Biotechnology for Biofuels

RESEARCH

Open Access

Engineering *Geobacillus thermoglucosidasius* for direct utilisation of holocellulose from wheat straw

Zeenat Bashir[1], Lili Sheng[1], Annamma Anil[2], Arvind Lali[2], Nigel P. Minton[1] and Ying Zhang[1*]

Abstract

Background: A consolidated bioprocessing (CBP), where lignocellulose is converted into the desired product(s) in a single fermentative step without the addition of expensive degradative enzymes, represents the ideal solution of renewable routes to chemicals and fuels. Members of the genus *Geobacillus* are able to grow at elevated temperatures and are able to utilise a wide range of oligosaccharides derived from lignocellulose. This makes them ideally suited to the development of CBP.

Results: In this study, we engineered *Geobacillus thermoglucosidasius* NCIMB 11955 to utilise lignocellulosic biomass, in the form of nitric acid/ammonia treated wheat straw to which expensive hydrolytic enzymes had not been added. Two different strains, BZ9 and BZ10, were generated by integrating the *cglT* (β-1,4-glucosidase) gene from *Thermoanaerobacter brockii* into the genome, and localising genes encoding different cellulolytic enzymes on autonomous plasmids. The plasmid of strain BZ10 carried a synthetic cellulosomal operon comprising the *celA* (Endoglucanase A) gene from *Clostridium thermocellum* and *cel6B* (Exoglucanase) from *Thermobifida fusca*; whereas, strain BZ9 contained a plasmid encoding the *celA* (multidomain cellulase) gene from *Caldicellulosiruptor bescii*. All of the genes were successfully expressed, and their encoded products secreted in a functionally active form, as evidenced by their detection in culture supernatants by Western blotting and enzymatic assay. In the case of the *C. bescii* CelA enzyme, this is one of the first times that the heterologous production of this multi-functional enzyme has been achieved in a heterologous host. Both strains (BZ9 and BZ10) exhibited improved growth on pre-treated wheat straw, achieving a higher final OD600 and producing greater numbers of viable cells. To demonstrate that cellulosic ethanol can be produced directly from lignocellulosic biomass by a single organism, we established our consortium of hydrolytic enzymes in a previously engineered ethanologenic *G. thermoglucosidasius* strain, LS242. We observed approximately twofold and 1.6-fold increase in ethanol production in the recombinant *G. thermoglucosidasius* equivalent to BZ9 and BZ10, respectively, compared to *G. thermoglucosidasius* LS242 strain at 24 h of growth.

Conclusion: We engineered *G. thermoglucosidasius* to utilise a real-world lignocellulosic biomass substrate and demonstrated that cellulosic ethanol can be produced directly from lignocellulosic biomass in one step. Direct conversion of biomass into desired products represents a new paradigm for CBP, offering the potential for carbon neutral, cost-effective production of sustainable chemicals and fuels.

Keywords: *Geobacillus thermoglucosidasius*, Biomass, Consolidated bioprocessing (CBP), Glycoside hydrolases, Cellulases, Endo/exoglucanases, β-Glucosidase

*Correspondence: Ying.Zhang@nottingham.ac.uk
[1] Clostridia Research Group, BBSRC/EPSRC Synthetic Biology Research Centre (SBRC), School of Life Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, UK
Full list of author information is available at the end of the article

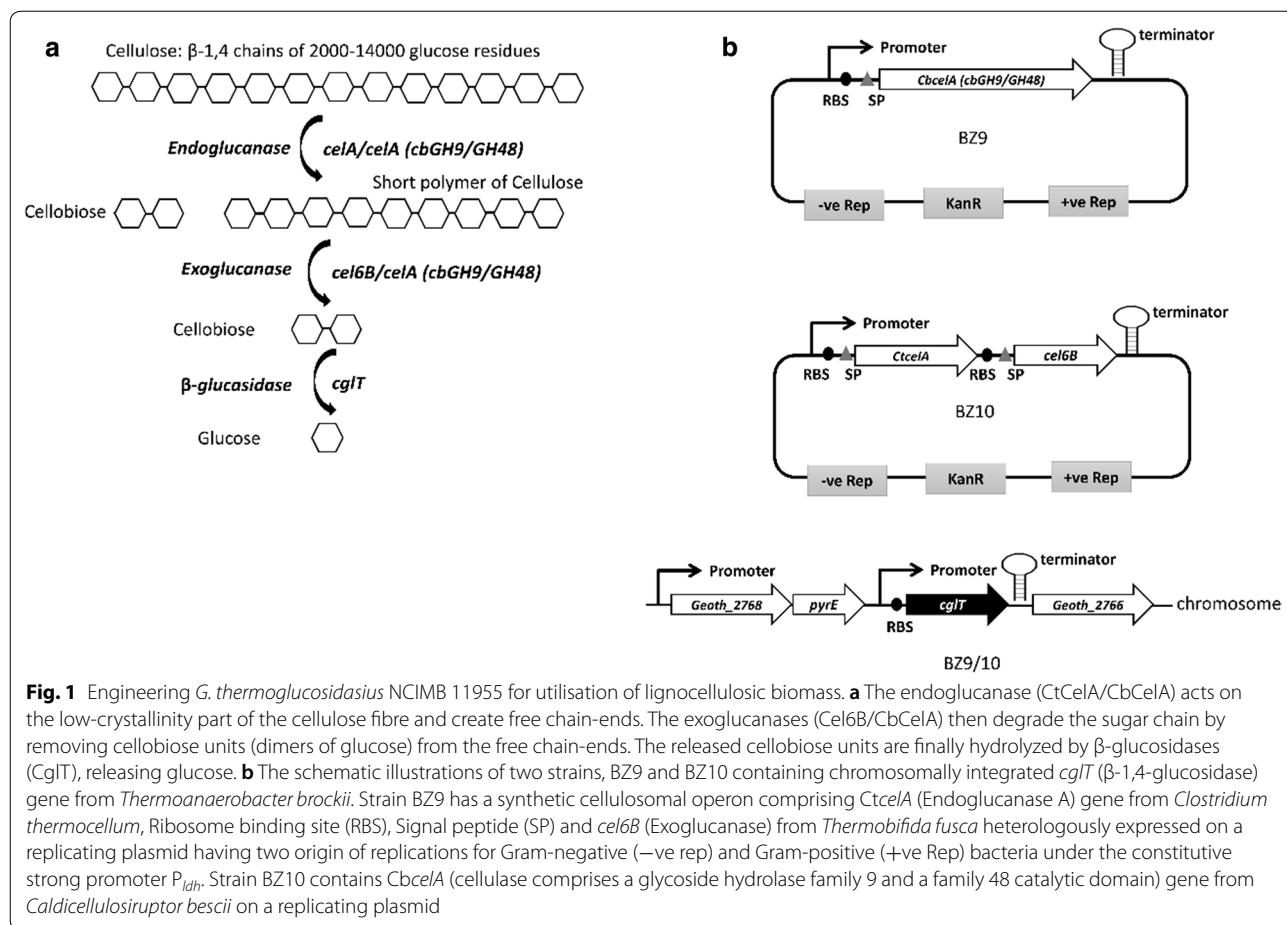Bashir *et al. Biotechnol Biofuels* (2019) 12:199

