et al. 2021). Phylogenetic trees were visualized using the iTOL web server (Letunic and Bork 2024).

Results

Identification of genes related to xylose utilization in *P. thermoglucosidasius* DSM 2542

To verify the putative genes related to xylose metabolism, homologs of reported proteins involved in xylose uptake and utilization were searched in the genome of *P.*

thermoglucosidasius DSM 2542 (CP012712.1). Besides previous reported genes *xylAB* (*AOT13_11570* - *AOT13_11575*, encoding xylose isomerase and xylulokinase, respectively) and *rbsACB* (*AOT13_01805* - *AOT13_01795*, encoding D-ribose ABC transporter) (Liang et al. 2022a; Wang et al. 2022a), homologic genes encoding an ROK family transcriptional regulator (*AOT13_16590*) and a set of sugar ABC transporter (*AOT13_11435*, *AOT13_11440*, and *AOT13_11445*) were identified in the genome of *P. thermoglucosidasius* strains DSM 2542 (Fig. 1a and Table S4).

The gene *AOT13_16590* encoded an ROK family transcriptional regulator, which was not co-localized with *xylAB* as observed in *Geobacillus kaustophilus* HTA426 (Gu et al. 2010). Its protein sequence showed a sequence identity of 68.1% to XylR of *G. kaustophilus* HTA426, and 37% to 70% to previously identified XylR orthologs from *Bacillus subtilis* 168, *Bacillus cereus* ATCC 10987, *Bacillus licheniformis* ATCC 14580, *Bacillus halodurans* C-125, *Bacillus clausii* KSM-K16, and *Geobacillus thermodenitrificans* NG80-2 (Fig. S1 and S2) (Gu et al. 2010). Phylogenetic analysis further confirmed that AOT13_16590 clustered within the clade of XylRs from *Geobacillus* (Fig. S2). Furthermore, our previous RNA-seq analysis revealed that the transcription level of *AOT13_16590* showed 9.7-fold increase when cultured with 1% xylose compared to those without xylose (Table S4) (Wang et al. 2022a). These findings suggested that AOT13_16590, designated as XylR, was involved in the regulation of xylose metabolism.

A set of sugar ABC transporter encoding genes, including *AOT13_11435*, *AOT13_11440*, and *AOT13_11445* (encoding ABC transporter substrate-binding protein, ATP-binding protein, and permease, respectively), were found to be directly upstream of the *ara* operon (*AOT13_11450* to *AOT13_11485*) (Liang et al. 2022a), which showed moderate sequence identities to XylF, XylG, and XylH of *G. kaustophilus* HTA426 (35.8%, 54.6%, and 43.1%, respectively, Fig. S3 and Table S4) (Gu et al. 2010). Furthermore, these genes showed 118- to 149-fold increase in their transcription levels when cultured with 1% xylose compared to those without xylose, which were about an order of magnitude higher than those of predicted ABC transporters for ribose and arabinose (Table S4) (Liang et al. 2022a; Wang et al. 2022a).

To verify the role of these sugar ABC transporter genes, we generated the mutant strain DSM 2542 ΔABCT (Fig. S4). We devised an efficient reporter-guided gene knockout method for *P. thermoglucosidasius*, which used sfGFP and mCherry as selecting markers to enable fast screening of desired crossover strains and then eliminated the need for additional screening methods, such as colony PCR. This gene knockout method used fluorescent proteins sfGFP (at temperatures of 60−70 °C) and mCherry (at temperatures lower than 50 °C) as selecting markers in *P. thermoglucosidasius* (Fig. S4A) with optimized expression of mCherry using various combinations of promoters and RBSs (Fig. S4B and S4C). In brief, this method included three steps, (1) selecting transformant with the gene knockout plasmid using green fluorescence of sfGFP; (2) selecting first double crossover strains using red fluorescence of mCherry and the loss of green fluorescence of sfGFP, in which the target DNA segment was replaced with the segment containing *mCherry*; (3) selecting second crossover strains using the loss of red fluorescence of mCherry, in which the inserting segment containing *mCherry* was removed to generate a markerless gene knockout strain (Fig. S4D and S4E). With this method, we have constructed the *xylR* gene knockout strain DSM 2542 Δ*xylR* and the sugar ABC transporter (*AOT13_11435*, *AOT13_11440*, and *AOT13_11445*) gene knockout strain DSM 2542 ΔABCT in a highly convenient and efficient manner (Fig. S4F and S4G).

Compared to the wild-type strain DSM 2542, strain DSM 2542 ΔABCT exhibited a significant impairment when

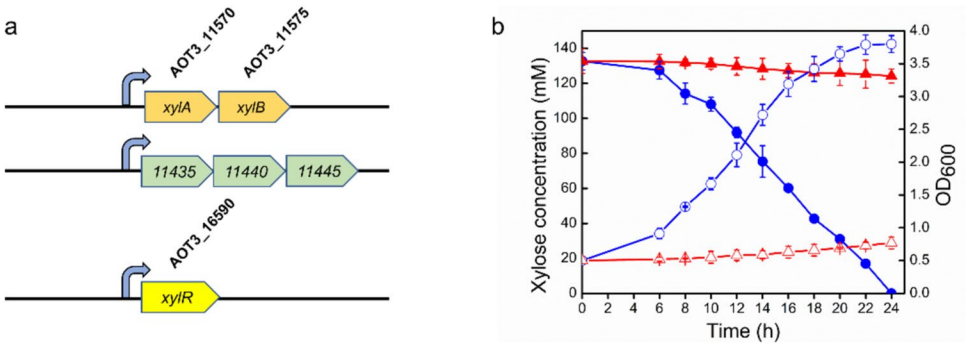

Fig. 1 Identification of xylose uptake and utilization genes in *P. thermoglucosidasius* DSM 2542. **a** Schematic representation of predicted xylose uptake and utilization genes in *P. thermoglucosidasius* DSM 2542. **b** Growth profile (blue open circle) and xylose consumption (blue closed circle) of DSM 2542 wild-type strain, and growth profile (red open triangle) and xylose consumption (red closed triangle) of DSM 2542 strain ΔABCT, using xylose as the sole carbon source

growth with xylose as the solo carbon source (Fig. 1b), while its growth on glucose remained unaffected (Fig. S5). These findings strongly supported that the sugar ABC transporter encoded by *AOT13_11435*, *AOT13_11440*, and *AOT13_11445* played a crucial role in xylose uptake in *P. thermoglucosidasius*. However, our results didn't completely excluded the involvement of other pentose uptake genes in the uptake of xylose, as revealed by previous studies (Liang et al. 2022a, b).

Identification and characterization the core sequence of the promoter of *xyl* operon

In our previous study, we have identified a set of xylose-inducible promoters in *P. thermoglucosidasius* DSM 2542 (Wang et al. 2022a). Notably, the promoter of the xylose operon (*xyl* operon), which includes *xylA* (*AOT13_11570*) and *xylB* (*AOT13_11575*) genes (Fig. 1a), demonstrated its robust activity when induced by xylose (Wang et al. 2022a). Thus, we have chosen this promoter as the basis for developing a xylose-inducible expression system in *P. thermoglucosidasius*.

To identify the minimal region exhibiting promoter activity, we systematically truncated the upstream DNA sequence of the *xylA* gene. PCR amplification resulted in DNA fragments *xylA500*p, *xylA400*p, *xylA300*p, *xylA200*p, and *xylA100*p, corresponding to the 500-bp, 400-bp, 300-bp, 200-bp, and 100-bp upstream sequences preceding the *xylA* gene start codon, respectively (see Fig. S6). These fragments were then inserted into the pUCG18-*sfgfp* vector and transformed into DSM 2542, generating strains DSM 2542-*xylA500*p, DSM 2542-*xylA400*p, DSM 2542-*xylA300*p, DSM 2542-*xylA200*p, and DSM 2542-*xylA100*p (Fig. 2a). Their promoter activities were evaluated by the fluorescence intensities of superfolder green fluorescent protein (sfGFP). As shown in Fig. 2b, while *xylA500*p exhibited robust promoter activity upon addition of 1% xylose, the truncated fragments (*xylA400*p, *xylA300*p, *xylA200*p, and *xylA100*p) displayed significantly weaker promoter activity. This suggests that the core promoter sequence is located within the region spanning −500 to −400 bp upstream of *xylA*, hereafter referred to as *xylA1*p.

