Applied Microbiology and Biotechnology (2025) 109:163
https://doi.org/10.1007/s00253-025-13508-z

MINI-REVIEW

The engineering toolbox of *Parageobacillus thermoglucosidasius*

Marie Millgaard[1] · Belinda Escher[1] · Ivan Pogrebnyakov[1] · Alex Toftgaard Nielsen[1]

Received: 16 December 2024 / Revised: 28 April 2025 / Accepted: 29 April 2025 / Published online: 9 July 2025

Abstract

The thermophilic bacterium *Parageobacillus thermoglucosidasius* is considered a promising host for industrial microbial production. However, its genetic engineering toolbox is still under development. The focus of this review is to provide an organised summary of the currently available resources to facilitate the use of this organism. The article offers a comprehensive overview of regulatory parts used for the construction of genetic circuits and plasmids, including promoters, ribosome binding sites, terminators, antibiotic resistance genes, reporter genes, and other genetic elements of interest. Additionally, it examines the developments in *P. thermoglucosidasius* vector designs over the years. Here, vectors were categorised either as replicative vectors intended for gene expression, or integrative vectors intended for genomic engineering. The functionality of each vector was described, and their contributions to the progress of molecular tools available for *P. thermoglucosidasius* were evaluated. The review also summarises recent advancements in CRISPR-based systems relevant to this organism. Finally, this review discusses potential improvements that could further contribute to the expanding engineering toolbox of *P. thermoglucosidasius*, paving the way for more advanced applications.

Key points

- *The existing engineering toolbox for P. thermoglucosidasius contributes to the growing interest in using it as a thermophilic production host.*
- *The engineering toolbox for P. thermoglucosidasius has expansion potential in genetic circuit parts and CRISPR-based tools.*
- *Thermophilic hosts like P. thermoglucosidasius are in need of more thermostable reporter genes and thermostable selection markers.*

Keywords *Parageobacillus thermoglucosidasius* · Genetic engineering · Thermophilic plasmid vectors · Genome integration methods · Genetic circuit parts

Introduction

Parageobacillus thermoglucosidasius is a thermophilic Gram-positive bacterium that has been gaining traction as a platform for microbial chemical production (Paredes-Barrada et al. 2024). In contrast to the petrochemical-based production strategies currently employed by the chemical industry, biobased production potentially presents a more sustainable approach where chemicals and materials can be synthesised from renewable feedstocks, such as lignocellulosic biomass (Krüger et al. 2018; Ewing et al. 2022). While industrial research often favours well-characterised mesophilic model organisms like *Escherichia coli* and *Bacillus subtilis*, interest in thermophilic host species has grown due to the many advantages associated with high-temperature fermentation, such as a reduced risk of contamination from mesophilic species and reduced cooling costs of large-scale fermentations (Zeldes et al. 2015; Calero and Nikel 2019; Paredes-Barrada et al. 2024). With an optimal growth temperature of around 60 °C, *P. thermoglucosidasius* has proven itself to be an appealing candidate for the development of moderately thermophilic productions (Paredes-Barrada et al. 2024). One key feature that distinguishes *P. thermoglucosidasius* from many other thermophiles is its receptiveness to genetic engineering approaches. While the engineering

Marie Millgaard and Belinda Escher shared first authorship.

✉ Alex Toftgaard Nielsen
atn@biosustain.dtu.dk

[1] The Novo Nordisk Foundation Center for Biosustainability, Building 220, Kemitorvet, 2800 Kgs, Lyngby, Denmark

toolbox of *P. thermoglucosidasius* is not as well developed as those of conventional mesophilic model organisms, the tools currently available have allowed for the development of *P. thermoglucosidasius* strains for the production of both native and non-native compounds including ethanol (Cripps et al. 2009; Zhou et al. 2016), τ-muurolol (Styles et al. 2021), riboflavin (Yang et al. 2020), and 2,3-butanediol (Zhou et al. 2020; Sheng et al. 2023). However, to establish *P. thermoglucosidasius* as a viable candidate for industrial-scale production, a more robust and comprehensive engineering toolbox is required. To facilitate future developments, the following review provides a summary of engineering tools and methodologies currently available for *P. thermoglucosidasius* alongside a discussion of desirable advancements.

Parts for genetic circuits and components for vector construction

The smallest elements in an engineering toolbox are regulatory parts like promoters, ribosome binding sites (RBSs), and terminators, which are used to build genetic circuits. These parts allow for the regulation of gene expression on both a transcriptional and translational level. This can be achieved through the induced or constitutive expression of mRNA, the modulation of mRNA levels through promoter strength, termination of transcription, or the modulation of protein amounts through RBS strength. It is desirable to be able to choose from a variety of parts, as it allows for fine-tuning of gene expression, or the activation of gene expression under the right conditions. To express genetic circuits, vectors are a commonly used platform. In turn, these vectors carry components that facilitate replication and selection. The following section presents a comprehensive review of all regulatory parts and vector components commonly applied in *P. thermoglucosidasius*. In addition, all reviewed parts and components have been listed in Tables 1 and 2, respectively.

Regulatory parts

Constitutive promoters initiate gene transcription in a stable and continuous manner, independent of environmental conditions. In *P. thermoglucosidasius,* the most widely used promoter is derived from the lactate dehydrogenase *ldh* and termed P_{ldh} (Thompson et al. 2008; Cripps et al. 2009; Bartosiak-Jentys et al. 2012, 2013; Lin et al. 2014; Zhou et al. 2020; Lau et al. 2021). While P_{ldh} is often used as a constitutive promoter, it should be noted that its activity is influenced by oxygen availability (Bartosiak-Jentys et al. 2012; Reeve 2016), and it can therefore be argued that it should not be considered constitutive. However, a variety of other constitutive promoters of natural or semi-synthetic origin have been characterised. Historically, individual constitutive promoters were imported from other Gram-positive bacteria to express individual genes in *P. thermoglucosidasius,* most notably from *Geobacillus stearothermophilus* and *Bacillus cereus* (Thompson et al. 2008; Cripps et al. 2009; Bartosiak-Jentys et al. 2012, 2013). One study tested a range of *G. thermodenitrificans* and *P. thermoglucosidasius*-derived constitutive promoters at 50 °C by expressing LacZ in *P. thermoglucosidasius*, but found all other tested promoters to be lower in activity compared to P_{ldh} (Lin et al. 2014). In addition, a study by Zhou et al. followed a similar approach, testing green fluorescent protein (GFP) expression with promoters from *P. thermoglucosidasius*, *B. subtilis*, and *G. thermodenitrificans*, but again found P_{ldh} to have the highest activity (Zhou et al. 2020). Similarly, Lau et al. tested promoters native to *P. thermoglucosidasius* and *G. thermodenitrificans* at 52 °C, where they instead found P_{gapD} to be the strongest (Lau et al. 2021). Studies testing semi-synthetic promoter libraries based on P_{rplS} (Reeve et al. 2016), or P_{groES} (Jensen et al. 2017; Pogrebnyakov et al. 2017), and a library of bioinformatically identified promoters with their respective RBSs (Gilman et al. 2019) expanded on constitutive promoters available for engineering in *P. thermoglucosidasius*. Most of the semi-synthetic promoters were characterised for their activity via superfolder GFP (sfGFP) expression, either at 55 °C (Reeve et al. 2016) or 60 °C (Pogrebnyakov et al. 2017; Gilman et al. 2019). Gilman et al. also tested all promoters in their study for mOrange expression (Gilman et al. 2019). On the other hand, Jensen et al. characterised their promoters through the expression of a thermostable β-galactosidase (Jensen et al. 2017).

While there is an expansive list of constitutive promoters characterised for *P. thermoglucosidasius,* only a handful of inducible promoters have currently been demonstrated. Most are induced through the presence of sugars, including cellobiose (Bartosiak-Jentys et al. 2013), xylose (Pogrebnyakov et al. 2017; Wang et al. 2022, 2024; Yang et al. 2023), and maltose (Kurashiki et al. 2020; Styles et al. 2021).