Page 2 of 16

Background

The sustainable production of chemicals and fuels from lignocellulosic biomass using microbial fermentation requires its deconstruction into simple sugars. This conversion is dependent on an initial pre-treatment step followed by the addition of hydrolytic enzymes. Pre-treatment can be accomplished using a variety of methods such as biological, chemical, mechanical and thermal processes, either alone or in combination [1, 2]. The subsequent complete hydrolysis of the cellulose in pre-treated biomass into glucose, however, requires the synergetic action of at least three types of glycoside hydrolases (GHs) that belong to different sequence-based families, as classified in the Carbohydrate-Active Enzyme database (CAZy) [3]. Endoglucanases hydrolyse internal bonds in the cellulose, generating short polymers which provide the substrate for cellobiohydrolases (or exoglucanases). These act in a unidirectional manner, either from non-reducing or reducing ends of cellulose polysaccharide chains, liberating cellobiose as the major product. Lastly, β-glucosidases convert cellobiose into glucose, relieving the system from end product inhibition [4]. The relative expense of pre-treatment steps, coupled with high cost of hydrolytic enzymes, has made the development of cost-effective strategies to produce chemicals and fuels from biomass very challenging.

One solution to improve the economics is consolidated bioprocessing (CBP), in which lignocellulose is converted into products in one step without added enzymes. In essence, CBP represents the simultaneous enzyme-mediated saccharification and fermentation of biomass into desired products by a single microorganism. Although, no natural microorganism possesses all of the desired features of CBP, certain bacteria and fungi possess at least some of the properties required [5–7]. Of particular interest are thermophilic microorganisms because of the following features: (i) they produce thermostable enzymes that can survive under the required harsh, bioprocessing conditions [8]; (ii) they grow at high temperatures that promotes higher rates of feedstock conversion [9]; (iii) the risk of contamination is eliminated at higher processing temperatures, which also reduce viscosity of substrate and product streams [10, 11], and; (iv) high temperature assists product removal and recovery, potentially lowering separation cost [12]. Among the thermophiles, the anaerobic, cellulolytic organism *Clostridium thermocellum* is one of the most extensively researched CBP organisms and effectively degrades cellulose using a multiprotein complex called the cellulosome. In its wild-type form, *C. thermocellum* is unable to catabolize pentose sugars resulting from the hydrolysis of hemicellulose, thereby reducing the overall biomass conversion. Its low ethanol tolerance also limits the levels of this

biofuel that can be produced [6, 13–15]. Other hemicellulose-utilising thermophilic microorganisms belonging to the genus, *Thermoanaerobacterium* and *Thermoanaerobacter*, have been developed as a CBP organism but are similarly limited in ethanol yields [16–18]. Recently, an extremely cellulolytic, thermophilic organism, *Caldicellulosiruptor bescii*, was engineered to produce an ethanol [19]. Although an important step towards establishing *C. bescii* as a platform for CBP, the high (78 °C) optimum growth temperature of *C. bescii* present a number of challenges to further metabolic engineering the organism for extend product range. These include the existence of a very limited number of antibiotics that are stable above 50 °C, and the lack of available genetic tools [20, 21].

Members of the genus *Geobacillus* are moderate thermophiles (optimum growth temperature is 55–60 °C) that have also been promoted as potential platforms for CBP. They can grow to a high cell density and can utilise a wide range of polymeric or short oligomeric carbohydrates for growth [22]. Importantly, in recent years, genetic tools have been developed that have enabled them to be engineered to improve the production of natural products, such as ethanol [23–26], as well as non-native products, such as isobutanol [27]. However, while most of the *Geobacillus* spp. can effectively degrade hemicellulose, they are unable to break down crystalline or the amorphous cellulose [28], despite their possession of numerous endoglucanases and endoxylanases [29–31]. This has led to the introduction of heterologous cellulase genes [32, 33] but the encoded GHs were only shown to be active against synthetic substrates such as carboxymethyl cellulose (CMC), phosphoric acid-swollen cellulose (PASC) and xylan. The introduction of GHs that are better able to degrade natural cellulosic substrate is essential if *Geobacillus* is to form the basis of CBP.

The standard commercial cellulosic enzyme cocktails for biomass deconstruction contains cellobiohydrolase/exoglucanases I (CBH I), cellobiohydrolase/exoglucanases II (CBH II) and β-glucosidase that act synergistically to release sugars for microbial conversion to products [34]. We, thus, anticipated that *G. thermoglucosidasius* recombinant strains endowed with the ability to produce an enzyme cocktail that could mimic the activity of the above would be potentially able to breakdown hemicellulose/cellulose and support growth of organism. Accordingly, in the present study, *G. thermoglucosidasius* NCIMB 11955 was endowed with the necessary thermostable exoglucanase, endoglucanase and β-glucosidase activities to allow its growth on pre-treated wheat straw that had received no enzymatic treatment (Fig. 1). Having established these defined modifications supported growth on a natural cellulosic substrate, they