Next, we aimed to identify the core elements of the *xylA1*p promoter. To pinpoint the transcriptional start site (TSS) of *xylA1*p, we utilized rapid amplification of 5'-cDNA ends (5'-RACE) analysis (Fig. 2c). With the TSS determined, we identified the putative −10 and −35 regions of *xylA1*p (Fig. 2c). Furthermore, we identified a conserved motif (CTTTGTTTG TACACTAGACAAACAAAT) within *xylA1*p, displaying high consistency with the reported XylR recognition sequence of *Parageobacillus* and *Bacillus* (Fig. 2c) (Liang et al. 2022a).

To identify the minimal promoter region capable of driving reporter gene expression in response to xylose, we systematically truncated the 100-bp promoter *xylA1*p while retaining its promoter core region. This resulted in *xylA2*p, *xylA3*p, *xylA4*p, and *xylA5*p variants, while *xylA6*p was used as a negative control in which the −10 region and the TSS were deleted (Fig. 2d). As assessed by the fluorescence intensity of sfGFP, all variants except *xylA6*p exhibited promoter activities similar to that of *xylA1*p (Fig. 2e). These findings established *xylA5*p as the minimal promoter sequence (Fig. 2e), which we subsequently employed in the subsequent investigations.

Characterization of the regulatory mechanism of XylR

Then we try to evaluate the interaction between XylR and *xylA5*p. To eliminate any naturally occurring regulation mediated via XylR, its encoding gene *xylR* was knockout in *P. thermoglucosidasius* DSM 2542, resulting strain DSM 2542 Δ*xylR* (Fig. S4F). Under the control of the promoter *xylA5*p, the *sfgfp* gene was expressed in DSM 2542 Δ*xylR*, resulting in strain DSM 2542 Δ*xylR*-*xylA5*p-*sfgfp* displaying fluorescence (Fig. 3a and b). Upon the addition of the *xylR* gene controlled by promoter p43 (a sequence from *B. subtilis* with promoter activity) (Wang et al. 2019; Wang and Doi 1984) amplified from plasmid puBS4.4-p43 (see Table S2), the resulting strain DSM 2542 Δ*xylR*-p43-*xylR*-*xylA5*p-*sfgfp* exhibited a significant reduction in fluorescence in the absence of xylose (Fig. 3a and b). Notably, upon the addition of xylose, the fluorescence was restored in a dose-dependent manner (Fig. 3b). These findings demonstrate that the *xylA5*p-*xylR* system is capable of sensing and responding to xylose, thereby derepressing the transcription of *xylA5*p.

The direct interaction between XylR and *xylA5*p was evaluated by electrophoretic mobility shift assay (EMSA) in vitro. XylR was expressed in *E. coli* and purified to homogeneous via Ni-NTA chromatography (Fig. S7). EMSA was performed using purified XylR and a probe containing promoter *xylA5*p (PxylA5), revealing that XylR could bind to PxylA5 in a concentration-dependent manner (Fig. 3c). No retardation was observed for the probe containing promoter p43 (Fig. S8), suggesting the binding between XylR and *xylA5*p was specific. The responses of XylR/PxylA5 complexes to xylose were also evaluated by EMSA. The XylR/PxylA5 complexes dissociated gradually with the increased xylose concentration (Fig. 3d). These results, together with the in vivo results of sfGFP reporter system, confirmed that XylR could directly bind to promoter *xylA5*p and repress its promoter activity. This repression effect could be relieved by xylose, thus inducing the transcription of genes downstream to promoter *xylA5*p.

Development of a xylose-inducible expression system IExyl*

The xylose-concentration-dependent GFP fluorescence observed in DSM 2542 Δ*xylR*-p43-*xylR*-*xylA5*p-*sfgfp*

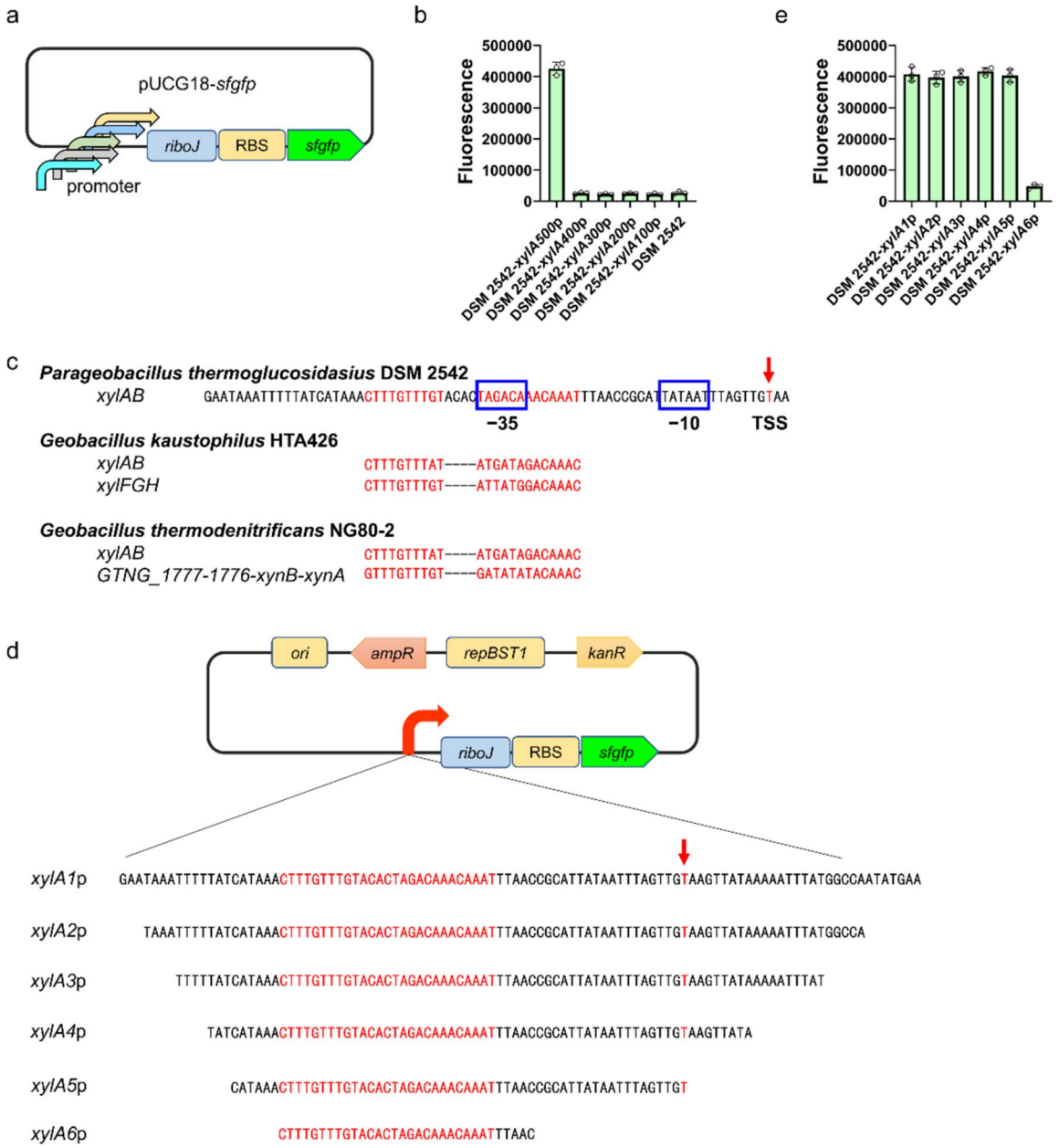

Fig. 2 Identification and characterization of a xylose-inducible promoter from *P. thermoglucosidasius*. **a** Schematic representation of the sfGFP reporter system. **b** Fluorescence intensity levels of DSM 2542 strains carrying plasmids with different promoters at 24 h. **c** Nucleotide sequences of the core region of the promoter of *xyl* operon. The transcriptional start site (TSS) is indicated by a red arrow. The putative −10 and −35 regions of the promoter are marked with boxes. Putative XylR binding sequence are labeled in red, together with XylR binding sites from *Geobacillus kaustophilus* HTA426 and *Geobacillus thermodenitrificans* NG80-2. **d** Construction of plasmids containing *sfgfp* and truncated promoters derived from *xylA1*p. **e** Fluorescence intensity levels of DSM 2542 strains harboring plasmids with different promoters at 24 h. The standard deviations were calculated from at least three independent experiments and were represented as error bars. The insulator sequence *riboJ* was reported to eliminate the interference between promoters and RBSs and improve the modularity of regulatory elements (Lou et al. 2012)