RBSs control gene expression at the translational level. They are 5'-untranslated regions on the mRNA upstream of the start-codon and initiate translation through ribosome recruiting. The first studies overexpressing genes in *P. thermoglucosidasius* often relied on native RBSs (Thompson et al. 2008; Cripps et al. 2009). However, later studies by Reeve et al., Pogrebnyakov et al., and Lau et al. together tested a total of 25 RBSs exhibiting a broad range of activities, from low to high (Reeve et al. 2016; Pogrebnyakov et al. 2017; Lau et al. 2021). Here, Pogrebnyakov et al. and Reeve et al. used the RBS Calculator (Salis et al. 2009) for the construction of synthetic RBSs. In both studies, the calculator was found to successfully generate RBS libraries with a varied range of activities, though the predictions generally

Table 1 Regulatory parts for *P. thermoglucosidasius* genetic circuits. Species abbreviations: Bce — *B. cereus*; Bsu — *B. subtilis*; Cbo — *C. botulinum*; Cdi — *C. difficile*; Cpa — *C. pasteurianum*; Eco — *E. coli*; Gst — *G. stearothermophilus*; Gsu — *G. subterraneus*; Gka — *G. kaustophilus*; Gth — *G. thermodenitrificans*; Lac — *L. acidophilus*; Lla — *L. lactis*; Pth — *P. thermoglucosidasius*; Tsa — *T. saccharolyticum*

Constitutive promoters

Part	Description	Origin	Source
P_{lct}	Derived from *lctA*	Gst NCA1503	(Thompson et al. 2008)
P_{ldh}	Derived from different *ldh* variants, influenced by oxygen levels	1) Gst NCA1503 2) Gst Strain 10 3) Gth DSM465 4) Gth 94 A1 5) Pth DSM2542 6) Gth	1) (Cripps et al. 2009) 2) (Cripps et al. 2009) 3) (Lin et al. 2014) 4) (Zhou et al. 2020) 5) (Zhou et al. 2020) 6) (Lau et al. 2021)
P_{pfl}	Derived from *pfl*, influenced by oxygen levels	Bce ATCC14579	(Cripps et al. 2009)
P_{Up2n38}	Based on native promoter from Pth NCIMB 11955	Synthetic	(Bartosiak-Jentys et al. 2013)
P [a]	Collection of 10 characterised promoters derived from various genes	Pth DSM2542, Gth DSM465, Bsu	(Lin et al. 2014)
p_{rplS} [a]	Library based on native promoter of *rplS* (Pth C56-YS93)	Synthetic	(Reeve et al. 2016)
p_{groES} [a]	Libraries based on native promoter of the *groES* (*Geobacillus* sp. GHH01)	Synthetic	(Jensen et al. 2017; Pogrebnyakov et al. 2017)
P_{pta}	Derived from *pta*	Pth DSM2542	(Frenzel et al. 2018)
P [a]	Library based on putative promoters identified from various species	Synthetic	(Gilman et al. 2019)
P [a]	Collection of 10 characterised promoters derived from various genes	Pth DSM2542, Gth 94 A1, Bsu 168	(Zhou et al. 2020)
P [a]	Collection of 8 characterised promoters derived from various genes	Pth NCIMB11955, Gth	(Lau et al. 2021)
P [a]	Library of 80 native 100 bp long promoters	Pth, Gth, Gst, Gka	(Caro-Astorga et al. 2025)

Inducible promoters

Part	Description	Origin	Source
$P_{ßglu}$	Cellobiose-inducible, derived from cellobiose-specific PTS operon	Pth NCIMB 11955	(Bartosiak-Jentys et al. 2013)
P_{xylA}	Xylose-inducible, derived from *xylA*	Pth C56-YS93	(Pogrebnyakov et al. 2017)
P_{gk704} [b]	Maltose-inducible, derived from potential starch-utilisation operon	Gka HTA426	(Kurashiki et al. 2020)
P_{glv}	Maltose-inducible, derived from *glvA*	Pth NCIMB 11955	(Styles et al. 2021)
$P_{(xylose)}$ [a]	Collection of 20 xylose-inducible promoters (9 of which are also glucose-insensitive), derived from various xylose-induced genes or operons	Pth DSM2542	(Wang et al. 2022)
IExyl*	Xylose-inducible expression system, based on the *xylA* promoter and regulator *xylR*	Pth DSM2542	(Wang et al. 2024)

Ribosome binding sites

Part	Description	Origin	Source
RBS [a]	Library designed using Salis Lab RBS calculator	Synthetic	(Salis et al. 2009; Reeve et al. 2016)
RBS [a]	Library designed using Salis Lab RBS calculator	Synthetic	(Salis et al. 2009; Pogrebnyakov et al. 2017)
RBS_{pheB}	Derived from *pheB* gene	Gst DSMZ 6285	(Reeve et al. 2016; Sheng et al. 2023)
RBS [a]	Collection of 15 characterised RBSs	Cbo, synthetic	(Lau et al. 2021)

Table 1 (continued)

Terminators			
Part	**Description**	**Origin**	**Source**
T_{rho1}	rho-independent terminator	pUCG18	(Reeve et al. 2016)
T_{rho2}	rho-independent terminator		(Bartosiak-Jentys et al. 2013; Reeve et al. 2016)
T_{S718}	Derived from native gene	Gth NG80	(Gilman et al. 2019)
T [a]	Collection of 9 rho-independent terminators from various species	Lla, Lac, Cdi, Eco, Cpa, Bsu, Bsu Φ29 phage	(Lau et al. 2021)
$T_{EcoT1\ T2}$	Double terminator derived from *rrnB*	Eco	(Sheng et al. 2023)
T [a]	Collection of 9 rho-independent terminators from various species	Gth, Gst, Gsu, Pth	(Caro-Astorga et al. 2025)
Other parts			
Part	**Description**	**Origin**	**Source**
SP_{GtXyl}	Secretion signal	Pth C56-YS93	(Bartosiak-Jentys et al. 2013)
Sec/SPI $SP_{WP\_062678531.1}$	Secretion signal	Pth DSM2542	(Caro-Astorga et al. 2025)
Tsac_2588 Ribo	Purine-responsive riboswitch	Tsa	(Marcano-Velazquez et al. 2019)
Tsac_2584 Ribo	Purine responsive riboswitch	Tsa	(Marcano-Velazquez et al. 2019)
RiboJ	Ribozyme, insulator sequence		(Lou et al. 2012; Wang et al. 2024; Caro-Astorga et al. 2025)
RbxE/F/G/K	Theophylline-responsive riboswitches	Synthetic	(Lau et al. 2024)

[a]Libraries or collections

[b]Induction with maltose not demonstrated in *P. thermoglucosidasius*

lacked precision in *P. thermoglucosidasius* (Reeve et al. 2016; Pogrebnyakov et al. 2017). However, the RBS calculator (Salis et al. 2009) has since been updated, and may therefore now prove more reliable. In contrast, Lau et al. screened 14 RBSs isolated from *Clostridium botulinum,* alongside one synthetically designed RBS (Lau et al. 2021).

Terminators are important for accurate gene expression. They prevent RNA Polymerase read-through, maintain proper mRNA length, and influence expression levels. Historically, terminators have been poorly described in *P. thermoglucosidasius* making it hard to understand how gene expression was terminated. Studies by Lau et al. and Caro-Astorga et al. are the first to have tested a variety of rho-independent terminators in this organism (Lau et al. 2021; Caro-Astorga et al. 2025). Most of these terminators were derived from other Gram-positive organisms, though even the *rrnB* terminator of *E. coli* has been demonstrated to sufficiently terminate transcription in *P. thermoglucosidasius* (Lau et al. 2021).

In addition to more well-established regulatory parts several riboswitches, a genetic insulator, and two signal peptides have been successfully applied in *P. thermoglucosidasius*. Riboswitches are mRNA structures that can bind a ligand causing a conformational change that allows or inhibits translation. Marcano-Velazquez et al. found two riboswitches sensitive to purine-derivatives to be functional in *P. thermoglucosidasius* (Marcano-Velazquez et al. 2019). A study by Lau et al. also tested a set of synthetic theophylline-responsive riboswitches for induced expression of stCas93 (Lau et al. 2024). The genetic insulator RiboJ is a self-cleaving ribozyme that has been reported to prevent interference from adjacent genetic elements in *E. coli* (Lou et al. 2012). Wang et al. were the first to incorporate RiboJ in their *P. thermoglucosidasius* vectors (Wang et al. 2024), while Caro-Astorga et al. also recently employed RiboJ as a part of their *Geobacillus* toolkit (Caro-Astorga et al. 2025). However, neither study concluded on its efficiency. Finally, secretion signals can be used to export expressed proteins to the extracellular space. To date, two different secretion signals have been tested in *P. thermoglucosidasius*. The first, SP_{GtXyl}, was harnessed to facilitate the secretion of endoglucanases into the culture medium (Bartosiak-Jentys et al. 2013). The second, Sec/SPI $SP_{WP\_062678531.1}$, similarly facilitated the export of mOrange and GnaA (Caro-Astorga et al. 2025).