Fig. 1 Engineering *G. thermoglucosidasius* NCIMB 11955 for utilisation of lignocellulosic biomass. **a** The endoglucanase (CtCelA/CbCelA) acts on the low-crystallinity part of the cellulose fibre and create free chain-ends. The exoglucanases (Cel6B/CbCelA) then degrade the sugar chain by removing cellobiose units (dimers of glucose) from the free chain-ends. The released cellobiose units are finally hydrolyzed by β-glucosidases (CglT), releasing glucose. **b** The schematic illustrations of two strains, BZ9 and BZ10 containing chromosomally integrated *cglT* (β-1,4-glucosidase) gene from *Thermoanaerobacter brockii*. Strain BZ9 has a synthetic cellulosomal operon comprising Ct*celA* (Endoglucanase A) gene from *Clostridium thermocellum*, Ribosome binding site (RBS), Signal peptide (SP) and *cel6B* (Exoglucanase) from *Thermobifida fusca* heterologously expressed on a replicating plasmid having two origin of replications for Gram-negative (−ve rep) and Gram-positive (+ve Rep) bacteria under the constitutive strong promoter P$_{ldh}$. Strain BZ10 contains Cb*celA* (cellulase comprises a glycoside hydrolase family 9 and a family 48 catalytic domain) gene from *Caldicellulosiruptor bescii* on a replicating plasmid

were introduced into a strain previously engineered for ethanol production. This enabled the one-step production of ethanol from the lignocellulose using single recombinant *G. thermoglucosidasius* strains, representing a significant step towards the generation of a platform for CBP.
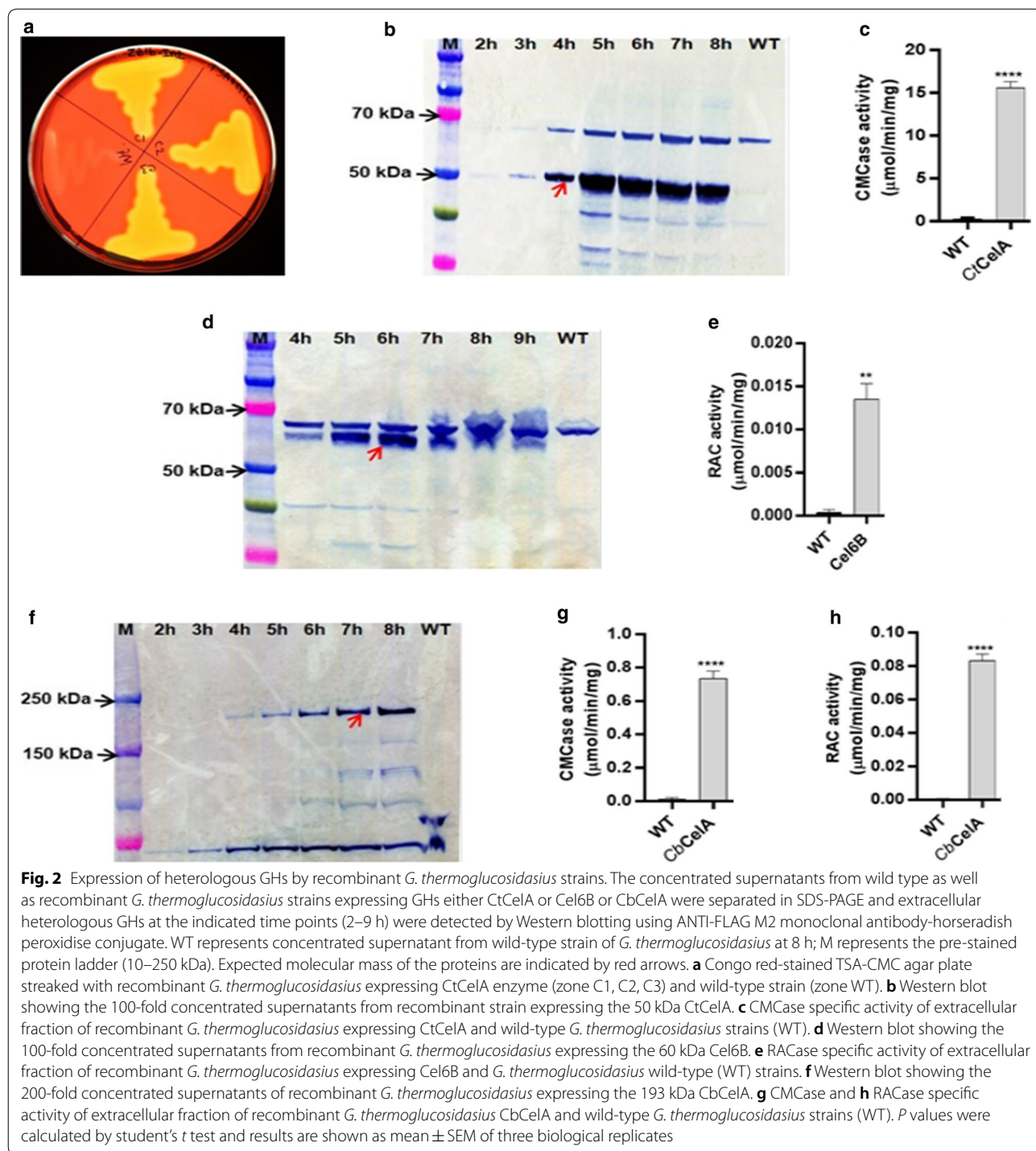
Results

Heterologous expression of extracellular GHs

The codon-optimised Ct*celA, cel6B* and Cb*celA* genes from *C. thermocellum, T. fusca* YX and *C. bescii*, respectively, were cloned into the *G. thermoglucosidasius* expression vector pMTLgSlimS under the transcriptional control of constitutive P$_{ldh}$ promoter, and the resulting plasmids, pMTLgSlimS−Ct*celA*, pMTLgSlimS−*cel6B* and pMTLgSlimS−Cb*celA* were transformed into *G. thermoglucosidasius* NCIMB 11955. The extracellular CtCelA, Cel6B and CbCelA were produced along with their native secretion signals containing a FLAG-epitope at the carboxy-terminus.

