Genetic toolbox for controlled expression of functional proteins in Geobacillus spp.

Pogrebnyakov, Ivan; Jendresen, Christian Bille; Nielsen, Alex Toftgaard

Published in:
PLOS ONE

Link to article, DOI:
10.1371/journal.pone.0171313

Publication date:
2017

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Pogrebnyakov, I., Jendresen, C. B., & Nielsen, A. T. (2017). Genetic toolbox for controlled expression of functional proteins in Geobacillus spp. PLOS ONE, 12(2), Article e0171313.
https://doi.org/10.1371/journal.pone.0171313
Genetic toolbox for controlled expression of functional proteins in Geobacillus spp.

Ivan Pogrebnyakov, Christian Bille Jendresen, Alex Toftgaard Nielsen*

The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark

* atn@biosustainability.dtu.dk

Abstract

Species of genus Geobacillus are thermophilic bacteria and play an ever increasing role as hosts for biotechnological applications both in academia and industry. Here we screened a number of Geobacillus strains to determine which industrially relevant carbon sources they can utilize. One of the strains, G. thermoglucosidasius C56-YS93, was then chosen to develop a toolbox for controlled gene expression over a wide range of levels. It includes a library of semi-synthetic constitutive promoters (76-fold difference in expression levels) and an inducible promoter from the xylA gene. A library of synthetic in silico designed ribosome binding sites was also created for further tuning of translation. The PxylA was further used to successfully express native and heterologous xylanases in G. thermoglucosidasius. This toolbox enables fine-tuning of gene expression in Geobacillus species for metabolic engineering approaches in production of biochemicals and heterologous proteins.

Introduction

For decades, thermophilic bacteria have been used in biotechnology. Their applications have mainly been confined to their thermostable enzymes, one of the most prominent examples being the Taq polymerase from Thermus aquaticus [1]. The global market for industrial enzymes in 2015 was estimated to be US$ 4.4 billion [2] with thermostable enzymes playing an ever increasing role [3]. The genus Geobacillus, comprising some thermophilic species previously belonging to genus Bacillus [4], has also been used in this regard. Examples of industrially relevant enzymes isolated from Geobacillus species include proteases [5], amylases [6], lipases, [7] and xylanases [8], to mention just a few.

Recently, however, there has been a growing interest in thermophiles as biotechnological hosts [9]. Thermophilic species of Bacteria and Archaea are promising candidates for a number of applications, from production of chemicals [10]; [11]; [12] to extraction of metals from mineral ores [13]. Geobacillus species are also keeping up with this trend. For example, successful metabolic engineering of ethanol [14]; [15] and isobutanol [16] production in G. thermoglucosidasius was achieved. Some Geobacillus strains have also been used for heterologous protein expression. A protein from an archaeon, which was insoluble when expressed in Escherichia coli, was successfully folded in G. kaustophilus [17]. This illustrates an important
Some thermostable proteins may need thermophilic rather than mesophilic hosts for proper expression. Apart from that, thermophilic bacteria may offer other interesting advantages over commonly used hosts, like *E. coli* and *B. subtilis*. These include higher reaction rates at elevated temperatures, decreased risk of contamination with mesophilic organisms, as well as ease of recovery of volatile compounds, such as ethanol [18]. In addition to that, *Geobacillus* species can utilize a wider range of carbon sources, notably C5 sugars like xylose and arabinose, and their polymers, xylan and arabinan [19]. These polymers are the most abundant parts of hemicellulose, which constitutes a considerable fraction of plant biomass [20]. Strains that are able to degrade cellulose have also been reported [21]. All these properties point to a high potential of geobacilli for biotechnology, e.g. for plant biomass conversion into value-added chemicals. Some strains have already been employed by commercial companies, e.g. TMO Renewables (www.tmo-group.com), ReBio Technologies (www.rebio.co.uk) and C5-6 Technologies (www.c56technologies.com).

A number of tools for genetic engineering of *Geobacillus* spp. have been developed, including transformation techniques [22]; [23]; episomal [24]; [25] and integration [14] vectors; selection markers [26]. A more detailed overview of current tools for engineering of *Geobacillus* strains can be found in [18].

Although metabolic engineering and heterologous protein expression has been achieved in *Geobacillus*, there have been few attempts to systematically characterize genetic parts, i.e. promoters and ribosome binding sites (RBS) for these species [18]; [27]. Production of biochemicals through metabolic engineering often requires the expression of a number of pathway enzymes, and the best production yields are not always achieved by the highest amount of the target pathway enzymes in the cell, but rather by fine-tuning of the expression levels of the individual enzymes [28]. Similarly, expression levels of heterologous proteins sometimes should be experimentally adjusted for optimal yields [29]. To achieve this goal, a set of both constitutive and inducible promoters as well as RBS’s of various strengths are needed. Recently, Reeve et al. described construction of a library of promoters using mutagenic PCR of a parent strong promoter from *G. thermoglucosidasius* [30]. The mutation rate in this method was approximately 10%, which makes the resulting promoters highly similar. This may cause homologous recombination between the promoters from this library and their wild type parent on the chromosome.

In this study we have generated a toolbox for controlled gene expression in *G. thermoglucosidasius*. The toolbox includes libraries of promoters and RBS sequences of various strengths that were screened using the superfolder green fluorescent protein (sfGFP) as a reporter. Importantly, much of the promoters’ sequences were randomized, so that they have low similarity. Additionally, a xylose-inducible promoter that enables strong and titratable expression was characterized. The developed tools were used to demonstrate the successful expression of two xylanases from *Geobacillus* species.