Vector components

Antibiotic resistance genes are a crucial component in vector design as they enable simple selection of transformants and support plasmid retention within the transformed cells. When working with *P. thermoglucosidasius*, a major challenge is the relatively low number of viable thermostable antibiotics and resistance genes. Kanamycin is the most widely utilised antibiotic in *P. thermoglucosidasius* due to its high thermostability and two well-established kanamycin resistance markers (kan^R). The first marker is derived from the *Staphylococcus aureus* pUB110 plasmid and is reported to be stable up to around 60 °C (Cripps et al. 2009; Madika et al. 2022). In turn,

Table 2 Antibiotic resistance genes and reporter genes tested in *P. thermoglucosidasius*

Antibiotic resistance genes			
Name	**Description**	**Conc.**[a]	**Source**
kan (TK101)	Kanamycin resistance, optimised in *G. stearothermophilus*	12.5 µg/mL	(Liao et al. 1986; Taylor et al. 2008)
cat [c]	Chloramphenicol resistance, originates from *S. aureus* pC194	7–15 µg/mL	(Zeigler 2001; Lin et al. 2014)
kan (pUB110)	Kanamycin resistance, originates from *S. aureus* pUB110	12.5 µg/mL	(Cripps et al. 2009)
cat	Chloramphenicol resistance, optimised in *G. kaustophilus*	5–10 µg/mL	(Kobayashi et al. 2015; Reeve et al. 2016)
spc^R	Spectinomycin resistance, originates from pDG1664	7–12.5 µg/mL	(Zhou et al. 2016)
aad9	Spectinomycin resistance	12.5 µg/mL	(Madika et al. 2022)
Reporter genes			
Name	**Description**	**Max Temp.**[b]	**Source**
GFPmut3*	Variant of green fluorescent protein	-	(Andersen et al. 1998; Zhou et al. 2020)
mOrange	Orange fluorescent protein	60 °C	(Shaner et al. 2004; Gilman et al. 2019)
mCherry [c]	Red fluorescent protein	50 °C	(Shaner et al. 2004; Reeve et al. 2016)
sfGFP	Superfolder green fluorescent protein	68 °C	(Pédelacq et al. 2006; Wang et al. 2024)
pheB	Oxygen-independent, produces yellow colour when cleaving catechol	55 °C[d]	(Bartosiak-Jentys et al. 2012)
BgaB	Oxygen-independent thermostable β-galactosidase from *G. stearothermophilus* DSM2027, produces blue colour with X-gal (oxygen), black colour with S-gal (No oxygen)	60 °C	(Lin et al. 2014; Jensen et al. 2017)
sfGFP(N39D/A179A)	Superfolder green fluorescent protein, improved fluorescence intensity	65 °C	(Frenzel et al. 2018)
sfCFP(N39D/A179A)	Superfolder cyan fluorescent protein, derived from sfGFP (N39D/A179 A)	60 °C	(Frenzel et al. 2018)
sfYFP(N39D/A179A)	Superfolder yellow fluorescent protein, derived from sfGFP (N39D/A179 A)	60 °C	(Frenzel et al. 2018)
Dasher GFP	Variant of green fluorescent protein	60 °C	(Gilman et al. 2019)
eCGP123	Thermophilic green fluorescent protein	75 °C[d]	(Lau et al. 2021)

[a]Antibiotic concentrations commonly used in *P. thermoglucosidasius* for each respective resistance gene

[b]The highest temperatures at which each reporter has been demonstrated to function in *P. thermoglucosidasius*

[c]Genes that are generally less thermostable than other options

[d]Activity measured in vitro from *P. thermoglucosidasius* cell extracts

the second marker, TK101, is the most widely used in *P. thermoglucosidasius* engineering and was developed by adapting the pUB110 marker in *G. stearothermophilus*, enabling selective growth at 70 °C (Liao et al. 1986; Liao and Kanikula 1990; Madika et al. 2022). Spectinomycin is another applicable candidate for selection at higher temperatures. Notably, a study by Zhou et al. incorporated a pDG1664-derived spectinomycin resistance marker (spc^R) in their genetic engineering vector pJZ04 and were thereby the first to demonstrate that spectinomycin is viable for antibiotic selection in *P. thermoglucosidasius* (Zhou et al. 2016). In addition, the *P. thermoglucosidasius* pMTL60000 modular vector series introduced and demonstrated the functionality of a spc^R variant *aad9*, though it has yet to be employed in other studies (Madika et al. 2022). Finally, the chloramphenicol resistance marker (cm^R) has been applied for selection in *P. thermoglucosidasius* in multiple studies, often using a cm^R derived from the *S. aureus* plasmid pC194 (Lin et al. 2014; Zhou et al. 2016, 2020; Frenzel et al. 2018). Similarly to TK101, cm^R has been adapted in *Geobacillus kaustophilus* to tolerate up to 60 °C (Kobayashi et al. 2015). Although the resulting variant was included in a *Geobacillus* plasmid set, the use of cm^R is generally less favoured due to its reduced plasmid transformation efficiencies, potentially a result of lower thermostability (Hussein et al. 2015; Kananavičiute and Čitavičius 2015; Reeve et al. 2016). Nevertheless, chloramphenicol resistance still offers a viable selection strategy, particularly for studies that wish to employ multiple selection markers.

Reporter genes code for proteins that can generate a visible or quantifiable signal. Fluorescent proteins (FPs) are a commonly used option because they generate a fast and measurable visual output. It should be noted that FPs are generally oxygen-dependent, and only a few FPs are fluorescent at higher temperatures. Among the FPs tested above 50 °C in *P. thermoglucosidasius* are different versions of GFP, like Dasher GFP (Gilman et al. 2019; Caro-Astorga et al. 2025), GFPmut3* (Andersen et al. 1998; Zhou et al. 2020), enhanced consensus green protein

variant 123 (eCGP123) (Kiss et al. 2009; Don Paul et al. 2011; Lau et al. 2021), and sfGFP (Pédelacq et al. 2006; Frenzel et al. 2018). Notably, Frenzel et al. engineered a variant of sfGFP stable until 65 °C, and developed cyan fluorescent (sfCFP) and yellow fluorescent (sfYFP) variants (Frenzel et al. 2018) both stable at 60 °C. Other FPs tested above 50 °C include mOrange (Shaner et al. 2004; Gilman et al. 2019; Caro-Astorga et al. 2025) and mCherry (Shaner et al. 2004; Reeve et al. 2016). Another category of reporter genes converts molecules into a detectable output. In *P. thermoglucosidasius*, this type of reporter gene has been established in the form of catechol 2,3-dioxygenase (Bartosiak-Jentys et al. 2012) and variants of β-galactosidase (Lin et al. 2014; Jensen et al. 2017).

Finally, origins of replication (*ori*) are an essential vector component that dictate the replication process, the copy number, and overall stability of plasmids. The role of different *ori*s in the plasmid vectors of *P. thermoglucosidasius* will be addressed in the following.

Plasmid vectors used in *P. thermoglucosidasius*

As one of the most important tools in bioengineering, plasmids are utilised for a variety of biotechnological applications — including DNA cloning, gene expression, and genetic engineering. For transformation, the two strategies most commonly used in *P. thermoglucosidasius* include electroporation and conjugation, as exemplified by Taylor et al. (2008) and Tominaga et al. (2016), respectively. While methods exist for protoplast transformation of *G. stearothermophilus*, a protocol has yet to be established for *P. thermoglucosidasius* (Wu and Welker 1989; Kananavičiute and Čitavičius 2015). Since transformation efficiency is a persistent challenge with *P. thermoglucosidasius*, a major focus in vector optimisation is the general reduction of plasmid size to better facilitate transformation, especially with larger cargo sequences.

Typically, plasmid cloning is performed in *E. coli* due to its higher transformation efficiency. For this reason, all vectors reviewed in this article have been designed in a shuttle vector format, carrying both a Gram-positive *ori* for *P. thermoglucosidasius* and most often a high-copy Gram-negative *ori* for *E. coli*, typically derived from ColE1. The choice of Gram-positive *ori* often plays a role in determining vector functionality as it influences the stability of the plasmid within *P. thermoglucosidasius*. Replicative vectors are designed to autonomously replicate in a stable manner allowing them to test factors such as gene expression while maintaining genome integrity. In contrast, integrative vectors are specifically designed to be incorporated into the genome and facilitate genetic modifications. To provide more insight into their functionality, this review broadly divided *P. thermoglucosidasius* plasmid vectors into the two categories. The remainder of this section explores past and present developments for both categories. A summary of all reviewed plasmids can be found in Table 3.

Replicative vectors

Replicative vectors are the favoured tool to test gene expression, protein production, or secretion in *P. thermoglucosidasius*. Usually, the effects of gene overexpression can be assessed faster using replicative vectors as genomic integration of genes can be time-consuming. Furthermore, plasmid-carried genes tend to exhibit higher expression levels because they are present on multiple plasmid copies. Consequently, higher copy numbers reduce the risk of mutations affecting the overall expression of the gene of interest. For successful gene expression in *P. thermoglucosidasius*, two common requirements imposed on replicative vector design are vector stability at higher cultivation temperatures and vector stability over time.

The development of replicative shuttle vectors for *P. thermoglucosidasius* began with an early precursor named pBST22 (Liao and Kanikula 1990). Besides carrying a thermostable kan^R, the pBST22 vector borrows the thermoresistant *ori* from pBST1, a native plasmid occurring in *G. stearothermophilus* (Liao and Kanikula 1990). Plasmids have previously been found to be stably replicated in *G. stearothermophilus* even at 68 °C, when the *ori* of pBST1 (repBST1) is present (Liao et al. 1986). pBST22 was first used as replicative vector in *P. thermoglucosidasius* by Thompson et al. for the expression of a pyruvate decarboxylase (Thompson et al. 2008). Though functional, pBST22 quickly became outdated, most likely due to its large size of 7.6 kb, and the lack of common shuttle vector features, such as a multiple cloning site (MCS). However, most current *P. thermoglucosidasius* replicative vectors still employ the repBST1 and kan^R derived from pBST22.