To examine CtCelA production and its endoglucanase activity, single recombinant *G. thermoglucosidasius*

colonies, as well as the wild-type strain, were tested for CMC hydrolysis on appropriate agar media. The clear zone around the recombinant strains demonstrated that CMC hydrolysis as a result of the secretion of endoglucanase was occurring. In contrast, parent strain lacked any activity by the same measure (Fig. 2a). To analyze the expression of CtCelA, Cel6B and CbCelA in recombinant *G. thermoglucosidasius* strains, anti-FLAG M2-peroxidase (HRP) antibody antibody was used in Western blots of concentrated culture supernatants collected every hour until the stationary phase of growth (OD600 > 7). The heterologous CtCelA protein at an estimated molecular weight of 50 kDa was clearly visible in the 100-fold concentrated supernatant after 3 h, reaching a maximum in 5 h. No such band was observed in the supernatant derived from the wild-type strain (Fig. 2b). An electrophoretic band that bound to the anti-FLAG antibodies equivalent to a protein of 65 kDa was evident in the supernatants derived from both the recombinant and wild-type strains suggesting that strain NCIMB 11955 contains a polypeptide that also binds to the anti-FLAG-antibody used. A few weak bands of a size smaller than

Bashir *et al. Biotechnol Biofuels* (2019) 12:199

Page 4 of 16

Fig. 2 Expression of heterologous GHs by recombinant *G. thermoglucosidasius* strains. The concentrated supernatants from wild type as well as recombinant *G. thermoglucosidasius* strains expressing GHs either CtCelA or Cel6B or CbCelA were separated in SDS-PAGE and extracellular heterologous GHs at the indicated time points (2–9 h) were detected by Western blotting using ANTI-FLAG M2 monoclonal antibody-horseradish peroxidise conjugate. WT represents concentrated supernatant from wild-type strain of *G. thermoglucosidasius* at 8 h; M represents the pre-stained protein ladder (10–250 kDa). Expected molecular mass of the proteins are indicated by red arrows. **a** Congo red-stained TSA-CMC agar plate streaked with recombinant *G. thermoglucosidasius* expressing CtCelA enzyme (zone C1, C2, C3) and wild-type strain (zone WT). **b** Western blot showing the 100-fold concentrated supernatants from recombinant strain expressing the 50 kDa CtCelA. **c** CMCase specific activity of extracellular fraction of recombinant *G. thermoglucosidasius* expressing CtCelA and wild-type *G. thermoglucosidasius* strains (WT). **d** Western blot showing the 100-fold concentrated supernatants from recombinant *G. thermoglucosidasius* expressing the 60 kDa Cel6B. **e** RACase specific activity of extracellular fraction of recombinant *G. thermoglucosidasius* expressing Cel6B and *G. thermoglucosidasius* wild-type (WT) strains. **f** Western blot showing the 200-fold concentrated supernatants of recombinant *G. thermoglucosidasius* expressing the 193 kDa CbCelA. **g** CMCase and **h** RACase specific activity of extracellular fraction of recombinant *G. thermoglucosidasius* CbCelA and wild-type *G. thermoglucosidasius* strains (WT). *P* values were calculated by student's *t* test and results are shown as mean ± SEM of three biological replicates

50 kDa were also apparent in the recombinant strain but absent in wild type. It is possible that this was due to degradation of CtCelA protein in extracellular medium.

Cell growth is often inhibited because of the metabolic burden imposed by the expression of heterologous protein genes [35]. However, in the present study, the growth profile of the recombinant *G. thermoglucosidasius* strains harbouring plasmid pMTLgSlimS−Ct*celA* was similar to that of wild type (Additional file 1: Figure S1). CMCase/endoglucanase activity of the recombinant *G. thermoglucosidasius* strains expressing CtCelA

Bashir *et al. Biotechnol Biofuels*   (2019) 12:199

Page 5 of 16

was found to be 15.6 ± 1 µmol min^{-1} mg^{-1} of protein confirming that the protein is functional (Fig. 2c).

The presence of recombinant Cel6B protein of an estimated molecular weight of 60 kDa was confirmed from the Western blotting of 100-fold concentrated supernatant derived from recombinant *G. thermoglucosidasius* harbouring plasmid pMTLgSlimS-*cel6B* (Fig. 2d). This observation confirmed the successful expression and secretion of Cel6B. To assess the exoglucanase activity of the heterologously produced Cel6B, enzyme assays using RAC (Regenerated Phosphoric Acid-Swollen Cellulose) as substrate were performed. From this, specific activity of Cel6B was estimated to be 0.0135 ± 0.003 µmol min^{-1} mg^{-1} of protein. No measurable activity was detected in supernatants derived from the wild type (Fig. 2e).

To confirm the full-length expression of CbCelA protein, 200-fold concentrated supernatant was analysed. Using Western Blotting and anti-FLAG antibodies, we detected a protein with an estimated molecular mass of 193 kDa in the supernatant of the recombinant *Geobacillus* culture after 4 h of growth. This corresponded to the estimated molecular weight of CbCelA and confirmed the successful production of this glycosyl hydrolase in a soluble form (Fig. 2f). Although the calculated molecular mass of CbCelA protein is 193 kDa, other studies have found that the native protein from *C. bescii* migrates on SDS-PAGE at an apparent molecular weight of 230 kDa as a consequence of the glycosylation of native enzyme [36, 37]. However, our results showed a band size of 193 kDa indicating the absence of glycosylation in the heterologously expressed CbCelA protein. Endoglucanase and exoglucanase functionality assays of the recombinant enzyme produced in *Geobacillus* indicated that specific activity of CbCelA against CMC and RAC was 0.735 ± 0.07 and 0.0831 ± 0.006 µmol min^{-1} mg^{-1} of protein, respectively, suggesting that CbCelA protein is functional (Fig. 2g, h).

In total, all the three enzymes (CtCelA, Cel6B and CbCelA) were, therefore, produced successfully in their soluble form in the *G. thermoglucosidasius* chassis, with measurable enzyme activity.

Genome integration of the *cglT* gene to improve cellobiose utilisation

The *G. thermoglucosidasius* NCIMB 11955 genome sequence contains an operon encoding a predicted cellobiose specific phosphotransferase system (PTS) followed by a gene annotated as β-1,4-glucosidase indicating that the organism can grow on cellobiose. However, the β-glucosidase operon is tightly regulated by *G. thermoglucosidasius* NCIMB 11955, which leads to low production of enzymes and inefficient degradation of cellobiose.