**Materials and methods**

**Strains, plasmids and media**

Bacterial strains and plasmids used in this study are listed in Table 1. *E. coli* cells were grown in lysogeny broth (LB) [30] with 100 μg/mL ampicillin added when needed for selection of plasmids. *Geobacillus* strains were grown in either of several media. The mTGP (modified from [25]) medium contained per liter: 17 g tryptone, 3 g soy peptone, 5 g NaCl, 2.5 g K₂HPO₄. After autoclaving, sterile solutions were added to final concentrations: 4 mL/L glycerol, 4 g/L sodium pyruvate, 0.59 mM MgSO₄, 0.91 mM CaCl₂ and 0.04 mM FeSO₄; agar to 1.5% (w/v) was added to solidify the medium when needed. Thermophile minimal medium...
(TMM) was adapted from [31] with some modifications. It contained, per liter: Six salts solution (SSS), 930 mL; 1 M MOPS (pH 8.2), 40 mL; 1 mM FeSO$_4$ in 0.4 M tricine, 10 mL; 0.132 M K$_2$HPO$_4$, 10 mL; 0.953 M NH$_4$Cl, 10 mL; 1 M CaCl$_2$, 0.5 mL; trace elements solution, 0.5 ml; Wolfe’s vitamin solution, 10 mL. SSS contained, per 930 mL: 4.6 g NaCl, 1.35 g Na$_2$SO$_4$, 0.23 g KCl, 0.037 g KBr, 1.72 g MgCl$_2$·6H$_2$O, 0.83 g NaNO$_3$. Trace elements solution contained, per liter: 1 g FeCl$_3$·6H$_2$O, 0.18 g ZnSO$_4$·7H$_2$O, 0.12 g CuCl$_2$·2H$_2$O, 0.12 g MnSO$_4$·H$_2$O, 0.18 g CoCl$_2$·6H$_2$O. Yeast extract in final concentrations of 0.02% (w/v) or 0.05% (w/v) was added when indicated. For Geobacillus spp. selections were done using kanamycin at a final concentration of 12.5 μg/mL.

| Strain or plasmid | Description | Reference |
|-------------------|-------------|-----------|
| **Strains**       |             |           |
| *E. coli* NEB5-alpha | *fhu*A2 Δ(*argF-lacZ*)U169 *phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17 | New England Biolabs |
| *G. thermoglucosidasius* C56-YS93 | Hot spring isolate of *G. thermoglucosidasius* | [32] |
| **Plasmids**      |             |           |
| pUCG18            | pMB1 and pBST22 ori, Amp$^R$, Km$^R$. *E. coli*-Geobacillus shuttle vector | [25] |
| plIPGE            | pMB1 and pBST22 ori, Amp$^R$, Km$^R$. P$_{groES}$::sfGFP with P$_{groES}$ having the CIRCE sequence deleted | This study |
| plP1              | Modified plIPGE with P$_{groES}$ derivative P1 | This study |
| plP2              | Modified plIPGE with P$_{groES}$ derivative P2 | This study |
| plP3              | Modified plIPGE with P$_{groES}$ derivative P3 | This study |
| plP4              | Modified plIPGE with P$_{groES}$ derivative P4 | This study |
| plP5              | Modified plIPGE with P$_{groES}$ derivative P5 | This study |
| plP6              | Modified plIPGE with P$_{groES}$ derivative P6 | This study |
| plP7              | Modified plIPGE with P$_{groES}$ derivative P7 | This study |
| plP8              | Modified plIPGE with P$_{groES}$ derivative P8 | This study |
| plP9              | Modified plIPGE with P$_{groES}$ derivative P9 | This study |
| plP10             | Modified plIPGE with P$_{groES}$ derivative P10 | This study |
| plP11             | Modified plIPGE with P$_{groES}$ derivative P11 | This study |
| plP12             | Modified plIPGE with P$_{groES}$ derivative P12 | This study |
| plP13             | Modified plIPGE with P$_{groES}$ derivative P13 | This study |
| plP14             | Modified plIPGE with P$_{groES}$ derivative P14 | This study |
| plP15             | Modified plIPGE with P$_{groES}$ derivative P15 | This study |
| plP16             | Modified plIPGE with P$_{groES}$ derivative P16 | This study |
| plP17             | Modified plIPGE with P$_{groES}$ derivative P17 | This study |
| plP18             | pMB1 and pBST22 ori, Amp$^R$, Km$^R$. P$_{pfl}$::sfGFP | This study |
| plP19             | Derivative of plPRL with RBS replaced with R1 | This study |
| plP20             | Derivative of plPRL with RBS replaced with R2 | This study |
| plP21             | Derivative of plPRL with RBS replaced with R3 | This study |
| plP22             | Derivative of plPRL with RBS replaced with R4 | This study |
| plP23             | Derivative of plPRL with RBS replaced with R5 | This study |
| plP24             | Derivative of plPRL with RBS replaced with R6 | This study |
| plP25             | pMB1 and pBST22 ori, Amp$^R$, Km$^R$. P$_{xylA}$::sfGFP | This study |
| plP26             | pMB1 and pBST22 ori, Amp$^R$, Km$^R$. *xylR*, P$_{xylA}$::sfGFP | This study |
| plP27             | pMB1 and pBST22 ori, Amp$^R$, Km$^R$. P$_{xylA}$::Geoth_2264 | This study |
| plP28             | pMB1 and pBST22 ori, Amp$^R$, Km$^R$. P$_{xylA}$::GtnG_1761 | This study |

doi:10.1371/journal.pone.0171313.t001
DNA manipulations

Genomic DNA was extracted using the Wizard® Genomic DNA Purification Kit (Promega) according to producer’s specifications. Plasmid extractions were performed using NucleoSpin® Plasmid EasyPure kit (Macherey-Nagel).

PCR and cloning

Primers used in this study are listed in Table 2. PCR of DNA fragments for USER cloning was performed with primers containing uracil using the Phusion U Hot Start DNA Polymerase (Thermo Fisher Scientific). Colony PCR was performed with Taq 2x Master Mix (New England Biolabs) in order to detect positive colonies. Reactions were done according to manufacturers’ recommendations with elongation times and annealing temperatures adjusted for specific targets and primers. In most cases annealing temperature was 60˚C and elongation time was programmed at 30 seconds per 1 kb. DNA cloning was performed using USER (uracil-specific excision reagent) technology. It is a simple and robust method, allowing seamless DNA insertions [33]. PCR-amplified DNA fragments containing a primer-incorporated uracil close to both of their 5’-ends were mixed (purification after PCR was not necessary) and treated with DpnI enzyme (Thermo Fisher Scientific) for 30 min at 37˚C to digest template DNA. USER™ enzyme (New England Biolabs) was then added, and the mixture was incubated in three steps: 1) 37˚C for 15 min; 2) 12˚C for 15 min; 3) 10˚C for 10 min. It was then transferred on ice and mixed with chemically competent E. coli cells.