Another notable early replicative vector is the pNW33N, as released by the Bacillus Genetic Stock Center (Zeigler 2001). Compared to pBST22, this vector features an *ori* from a plasmid native to *Bacillus coagulans* called pBC1, a cm^R, and an MCS derived from the *E. coli* vector pUC19 (Yanisch-Perron et al. 1985; Zeigler 2001). However, this vector has not seen any further development, compared to the efforts put into the line of repBST1-based vectors. Examples of the use of pNW33N as a replicative vector include the stable expression of isobutanol pathway enzymes (Lin et al. 2014) and the overexpression of a pyruvate decarboxylase (Zhou et al. 2016).

The first modern replicative vector specifically developed for *P. thermoglucosidasius* is pUCG18, introduced in a study by Taylor et al. (2008). To construct pUCG18, the repBST1

Table 3 Summary of *P. thermoglucosidasius* vectors of interest

Gram+ *ori*	Vector	Resistance	Features	Size (kb)	Source
Replicative vectors					
repBST1	pBST22	*kan*^{R} (TK101) *cm*^{R} (pC194) *amp*^{R} (pUC19)	-	7.6	(Liao and Kanikula 1990)
	pUCG18	*kan*^{R} (TK101) *amp*^{R} (pUC18)	MCS (pUC18)	6.3	(Taylor et al. 2008)
	pUCG3.8	*kan*^{R} (TK101)	MCS (pUC18)	3.8	(Bartosiak-Jentys et al. 2013)
	pG1K	*kan*^{R} (TK101)	MCS	3.7	(Reeve et al. 2016)
	pG1C	*cm*^{R} (cat)	MCS	3.9	(Reeve et al. 2016)
	pG1AK	*kan*^{R} (TK101) *amp*^{R} (pUC18)	MCS	4.7	(Reeve et al. 2016)
	pS797	*amp*^{R} (pUC18) *kan*^{R} (TK101)	MCS	5.4	(Gilman et al. 2019)
	pGKE119	*kan*^{R} (TK101)	MCS	5.8	(Kurashiki et al. 2020)
	pMTL66000[a]	-	-	-	(Madika et al. 2022)
pBC1	pNW33N	*cm*^{R} (pC194)	MCS (pUC19)	4.2	(Zeigler 2001)
pNCI001	pMTL64000[a]	-	-	-	(Madika et al. 2022)
pNCI002	pMTL65000[a]	-	-	-	(Madika et al. 2022)
pGEOTH02	pMTL67000[a]	-	-	-	(Madika et al. 2022)
Integrative vectors					
pUB110	pTMO31	*kan*^{R} (pUB110) *amp*^{R} (pUC19)	MCS (pUC19)	5.1	(Cripps et al. 2009)
	pJZ04	*spc*^{R} (pDG1644) *amp*^{R} (pUC19)	-	-	(Zhou et al. 2016)
	pG2K	*kan*^{R} (TK101)	MCS	3.8	(Reeve et al. 2016)
	pMTL61110 [a]	*kan*^{R} (pUB110)	MCS (pMTL81551)	4.8	(Sheng et al. 2017)
	pUB-sfGFP	*kan*^{R} (TK101) *amp*^{R} (pUC19)	sfGFP selection	-	(Yang et al. 2020)
	pG2ACT_Int [b]	*cm*^{R} (cat)	Serine integrase *Int1*	-	(Styles et al. 2021)
	pMTL575555	*kan*^{R} (TK101)	stCas91	7.2	(Lau et al. 2021)
	pMTL675555	*kan*^{R} (TK101)	stCas93	8.0	(Lau et al. 2021)
	pMTL-RiboCas93	*kan*^{R} (TK101)	stCas93 inducible by theophylline via P_{gapdh}-RbxE	8.0	(Lau et al. 2024)
	pMTL61000 [a]	-	-	-	(Madika et al. 2022)
	pMM7	*kan*^{R} (TK101)	sfGFP, *spo0 A*^{KO} cassette[c]	6.1	(Millgaard et al. 2023)
	pUB31-*xylR-mCherry-sfgfp*	*kan*^{R} (TK101)	sfGFP, mCherry, *xylR*^{KO} cassette [c]	-	(Wang et al. 2024)
pUB110.2	pMTL-LS5	*kan*^{R} (pUB110)	MCS (pMTL81551), *pyrE* counter-selection	5.2	(Sheng et al. 2017)
	pMTL62000 [a]	-	-	-	(Madika et al. 2022)
pUB110.3	pMTL63000 [a]	-	-	-	(Madika et al. 2022)
repBST1	pUCG3.8Bgl	*kan*^{R} (TK101)	*bgl* counter-selection	-	(Bacon et al. 2017)

[a]Part of the modular pMTL60000 shuttle vector series for *P. thermoglucosidasius* (Sheng et al. 2017; Madika et al. 2022). The following parts are available: Gram + *ori*: ColE1 ± *oriT-traJ*, p15 A ± *oriT-traJ*; Resistance genes; *kan*^{R} (pUB110), *kan*^{R} (TK101), *aad9*; Application-modules: MCS, P_{ldh} + MCS, MCS + sfGFP, MCS + eCGP123

[b]While not designed to integrate, this vector facilitates the genomic integration of other plasmids

[c]Carries cassette for deletion of the specified gene in *P. thermoglucosidasius*

and *kan*^{R} were harnessed from pBST22 (Liao and Kanikula 1990) and combined with an MCS and ampicillin resistance marker (*amp*^{R}) both harnessed from pUC18 (Yanisch-Perron et al. 1985). Advantageously, pUCG18 is notably smaller compared to pBST22. The functionality of pUCG18 has been demonstrated through the overexpression of a pyruvate decarboxylase (Taylor et al. 2008), a β-galactosidase (Bartosiak-Jentys et al. 2012), and the *rib* cluster from *P.*

thermoglucosidasius DSM2542 (Wang et al. 2022). While pUCG18 is stable until 68 °C and selected for with the most heat-stable antibiotic (Taylor et al. 2008), it is still a comparatively large base vector with 6.3 kb, which limits cargo size with respect to transformation efficiency.

Since its development, pUCG18 has served as a starting point for multiple other replicative vectors. One notable derivative of pUCG18 is pS797, which was employed for a larger promoter study by Gilman et al. (2019). Another example is pGKE119 which was employed for protein production in *P. thermoglucosidasius* (Kurashiki et al. 2020). Furthermore, a study by Bartosiak-Jentys et al. introduced an optimised derivative of pUCG18 (Bartosiak-Jentys et al. 2013). Here, they reduced the vector size by removing amp^R, as the thermostable kan^R is also a viable selection marker for *E. coli* (Bartosiak-Jentys et al. 2013). This led to the development of the 3.8 kb vector pUCG3.8 (Bartosiak-Jentys et al. 2013). Derivatives of pUCG3.8 were shown to allow expression of endoglucanases and a catechol 2,3-dioxygenase (Bartosiak-Jentys et al. 2013). Three years later, Reeve et al. expanded on replicative vectors through the development of a modular plasmid set (Reeve et al. 2016). This plasmid set includes a derivative called pG1C with a cm^R, another derivate called pG1K which advances upon previous designs through the incorporation of a terminator downstream of kan^R leading to improved expression, and the replicative vectors pG1AK and pG1AC which additionally feature exchangeable reporter genes and amp^R (Reeve et al. 2016). To date, pG1K is the smallest available replicative vector for *P. thermoglucosidasius* with just 3.7 kb (Reeve et al. 2016), although this vector has not yet been reported for use in plasmid-based gene expression. However, the related vector pG1AK-sfGFP was employed in a study testing the functionality of purine-sensing riboswitches in *P. thermoglucosidasius* (Marcano-Velazquez et al. 2019).

In recent times, Madika et al. introduced the novel modular pMTL60000 vector series for *P. thermoglucosidasius* featuring replicative plasmids (Madika et al. 2022). Most notably, the vector series encompasses three new replicative low copy *ori*s pNCI001, pNCI002, and pGEOTH02, all of which are derived from plasmids native to *P. thermoglucosidasius*. These new *ori*s were confirmed to be compatible with all alternative *ori*s featured in the vector series (Madika et al. 2022). As such, this could allow for the development of multi-plasmid expression systems in *P. thermoglucosidasius*, although this has yet to be tested.

Integrative vectors

Integrative plasmids are currently the primary tool used for targeted genome engineering in *P. thermoglucosidasius*. These plasmids most commonly operate through the execution of double crossover events, as facilitated by homologous recombination (HR) of matching sequences present in both genome and plasmid. Thus, a plasmid carrying a fragment with the target homologous sequences and intended modification can be fully integrated into the target site and subsequently looped out, leaving only the modification behind.