Accumulation of this disaccharide results in feedback inhibition of many exo- and endo-glucanases [38, 39]. To ensure the maximum rate of cellobiose conversion to glucose, an additional β-glucosidase gene (*cglT* from *T. brockii*) was integrated into the chromosome of *G. thermoglucosidasius* via ACE (Allele-Coupled Exchange) [26]. Integration of the *cglT* at the *pyrE* locus was achieved by transforming plasmid pMTLgSlimS–LS3–*cglT* into electro-competent cells of a Δ*pyrE* deletion mutant of *G. thermoglucosidasius* NCIMB 11955 and subsequentially selecting for uracil prototrophs [26]. Prototrophic colonies were further screened for the presence of the correctly integrated P$_{ldh}$–*cglT* cassette by colony-PCR analysis using oligonucleotide primers pyrE_C1_F and pyrE_C2_R which anneal to the genomic sequences up- and downstream of the *pyrE* locus. Correct integration of the P$_{ldh}$–*cglT* cassette should result in the amplification of a DNA fragment of 3.8 kb in size; while in the case of the parental strain, the same primers generate a 2.0-kb fragment. All clones tested were found to generate the expected 3.8-kb fragment (Additional file 1: Figure S2) confirming the integration of *cglT* gene into the chromosome. The recombinant, integrant strain obtained was designated *G. thermoglucosidasius* ZB3bInt.

To investigate the expression and functionality of intracellular CglT in *G. thermoglucosidasius* ZB3bInt, 80 µg of total protein from a cell-free extract (CFE) was resolved on SDS-PAGE followed by anti-FLAG-tag mediated detection of protein by Western blotting. A distinct protein band estimated to be 50 kDa in size, corresponding to the predicted molecular mass of CglT protein, was observed in supernatants derived from the integrant strain (Fig. 3a). An equivalently size protein was absent in supernatants derived from the wild-type strain. Further, a CFE of *G. thermoglucosidasius* ZB3bInt expressing *cglT* gene showed β-glucosidase activity of 1.21 ± 0.07 µmol min^{-1} mg^{-1} of protein against the colorimetric substrate pNPG (4-nitrophenyl β-D-glucopyranoside) at 60 °C (Fig. 3b).

To determine if the integration of *cglT* gene can enhance cellobiose utilisation, the recombinant *G. thermoglucosidasius* ZB3bInt and wild-type strains were grown at 55 °C in ASYE-modified medium supplemented with 30 g/l of cellobiose (87.7 mM) as the sole carbon source. Cellobiose consumption and growth profiles showed that recombinant *G. thermoglucosidasius* ZB3bInt strains utilised 100% of cellobiose available compared to the wild-type strain which utilised only $45 \pm 0.02\%$ of total cellobiose after 10 h of growth (Fig. 3c). Although the recombinant ZB3bInt strain exhibited a slight lag in growth compared to wild type in the early exponential phase of growth, after 7 h, its cell density increased significantly and had reached a much
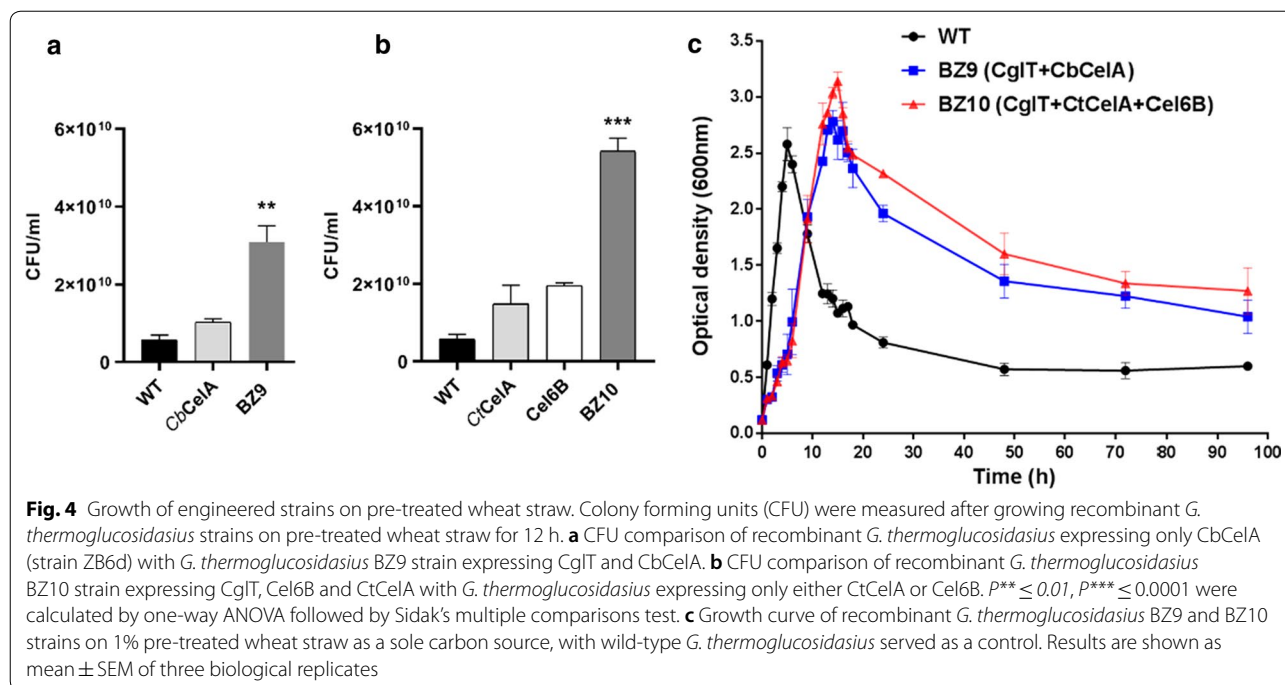
Fig. 3 Intracellular expression of CglT enhanced cellobiose utilisation. **a** Western blot showing FLAG-tagged CglT using ANTI-FLAG M2 monoclonal antibody-horseradish peroxidise conjugate. Lane Wt. and CglT represents soluble fraction of cell lysate from wild-type *G. thermoglucosidasius* and the recombinant *G. thermoglucosidasius* ZB3bInt strain, respectively. Lane M is pre-stained protein ladder (10–250 kDa); Expected molecular mass of 50 kDa for CglT is indicated by red arrow. **b** pNPGase specific activity of recombinant *G. thermoglucosidasius* ZB3bInt strains (ZB3bInt) and wild-type *G. thermoglucosidasius* strains (WT). *P* values were calculated by student's t test and results are shown as mean ± SEM of three biological replicates. **c** Cellobiose consumption and growth profiles by recombinant *G. thermoglucosidasius* ZB3bInt (ZB3bInt) and wild-type *G. thermoglucosidasius* strains (WT) when grown on 3.0% cellobiose as the carbon source at 55 °C. Black circles, *G. thermoglucosidasius*_ZB3bInt (OD600 nm); Red circles, wild-type (OD600 nm); blue square, *G. thermoglucosidasius*_Zb3bInt (remaining cellobiose); green squares, wild-type (remaining cellobiose). Results are shown as mean ± SEM of three biological replicates

higher OD600 at the stationary phase of growth than the wild type (Fig. 3c). These results suggest that the addition of the heterologous *cglT* gene significantly enhanced the ability of *G. thermoglucosidasius* to utilise cellobiose, justifying the use of ZB3bInt as the background strain for the introduction of heterologous exo- and endo-glucanases.