Table 2. Oligonucleotides used in this study.

| Name    | Sequence 5’→3’                  | Target                                           |
|---------|---------------------------------|-------------------------------------------------|
| PNJ24b  | AATTCGGAAATCATGGTATAGCTGGTTTCC | sfGFP with pUCG18 backbone                       |
| PNJ94   | AGGAUCAGCAGGGAAGAGAGGGAGATGAAAGGGAAGAGG |                                                |
| PNJ97   | ACGAATAGTACATAGCTGAAAGGATGAAAGGGAAGAGG |                                                |
| PNJ98   | ACGAATTACAGATCCCTCATATCTGCAATGAAAGGGAAGAGG |                                                |
| PNJ205  | ACGAATAGTACATAGCTGAAAGGATGAAAGGGAAGAGG |                                                |
| PNJ215  | ACGAATAGTACATAGCTGAAAGGATGAAAGGGAAGAGG |                                                |
| PNJ207  | ACGAATAGTACATAGCTGAAAGGATGAAAGGGAAGAGG |                                                |
| PNJ210  | ACGAATAGTACATAGCTGAAAGGATGAAAGGGAAGAGG |                                                |
| PNJ212  | ACGAATAGTACATAGCTGAAAGGATGAAAGGGAAGAGG |                                                |
| PNJ213  | ACGAATAGTACATAGCTGAAAGGATGAAAGGGAAGAGG |                                                |
| PNJ215  | ACGAATAGTACATAGCTGAAAGGATGAAAGGGAAGAGG |                                                |
| PNJ217  | ACGAATAGTACATAGCTGAAAGGATGAAAGGGAAGAGG |                                                |
| PNJ267  | ACGAATAGTACATAGCTGAAAGGATGAAAGGGAAGAGG |                                                |
| PNJ268  | ACGAATAGTACATAGCTGAAAGGATGAAAGGGAAGAGG |                                                |
| PNJ269  | ACGAATAGTACATAGCTGAAAGGATGAAAGGGAAGAGG |                                                |
| PNJ388  | ACGAATAGTACATAGCTGAAAGGATGAAAGGGAAGAGG |                                                |
| PNJ389  | ACGAATAGTACATAGCTGAAAGGATGAAAGGGAAGAGG |                                                |
| PNJ458  | ACGAATAGTACATAGCTGAAAGGATGAAAGGGAAGAGG |                                                |
| PNJ459  | ACGAATAGTACATAGCTGAAAGGATGAAAGGGAAGAGG |                                                |
| PNJ474  | ACGAATAGTACATAGCTGAAAGGATGAAAGGGAAGAGG |                                                |
| PNJ475  | ACGAATAGTACATAGCTGAAAGGATGAAAGGGAAGAGG |                                                |
| PNJ550  | ACGAATAGTACATAGCTGAAAGGATGAAAGGGAAGAGG |                                                |
| PNJ551  | ACGAATAGTACATAGCTGAAAGGATGAAAGGGAAGAGG |                                                |
| PNJ552  | ACGAATAGTACATAGCTGAAAGGATGAAAGGGAAGAGG |                                                |

Table 2. Oligonucleotides used in this study.

| Name    | Sequence 5’→3’                  | Target                                           |
|---------|---------------------------------|-------------------------------------------------|
| PNJ24b  | AATTCGGAAATCATGGTATAGCTGGTTTCC | sfGFP with pUCG18 backbone                       |
| PNJ94   | AGGAUCAGCAGGGAAGAGAGGGAGATGAAAGG |                                                |
| PNJ97   | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ98   | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ205  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ215  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ207  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ210  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ212  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ213  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ215  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ217  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ267  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ268  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ269  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ388  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ389  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ458  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ459  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ474  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ475  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ550  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ551  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ552  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |

Table 2. Oligonucleotides used in this study.

| Name    | Sequence 5’→3’                  | Target                                           |
|---------|---------------------------------|-------------------------------------------------|
| PNJ24b  | AATTCGGAAATCATGGTATAGCTGGTTTCC | sfGFP with pUCG18 backbone                       |
| PNJ94   | AGGAUCAGCAGGGAAGAGAGGGAGATGAAAGG |                                                |
| PNJ97   | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ98   | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ205  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ215  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ207  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ210  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ212  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ213  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ215  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ217  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ267  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ268  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ269  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ388  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ389  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ458  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ459  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ474  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ475  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ550  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ551  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |
| PNJ552  | ACGAATAGTACATAGCTGAAAGGATGAAAGG |                                                |

doi:10.1371/journal.pone.0171313.t002
Transformation of *E. coli* and *G. thermoglucosidasius*

The procedure was based on the protocol described by [25] with some steps modified. *G. thermoglucosidasius* was grown overnight on an mTGP agar plate at 60°C. A single colony was inoculated into 50 mL of pre-warmed liquid mTGP in a 250 mL flask and incubated at 60°C and 250 rpm until the culture reached OD_{600} of 1.6–2. Cells were cooled down on ice for 10 min and harvested by centrifugation at 4000 g for 10 min. They were washed three times (4000 g for 10 min) with freshly prepared ice-cold electroporation buffer. The buffer contained, per 100 mL: 17.12 g sucrose, 0.042 g MgCl$_2$·6H$_2$O, 5 mL glycerol. After the last washing step, the cell pellet was suspended in 1 mL of electroporation buffer, distributed in 60 μL aliquots and stored at -80°C until further use.

For the transformation, an aliquot was thawed on ice and mixed with DNA. It was transferred into an electroporation cuvette with a 2 mm gap between electrodes and subjected to a discharge at 2 kV, with a typical time constants of 4–5 ms using the MicroPulser™ (Bio-Rad). Cells were dissolved in 3 mL mTGP and recovered at 52°C for 2 hours at 200 rpm. Afterwards they were spun down and seeded on selective agar media plates. Transformation efficiencies typically were 10^{1–10^2} colonies per microgram of DNA.

sfGFP measurement

The sfGFP [34] was used as a reporter to assess the expression levels. It was previously shown to be active in *Geobacillus* species [27]. For quantification of sfGFP expression driven by different promoters and RBS's, *Geobacillus* strains carrying the respective constructs were grown overnight at 60°C in TMM with 0.05% yeast extract and 0.2% glucose. 2 μL of these cultures were inoculated into 100 μL of fresh pre-heated media in flat-bottom 96-well microtiter plate (Greiner Bio-One) and sealed airtight with VIEWSeal (In Vitro) to prevent water evaporation. Plates were incubated at 60°C and 200 rpm. Periodically fluorescence was measured with the ELx808™ Microplate Reader (BioTek) with the excitation at 485 nm and emission at 535 nm. Values at the middle of log phase were taken for analysis. Fluorescence was normalized to OD_{600} measured at the same time.