This integration-based engineering approach was pioneered in *P. thermoglucosidasius* through a study by Cripps et al. in which a series of integrative plasmids was employed for the deletion of *ldhA* and *pflB* as well as the insertion of promoter elements for the *pdh* operon (Cripps et al. 2009). In this study, shuttle vectors pUB190, pTMO31, and pTMO19 were developed, from which all integrative plasmids were then constructed. Interestingly, pTMO19 is derived from pNW33N, which is more commonly used as a replicative plasmid. While all vectors were demonstrated to be functional, Cripps et al. considered pTMO31 to be the most optimal platform. It carries a MCS for insertion of recombination fragments, a pUC19-derived Gram-negative *ori* and amp^R for selection in *E. coli*, and a pUB110-derived Gram-positive *ori* and kan^R for *P. thermoglucosidasius*. Core to the integrative strategy is the relatively low thermostability of the pUB110 *ori,* estimated at up to 52 °C, coupled with the comparatively high thermostability of kanamycin selection. Genomic integration of the plasmid is necessary to retain kanamycin resistance at higher temperatures, thereby allowing for selection of cells that underwent the first recombinant event. Plasmid loop out was then promoted through repeated subculturing without kanamycin, and the resulting second recombinants were identified via replica plating, as based on their loss of kanamycin resistance. However, since this second crossover event may either produce a modified strain or a wildtype revertant, PCR was necessary to confirm the modification (Cripps et al. 2009).

Though the approach to genetic engineering of *P. thermoglucosidasius* has seen several developments throughout the years, many current gene editing strategies still largely adhere to the principles of the temperature-based HR strategy established by Cripps et al. (2009). Many advancements are therefore more focused on optimising and diversifying the process to reduce labour and increase success rates. As such, a subsequent addition to the *P. thermoglucosidasius* engineering toolbox was presented when Reeve et al. published their modular *Geobacillus* plasmid set (Reeve et al. 2016). Here, the pG2K vector was included which also carries the pUB110 *ori* and thermostable kan^R enabling integrative functionality in *P. thermoglucosidasius* (Reeve et al. 2016). In addition to its customisable nature, a favourable trait of pG2K is its comparatively smaller size, improving transformation efficiency (Reeve et al. 2016).

Interestingly, an integration system that does not rely on HR has since been developed from the Reeve et al. vectors. This was achieved in a study by Styles et al. which demonstrated a functional serine integrase system by integrating a

sizable τ-muurolol production pathway into the *P. thermoglucosidasius* genome (Styles et al. 2021). To avoid co-transferring the integrase system into the genome, it was placed on a separate pG2K-derived helper plasmid p2ACT_Int which then enabled the integration of the pathway plasmid. The thermosensitive pUB110 *ori* was subsequently used to cure both plasmids and select for successful integrants. As such, this study presents a promising alternative approach to gene modification in *P. thermoglucosidasius* particularly for the integration of larger DNA fragments (Styles et al. 2021). In addition, this emphasises the potential for further development of more versatile serine integrase systems, as known from other species (Merrick et al. 2018).

Many recent advancements have focused on improving isolation of second recombinants, as replica plating and sequence verification are both time-consuming. One strategy is based on the development of counter-selection methods. Here, a notable example is the *bgl* method which takes advantage of the toxic effects generated when ß-galactosidase (Bgl) cleaves 5-bromo-4-chloro-3-indolyl substrates. While first established in *Micrococcus luteus* and *Thermus thermophilus* (Angelov et al. 2013), this method was subsequently implemented in *P. thermoglucosidasius* by Bacon et al. (2017). Here, they incorporated the *bgl* gene as a counter-selection marker into the backbone of their engineering vector pUCG3.8Bgl. As the name indicates, this vector is derived from the replicative vector pUCG3.8 making it the only entry in the "integrative" category that does not carry a thermosensitive *ori*. Instead, Bacon et al. employed a combination of positive selection (kan^R) and negative selection (*bgl*) to isolate the desired secondary recombinants. Notably, this selection strategy enabled the development of an integrative system that effectively eliminates the chance of wildtype reversion, thereby reducing verification time (Bacon et al. 2017).

As an alternative, Sheng et al. implemented a counter-selection method in *P. thermoglucosidasius*, which is instead based on orotate phosphoribosyltransferase (PyrE) activity (Sheng et al. 2017). This study explored *pyrE* for both positive and negative selection, as PyrE activity is required to maintain uracil prototrophy and yet generates a toxic effect when metabolising 5-fluoroorotic acid. The target strain must lack PyrE activity for either selection strategy to function. To develop the *pyrE* system, Sheng et al. first introduced the pMTL61110 vector. They investigated how truncation of the pUB110 *ori* impacted the stability of pMTL61110, resulting in two shortened variants, pUB110.2 and pUB110.3, which both displayed dramatic decreases in thermostability, thereby expediting the isolation of plasmid integrants (Sheng et al. 2017; Madika et al. 2022). The *pyrE* counter-selection vector pMTL-LS5 was generated by incorporating *pyrE* into the backbone of a pUB110.2-carrying plasmid. Using the *pyrE* counter-selection system, Sheng et al. reported an average mutant construction time of 5 days (Sheng et al. 2017).

The study by Sheng et al. served as an early introduction to the pMTL60000 series of *P. thermoglucosidasius* modular vectors, later published by Madika et al. (Sheng et al. 2017; Madika et al. 2022). Here, the pUB110 *ori* and its derivates were found to be incompatible with the pBST1 *ori* and each other, but not with the remaining *ori*s in the series (Madika et al. 2022). Both truncated pUB110 *ori*s are part of the vector series, enabling their use in future integrative vector development (Madika et al. 2022). Notably, the original pMTL61110 plasmid eventually served as the basis for establishing a CRISPR/Cas9-mediated integrative system in *P. thermoglucosidasius* (Lau et al. 2021), as discussed in a later section.

Finally, another avenue of advancement involves the implementation of a visual cue that allows for fast differentiation of first and second recombinants. Here, two studies explored similar approaches in *P. thermoglucosidasius*, by incorporating sfGFP into the backbone of an integrative vector (Yang et al. 2020; Millgaard et al. 2023). Following second recombination, cells without the integrated backbone will lose their green fluorescence, making them easily identifiable without replica plating. The non-fluorescent colonies are subsequently screened to discern mutants from wildtype revertants. The pUB-sfGFP vector of the Yang et al. study is derived from a variant of pTMO31 and carries both kan^R and amp^R (Hills 2014; Yang et al. 2020). In turn, the pMM7 vector of the Millgaard et al. study is derived from pMTL61110 and thus relies solely on kanamycin selection (Millgaard et al. 2023). Finally, a recent study by Wang et al. demonstrated another advancement by incorporating both sfGFP and mCherry in their pUB31-*xylR-mCherry-sfgfp* knockout vector, enabling efficient selection of both first and second recombinants in their HR-based DNA replacement engineering strategy (Wang et al. 2024). These studies successfully demonstrated the efficiency of these types of systems for genetic engineering in *P. thermoglucosidasius*.

Thermostable CRISPR-Cas systems

Clustered regularly interspaced short palindromic repeats (CRISPR) together with CRISPR associated protein (Cas) are known to be one of the bacterial defence mechanisms against foreign DNA. CRISPR-Cas systems consist of one or several Cas enzymes and a CRISPR array, a set of RNA spacers that guide Cas to the protospacer. Combined, the Cas and RNA can introduce a double strand break which renders foreign DNA inactive. A more detailed description of bacterial CRISPR-Cas systems can be found in a review by Arroyo-Olarte et al. (2021). Genome editing methods harnessing CRISPR-Cas achieve selection for mutants

through the expression of guide RNAs (gRNA) targeting the wildtype, thereby allowing for efficient mutation generation. Furthermore, CRISPR-Cas methods also include applications like CRISPR activation and CRISPR interference (CRISPRi), which both involve screening of multi-target libraries by either increasing or inhibiting target gene expression. Therefore, the development of CRISPR-Cas systems for *P. thermoglucosidasius* presents a potentially powerful addition to its engineering toolbox.

A major hurdle in importing established CRISPR-Cas systems to *P. thermoglucosidasius* usually stems from their mesophilic origin. Indeed, a mesophilic CRISPR-Cas system was reported to be inactive at the growth temperatures of *P. thermoglucosidasius* (Mougiakos et al. 2017a). Although thermostable CRISPR-Cas systems were identified and reviewed (Le and Sun 2022), most of them are yet to be tested for their functionality in *P. thermoglucosidasius.* For example, Mougiakos et al. characterised a Cas9 enzyme native to *P. thermodenitrificans* named ThermoCas9 (Mougiakos et al. 2017b). Recently, Paredes-Barrada et al. introduced two gene knockouts in *P. thermoglucosidasius* facilitated by the expression of ThermoCas9 upon induction with cellobiose Paredes-Barrada et al. (2025).

On the other hand, Lau et al. assembled a plasmid-based CRISPR-Cas system by cloning the gene encoding for either Cas9.1 or Cas9.3, both from *Streptococcus thermophilus* (st) (Hao et al. 2018), into pMTL61110 and used it for gene deletion in *P. thermoglucosidasius* (Lau et al. 2021). With their work, Lau et al. demonstrate that when utilising a stCas9.1 or stCas9.3-based CRISPR-Cas system, successful edits can be identified as early as from colonies restreaked post transformation, afterwards only requiring heat-induced plasmid loss to obtain a clean mutant (Lau et al. 2021). Sheng et al. used this system to engineer the strain NCIMB 11955 for acetoin and 2,3-butanediol production (Sheng et al. 2023).