indicated the potential of XylR and *xylA5*p as an inducible expression system in *P. thermoglucosidasius*. Full induction was achieved at a xylose concentration of 10 mM (Fig. 3b). Interestingly, the fully induced fluorescence surpassed that driven by the commonly used promoters pLdh (the promoter of lactate dehydrogenase encoding gene) (Cripps et al. 2009) and pRplsWT (a constitutive promoter selected from a promoter library) (Reeve et al. 2016), while the fluorescence of pLdh is significantly lower than that of pRplsWT, indicating the remarkably strong activity of *xylA5*p (Fig. 4a).

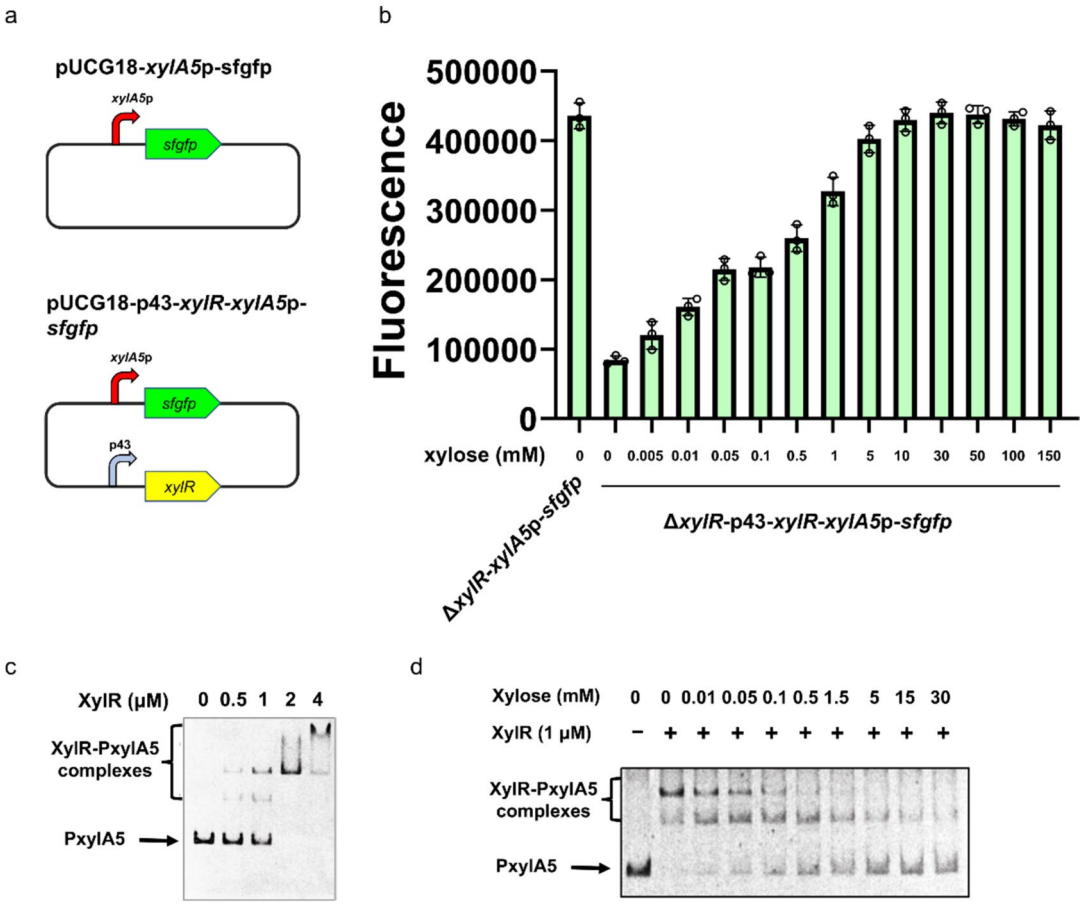

Fig. 3 The regulatory mechanism of XylR. **a** Schematic representation of the sfGFP reporter system. **b** Fluorescence levels of strain Δ*xylR*-p43-*xylR*-*xylA5*p-*sfgfp* with varying concentrations of xylose at 24 h. Strain DSM 2542 Δ*xylR*-*xylA5*p-*sfgfp* was used as a positive control. **c** EMSA depicting the interaction between XylR and a probe containing *xylA5*p (PxylA5). **d** Dissociation effect of xylose to the complex of XylR and probe PxylA5 at different concentration of xylose. Each lane contains 6 ng of PxylA5, 1 μM XylR and different amounts of xylose. The standard deviations were calculated from at least three independent experiments and were represented as error bars

However, a higher level of leaky expression was observed in the absence of xylose induction compared to the control (Figs. 3b and 4a). To address this issue, we explored the inducible expression system by replacing promoter p43 with promoters *xylR*p, pLdh, and pRplsWT, respectively, with the aim of achieving more stringent control over the promoter activity of *xylA5*p by increased expression level of XylR (Fig. 4b). Through a comparison of GFP fluorescence in strains DSM 2542 Δ*xylR*-*xylR*p-*xylR*-*xylA5*p-*sfgfp*, DSM 2542 Δ*xylR*-pLdh-*xylR*-*xylA5*p-*sfgfp*, and DSM 2542 Δ*xylR*-pRplsWT-*xylR*-*xylA5*p-*sfgfp* in the absence of xylose, we observed a significant reduction in leaky expression, reaching a level similar to that of the control (Fig. 4b and c). Subsequently, we selected DSM 2542 Δ*xylR*-*xylA5*p-*xylR*p-*xylR* for further investigations.

As shown in Fig. 4d, the fluorescence intensity of the strain Δ*xylR*-*xylR*p-*xylR*-*xylA5*p-*sfgfp* exhibited an increase in response to escalating xylose concentrations. To assess the transcriptional activity of *xylA5*p in strain Δ*xylR*-*xylR*p-*xylR*-*xylA5*p-*sfgfp* under different xylose doses, RT-qPCR was performed to measure the transcript levels of the *sfGFP* gene. The results demonstrated a xylose concentration-dependent enhancement in *sfGFP* transcript abundance (Fig. 4e), indicating the sensitive and controlled induction of promoter *xylA5*p in strain Δ*xylR*-*xylR*p-*xylR*-*xylA5*p-*sfgfp* at the transcriptional level in *P. thermoglucosidasius*. Notably, we observed a significant inhibitory effect of glucose, a commonly utilized fermentation substrate, on the transcriptional activity of promoter *xylA5*p in strain Δ*xylR*-*xylR*p-*xylR*-*xylA5*p-*sfgfp* (Fig. 4f). This observation posed a significant challenge to the practical application of the system when employing lignocellulosic biomass or other glucose-containing substrates as carbon sources. The detrimental effect of glucose, known as carbon catabolite repression (CCR), has to be addressed to overcome this limitation.