Xylanase assay

Xylanase activity was measured with EnzChek® Ultra Xylanase Assay Kit (Life Technologies) according to manufacturer’s instructions. Briefly, cells were grown for 21 hours, reaching similar densities, harvested and lysed with CellLytic™ B Plus Kit (Sigma-Aldrich). Cell lysate and supernatant from cultures were diluted and 50 μL of dilutions were mixed with 50 μL of xylanase substrate working solution in flat-bottom 96-well microtiter plate (Greiner Bio-One). They were incubated at room temperature for 40 min and the release of reaction products was measured with the ELx808™ Microplate Reader (BioTek) with the excitation at 360 nm and emission at 460 nm. Total protein content was measured with Novagen® BCA Protein Assay Kit (Merck) and xylanase activity was normalized to it.

Results

Growth of *Geobacillus* strains on various carbon sources

In order to assess biotechnological potential of *Geobacillus* spp. we analyzed the ability of four strains to utilize a number of carbon sources: *G. thermoglucosidasius* 2542^T which was previously used in metabolic engineering of isobutanol production [16]. *G. thermoglucosidasius* M10EXG is a natural isolate with high tolerance to ethanol and which is thus a promising host for its production [31]. *G. thermoglucosidasius* CS6-YS93 is another strain of the same species
which genome is sequenced and annotated [32]. G. stearothermophilus NUB3621 has been used in a number of applications [35] and its genome has been also recently sequenced [27]. Bacterial cultures were grown in minimal medium (TMM) supplemented with a number of different carbon sources. These included sugars (glucose, xylose, arabinose) and more complex carbohydrates (cellobiose and xylan), which constitute a major part of lignocellulosic biomass. Glycerol and acetate were also included in the screening because they are cheap and suitable as industrial carbon sources. Most strains utilized a number of investigated carbon sources, but showed poor growth on xylan (Fig 1). Of these, G. thermoglucosidasius C56-YS93 had the highest growth yields on glucose combined with good growth on glycerol, acetate and cellobiose. In addition, its genome has been sequenced and annotated and is available online [32], which makes it easier to design and manipulate genetic changes in this strain. Therefore, we chose it for further studies.

Promoter library

In order to facilitate effective metabolic engineering strategies in Geobacillus, it is desirable to have access to a number of promoters with different strengths. A library of semi-synthetic promoters was therefore constructed using a method described by Jensen and Hammer [36]. It includes the randomization of promoter regions between -35 and -10 elements, while leaving these elements intact, as a way to vary promoter strength. This method has been used to construct promoter libraries for E. coli [37], Lactococcus lactis [36], and Saccharomyces cerevisiae [38]. Its advantages include the ease of library construction and gradual increments in strength among the resulting promoters [37].

Here we created a library of synthetic promoters for Geobacillus spp, based on the native and strong promoter of the groESL operon from Geobacillus sp. GHH01 (locus tag GHH_c02820, RefSeq GHH_RS01420). Its regulatory CIRCE sequence [39]; [40] was deleted and the sequences between and around its -35 (TTGCAA) and -10 (TAATAT) elements were randomized using a degenerate oligonucleotide sequence (PNJ388 in Table 1). The ribosome binding site (RBS) was left intact. Fusion of these constructs with sfGFP produced a library that was transformed into G. thermoglucosidasius C56-YS93. To evaluate the strength of the different promoters, superfolder GFP (sfGFP) fluorescence was measured at the middle of log phase. Low transformation efficiency of G. thermoglucosidasius limited the library to 17 constructs,
which nevertheless covered a 76-fold range of expression levels (Fig 2). Two promoters in the library exhibited higher expression when compared to the groESL promoter, while two gave comparable expression levels as to that of the native, and the rest were weaker.

**RBS library**

Modulation of translation initiation is often used as a tool to regulate the level of protein production. An array of ribosome binding sites (RBS’s) was therefore constructed using the RBS Calculator [41]; [42]. This software calculates the thermodynamics of interactions between the ribosome and the mRNA. Based on this model, it generates an RBS sequence with a given theoretical translation initiation rate. The model takes into account not only the Shine-Dalgarno sequence, but also sequences flanking it. Since the consensus sequence of bacterial RBS consists of six nucleotides, it is problematic to use RBS Calculator to compare its strength to that of the resulting RBS.

![Fig 2. Library of semi-synthetic promoters.](image)
In an alternative randomization approach, Bonde et al. [43] constructed a comprehensive library of almost all possible permutations of six nucleotides acting as RBS (the consensus sequence in *E. coli* being AGGAGG) and studied their effect on protein expression. We hypothesized that for a thermophilic organism like *Geobacillus* sp. it is worthwhile using ratio-nal RBS design because several of its strains are available in the database of RBS Calculator. RBS libraries described by Bonde et al. were created for *E. coli* and might not work in a different genetic context (and at different temperature) of Gram positive bacteria.

A set of RBS’s with a range of different predicted translation initiation rates was created and fused with the promoter $P_{pfl}$ of the pyruvate-formate lyase gene (*pfl*) of *G. thermoglucosidasius* C56-YS93 (locus tag Geoth_3895, RefSeq GEOTH_RS19245), where the native RBS was replaced by a synthetic one. sfGFP was again used as a reporter for the screening. Two of the tested RBS’s showed low expression levels, while the rest resulted in middle to high expression levels (Fig 3).

**Inducible promoter**

Inducible promoters are valuable tools for various applications in molecular biology, because they enable the modulation of gene expression as a function of the concentration of the induc-ting factor. Here we investigated a xylose-inducible promoter of the xylose isomerase gene (*xylA*), because its homologues in *Bacillus* species have been extensively studied [44]; [45] and

---

**Fig 3. Library of synthetic RBS sequences.** The library was constructed using the RBS Calculator and fused to the $P_{pfl}$ promoter of *G. thermoglucosidasius* C56-YS93 and sfGFP. (A) Sequences of RBS (grey background) and their strengths (standard deviations in parentheses) as measured by sfGFP fluorescence. (B) RBS activity compared to the predicted translation initiation rate. RBS’s on the horizontal axis of the graph appear in the same order as in the table.

doi:10.1371/journal.pone.0171313.g003
used for protein production [46]. The operator sequence of \textit{xylA} gene in \textit{G. thermoglucosidasiu s} (5’-TTAGTTTATATGATAGACAAAC-3’) shares 73% similarity with that of \textit{B. subtilis}.