As an alternative to non-native CRISPR-Cas systems, native *P. thermoglucosidasius* systems can be recruited through plasmid-based gRNA expression. Yang et al. reported *P. thermoglucosidasius* to possess a class Ib and a class III CRISPR-Cas system (Yang et al. 2023). They characterised the type Ib system and discovered an inherent editing and silencing dualism determined by spacer length. Briefly, the reported CRISPR-Cas system not only supports genome editing but also displays gene silencing capabilities required for applications like CRISPRi. The authors demonstrate both qualities by engineering the strain NCIMB 11955 for riboflavin production (Yang et al. 2023). Recently, a different study by Yang et al. combined the native CRISPR-Cas Ib system with simultaneous expression of non-homologous end joining enzymes and achieved long-range deletions of random lengths (Yang et al. 2024). Though this set-up is of limited use for precise genome engineering, the strains obtained using this tool could help understand gene essentiality. In addition, Yang et al. report that the studied CRISPR-Cas system has a bias for low GC-content spacers in editing efficiency and provide a computational tool for gRNA selection (Yang et al. 2024). While these accomplishments have added CRISPR-Cas systems to the *P. thermoglucosidasius* engineering toolbox, the exploration of the whole potential of CRISPR-Cas applications is still at an early stage.

Limitations and extensions of the *P. thermoglucosidasius* toolbox

Though advancements have been made, the *P. thermoglucosidasius* toolbox still holds potential for further developments. A major challenge lies in the currently limited availability of genetic elements, including promoters, RBSs, and terminators. While studies such as those conducted by Pogrebnyakov et al. and Gilman et al. have expanded the collection (Pogrebnyakov et al. 2017; Gilman et al. 2019), the toolbox would greatly benefit from a wider variety of well-characterised libraries allowing for greater control over gene expression. A key issue is the lack of available inducible promoters which limits options for choosing inducers and adjusting dynamic ranges. Through screenings, it is likely that more inducible promoters could be identified in the genome of *P. thermoglucosidasius*. Alternatively, it may be possible to adopt inducible systems from other thermophilic species. Potential candidates could for example include the maltose or *myo*-Inositol inducible systems demonstrated in *G. kaustophilus* (Suzuki et al. 2013), or a mannitol inducible promoter employed in *Bacillus methanolicus* (Irla et al. 2016). Here, we would also like to highlight a study by Mol et al. (2022) who worked on a modified TetR-P_{tet} system that allows for inducible expression of mRuby2 and BgaB in *P. thermoglucosidasius* (work in review).

Antibiotic resistance genes currently represent a major bottleneck in *P. thermoglucosidasius* engineering, largely due to their generally poor thermostability. While several potentially thermostable antibiotics have been proposed, only kanamycin, chloramphenicol, and spectinomycin have currently been shown to work in *P. thermoglucosidasius* (Hussein et al. 2015; Kananavičiute and Čitavičius 2015; Drejer et al. 2018). It is possible that more thermostable variants could be adapted from mesophilic resistance genes, as exemplified by the development of kan^R (TK101) and cm^R (*cat*) (Liao et al. 1986; Kobayashi et al. 2015). Notably, thermostable variants of bleomycin and hygromycin resistance genes have already been developed in *T. thermophilus* (Brouns et al. 2005; Nakamura et al. 2005), while a thiostrepton resistance gene similarly was developed in

G. kaustophilus (Wada et al. 2016). Though currently not tested in *P. thermoglucosidasius*, these could prove to be viable for selection. Erythromycin is another thermostable antibiotic that could potentially be used for selection if an appropriate selection marker is found (Gregory et al. 2001). Further exploration of thermophilic genomes may yet reveal more naturally thermostable selective markers. Finally, it should be noted that antibiotic-free selection systems could be another attractive addition to the *P. thermoglucosidasius* toolbox (Kananavičiute and Čitavičius 2015). One example is through the development of auxotrophic strains, as previously demonstrated with the *P. thermoglucosidasius* Δ*pyrE* strain (Sheng et al. 2017). Another option could be harnessing toxin-antitoxin systems native to *P. thermoglucosidasius* for plasmid stabilisation, as recently explored by Lau et al. (2021).

Reporter genes encounter many of the same challenges as antibiotic resistance genes. Although several FPs have been shown to function in *P. thermoglucosidasius*, many share overlapping spectra, and some struggle to fluoresce at higher temperatures. As with resistance genes, viable reporters could potentially be developed from less thermostable counterparts, as seen with sfGFP (Frenzel et al. 2018), or sourced from other thermophilic species. Considering the few anaerobic reporter genes currently available, it may be of interest to adapt ligand-based fluorescent systems to *P. thermoglucosidasius*. Here, potential options include the pFAST system established in the anaerobic thermophile *Thermoanaerobacter kivui* (Hocq et al. 2023), as well as a series of flavin-binding FPs identified in various thermophilic species (Wingen et al. 2017).

From literature, it is apparent that there is a strong tendency for replicative vectors to employ relatively thermostable *ori*s and integrative vectors to employ relatively thermosensitive *ori*s. It should, however, be mentioned that plasmids incorporating a thermosensitive *ori* can still be used for plasmid-based gene expression. This was for example demonstrated by Bashir et al. who used pMTL61110-derived plasmids for overexpression of various genes at 52 °C (Bashir et al. 2019). However, studies seeking stable plasmid-based gene expression are inevitably constrained to lower cultivation temperatures should they choose a thermosensitive vector. Conversely, integrative vectors are not required to carry a thermosensitive *ori*, as exemplified by Bacon et al. utilising counter-selection, rather than temperature, to isolate the desired integrants (Bacon et al. 2017). Yet, many studies still employ thermosensitive vectors due to their well-established and reliable use in literature.

There are still many opportunities to innovate and improve the metabolic engineering methodologies of *P. thermoglucosidasius*. To start, low plasmid transformation efficiencies are a reoccurring challenge, making the optimisation of transformation protocols vital to enable plasmid library screenings, for example. One way to unlock higher transformation efficiencies could be through the investigation of the restriction and methylation systems native to *P. thermoglucosidasius.* Here, Yang et al. have already laid the foundation by deleting genes encoding for native restriction enzymes and achieving higher transformation efficiencies (Yang et al. 2023). We would also like to highlight the work of Heidelbach et al. (2025) who explored the DNA methylation patterns of a variety of species, including *P. thermoglucosidasius* (work under review). A future development could be a strain of *E. coli* that methylates shuttle plasmids in a pattern close to the methylation pattern of *P. thermoglucosidasius,* hopefully reducing plasmid degradation post transformation. The development of techniques for *in vivo* plasmid assembly in *P. thermoglucosidasius* is also desirable, as this would avoid the cloning step in *E. coli* altogether, thus enabling more streamlined vector designs outside the shuttle-vector format.

Although the available genome engineering tools were refined over the last years, there may still be limitations to achieving genomic alterations. Notably, many strategies currently rely solely on HR to facilitate genetic modifications, which in turn may complicate the selection of desired recombinants if the target gene is important to the fitness of the cell. In these cases, continued development and implementation of genetic engineering systems with more stringent selection, such as counter-selection, could help enforce the generation of the desired recombinants. In addition, metabolic characterisation of *P. thermoglucosidasius* could be improved through omics-based research and gene deletion libraries, enabling more precise rational engineering strategies. For heterologous production, the characterisation of novel integration sites across the genome of *P. thermoglucosidasius* would also provide more options for pathway integration. Finally, the implementation of alternative recombineering strategies like Multiplex Automated Genome Engineering (MAGE) (Wannier et al. 2021) or transposon mutagenesis could help avoid method-based limitations. Here, the introduction of integrase- and CRISPR-based systems already heralds the development of more diverse strategies for genetic manipulation. However, library screenings associated to non-native CRISPR-Cas systems might still be limited by transformation efficiency, due to larger plasmid sizes caused by the coding sequence length of the available Cas9 enzymes. A way to circumvent this limitation could be the stable integration of the *cas9* gene in question, which in turn requires careful choice of the integration site for optimal expression. Alternatively, one could opt for the CRISPR Ib system native to *P. thermoglucosidasius*, although this system might be subjected to internal regulation in expression, which currently remains uncharacterised.

While the toolbox of *P. thermoglucosidasius* is still in development, the considerable progress achieved in recent years highlights the growing interest in this thermophilic species. Future research into genetic engineering and strain development could manifest the potential of *P. thermoglucosidasius* as a future platform for more sustainable chemical production.

Author contribution MM, BE, IP and ATN conceived and designed the review. MM and BE conducted a literature study and analysis. MM and BE wrote the manuscript. IP and ATN revised the manuscript. MM, BE, IP and ATN approved the final manuscript.

Funding This work was financially supported by European Union's Horizon 2020 research and innovation programme under grant agreement number 101037009 (PyroCO2). We further acknowledge a grant to DTU Biosustain, with Grant Number NNF20CC0035580. We further acknowledge funding from the Villum Fonden (grant number 40986) and the Danish Research Council (grant number 103200448B). The funding bodies have not been involved in the design, data collection, analysis, and manuscript writing of this article.

Declarations

Ethics approval and consent to participate Not applicable.

Conflict of interest The authors declare no competing interests.