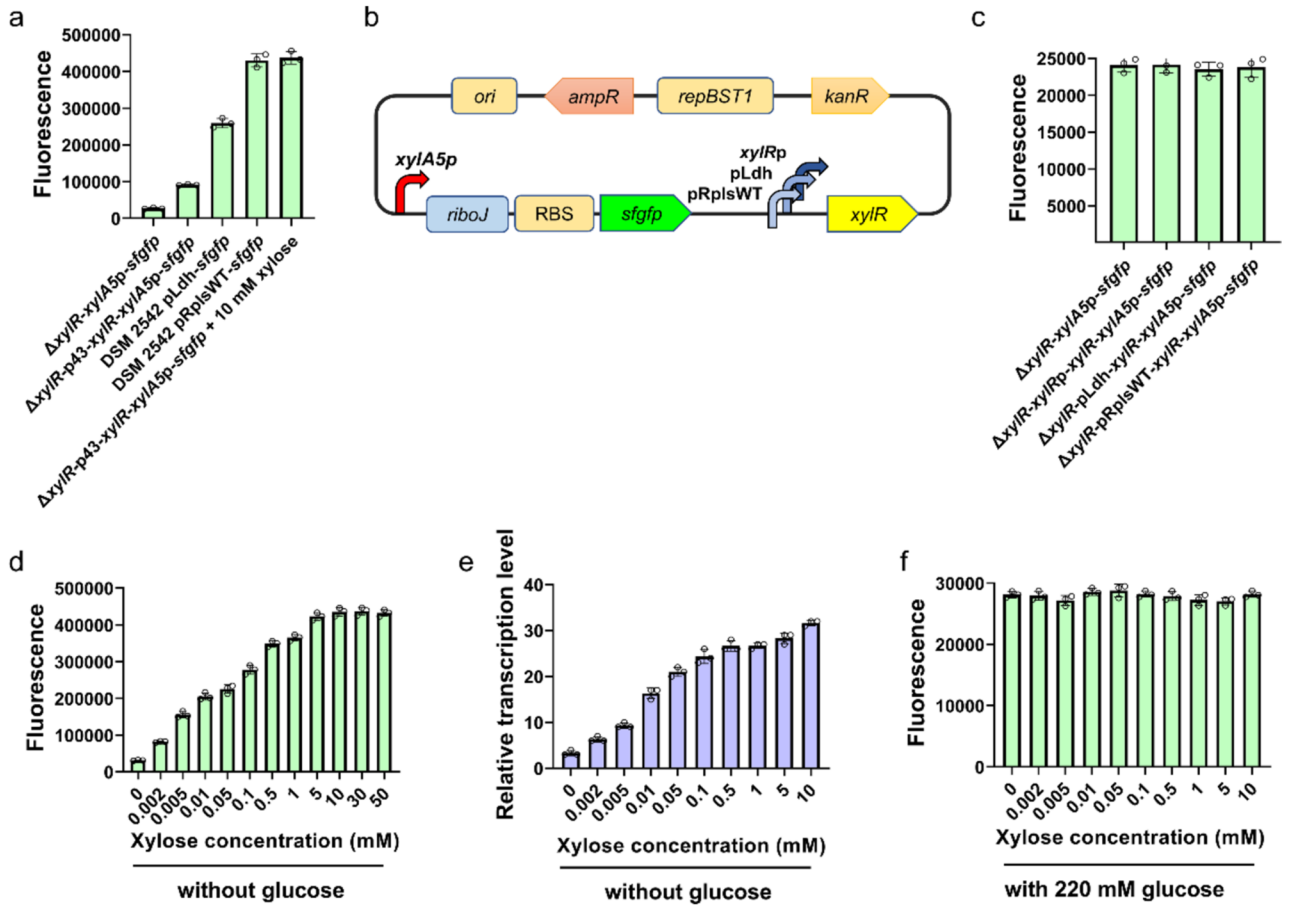

Fig. 4 Development and characterization of the inducible expression system Δ*xylR*-*xylR*p-*xylR*-*xylA5*p-*sfgfp*. **a** GFP fluorescence of strains with and without xylose induction. **b** Plasmid map depicting the components of the inducible expression system. **c** GFP fluorescence of DSM 2542 strains harboring inducible expression system with different promoters for *xylR* when cultured without xylose at 24 h. **d** GFP fluorescence of strain Δ*xylR*-*xylR*p-*xylR*-*xylA5*p-*sfgfp* in response to different doses of xylose at 24 h. The concentrations of xylose were 0, 0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1, 5, 10, 30, and 50 mM. **e** Quantitative analysis of the relative transcript levels of *sfgfp* determined by RT-qPCR in strain Δ*xylR*-*xylR*p-*xylR*-*xylA5*p-*sfgfp* in response to different doses of xylose at 24 h. The transcript strength of *sfgfp* in strain Δ*xylR*-p43-*xylR*-*xylA5*p-*sfgfp* was arbitrarily assigned a value of 1. **f** GFP fluorescence in strain Δ*xylR*-*xylR*p-*xylR*-*xylA5*p-*sfgfp* cultured in medium with glucose as the sole carbon source at 24 h. The standard deviations were calculated from at least three independent experiments and were represented as error bars. The control referred to a DSM 2542 Δ*xylR*-*sfgfp* strain containing a plasmid (pUCG18-*sfgfp*) with a promoterless *sfgfp*

To develop a xylose-inducible expression system that is not repressed by glucose, we constructed an expression system named IExyl3, in which the transcription of the xylose transporter genes was controlled by a constitutive promoter pLdh (Fig. 5a). The performance of IExyl3 was evaluated in *P. thermoglucosidasius*. Strain DSM 2542 IExyl3 was cultured in medium with or without 4% glucose, supplemented with xylose concentrations ranging from 1 µM to 10 mM. Unfortunately, only weak fluorescence was observed in the presence of glucose after 10 hours of culture regardless of the xylose concentration added, indicating the promoter activity was repressed by remaining glucose in the fermentation medium (Fig. 5b and S9). In contrast, varying fluorescence intensities were observed in glucose-free medium, indicating that the transcription of these xylose transporter genes was still inhibited by glucose (Fig. 5b).

Previous studies suggested that the CCR-mediated regulation of xylose metabolism might primarily involve the expression control of pentose transporter operons (Li et al. 2018; Liang et al. 2022a). As depicted in Fig. 5c, an anticipated catabolite responsive element (*cre*) site, TGGAAC CGGTTTCA, was identified at position +991 relative to the ATG start codon of *AOT13_11435*, which aligns with the reported consensus *cre* sequence (TGNAANCGNNNNCN) observed in *B. subtilis* and *P. thermoglucosidasius* NCIMB 11955 (Liang et al. 2022a; Weickert and Chambliss 1990). Consequently, it can be inferred that this element is subject to CcpA-mediated CCR for the regulation of these xylose