Assembled GHs enable *G. thermoglucosidasius* to grow on pre-treated wheat straw

To develop *G. thermoglucosidasius* chassis expressing all the enzymes encoding endoglucanase, exoglucanase and β-glucosidase for complete degradation of cellulose, pMTLgSlimS–Cb*celA* and pMTLgSlimS–Ct*celA–cel6B* plasmids were individually transformed into electro-competent cells of *G. thermoglucosidasius* ZB3bInt (expressing the integrated *cglT*), thereby generating recombinant *G. thermoglucosidasius* strain BZ9 and BZ10, respectively. The combined effect of GHs for the degradation of pre-treated wheat straw was investigated by analysing the growth of recombinant *G. thermoglucosidasius* BZ9 and BZ10 with the recombinant strains expressing only either CtCelA or Cel6B or CbCelA, along with wild type as a control. Pre-treated wheat straw was prepared in a two-step process using nitric acid followed by ammonia at the DBT-ICT 2G (Department of Biotechnology—Institute of Chemical Technology Second Generation) Ethanol Demonstration Facility and was

kindly gifted by ICT, Mumbai for the study (Additional file 1: Figure S3). It is composed of 85.15% cellulose and 4.35% xylose of the total carbohydrates, as reported in the patent [40]. Cultures of the two recombinant strains BZ9 and BZ10, alongside the wild-type and ZB3bInt controls, prepared in ASYE-modified medium supplemented with 10 g/l pre-treated wheat straw, were grown at 55 °C for approximately 100 h. No growth difference was observed between wild-type and ZB3bInt strains; hence, the wild-type data are included in Fig. 4. Estimates of the number of viable cells present were made throughout growth by counting the number of colony forming units (CFU) obtained when serially dilutions were plated on TSA agar plates. The CFU obtained after 12 h of growth from strain BZ9 was twofold higher compared to strain expressing only CbCelA (Fig. 4a); while, a threefold difference was seen with BZ10 compared to the strains expressing either only CtCelA (strain ZB1b) or Cel6B (strain ZB9a) (Fig. 4b). These results strongly suggest that the combined action of GHs; CtCelA, Cel6B, CglT, and CbCelA has resulted in the production of sufficient simple sugars from pre-treated wheat straw to support improved cell growth and biomass accumulation in recombinant strains BZ9 and BZ10.

Overall, the recombinant BZ9 and BZ10 strains exhibited higher growth than the wild type which was sustained during the whole 96-h period monitored. Although during the first 6 h of growth phase, the rate of

Bashir *et al. Biotechnol Biofuels* (2019) 12:199

Page 7 of 16

Fig. 4 Growth of engineered strains on pre-treated wheat straw. Colony forming units (CFU) were measured after growing recombinant *G. thermoglucosidasius* strains on pre-treated wheat straw for 12 h. **a** CFU comparison of recombinant *G. thermoglucosidasius* expressing only CbCelA (strain ZB6d) with *G. thermoglucosidasius* BZ9 strain expressing CglT and CbCelA. **b** CFU comparison of recombinant *G. thermoglucosidasius* BZ10 strain expressing CglT, Cel6B and CtCelA with *G. thermoglucosidasius* expressing only either CtCelA or Cel6B. $P** \leq 0.01$, $P*** \leq 0.0001$ were calculated by one-way ANOVA followed by Sidak's multiple comparisons test. **c** Growth curve of recombinant *G. thermoglucosidasius* BZ9 and BZ10 strains on 1% pre-treated wheat straw as a sole carbon source, with wild-type *G. thermoglucosidasius* served as a control. Results are shown as mean ± SEM of three biological replicates

growth of the wild type appeared to be greater than BZ9 and BZ10, the wild-type cell density decreased drastically after 8 h, suggesting that the initial growth was supported by the small amount of xylose presented in pre-treated wheat straw. Once this had been consumed, wild-type *G. thermoglucosidasius* could not sustain growth further, as evidenced by the rapid decline in OD600. While the recombinant *G. thermoglucosidasius* BZ9 and BZ10 strains showed a longer lag phase up until the 6-h time point, they went on to achieve a much higher OD600, peaking after the 16-h time point after which there was a gradual decline (Fig. 4c).

To confirm that the recombinant BZ9 and BZ10 strains degrade pre-treated wheat straw by producing GHs, cultures were serially diluted at 24, 48, 72 and 96 h of growth and then plated on TSA-CMC plates. We observed a greater number of colonies, as well as clear zones, on Congo red-stained plates up to the 96-h timepoint in the samples derived from the recombinant BZ9 and BZ10 cultures compared to the wild-type control, where there were less colonies and no indication of CMC degradation. These data confirmed that the recombinant cultures were producing active endo-glucanases (CtCelA/CbCelA) (Additional file 1: Figure S4). In addition, we found measurable cellobiohydro-lase activity against chromogenic substrate pNPC only in the recombinant *G. thermoglucosidasius* BZ9 and BZ10 strains, confirming the integrity of Cel6B and CbCelA. The specific activity at 24, 48, 72 and 96 h of

growth was 0.0279 ± 0.002, 0.0255 ± 0.003, 0.020 ± 0.001 and 0.0187 ± 0.002 µmol min^{-1} mg^{-1} of protein in *G. thermoglucosidasius* BZ9; in *G. thermoglucosidasius* BZ10, the activities were 0.0308 ± 0.01, 0.0226 ± 0.002, 0.0216 ± 0.002, 0.186 ± 0.002 µmol min^{-1} mg^{-1} of protein, respectively (Additional file 1: Figure S5a, b). Overall, these data demonstrated that the two engineered *G. thermoglucosidasius* strains, BZ9 and BZ10, were capable of degrading pre-treated wheat straw as the sole carbon source without the addition of any hydrolytic enzymes.