References

Andersen JB, Sternberg C, Poulsen LK, Bjørn SP, Givskov M, Molin S (1998) New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. Appl Environ Microbiol 64:2240–2246. https://doi.org/10.1128/aem.64.6.2240-2246.1998

Angelov A, Li H, Geissler A, Leis B, Liebl W (2013) Toxicity of indoxyl derivative accumulation in bacteria and its use as a new counterselection principle. Syst Appl Microbiol 36:585–592. https://doi.org/10.1016/j.syapm.2013.06.001

Arroyo-Olarte RD, Bravo Rodríguez R, Morales-Ríos E (2021) Genome editing in bacteria: CRISPR-Cas and beyond. Microorganisms 9. https://doi.org/10.3390/microorganisms9040844

Bacon LF, Hamley-Bennett C, Danson MJ, Leak DJ (2017) Development of an efficient technique for gene deletion and allelic exchange in *Geobacillus* spp. Microb Cell Fact 16:58. https://doi.org/10.1186/s12934-017-0670-4

Bartosiak-Jentys J, Eley K, Leak DJ (2012) Application of pheB as a reporter gene for *Geobacillus* spp., enabling qualitative colony screening and quantitative analysis of promoter strength. Appl Environ Microbiol 78:5945–5947. https://doi.org/10.1128/AEM.07944-11

Bartosiak-Jentys J, Hussein AH, Lewis CJ, Leak DJ (2013) Modular system for assessment of glycosyl hydrolase secretion in *Geobacillus thermoglucosidasius*. Microbiol (United Kingdom) 159:1267–1275. https://doi.org/10.1099/mic.0.066332-0

Bashir Z, Sheng L, Anil A, Lali A, Minton NP, Zhang Y (2019) Engineering *Geobacillus thermoglucosidasius* for direct utilisation of holocellulose from wheat straw. Biotechnol Biofuels 12:199. https://doi.org/10.1186/s13068-019-1540-6

Brouns SJJ, Wu H, Akerboom J, Turnbull AP, De Vos WM, Van Der Oost J (2005) Engineering a selectable marker for hyperthermophiles. J Biol Chem 280:11422–11431. https://doi.org/10.1074/jbc.M413623200

Calero P, Nikel PI (2019) Chasing bacterial chassis for metabolic engineering: a perspective review from classical to non-traditional microorganisms. Microb Biotechnol 12:98–124. https://doi.org/10.1111/1751-7915.13292

Caro-Astorga J, Rogan M, Malcı K, Ming H, Debenedictis E, James P, Ellis T (2025) SubtiToolKit: a bioengineering kit for *Bacillus subtilis* and Gram-positive bacteria. Trends Biotechnol 0–23. https://doi.org/10.1016/j.tibtech.2025.02.004

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 11:398–408. https://doi.org/10.1016/j.ymben.2009.08.005

Don Paul C, Traore DAK, Byres E, Rossjohn J, Devenish RJ, Kiss C, Bradbury A, Wilce MCJ, Prescott M (2011) Expression, purification, crystallization and preliminary X-ray analysis of eCGP123, an extremely stable monomeric green fluorescent protein with reversible photoswitching properties. Acta Crystallogr Sect F Struct Biol Cryst Commun 67:1266–1268. https://doi.org/10.1107/S1744309111028156

Drejer EB, Hakvåg S, Irla M, Brautaset T (2018) Genetic tools and techniques for recombinant expression in thermophilic *Bacillaceae*. Microorganisms 6:42. https://doi.org/10.3390/microorganisms6020042

Ewing TA, Nouse N, van Lint M, van Haveren J, Hugenholtz J, van Es DS (2022) Fermentation for the production of biobased chemicals in a circular economy: a perspective for the period 2022–2050. Green Chem 24:6373–6405. https://doi.org/10.1039/d1gc04758b

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

Gilman J, Singleton C, Tennant RK, James P, Howard TP, Lux T, Parker DA, Love J (2019) Rapid, heuristic discovery and design of promoter collections in non-model microbes for industrial applications. ACS Synth Biol 8:1175–1186. https://doi.org/10.1021/acssynbio.9b00061

Gregory ST, Cate JHD, Dahlberg AE (2001) Spontaneous erythromycin resistance mutation in a 23S rRNA gene, *rrlA*, of the extreme thermophile *Thermus thermophilus* IB-21. J Bacteriol 183:4382–4385. https://doi.org/10.1128/JB.183.14.4382-4385.2001

Hao M, Cui Y, Qu X (2018) Analysis of CRISPR-cas system in *Streptococcus thermophilus* and its application.

Heidelbach S, Dall SM, Bøjer JS, Nissen J, van der Maas LNL, Sereika M, Kirkegaard RH, Jensen SI, Kousgaard SJ, Thorlacius-Ussing O, Hose K, Nielsen TD, Albertsen M (2025) Nanomotif: leveraging DNA methylation motifs for genome recovery and host

association of plasmids in metagenomes from complex microbial communities. bioRxiv **[Preprint].** https://doi.org/10.1101/2024.04.29.591623

Hills CA (2014) Acetate metabolism in *Geobacillus thermoglucosidasius* and strain engineering for enhanced bioethanol production. University of Bath

Hocq R, Bottone S, Gautier A, Pflügl S (2023) A fluorescent reporter system for anaerobic thermophiles. Front Bioeng Biotechnol 11:1–12. https://doi.org/10.3389/fbioe.2023.1226889

Hussein AH, Lisowska BK, Leak DJ (2015) The genus *Geobacillus* and their biotechnological potential. Adv Appl Microbiol 92:1–48. https://doi.org/10.1016/bs.aambs.2015.03.001

Irla M, Heggeset TMB, Nærdal I, Paul L, Haugen T, Le SB, Brautaset T, Wendisch VF (2016) Genome-based genetic tool development for *Bacillus methanolicus*: theta-and rolling circle-replicating plasmids for inducible gene expression and application to methanol-based cadaverine production. Front Microbiol 7:1–13. https://doi.org/10.3389/fmicb.2016.01481

Jensen TØ, Pogrebnyakov I, Falkenberg KB, Redl S, Nielsen AT (2017) Application of the thermostable β-galactosidase, BgaB, from *Geobacillus stearothermophilus* as a versatile reporter under anaerobic and aerobic conditions. AMB Express 7. https://doi.org/10.1186/s13568-017-0469-z

Kananavičiute R, Čitavičius D (2015) Genetic engineering of *Geobacillus* spp. J Microbiol Methods 111:31–39. https://doi.org/10.1016/j.mimet.2015.02.002

Kiss C, Temirov J, Chasteen L, Waldo GS, Bradbury ARM (2009) Directed evolution of an extremely stable fluorescent protein. Protein Eng des Sel 22:313–323. https://doi.org/10.1093/protein/gzp006

Kobayashi J, Furukawa M, Ohshiro T, Suzuki H (2015) Thermoadaptation-directed evolution of chloramphenicol acetyltransferase in an error-prone thermophile using improved procedures. Appl Microbiol Biotechnol 99:5563–5572. https://doi.org/10.1007/s00253-015-6522-4

Krüger A, Schäfers C, Schröder C, Antranikian G (2018) Towards a sustainable biobased industry – highlighting the impact of extremophiles. N Biotechnol 40:144–153. https://doi.org/10.1016/j.nbt.2017.05.002

Kurashiki R, Mizuno T, Murata K, Ohshiro T, Suzuki H (2020) A plasmid vector that directs hyperproduction of recombinant proteins in the thermophiles *Geobacillus* species. Extremophiles 24:147–156. https://doi.org/10.1007/s00792-019-01142-3

Lau MSH, Madika A, Zhang Y, Minton NP (2024) *Parageobacillus thermoglucosidasius* strain engineering using a theophylline responsive RiboCas for controlled gene expression. ACS Synth Biol 13:1237–1245. https://doi.org/10.1021/acssynbio.3c00735

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 10:1739–1749. https://doi.org/10.1021/acssynbio.1c00138

Le Y, Sun J (2022) CRISPR/Cas genome editing systems in thermophiles: current status, associated challenges, and future perspectives. Adv Appl Microbiol 118:1–30. https://doi.org/10.1016/bs.aambs.2022.02.001

Liao H, McKenzie T, Hageman R (1986) Isolation of a thermostable enzyme variant by cloning and selection in a thermophile. Proc Natl Acad Sci U S A 83:576–580. https://doi.org/10.1073/pnas.83.3.576

Liao HH, Kanikula AM (1990) Increased efficiency of transformation of *Bacillus stearothermophilus* by a plasmid carrying a thermostable kanamycin resistance marker. Curr Microbiol 21:301–306. https://doi.org/10.1007/BF02092095

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

Lou C, Stanton B, Chen YJ, Munsky B, Voigt CA (2012) Ribozyme-based insulator parts buffer synthetic circuits from genetic context. Nat Biotechnol 30:1137–1142. 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 8:633–640. https://doi.org/10.1021/acssynbio.8b00487

Merrick CA, Zhao J, Rosser SJ (2018) Serine integrases: advancing synthetic biology. ACS Synth Biol 7:299–310. https://doi.org/10.1021/acssynbio.7b00308