The promoter from the \textit{G. thermoglucosidasius} C56-YS93 xylose isomerase (\textit{xylA}, locus tag Geoth_2243) was examined by fusing a 160 bp region immediately upstream from the \textit{xylA} gene to a gene encoding sfGFP on a plasmid. The expression of sfGFP was measured for cells exposed to a range of xylose concentrations from 0 to 0.5\% (w/v) with either 0.5\% (w/v) glucose or 0.5\% (w/v) glycerol as a main carbon source. For the glycerol medium, a step-wise increase in sfGFP expression was observed as a function of increasing xylose concentration, while the level of induction was less pronounced when glucose was present in the medium (Fig 4). The dynamic range of expression also varied significantly, where 2-fold difference was observed in glucose medium compared to 6.5-fold when cells were grown on glycerol medium. The basal expression from the non-induced promoter in glucose medium was lower when compared to the one with glycerol.

A considerable basal expression from the uninduced \textit{P}_{xylA} was observed for both carbon sources. We hypothesized that it might be due to a repressor protein being titrated out by multiple copies of the extrachromosomal \textit{P}_{xylA}-sfGFP construct. Based on the homology of \textit{xylA} and its operator to those in \textit{B. subtilis} [45], the regulation mechanism of \textit{xylA} expression may likely be similar in \textit{G. thermoglucosidasius} as it is in \textit{B. subtilis}, where XylR is a repressor of \textit{xylA} gene expression [44]. Hence, in order to make a tighter promoter system, we expressed a putative \textit{xylR} gene (Geoth_1256) with its native promoter and terminator on the same plasmid. This resulted in a decrease in basal sfGFP expression, although some expression still remained (Fig 4). At zero or low inducer concentrations, additional copies of \textit{xylR} decreased sfGFP expression. However, the effect was reversed at higher concentrations (Fig 4). Under these conditions overexpression of XylR surprisingly resulted in higher expression from \textit{P}_{xylA}. 

![Fig 4. Inducible protein expression. Expression of Superfolder GFP is controlled by different concentrations of xylose, when its gene is expressed from the inducible promoter \textit{P}_{xylA}. The main carbon source was 0.5\% (w/v) glucose or 0.5\% (w/v) glycerol as indicated. Constructs further carrying the regulator gene \textit{xylR} on the same plasmid are depicted by dark color. doi:10.1371/journal.pone.0171313.g004](image-url)
The sfGFP expression levels in pIP26 (xylR + P\textsubscript{xylA}:sfGFP) differed almost 12-fold between uninduced and fully induced conditions when cells were grown in medium containing glycerol as a carbon source.

**Xylanase production using the P\textsubscript{xylA} expression system**

Many *Geobacillus* species possess a conserved cluster of about 200 kb within a genome containing the genes for xylan utilization, notably a number of xylanases [19]. Xylanases are widely used in paper mill industry, animal feed processing and bakery, and the use of thermostable enzymes is also advantageous in certain fields. Therefore, we sought to use the strain and tools characterized above to overexpress two enzymes: a endo-1,4-β-xylanase native to *G. thermoglucosidasius* C56-YS93 (locus tag Geoth\_2264, RefSeq GEOTH\_RS11140) and the xylanase T-6 encoded by *xynA* in *G. thermodenitrificans* NG80-2 (locus tag GTNG\_1761, RefSeq GTNG\_RS09220) as relevant models for homologous and heterologous protein expression. In addition, xylanase T-6 has a putative N-terminal 28-amino acid signal peptide (MLKRSKAIIVGFSFMLLLPLGMTNALA) predicted by SignalP 4.1 server [47] that potentially enables it to be secreted from the cell. Endo-1,4-β-xylanase (Geoth\_2264) lacks a signal peptide.

To demonstrate the applicability of the inducible P\textsubscript{xylA} for protein expression, two xylanases were put under control of this promoter and expressed in the presence of the inducer (xylose). As shown in Fig 5, most of the xylanase T-6 activity (70%) was observed in the supernatant, indicating that it was secreted from the cell. Thus, in this case, a signal peptide from one species (*G. thermodenitrificans* NG80-2) was active in the other (*G. thermoglucosidasius* C56-YS93). The endo-xylanase was also successfully overexpressed and showed relatively high intracellular activity.

![Fig 5. Application of induction system for expression of xylanases. Heterologous (GTNG\_1761) and homologous (Geoth\_2264) proteins overexpression in *G. thermoglucosidasius* under control of the P\textsubscript{xylA} promoter. Basal xylanase activity in wild type strain is shown in blue, activities of overexpressed enzymes in red; dark and light colors correspond to extracellular and intracellular activities, respectively.](image-url)

doi:10.1371/journal.pone.0171313.g005
Discussion

In this study we generated a set of tools for gene expression in *G. thermoglucosidasius* and characterized their use for homologous and heterologous protein production.

A library of ribosome binding sites was developed using the RBS Calculator [41]. It was previously shown that a computational model based on the thermodynamics of RNA binding to ribosome does not always accurately predict the actual translation efficiency [48]. Factors other than the strength of the Shine-Dalgarno sequence might play a role [49]; [50]. In this study the sfGFP expression levels generally correlated with predicted translation initiation rates, except for one outlier. One of the factors that may have influenced the accuracy of prediction in this study was the default settings of the RBS Calculator v1.1. In this version the default temperature is 37˚C and could not be adjusted for growth optimum of 60˚C of the thermophilic *G. thermodenitrificans* NG80-2, which was used as a model. Although the number of RBS sequences tested (six) may be too low to make general statements on predictability of the RBS Calculator for *Geobacillus* species, the library is large enough for practical purposes of controlling gene expression, as it covers a relatively wide range of translation efficiencies. Future research may use the RBS’ designed here in the context of the bicistronic architecture [51] to improve precision of protein biosynthesis, especially in cases of difficult-to-express proteins.