Engineering the bioethanol producing *G. thermoglucosidasius* LS242 for biomass utilisation

To demonstrate that cellulosic ethanol can be produced directly from lignocellulosic biomass by a single organism, we next established our consortium of hydrolytic enzymes in a previously engineered ethanologenic *G. thermoglucosidasius* strain, LS242 (Δldh, pdh^{up}, Δpfl) [26] and tested the ability of the resultant strain to utilise pre-treated wheat straw. To achieve this, *cglT* was integrated into the *pyrE* locus of the *G. thermoglucosidasius* LS242 genome using ACE. This was accomplished as before, through the electrotransformation of LS242 with the plasmid pMTLgSlimS–LS3–*cglT* into *G. thermoglucosidasius* LS004, a $\Delta pyrE$ deletion mutant of LS242, and selecting for uracil prototrophs. Prototrophic colonies were then screened by colony-PCR analysis using primers pyrE_C1_F and pyrE_C2_R. Correct integration of the P$_{ldh}$–*cglT* cassette resulted in the amplification of an

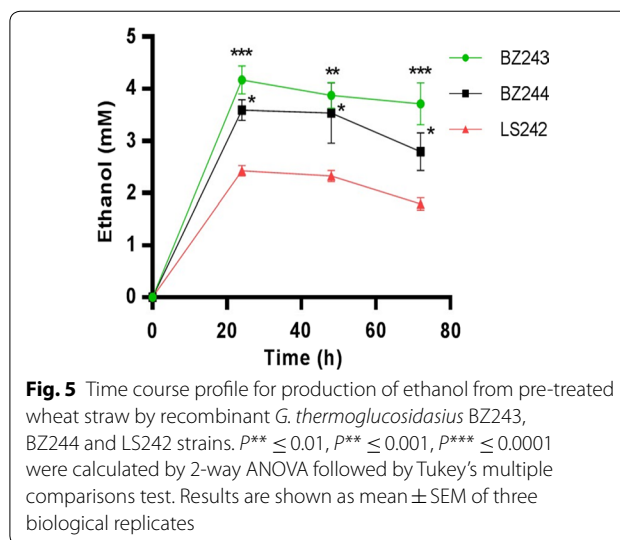Bashir *et al. Biotechnol Biofuels* (2019) 12:199

Page 8 of 16

expected 3.8-kb product size. Sanger sequencing further confirmed the presence of the predicted nucleotide sequence and the strain was designated *G. thermoglucosidasius* BZ242. Subsequently, to construct ethanologenic as well as cellulolytic strains, pMTLgSlimS–Cb*celA* and pMTLgSlimS–Ct*celA–cel6B* plasmids were individually transformed into electro-competent cells of *G. thermoglucosidasius* BZ242, thereby generating recombinant *G. thermoglucosidasius* strains BZ243 and BZ244, respectively.

To investigate the production of ethanol from pretreated wheat straw using the recombinant *G. thermoglucosidasius* strains BZ243 and BZ244, together with the engineered ethanologenic *G. thermoglucosidasius* strain LS242 (Δ*ldh*, *pdh*^up, Δ*pfl*) as a control, microaerobic fermentations were carried out as previously described [26]. Culture supernatants from these strains were used for measurement of ethanol production by HPLC at growth time points of 24, 48 and 72 h. We observed approximately twofold (4.2-mM ethanol) and 1.6-fold (3.7-mM ethanol) increase in ethanol production in the recombinant *G. thermoglucosidasius* BZ243 and BZ244 strains, respectively, compared to *G. thermoglucosidasius* LS242 strain (2-mM ethanol) at 24 h of growth. *G. thermoglucosidasius* LS242 strain produced 2-mM ethanol from wheat straw, most likely as a consequence of consumption of xylose present in pre-treated wheat straw. The levels of ethanol in these cultures remained constant up to 48 h and then gradually declined due to the volatile nature of ethanol at temperatures of 55 °C (Fig. 5).

Discussion

The goal of the study described here was to establish the principle of cellulolytic CBP chassis based on the thermophile *G. thermoglucosidasius*. To that end, we first equipped the otherwise non-cellulolytic, *G. thermoglucosidasius* wild-type strain with the ability to produce heterologous endoglucanase, exoglucanase and β-glucosidase and demonstrated that the engineered recombinant strains were capable of acquiring carbon and energy from pre-treated wheat straw for growth and accumulation of biomass. We then applied this strategy to an engineered ethanologenic *G. thermoglucosidasius* strain, allowing ethanol to be produced directly from lignocellulosic biomass in one step in a single organism.

Based on reported thermophilic activities and properties, the following enzymes were selected to create a cellulolytic *G. thermoglucosidasius*: *C. thermocellum* CtCelA (endoglucanase); *T. fusca* Cel6B (exoglucanase); *C. bescii* CbCelA (endo and exo-glucanase) and *T. brockii* CglT (β-glucosidases). CtCelA is a thermophilic, cellulosomal enzyme belonging to GH8 family encoding a polypeptide of 488 amino acids, consisting of a signal

Fig. 5 Time course profile for production of ethanol from pre-treated wheat straw by recombinant *G. thermoglucosidasius* BZ243, BZ244 and LS242 strains. $P^{**} \leq 0.01$, $P^{**} \leq 0.001$, $P^{***} \leq 0.0001$ were calculated by 2-way ANOVA followed by Tukey's multiple comparisons test. Results are shown as mean ± SEM of three biological replicates

peptide, a catalytic domain and a C-terminal dockering domain [41]. Cel6B has a similar activity to the fungal *T. reesei* exocellulase Cel6A, but with higher thermostability (optimal activity at 60 °C) and a much broader pH optimum. Its catalytic domain is linked via a linker-peptide to a carbohydrate binding module 2 (CBM2) at N-terminus, which confines the substrate and appears to modulate the activities or properties of the catalytic core [42]. CbCelA is a large multimodular enzyme complex containing N-terminal endo-β-1,4-glucanase domain belonging to GH9, three CBM3 domains and a C-terminal GH48 exo-β-1,4-glucanase domain. At elevated temperatures, CbCelA enzyme activity towards its crystalline substrate (avicel) has been reported to be seven times as high as that of the common exo- and endo-cellulase standard mixture, Cel7A and Cel5A [43, 44]. CglT encoding β-glucosidases is an intracellular protein belonging to GH1 family and studies have shown that it has a high glucose tolerance as well as thermostability [45].