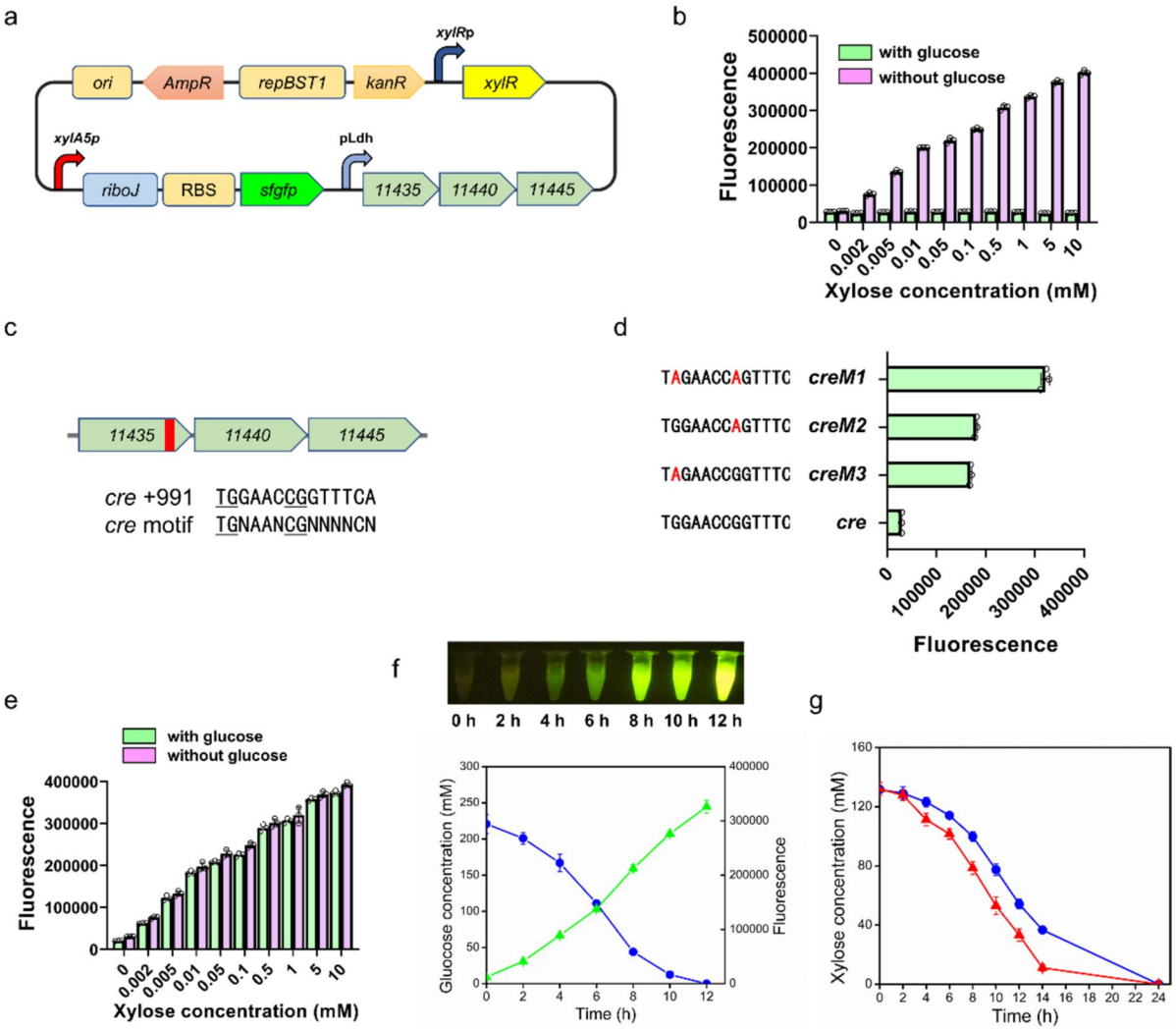

Fig. 5 Development of the inducible expression system IExyl*. **a** Plasmid map of the inducible expression system IExyl3. **b** Evaluation of IExyl3 in medium with or without 4% glucose as the sole carbon source at 10 h. **c** Genetic organization of the identified xylose transporter encoding genes (*AOT13_11435*, *AOT13_11440*, and *AOT13_11445*). The putative *cre* motif at +991 is highlighted in red and located within the *AOT13_11435* gene. The lower section presents a comparison of the *cre* sequence in *AOT13_11435* (*cre* +991) and the conserved *cre* sequence in *B. subtilis* (*cre* motif), in which highly conserved nucleotides are underlined. **d** GFP fluorescence and corresponding *cre* site sequences for DSM 2542 strains carrying mutated *cre* site sequences at 24 h. Mutated nucleotides are highlighted in red. **e** Evaluation of GFP fluorescence in DSM 2542 *creM3* in response to increasing concentrations of xylose, in the presence or absence of glucose, at 24 h. **f** GFP fluorescence (green) and glucose consumption (blue) of DSM 2542 *creM3* in the presence of glucose and 1 mM xylose. **g** Xylose consumption comparison between DSM 2542 IExyl3 (blue) and DSM 2542 *creM3* (red). The standard deviations were calculated from at least three independent experiments and were represented as error bars

transporter genes. Next, we mutated this *cre* site to afford three distinct mutated versions, namely *creM1*, *creM2*, and *creM3* (Fig. 5d). As illustrated in Fig. 5d, a substantial 5- to 10-fold enhancement in fluorescence intensity was observed for these mutated strains when glucose was employed as the carbon source. These findings underscore the crucial role of conserved nucleotides within the *cre* consensus sequence in CCR. Among the mutated sequences, *creM3* exhibited the highest increase in GFP fluorescence in response to varying concentrations of xylose, in both the presence and absence of glucose (Fig. 5e, 5f, and S10). Concurrently, the overexpression of xylose transporter genes resulted in an accelerated rate of xylose consumption (Fig. 5g), indicating that the *cre* mutant enables efficient xylose transport even in the presence of glucose. Accordingly, the corresponding expression system was designated as IExyl*.

Application of IExyl* in metabolic engineering for increased riboflavin yield in *P. thermoglucosidasius*

Previous studies on riboflavin-producing strains of *B. subtilis* have identified target genes for enhanced riboflavin production. Overexpression of glucose-6-phosphate dehydrogenase (encoded by *zwf* gene, Fig. 6a) increased the metabolic flow of the pentose phosphate pathway, resulting in 25% increase in riboflavin production (Duan et al. 2010). Another crucial enzyme, GTP cyclohydrolase II/3,4-dihydroxy-2-butanone-4-phosphate synthase (encoded by *ribA*, Fig. 6a), acted as the rate-limiting step in riboflavin biosynthesis in *B. subtilis* (You et al. 2022). Overexpression of *ribA* led to 25% increase in riboflavin production when glucose was used as the carbon source (You et al. 2022). Recent advances in synthetic biology have demonstrated that fine-tuning gene expression to optimize intracellular metabolic flow is a promising strategy for maximizing target product yield (Gupta et al. 2017).

In a previous study, the riboflavin biosynthetic gene cluster of *P. thermoglucosidasius DSM 2542* was overexpressed using pLdh promoter in this strain, resulting a riboflavin-producing strain *P. thermoglucosidasius* DSM 2542 Rib-Gtg (Yang et al. 2021). Therefore, we utilized the inducible expression system IExyl* to finely regulate the expression of *zwf* and *ribA* genes in DSM 2542 Rib-Gtg, resulting the engineered strains DSM 2542 Rib-Gtg-IExyl*-*zwf* and DSM