An inducible promoter of the *xylA* gene studied here showed a 12-fold dynamic range between uninduced and fully induced states, while at the same time demonstrating a significant basal activity. It is desirable for an inducible promoter to be tightly regulated, which means that it should have very low level of expression when not induced. In *B. subtilis*, the active repressor protein (XylR) binds to its motif in the promoter region. Additionally, the *xylA* gene is also negatively regulated by catabolite repression by CcpA in the presence of glucose-6-phosphate, where the cis acting element is the catabolite responsive element (CRE), a 14-bp sequence within the upper part of *xylA* [52]. In addition, glucose-6-phosphate can act on the activity of XylR itself [52]. However, CRE is absent in *xylA* gene in *G. thermoglucosidasius*, although the gene product is highly homologous (75% amino acid identity) to that of XylA from *B. subtilis*. Therefore, we could not use CRE to decrease the leakiness of *xylA* promoter (e.g. by fusing it to the heterologous gene). Importing the catabolite repression system from *B. subtilis* is hindered by its possibly lower thermostability. A homologue of the *B. subtilis* ccpA gene is present in *G. thermoglucosidasius* C56-YS93 genome (Geoth_0851). However, to the best of our knowledge, its target sequence is currently unknown.

Additional copies of the putative *xylR* gene did on the other hand reduce basal expression from *P*<sub>xylA</sub>. The repression by XylR was significantly more pronounced in the presence of glucose, which is in agreement with *B. subtilis* model. However, at higher concentrations of the inducer xylose, the presence of additional XylR resulted in increased *P*<sub>xylA</sub> activity, i.e. its repressor activity was reversed. This may be due to an unknown mechanism of XylR-mediated regulation in *Geobacillus* spp., so that at zero or low xylose concentrations XylR acts as a repressor, while at high concentrations in becomes an activator. Similar cases of such dual repressors/activators are known in some bacteria, as for example the Cra protein [53] and AraC regulator [54] in *E. coli*.

One possible way to decrease basal expression from an inducible promoter is to subject it to directed evolution. It involves applying error-prone PCR to a parent promoter in order to generate a library of promoters with random mutations. This library can then be screened for desirable properties. Apart from tighter promoter versions, a number of other useful properties could be searched for. These could include wider dynamic ranges, sensitivity (the rate at which induction increases with inducer), etc. [55].
Another possible candidate for an inducible system is the promoter of \textit{araD} gene. AraD is a part of arabinose utilization system and in \textit{B. subtilis} its expression and the expression of other genes in the same operon is induced by arabinose. It is controlled by the regulation protein AraR which binds to the operator sequence and acts as a repressor [56]. In the presence of arabinose it releases from DNA which makes the transcription possible. The arabinose utilization operons with regulatory and structural genes including \textit{araR} and \textit{araD} were characterized in at least one species of \textit{Geobacillus} [57]. We also found that putative \textit{araD} with a respective operator sequence (5'–ATTGTACGTACAA–3’) and \textit{araR} are present in \textit{G. thermodenitrificans} NG80-2 and \textit{G. kaustophilus} HTA426. Future work will be needed to characterize this and other inducible promoter systems in \textit{Geobacillus} strains.

Apart from the inducible \textit{xylA} promoter, a library of 17 constitutive promoters was created and quantified in this study. Importantly, the dynamic range of the inducible \textit{P_{xylA}} falls within the expression range of the library. This feature might find an application e.g. in cases where it is necessary to find an optimal expression level of a certain gene. It might be carried out by varying the activity of the inducible promoter, and afterwards placing the respective gene under the constitutive promoter of comparable strength.

This study provides a toolkit for controlled gene expression in \textit{G. thermoglucosidasius}. Since there is a growing interest in \textit{Geobacillus} spp. in both academia and industry, these tools would be valuable instruments for a number of different applications.

**Author contributions**

- **Conceptualization:** IP CBJ ATN.
- **Data curation:** IP CBJ.
- **Formal analysis:** IP CBJ.
- **Funding acquisition:** ATN.
- **Investigation:** IP.
- **Project administration:** ATN.
- **Supervision:** ATN.
- **Visualization:** IP.
- **Writing – original draft:** IP.
- **Writing – review & editing:** IP CBJ ATN.

**References**

1. Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, et al. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science. 1988; 239: 487–491. PMID: 2448875
2. Sarrouh B, Santos TM, Miyoshi A, Dias R, Azevedo V. Up-to-date insight on industrial enzymes applications and global market. J Bioprocess Biotech. 2012; S4: 1–10.
3. Haki GD, Rakshit SK. Developments in industrially important thermostable enzymes: A review. Biore- sour Technol. 2003; 89: 17–34. PMID: 12676497
4. Nazina TN, Tourova TP, Poltaraus AB, Novikova E V, Grigoryan AA, Ivanova AE, et al. Taxonomic study of aerobic thermophilic bacilli: descriptions of \textit{Geobacillus} subterraneus gen. nov., sp. nov. and \textit{Geobacillus} uzenensis sp. nov. from petroleum reservoirs and transfer of \textit{Bacillus} stearothermophilus, \textit{Bacillus} thermo- catenulatus, \textit{Bacillus}. Int J Syst Evol Microbiol. 2001; 51: 433–446. doi: 10.1099/00207719-51-2-433 PMID: 11321089
5. Chen XG, Stabnikova O, Tay JH, Wang JY, Tay STL. Thermoadaptive extracellular proteases of Geobacillus caldoproteolyticus, sp. nov., from sewage sludge. Extremophiles. 2004; 8: 489–498. doi: 10.1007/s00792-004-0412-5 PMID: 15322950

6. Mok SC, Teh AH, Saito JA, Najimudin N, Alam M. Crystal structure of a compact α-amylase from Geobacillus thermodenaturans. Enzyme Microb Technol. Elsevier Inc.; 2013; 53: 46–54.