Millgaard M, Bidart GN, Pogrebnyakov I, Nielsen AT, Welner DH (2023) An improved integrative GFP-based vector for genetic engineering of *Parageobacillus thermoglucosidasius* facilitates the identification of a key sporulation regulator. AMB Express 13. https://doi.org/10.1186/s13568-023-01544-9

Mol V, Falkenberg KB, De Almeida Will Á, Pogrebnyakov I, Beck C, Skøttrup AL, Nielsen AT, Jensen SI (2022) From empirical to data-driven host selection: a broad-host-range expression platform to facilitate chassis screening. bioRxiv **[Preprint].** https://doi.org/10.1101/2022.08.09.503317

Mougiakos I, Bosma EF, Weenink K, Vossen E, Goijvaerts K, Van Der Oost J, Van Kranenburg R (2017a) Efficient genome editing of a facultative thermophile using mesophilic spCas9. ACS Synth Biol 6:849–861. https://doi.org/10.1021/acssynbio.6b00339

Mougiakos I, Mohanraju P, Bosma EF, Vrouwe V, Finger Bou M, Naduthodi MIS, Gussak A, Brinkman RBL, Van Kranenburg R, Van Der Oost J (2017b) Characterizing a thermostable Cas9 for bacterial genome editing and silencing. Nat Commun 8. https://doi.org/10.1038/s41467-017-01591-4

Nakamura A, Takakura Y, Kobayashi H, Hoshino T (2005) In vivo directed evolution for thermostabilization of *Escherichia coli* hygromycin B phosphotransferase and the use of the gene as a selection marker in the host-vector system of *Thermus thermophilus*. J Biosci Bioeng 100:158–163. https://doi.org/10.1263/jbb.100.158

Paredes-Barrada M, Mathissen A, van der Molen RA, Jiménez-Huesa PJ, Polano ME, Donati S, Abele M, Ludwig C, van Kranenburg R, Claassens NJ (2025) Awakening of the RuMP cycle for partial methylotrophy in the thermophile *Parageobacillus thermoglucosidasius*. Metab Eng 91:145–157. https://doi.org/10.1016/j.ymben.2025.04.002

Paredes-Barrada M, Kopsiaftis P, Claassens NJ, van Kranenburg R (2024) *Parageobacillus thermoglucosidasius* as an emerging thermophilic cell factory. Metab Eng 83:39–51. https://doi.org/10.1016/j.ymben.2024.03.001

Pédelacq JD, Cabantous S, Tran T, Terwilliger TC, Waldo GS (2006) Engineering and characterization of a superfolder green fluorescent protein. Nat Biotechnol 24:79–88. https://doi.org/10.1038/nbt1172

Pogrebnyakov I, Jendresen CB, Nielsen AT (2017) Genetic toolbox for controlled expression of functional proteins in *Geobacillus* spp. PLoS ONE 12:1–15. https://doi.org/10.1371/journal.pone.0171313

Reeve B (2016) *Geobacillus thermoglucosidans* as a thermophile chassis for synthetic biology. Imperial College London

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 5:1342–1347. https://doi.org/10.1021/acssynbio.5b00298

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

Shaner NC, Campbell RE, Steinbach PA, Giepmans BNG, Palmer AE, Tsien RY (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. Nat Biotechnol 22:1567–1572. https://doi.org/10.1038/nbt1037

Sheng L, Kovács K, Winzer K, Zhang Y, Minton NP (2017) Development and implementation of rapid metabolic engineering tools for chemical and fuel production in *Geobacillus thermoglucosidasius* NCIMB 11955. Biotechnol Biofuels 10:5. https://doi.org/10.1186/s13068-016-0692-x

Sheng L, Madika A, Lau MSH, Zhang Y, Minton NP (2023) Metabolic engineering for the production of acetoin and 2,3-butanediol at elevated temperature in *Parageobacillus thermoglucosidasius* NCIMB 11955. Front Bioeng Biotechnol 11. https://doi.org/10.3389/fbioe.2023.1191079

Styles MQ, Nesbitt EA, Hoffmann TD, Queen J, Ortenzi MV, Leak DJ (2021) The heterologous production of terpenes by the thermophile *Parageobacillus thermoglucosidasius* in a consolidated bioprocess using waste bread. Metab Eng 65:146–155. https://doi.org/10.1016/j.ymben.2020.11.005

Suzuki H, Yoshid K, ichi, Ohshima T (2013) Polysaccharide-degrading thermophiles generated by heterologous gene expression in *Geobacillus kaustophilus* HTA426. Appl Environ Microbiol 79:5151–5158. https://doi.org/10.1128/AEM.01506-13

Taylor MP, Esteban CD, Leak DJ (2008) Development of a versatile shuttle vector for gene expression in *Geobacillus* spp. Plasmid 60:45–52. https://doi.org/10.1016/j.plasmid.2008.04.001

Thompson AH, Studholme DJ, Green EM, Leak DJ (2008) Heterologous expression of pyruvate decarboxylase in *Geobacillus thermoglucosidasius*. Biotechnol Lett 30:1359–1365. https://doi.org/10.1007/s10529-008-9698-1

Tominaga Y, Ohshiro T, Suzuki H (2016) Conjugative plasmid transfer from *Escherichia coli* is a versatile approach for genetic transformation of thermophilic *Bacillus* and *Geobacillus* species. Extremophiles 20:375–381. https://doi.org/10.1007/s00792-016-0819-9

Wada K, Kobayashi J, Furukawa M, Doi K, Ohshiro T, Suzuki H (2016) A thiostrepton resistance gene and its mutants serve as selectable markers in *Geobacillus kaustophilus* HTA426. Biosci Biotechnol Biochem 80:368–375. https://doi.org/10.1080/09168451.2015.1079478

Wang J, Li Z, Wang W, Pang S, Yao Y, Yuan F, Wang H, Xu Z, Pan G, Liu Z, Chen Y, Fan K (2022) Dynamic control strategy to produce riboflavin with lignocellulose hydrolysate in the thermophile *Geobacillus thermoglucosidasius*. ACS Synth Biol 11:2163–2174. https://doi.org/10.1021/acssynbio.2c00087

Wang J, Wang W, Chen Y, Liu Z, Ji X, Pan G, Li Z, Fan K (2024) Development of a xylose-inducible and glucose-insensitive expression system for *Parageobacillus thermoglucosidasius*. Appl Microbiol Biotechnol 108:493. https://doi.org/10.1007/s00253-024-13333-w

Wannier TM, Ciaccia PN, Ellington AD, Filsinger GT, Isaacs FJ, Javanmardi K, Jones MA, Kunjapur AM, Nyerges A, Pal C, Schubert MG, Church GM (2021) Recombineering and MAGE.

Wingen M, Jaeger KE, Gensch T, Drepper T (2017) Novel thermostable flavin-binding fluorescent proteins from thermophilic organisms. Photochem Photobiol 93:849–856. https://doi.org/10.1111/php.12740

Wu L, Welker NE (1989) Protoplast transformation of *Bacillus stearothermophilus* NUB36 by plasmid DNA. J Gen Microbiol 135:1315–1324. https://doi.org/10.1099/00221287-135-5-1315

Yang Z, Li B, Bu R, Wang Z, Xin Z, Li Z, Zhang L, Wang W (2024) A highly efficient method for genomic deletion across diverse lengths in thermophilic *Parageobacillus thermoglucosidasius*. Synth Syst Biotechnol 9:658–666. https://doi.org/10.1016/j.synbio.2024.05.009

Yang Z, Li Z, Li B, Bu R, Tan G-Y, Wang Z, Yan H, Xin Z, Zhang G, Li M, Xiang H, Zhang L, Wang W (2023) A thermostable type I-B CRISPR-Cas system for orthogonal and multiplexed genetic engineering. Nat Commun 14:6193. https://doi.org/10.1038/s41467-023-41973-5

Yang Z, Sun Q, Tan G, Zhang Q, Wang Z, Li C, Qi F, Wang W, Zhang L, Li Z (2020) Engineering thermophilic *Geobacillus thermoglucosidasius* for riboflavin production. Microb Biotechnol 14:363–373. https://doi.org/10.1111/1751-7915.13543

Yanisch-Perron C, Vieira J, Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mpl8 and pUC19 vectors. Gene 33:103–119. https://doi.org/10.1016/0378-1119(85)90120-9

Zeigler DR (2001) The Genus Geobacillus introduction and strain catalog. Bacillus Genet Stock Cent 3:1–25

Zeldes BM, Keller MW, Loder AJ, Straub CT, Adams MWW, Kelly RM (2015) Extremely thermophilic microorganisms as metabolic engineering platforms for production of fuels and industrial chemicals. Front Microbiol 6. https://doi.org/10.3389/fmicb.2015.01209

Zhou J, Lian J, Rao CV (2020) Metabolic engineering of *Parageobacillus thermoglucosidasius* for the efficient production of (2R, 3R)-butanediol. Appl Microbiol Biotechnol 104:4303–4311. https://doi.org/10.1007/s00253-020-10553-8

Zhou J, Wu K, Rao CV (2016) Evolutionary engineering of *Geobacillus thermoglucosidasius* for improved ethanol production. Biotechnol Bioeng 113:2156–2167. https://doi.org/10.1002/bit.25983