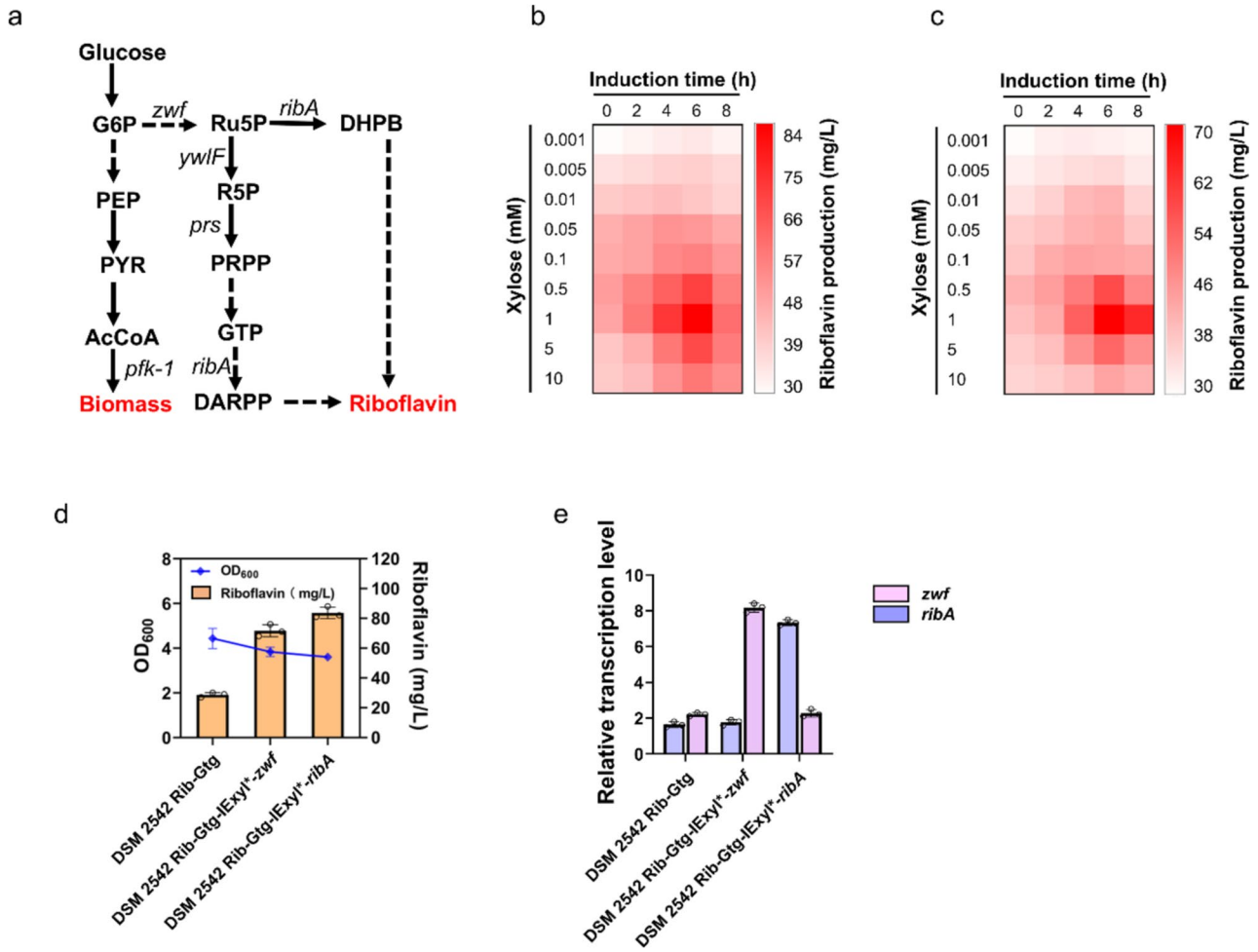

Fig. 6 Application of IExyl* for the improvement of riboflavin production in *P. thermoglucosidasius*. **a** Schematic representation of the riboflavin biosynthetic pathway in *P. thermoglucosidasius*. **b** Heat map illustrating the riboflavin yields of DSM 2542 Rib-Gtg-IExyl*-*ribA* under different xylose concentrations and induction time points. **c** Heat map illustrating the riboflavin yields of DSM 2542 Rib-Gtg-IExyl*-*zwf* under different xylose concentrations and induction time points. **d** Growth profiles and riboflavin yields of DSM 2542 Rib-Gtg, DSM 2542 Rib-Gtg-IExyl*-*zwf*, and DSM 2542 Rib-Gtg-IExyl*-*ribA* in shake-flask fermentation at 24 h. **e** Relative transcription levels of *zwf* and *ribA* genes determined by RT-qPCR in DSM 2542 Rib-Gtg, DSM 2542 Rib-Gtg-IExyl*-*zwf*, and DSM 2542 Rib-Gtg-IExyl*-*ribA* at 24 h, respectively. The standard deviations were calculated from at least three independent experiments and were represented as error bars

2542 Rib-Gtg-IExyl*-*ribA*. The expression levels of *zwf* and *ribA* were optimized by adding xylose at different doses and time points. Riboflavin production was subsequently evaluated under various induction conditions. Our results revealed that the optimal conditions for riboflavin production in both DSM 2542 Rib-Gtg-IExyl*-*zwf* and DSM 2542 Rib-Gtg-IExyl*-*ribA* were achieved by adding 1 mM xylose at 6 h (Fig. 6b and c).

Growth and riboflavin production were assessed through shake-flask fermentation of the fine-tuning strains DSM 2542 Rib-Gtg-IExyl*-*zwf* and DSM 2542 Rib-Gtg-IExyl*-*ribA*, and compared to the starting strain DSM 2542 Rib-Gtg. As shown in Fig. 6d, the biomass of the fine-tuning strains at 24 h was slightly lower than that of the starting strain DSM 2542 Rib-Gtg, while the riboflavin yields of the fine-tuning strains exhibited a 2.3- to 2.8-fold increase compared to the starting strain (Fig. 6d). The specific rates of glucose consumption for the engineered strains were similar to that of the parent strain DSM 2542 Rib-Gtg (Table S5). These findings showed that the inducible expression system IExyl* can serve as an effective tool for rational engineering in *Parageobacillus*, offering a promising approach for enhancing product yields.

The impact of fine-tuning the *zwf* and *ribA* genes using IExyl* was further investigated in the fine-tuning strains. As depicted in Fig. 6e, the transcription levels of both genes were significantly enhanced in the fine-tuning strains, as assessed by RT-qPCR. Additionally, the intracellular concentrations of intermediate metabolites of the pentose phosphate pathway were also measured (Table 1). The pentose phosphate pathway plays a crucial role in providing ribulose-5-phosphate, a direct precursor for riboflavin biosynthesis, as well as ribose-5-phosphate, which is converted to phosphoribosylpyrophosphate (PRPP) and subsequently to GTP, another biosynthetic precursor of riboflavin. In the fine-tuning strains, slight increase in the intracellular concentration of ribulose-5-phosphate was observed, while the concentration of ribose-5-phosphate exhibited more obvious increase (Table 1). Moreover, compared to the parent strain, the intracellular PRPP concentration in the fine-tuning strains was significantly elevated, reaching a maximum concentration of 3.99 nmol/mg (dry weight) (Table 1). These findings indicate that the inducible expression system successfully reconfigures the intracellular metabolite concentrations, thereby promoting riboflavin production. In conclusion, we have designed a plug-and-play plasmid utilizing the newly developed inducible expression system IExyl*, which has been employed for fine-tuning gene expression to coordinate intracellular metabolism and maximize the yield of target products.

Table 1 Intracellular Concentration of Metabolites in *P. thermoglucosidasius* strains

	DSM 2542 Rib-Gtg	DSM 2542 Rib-Gtg- IExyl*-*zwf*	DSM 2542 Rib-Gtg- IExyl*-*ribA*
ribose-5-P	3.69 ± 0.11	3.94 ± 0.13	3.97 ± 0.22
ribulose-5-P	1.23 ± 0.14	1.28 ± 0.07	1.31 ± 0.08
PRPP	1.11 ± 0.06	2.87 ± 0.11	3.99 ± 0.17

Discussion

Optimizing the levels of gene expression stands as a critical factor in attaining heightened yield and productivity for desired products (Gupta et al. 2017). The adverse impact on cell growth and product yields due to excessive expression of genes underscores the need for precision. To achieve the optimal results, precise control of the expression timing and strength of target genes is required. Especially, expression of target genes at proper timing are paramount to avoid competition with the cell growth for precursors, ensuring the balance between cell growth and the ultimate yield of the target product. Additionally, this precision in timing serves to alleviate the possible toxic effects associated with the premature accumulation of the target product (Wang et al. 2016). Employing an inducible expression system emerges as a valuable tool for ascertaining the optimal strength of expression and the ideal time point for such applications. This system not only facilitates controlled expression of heterologous proteins but also aids in constructing genetic tools for expressing genes of interest at specific stages, adding an extra layer of flexibility and control to the process.

Xylose-inducible promoters have been extensively applied and modified across diverse bacterial species including *B. subtilis*, *Clostridium perfringens*, *Brevibacillus choshinensis*, and *Lactococcus lactis* to attain precise control over gene expression (Bhavsar et al. 2001; D'Urzo et al. 2013; Guzman et al. 1995; Nariya et al. 2011). In our previous study, we developed a suite of xylose-responsive promoters for *P. thermoglucosidasius*, providing a platform for the dynamic control over the full pathway of riboflavin production (Wang et al. 2022a). However, these xylose-inducible promoters are susceptible to glucose repression, impeding their further application in mixed substrates containing glucose. In this study, we successfully provided an xylose-inducible and glucose-insensitive gene expression system for *Parageobacillus* strains, allowing for dynamic metabolic control using xylose as inducer while glucose as the primary carbon source.