7. Leow TC, Rahman RNZRA, Basri M, Salleh AB. A thermoadaptive lipase of Geobacillus sp. T1. Extremophiles. 2007; 11: 527–535. doi: 10.1007/s00792-007-0069-y PMID: 17426920

8. Mechaly A, Teplitsky A, Belakhov V, Baasov T, Shoham G, Shoham Y. Overproduction and characterization of selenomethionine xylanase T-6. J Biotechnol. 2000; 78: 83–86. PMID: 10702913

9. Bosma EF, Van Der Oost J, De Vos WM, Van Kranenburg R. Sustainable production of bio-based chemicals by extremophiles. Curr Biotechnol. 2013; 2: 360–379.

10. Taylor MP, Eley KL, Martin S, Tuffin MI, Burton SG, Cowan DA. Thermophilic ethanologenesis: future prospects for second-generation bioethanol production. Trends Biotechnol. 2009; 27: 398–405. doi: 10.1016/j.tibtech.2009.03.006 PMID: 19481826

11. Wittlich P, Themann A, Vorlop KD. Conversion of glycerol to 1,3-propanediol by a newly isolated thermophilic strain. Biotechnol Lett. 2001; 23: 463–466.

12. Chung D, Cha M, Guss AM, Westpheling J. Direct conversion of plant biomass to ethanol by engineered Caldicellulosiruptor bescii. Proc Natl Acad Sci U S A. 2014; 111: 8931–6. doi: 10.1073/pnas.1402210111 PMID: 24889625

13. Clark ME, Batty JD, van Buuren CB, Dew DW, Eamon MA. Biotechnology in minerals processing: Technological breakthroughs creating value. Hydrometallurgy. 2006; 83: 3–9.

14. Cripps RE, Eley K, Leak DJ, Rudd B, Taylor M, Todd M, et al. Metabolic engineering of Geobacillus thermodenaturans for high yield ethanol production. Metab Eng. Elsevier; 2009; 11: 399–408.

15. Van Zyl LJ, Taylor MP, Eley K, Tuffin M, Cowan DA. Engineering pyruvate decarboxylase-mediated ethanol production in the thermophilic host Geobacillus thermodenaturans. Appl Microbiol Biot. 2014; 98: 1247–1259.

16. Lin PP, Rabe KS, Takasumi JL, Kadisch M, Arnold FH, Liao JC. Isobutanol production at elevated temperatures in thermophilic Geobacillus thermodenaturans. Metab Eng. Elsevier; 2014; 24: 1–8.

17. Suzuki H, Yoshida KI, Ohshima T. Polysaccharide-degrading thermophiles generated by heterologous gene expression in Geobacillus kaustophilus HTA426. Appl Environ Microbiol. 2013; 79: 5151–5158.

18. Kanaanvičiute R, Citavičius D. Genetic engineering of Geobacillus spp. J Microbiol Methods. 2015; 111: 31–39. doi: 10.1016/j.mimet.2015.02.002 PMID: 25659824

19. Brumm PJ, De Maayer P, Mead DA, Cowan DA. Genomic analysis of six new Geobacillus strains reveals highly conserved carbohydrate degradation architectures and strategies. Front Microbiol. 2015; 6: 1–15.

20. Martin C, Puls J, Saake B, Schreiber a. Effect of Glycerol Pretreatment on Component Recovery and Enzymatic Hydrolysis of Sugarcane Bagasse. Cellul Chem Technol. 2011; 45: 487–494. Available: <Go to ISI>://WOS:000299143700008

21. Rastogi G, Bhalla A, Ashikari A, Bischoff KM, Hughes SR, Christopher LP, et al. Characterization of thermostable cellulases produced by Bacillus and Geobacillus strains. Biorevers Technol. Elsevier Ltd; 2010; 101: 8798–8806.

22. Imanaka T, Fuji M, Aramori I, Aiba S. Transformation of Bacillus steaothermophilus with plasmid DNA and characterization of shuttle vector plasmids between Bacillus steaothermophilus and Bacillus subtilis. J Bacteriol. 1982; 149: 824–830. PMID: 6277855

23. Narumi I, Sawakami K, Nakamoto S, Nakayama N, Yanagisawa T, Takahashi N, et al. A newly isolated Bacillus steaothermophilus K1041 and its transformation by electroporation. Biotechnol Biotechnol. 1992; 6: 83–86.

24. Zeigler DR. The genus Geobacillus. Bacillus genetic stock center catalog of strains Seventh Edition. 2001. http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Catalog+of+Strains#4

25. Taylor MP, Esteban CD, Leak DJ. Development of a versatile shuttle vector for gene expression in Bacillus spp. Plasmid. 2008; 60: 45–52. doi: 10.1016/j.plasmid.2008.04.001 PMID: 18501964

26. Liao H, McKenzie T, Hageman R. Isolation of a thermostable enzyme variant by cloning and selection in a thermophile. Proc Natl Acad Sci U S A. 1986; 83: 576–80. PMID: 3003740

27. Blanchard K, Robic S, Matsumura I. Transformable facultative thermophile Geobacillus steaothermophilus NUB3621 as a host strain for metabolic engineering. Appl Microbiol Biot. 2014; 98: 6715–6723.

28. Ingram LO, Jarboe LR, Zhang X, Wang X, Moore JC, Shannagum KT. Metabolic engineering for production of biorenewable fuels and chemicals: Contributions of synthetic biology. J Biomed Biotechnol. 2010;2010.
29. Makrides SC, Makrides SC. Strategies for Achieving High-Level Expression of Genes in Escherichia coli. Microbiol Rev. 1996; 60: 512–538. PMID: 8840785

30. Sambrook J, Fritsch EF, Maniatis T. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press. 1989. p. 626.

31. Fong JCN, Svenson CJ, Nakasugi K, Leong CTC, Bowman JP, Chen B, et al. Isolation and characterization of two novel ethanol-tolerant facultative-anaerobic thermophilic bacteria strains from waste compost. Extremophiles. 2006; 10: 363–372. doi: 10.1007/s00792-006-0507-2 PMID: 16532362

32. Brumm P, Land ML, Hauser LJ, Jeffries CD, Chang Y-J, Mead DA. Complete genome sequences of Geobacillus thermoglucosidasius C56-YS93, a novel biomass degrader isolated from obsidian hot spring in Yellowstone National Park. Stand Genomic Sci. Standards in Genomic Sciences; 2015; 10: 73. doi: 10.1186/s40793-015-0031-z PMID: 26442136