Our analysis of gene transcription levels in both the presence and absence of xylose revealed the promoter of the *xyl* operon as a robust candidate for a xylose-inducible promoter. These results, together with the bioinformatic analysis of the genome sequence of *P. thermoglucosidasius* DSM 2542, also suggested that a set of sugar ABC transporter and an ROK

family transcriptional regulator XylR were involved in the uptake of xylose and the regulation of xylose metabolism (Wang et al. 2022a). In this study, we deciphered the core sequence of the promoter of *xyl* operon, now designated as *xylA5*p. By combining *xylA5*p in tandem with its repressor XylR, subsequently fine-tuning to minimize leak expression and surmount glucose repression challenges, a xylose-inducible expression system applicable with glucose was developed for *Parageobacillus*, which was designated as IExyl*. Significantly, our xylose-inducible system effectively utilizes glucose as the primary carbon source while leveraging xylose as an inducer. Then, it was used to enhance riboflavin production in *P. thermoglucosidasius*, representing a noteworthy yield improvement in riboflavin in engineered *Parageobacillus* strains.

CCR poses a great challenges to the effective deployment of xylose-inducible expression systems in industrial fermentation scenarios, given that glucose stands as the favored carbon source for the majority of microorganisms (Liang et al. 2022a, b; You et al. 2022). Previous research has unveiled that glucose-6-phosphate, a metabolic byproduct of glucose, counteracts the induction of the xylose regulon by acting through the global regulator CcpA (Dahl et al. 1995; Faires et al. 1999). However, a straightforward elimination of CcpA via gene knockout is not a tenable solution. Due to its pivotal role in carbohydrate metabolism, the absence of CcpA showed serious adverse effects on cell growth and fermentation efficiency (Faires et al. 1999; Tobisch et al. 1999). Consequently, our study revealed that the fine tuning CCR-related components, as opposed to their complete deletion, was effective to remove CCR effect. By introducing mutations in the *cre* sequences of the xylose transporter gene *AOT13_11435*, we successfully weakened the binding affinity between CcpA and the *cre* sites, resulting in over ten-fold increase in fluorescence intensity of sfGFP induced by xylose with the accompany of glucose. Therefore, the engineering strategy implemented in this study significantly enhances the practical viability of xylose-inducible expression systems. Moreover, it provides valuable insights applicable to the refinement of other sugar-inducible expression systems.

Our xylose-inducible expression system IExyl* showed robust functionality in the concurrent presence of glucose and xylose. It strategically utilizes glucose as the primary carbon source favorable for cell growth, while harnessing xylose as a cost-effective inducer. This system presents significant potential for dynamic control over protein expression and the biosynthesis of targeted products, particularly when glucose is the preferred carbon source. In summary, our xylose-inducible and glucose-insensitive IExyl* system provides a new tool for future metabolic engineering and synthetic biology investigations in the thermophilic *Parageobacillus* strains.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00253-024-13333-w.

Author contribution J.W., K.F., Z.Li and Y.C. conceived and designed research. J.W. conducted experiments. W.W. and Z.Liu contributed new reagents or analytical tools. J.W., K.F., X.J. and G.P. analyzed data and wrote the manuscript. All authors contributed to critical revision of the manuscript and gave final approval of the manuscript for publication.

Funding This work was supported by the National Key R&D Program of China (2020YFA0907700 and 2021YFC2100600), and the National Natural Science Foundation of China (32070067).

Data availability All data generated and analyzed during this study are presented and made available in this article and its Supplementary Information.

Declarations

Ethics approval This article does not contain any studies with human participants or animals performed by any of the authors.

Competing interests The authors declare no competing interests.

References

Bhavsar AP, Zhao XM, Brown ED (2001) Development and characterization of a xylose-dependent system for expression of cloned genes in *Bacillus subtilis*: Conditional complementation of a teichoic acid mutant. Appl Environ Microbiol 67(1):403–410. https://doi.org/10.1128/AEM.67.1.403-410.2001

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. BMC bioinformatics 10:421. https://doi.org/10.1186/1471-2105-10-421

Canton B, Labno A, Endy D (2008) Refinement and standardization of synthetic biological parts and devices. Nat Biotechnol 26(7):787–93. https://doi.org/10.1038/nbt1413

Cripps RE, Eley K, Leak DJ, Rudd B, Taylor M, Todd M, Boakes S, Martin S, Atkinson T (2009) Metabolic engineering of *Geobacillus thermoglucosidasius* for high yield ethanol production. Metab Eng 11(6):398–408. https://doi.org/10.1016/j.ymben.2009.08.005

Dahl MK, Schmiedel D, Hillen W (1995) Glucose and glucose-6-phosphate interaction with Xyl repressor proteins from *Bacillus* spp may contribute to regulation of xylose utilization. J Bacteriol

177(19):5467–5472. https://doi.org/10.1128/jb.177.19.5467-5472.1995
Daily J (2016) Parasail: SIMD C library for global, semi-global, and local pairwise sequence alignments. BMC bioinformatics 17:81. https://doi.org/10.1186/s12859-016-0930-z
Duan YX, Chen T, Chen X, Zhao XM (2010) Overexpression of glucose-6-phosphate dehydrogenase enhances riboflavin production in *Bacillus subtilis*. Appl Microbiol Biotechnol 85(6):1907–1914. https://doi.org/10.1007/s00253-009-2247-6
D'Urzo N, Martinelli M, Nenci C, Brettoni C, Telford JL, Maione D (2013) High-level intracellular expression of heterologous proteins in *Brevibacillus choshinensis* SP3 under the control of a xylose inducible promoter. Microb Cell Fact 12. https://doi.org/10.1186/1475-2859-12-12
Faires N, Tobisch S, Bachem S, Martin-Verstraete I, Hecker M, Stulke J (1999) The catabolite control protein CcpA controls ammonium assimilation in *Bacillus subtilis*. J Mol Microbiol Biotechnol 1(1):141–8
Frenzel E, Legebeke J, van Stralen A, van Kranenburg R, Kuipers OP (2018) In vivo selection of sfGFP variants with improved and reliable functionality in industrially important thermophilic bacteria. Biotechnol Biofuels 11:8. https://doi.org/10.1186/s13068-017-1008-5
Gu Y, Ding Y, Ren C, Sun Z, Rodionov DA, Zhang W, Yang S, Yang C, Jiang W (2010) Reconstruction of xylose utilization pathway and regulons in Firmicutes. BMC Genomics 11:255. https://doi.org/10.1186/1471-2164-11-255
Guo X, Shi S (2023) De novo genome sequencing and assembly of *Rhodosporidium toruloides* strain deltadao1e. Microbiol Resour Announc 12(2):e0060022. https://doi.org/10.1128/mra.00600-22
Gupta A, Reizman IMB, Reisch CR, Prather KLJ (2017) Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. Nat Biotechnol 35(3):273–279. https://doi.org/10.1038/nbt.3796
Guzman LM, Belin D, Carson MJ, Beckwith J (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J Bacteriol 177(14):4121–30. https://doi.org/10.1128/jb.177.14.4121-4130.1995
Hector RE, Mertens JA (2017) A synthetic hybrid promoter for xylose-regulated control of gene expression in *Saccharomyces* yeasts. Mol Biotechnol 59(1):24–33. https://doi.org/10.1007/s12033-016-9991-5
Kim D, Woo HM (2018) Deciphering bacterial xylose metabolism and metabolic engineering of industrial microorganisms for use as efficient microbial cell factories. Appl Microbiol Biotechnol 102(22):9471–9480. https://doi.org/10.1007/s00253-018-9353-2
Lau MSH, Sheng L, Zhang Y, Minton NP (2021) Development of a suite of tools for genome editing in *Parageobacillus thermoglucosidasius* and their use to identify the potential of a native plasmid in the generation of stable engineered strains. ACS Synth Biol 10(7):1739–1749. https://doi.org/10.1021/acssynbio.1c00138
Letunic I, Bork P (2024) Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Res. https://doi.org/10.1093/nar/gkae268
Li Y, Jin K, Zhang L, Ding Z, Gu Z, Shi G (2018) Development of an inducible secretory expression system in *Bacillus licheniformis* based on an engineered xylose operon. J Agric Food Chem 66(36):9456–9464. https://doi.org/10.1021/acs.jafc.8b02857
Liang JH, Roberts A, van Kranenburg R, Bolhuis A, Leak DJ (2022a) Relaxed control of sugar utilization in *Parageobacillus thermoglucosidasius* DSM 2542. Microbiol Res 256:126957. https://doi.org/10.1016/j.micres.2021.126957
Liang JH, van Kranenburg R, Bolhuis A, Leak DJ (2022b) Removing carbon catabolite repression in *Parageobacillus thermoglucosidasius* DSM 2542. Front Microbiol 13. https://doi.org/10.3389/fmicb.2022.985465
Lin PP, Rabe KS, Takasumi JL, Kadisch M, Arnold FH, Liao JC (2014) Isobutanol production at elevated temperatures in thermophilic *Geobacillus thermoglucosidasius*. Metab Eng 24:1–8. https://doi.org/10.1016/j.ymben.2014.03.006
Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25(4):402–8. https://doi.org/10.1006/meth.2001.1262
Lou C, Stanton B, Chen Y-J, Munsky B, Voigt CA (2012) Ribozyme-based insulator parts buffer synthetic circuits from genetic context. Nature Biotechnol 30(11):1137–42. https://doi.org/10.1038/nbt.2401
Madika A, Spencer J, Lau MSH, Sheng L, Zhang Y, Minton NP (2022) pMTL60000: a modular plasmid vector series for *Parageobacillus thermoglucosidasius* strain engineering. J Microbiol Methods 202:106600. https://doi.org/10.1016/j.mimet.2022.106600
Marcano-Velazquez JG, Lo J, Nag A, Maness PC, Chou KJ (2019) Developing riboswitch-mediated gene regulatory controls in thermophilic bacteria. ACS Synth Biol 8(4):633–640. https://doi.org/10.1021/acssynbio.8b00487
Markley AL, Begemann MB, Clarke RE, Gordon GC, Pfleger BF (2015) Synthetic biology toolbox for controlling gene expression in the Cyanobacterium *synechococcus* sp strain PCC 7002. Acs Synthetic Biology 4(5):595–603. https://doi.org/10.1021/sb500260k
Mimee M, Tucker AC, Voigt CA, Lu TK (2015) Programming a human commensal bacterium, bacteroides thetaiotaomicron, to sense and respond to stimuli in the murine gut microbiota. Cell Syst 1(1):62–71. https://doi.org/10.1016/j.cels.2015.06.001
Nariya H, Miyata S, Kuwahara T, Okabe A (2011) Development and characterization of a xylose-inducible gene expression system for *Clostridium perfringens*. Appl Environ Microb 77(23):8439–8441. https://doi.org/10.1128/Aem.05668-11
Nikel PI, Martinez-Garcia E, de Lorenzo V (2014) Biotechnological domestication of pseudomonads using synthetic biology. Nat Rev Microbiol 12(5):368–379. https://doi.org/10.1038/nrmicro3253
Noguchi Y, Kashiwagi N, Uzura A, Ogino C, Kondo A, Ikeda H, Sota M (2018) Development of a strictly regulated xylose-induced expression system in *Streptomyces*. Microb Cell Fact 17. https://doi.org/10.1186/S12934-018-0991-Y
Reeve B, Martinez-Klimova E, de Jonghe J, Leak DJ, Ellis T (2016) The *Geobacillus* plasmid set: a modular toolkit for thermophile engineering. ACS Synth Biol 5(12):1342–1347. https://doi.org/10.1021/acssynbio.5b00298
Rodionov DA, Mironov AA, Gelfand MS (2001) Transcriptional regulation of pentose utilisation systems in the Bacillus/Clostridium group of bacteria. FEMS Microbiol Lett 205(2):305–314. https://doi.org/10.1111/j.1574-6968.2001.tb10965.x
Scheler A, Hillen W (1994) Regulation of xylose utilization in *Bacillus licheniformis*: xyl repressor-xyl-operator interaction studied by DNA modification protection and interference. Mol Microbiol 13(3):505–512. https://doi.org/10.1111/j.1365-2958.1994.tb00445.x
Schultz JC, Cao M, Zhao H (2019) Development of a CRISPR/Cas9 system for high efficiency multiplexed gene deletion in *Rhodosporidium toruloides*. Biotechnol Bioeng 116(8):2103–2109. https://doi.org/10.1002/bit.27001
Song YF, Nikoloff JM, Fu G, Chen JQ, Li QG, Xie NZ, Zheng P, Sun JB, Zhang DW (2016) Promoter screening from *Bacillus subtilis* in various conditions hunting for synthetic biology and industrial applications. PLoS One 11(7). https://doi.org/10.1371/journal.pone.0158447
Tamura K, Stecher G, Kumar S (2021) MEGA11: Molecular Evolutionary Genetics Analysis version 11. Mole Biol Evol 38(7):3022–3027. https://doi.org/10.1093/molbev/msab120

Taylor MP, Eley KL, Martin S, Tuffin MI, Burton SG, Cowan DA (2009) Thermophilic ethanologenesis: future prospects for second-generation bioethanol production. Trends Biotechnol 27(7):398–405. https://doi.org/10.1016/j.tibtech.2009.03.006

Tobisch S, Zuhlke D, Bernhardt J, Stulke J, Hecker M (1999) Role of CcpA in regulation of the central pathways of carbon catabolism in *Bacillus subtilis*. J Bacteriol 181(22):6996–7004. https://doi.org/10.1128/JB.181.22.6996-7004.1999

Wang PZ, Doi RH (1984) Overlapping promoters transcribed by *Bacillus subtilis* sigma 55 and sigma 37 RNA polymerase holoenzymes during growth and stationary phases. J Biol Chem 259(13):8619–8625

Wang W, Yang T, Li Y, Li S, Yin S, Styles K, Corre C, Yang K (2016) Development of a synthetic oxytetracycline-inducible expression system for Streptomycetes using de novo characterized genetic parts. ACS Synth Biol 5(7):765–73. https://doi.org/10.1021/acssynbio.6b00087

Wang J, Wang W, Wang H, Yuan F, Xu Z, Yang K, Li Z, Chen Y, Fan K (2019) Improvement of stress tolerance and riboflavin production of Bacillus subtilis by introduction of heat shock proteins from thermophilic bacillus strains. Appl Microbiol Biotechnol 103(11):4455–4465. https://doi.org/10.1007/s00253-019-09788-x

Wang JY, Li ZL, Wang WS, Pang S, Yao YP, Yuan F, Wang HZ, Xu Z, Pan GH, Liu ZH, Chen YH, Fan KQ (2022a) Dynamic control strategy to produce riboflavin with lignocellulose hydrolysate in the thermophile *Geobacillus thermoglucosidasius*. Acs Syn Biol 11(6):2163–2174. https://doi.org/10.1021/acssynbio.2c00087

Wang Z, Fang Y, Shi Y, Xin Y, Gu Z, Yang T, Li Y, Ding Z, Shi G, Zhang L (2022b) Analysis of xylose operon from *Paenibacillus polymyxa* ATCC842 and development of tools for gene expression. Int J Mol Sci 23(9). https://doi.org/10.3390/ijms23095024

Weickert MJ, Chambliss GH (1990) Site-directed mutagenesis of a catabolite repression operator sequence in *Bacillus subtilis*. Proc Natl Acad Sci U S A 87(16):6238–6242. https://doi.org/10.1073/pnas.87.16.6238

Yang Z, Sun Q, Tan G, Zhang Q, Wang Z, Li C, Qi F, Wang W, Zhang L, Li Z (2021) Engineering thermophilic *Geobacillus thermoglucosidasius* for riboflavin production. Microb Biotechnol 14(2):363–373. https://doi.org/10.1111/1751-7915.13543

You JJ, Du YX, Pan XW, Zhang X, Yang TW, Rao ZM (2022) Increased production of riboflavin by coordinated expression of multiple genes in operons in *Bacillus subtilis*. Acs Synthetic Biology 11(5):1801–1810. https://doi.org/10.1021/acssynbio.1c00640