33. Cavaleiro AM, Kim SH, Seppäla S, Nielsen MT, Narholm MHH. Accurate DNA Assembly and Genome Engineering with Optimized Uracil Excision Cloning. ACS Synth Biol. 2015; 4: 1042–1046. doi: 10.1021/acssynbio.5b00113 PMID: 26263045

34. Pe´ delacq J-D, Cabantous S, Tran T, Terwilliger TC, Waldo GS. Engineeri ng and characteriz ation of a superfolde r green fluorescen t protein. Nat Biotechn ol. 2006; 24: 79–88. doi: 10.1038/nbt.1172 PMID: 16369541

35. Pe´ delacq J-D, Cabantous S, Tran T, Terwilliger TC, Waldo GS. Engineeri ng and characteriz ation of a superfolde r green fluorescen t protein. Nat Biotechn ol. 2006; 24: 79–88. doi: 10.1038/nbt.1172 PMID: 16369541

36. Couhiago R, Chen S, Shamoy Y. In Vivo Molecular Evolution Reveals Biophysical Origins of Organismal Fitness. Mol Cell. 2006; 22: 441–499. doi: 10.1016/j.molcel.2006.04.012 PMID: 16713575

37. Jensen PR, Hammer K. The sequence of spacers between the consensus sequences modulates the strength of prokaryotic promoters. Appl Env Microbiol. 1998; 64: 82–87.

38. Braatsch S, Helmark S, Kranz H, Koebmann B, Jensen PR. Rapid fine tuning of Eschericia coli gene expression. Biotechniques. 2008; 45: 1–4.

39. Nevoigt E, Kohinke J, Fischer CR, Alper H, Stahl U, Stephanopoulos G. Engineering of promoter replacement cassettes for fine-tuning of gene expression in Saccharomyces cerevisiae. Appl Environ Microbiol. 2006; 72: 5266–5273. doi: 10.1128/AEM.00530-06 PMID: 16885275

40. Zuber U, Schumann W. CIRCE, a novel heat-shoc k element involved in regulation of heat-shock operon dnaK of Bacillus subtilis. J Bacteriol. 1994; 176: 1359–1363. Available: <Go to ISI://A1994MY35100020 PMID: 8113175

41. Salis HM, Mirsky EA, Voigt CA. Automated design of synthetic ribosome binding sites to control protein expression. Nat Biotechnol. 2009; 27: 946–950. doi: 10.1038/nbt.1568 PMID: 19801975

42. Espah Borujeni A, Channarasappa AS, Salis HM. Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. Nucleic Acids Res. 2014; 42: 2646–2659. doi: 10.1093/nar/gkt1139 PMID: 24234441

43. Bonde MT, Pedersen M, Klausen MS, Jensen SI, Wulff T, Harrison S, et al. Predictable tuning of protein expression in bacteria. Nat Methods. 2016; 13: 233–236. doi: 10.1038/nmeth.3727 PMID: 26752768

44. Gu Y, Ding Y, Ren C, Sun Z, Rodionov Da, Zhang W, et al. Reconstruction of xylose utilization pathway and regulons in Firmicutes. BMC Genomics. 2010; 11: 255. doi: 10.1186/1471-2164-11-255 PMID: 20406496

45. Stammen S, Müller BK, Korneli C, Biedendieck R, Gamer M, Franco-Lara E, et al. High-yield intra- And extracellular protein production using bacillus megaterium. Appl Environ Microbiol. 2010; 76: 4037–4046. doi: 10.1128/AEM.00431-10 PMID: 20435764

46. Petersen TN, Bruton B, Korneli C, Biedendieck R, Gamer M, Franco-Lara E, et al. High-yield intra- And extracellular protein production using bacillus megaterium. Appl Environ Microbiol. 2010; 76: 4037–4046. doi: 10.1128/AEM.00431-10 PMID: 20435764

47. Ploeser B, Bragoli A, Brinkmann V. Discriminating signal peptides from transmembrane regions. Nat Methods. Nature Publishing Group; 2011; 8: 785–786.

48. Li GW, Burkhardt D, Gross C, Weissman JS. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. Cell. Elsevier; 2014; 157: 624–635.

49. Nomura M, Gourse R, Baughman G. Regulation of the synthesis of ribosomes and ribosomal components. Annu Rev Biochem. 1984; 53: 75–117. doi: 10.1146/annurev.bi.53.070184.000451 PMID: 6206785

50. Mccarthy JEG, Gualerzt C. Translational control of prokaryotic gene expression. Trends Genet. 1990; 6: 78–85. PMID: 2183416
51. Mutalik VK, Guimaraes JC, Cambray G, Lam C, Christoffersen MJ, Mai Q-A, et al. Precise and reliable gene expression via standard transcription and translation initiation elements. Nat Methods. 2013; 10: 354–360. doi: 10.1038/nmeth.2404 PMID: 23474465

52. Kraus A, Hueck C, Gartner D, Hillen W. Catabolite repression of the Bacillus subtilis xyl operon involves a cis element functional in the context of an unrelated sequence, and glucose exerts additional xylR-dependent repression. J Bacteriol. 1994; 176: 1738–1745. PMID: 8132469

53. Saier MH, Ramseier TM. The Catabolite Repressor / Activator (Cra) Protein of Enteric Bacteria. Microbiology. 1996; 178: 3411–3417. Available: http://view.ncbi.nlm.nih.gov/pubmed/8655535

54. Martin RG, Rosner JL. The AraC transcriptional activators. Curr Opin Microbiol. 2001; 4: 132–137. PMID: 11282467

55. Tyo KEJ, Nevoigt E, Stephanopoulos G. Directed evolution of promoters and tandem gene arrays for customizing RNA synthesis rates and regulation [Internet]. 1st ed. Methods in Enzymology. Elsevier Inc.; 2011.

56. Mota LJ, Tavares P, Sá-Nogueira I. Mode of action of AraR, the key regulator of L-arabinose metabolism in Bacillus subtilis. Mol Microbiol. 1999; 33: 476–489. PMID: 10417639

57. Shulami S, Raz-Pasteur A, Tabachnikov O, Gilead-Gropper S, Shner I, Shoham Y. The L-arabinan utilization system of Geobacillus stearothermophilus. J Bacteriol. 2011; 193: 2838–2850. doi: 10.1128/JB.00222-11 PMID: 21460081