Heterologous expression of the GH enzymes was under the control of the well-established promoter (P_{ldh}) of the *G. stearothermophilus ldh* gene [23], while secretion was aided by native signal peptides identified on the cellulases used. The extracellular GHs were successfully expressed as soluble, functional proteins, including the full-length CbCelA from *C. bescii*. CbCelA has been prominent in GH research since its discovery as it effectively degrades the high as well as low-crystallinity substrates [43]. However, so far its production has only been possible in its native host and in a mesophilic heterologous host, *E. coli* [46, 47]. The production of secreted and full-length functional CbCelA in *G. thermoglucosidasius* is, therefore, one of the first time in a thermophile that heterologous production of this multi-functional

enzyme has been achieved. Nevertheless, despite the employment of a strong promoter, visible detection of a protein band via Western blot required a 100-fold concentration of *G. thermoglucosidasius* culture supernatant. One possible reason for this might be that the signal peptides used were native to *C. bescii* and not optimal for secretion in the host of *G. thermoglucosidasius* [48, 49]. Alternatively, as reported in *Bacillus* species [50–52], heterologous proteins are more susceptible to proteolysis by extracellular proteases. Interestingly, the level of CbCelA protein produced was lower than that observed with CtCelA and Cel6B despite the fact that the encoded genes were expressed from the same, P_{ldh} promoter. It has been shown that in *C. bescii*, CbCelA is extensively glycosylated when expressed extracellularly, either as a requirement for protein translocation, or as a necessary post translational modification to shield the complex architecture of CbCelA against proteases [47]. It is probable that such a mechanism is lacking in *Geobacillus* resulting in lower levels of expression or higher level of protease cleavage once the protein is produced and secreted. Furthermore, the high prevalence of proline–threonine–serine regions found in the linkers between the CBM domains in CbCelA might also be more susceptible to proteolysis, contributing to lower expression levels [37].

Significant enzyme activities were obtained for all the heterologous expressed GH enzymes without concentrating the culture supernatant. For CtCelA and Cel6B, a specific activity of 15 and 0.0135 µmol min^{-1} mg^{-1} of protein was observed against their model substrate CMC and RAC, respectively. Finally, for CbCelA, both endo- and exo-glucanase activities were confirmed as evident by the degradation of model substrates, CMC and RAC $(0.735 \pm 0.07$ and 0.0831 ± 0.006 µmol min^{-1} mg^{-1} of protein, respectively). It is notable that the exoglucanase activity of CbCelA was more than sevenfold higher compared to Cel6B, despite a much lower protein expression level as detected by Western blotting. This might be due to the fact that CbCelA enzyme has three CBM domains; while Cel6B has only one, which is reported to enhance the activity towards the substrates [53–55]. Several studies have reported the heterologous expression of cellulases in bacteria, fungi and yeast. However, the production of heterologous cellulases in heterologous hosts can be affected by numerous factors, including conditions of growth, heterologous secretion and the optimum temperature and pH at which the enzymes are active [56–58]. In the present study, although assays of the enzymes CtCelA, Cel6B and CbCelA were undertaken under conditions (pH 6.0 and temperature 60 °C) that were within the optimal temperature and pH ranges of these cellulases [37, 59], specific activities were low.

This was most likely a consequence of using crude supernatants as opposed to purified enzymes.

For the intracellular CglT, a significant level of protein expression was observed on Western blotting despite the chromosomal location of the gene in *G. thermoglucosidasius* ZB3bInt. It is apparent from previous studies that during cellulose saccharification, cellobiose accumulation causes feedback inhibition against many exo- and endoglucanases [60] and can often be the rate-limiting step in the hydrolytic process. Although, most of the *Geobacillus* species possess the ability to transport and metabolise disaccharides, oligosaccharides, sugar alcohols and sugar acids, *G. thermoglucosidasius* NCIMB 11955 is known to utilise cellobiose [28, 32]. However, the expression of β-glucosidase genes is invariably tightly regulated and controlled via several global regulators in most microorganisms and is often maintained in a silent state in the absence of cellobiose or other disaccharides [61–63]. As a consequence, cellobiose is utilised inefficiently. In our study, we demonstrated that the integration of CglT encoding β-glucosidase in the recombinant *G. thermoglucosidasius* ZB3Int strain led to complete degradation of 30 g/l cellobiose in 10 h; whereas, the wild-type strain failed to completely utilise the cellobiose provided.

In the present study, the exo–endo-glucanase properties of CtCelA plus Cel6B and CbCelA were combined with the β-glucosidase activity of CglT in recombinant strains BZ9 and BZ10, respectively. The growth of both strains on pre-treated wheat straw was considerably improved compared to wild type, although an initial lag phase was observed. The latter is plausible due to the imposition of a metabolic burden on the engineered strains, as a consequence of the production of the cellulose degrading enzymes. The slightly more vigorous growth of wild-type strain in the early phase of the cultures is most likely due to the utilisation of the small amount of xylose present in pre-treated wheat straw. Once this is exhausted, the cell density of wild-type cultures significantly declined. Monitoring of cellulolytic activities in BZ9 and BZ10 strains growing on pre-treated wheat straw as a sole carbon source confirmed that all the enzymes were expressed, secreted and functional.

To produce cellulosic ethanol, the biomass degrading ability of the recombinant *G. thermoglucosidasius* BZ9 and BZ10 was combined with the ethanologenic property of the *G. thermoglucosidasius* LS242 strain [26] resulting in strains BZ243 and BZ244, respectively. Production of ethanol from pre-treated wheat straw was approximately twofold and 1.6-fold higher compared to non-engineered LS242, demonstrating that at least some cellulose was converted into ethanol. Despite the fact that strains BZ243 and BZ244 were the first reported engineered *G. thermoglucosidasius* strains

producing ethanol directly using cellulose obtained from industry relevant process without enzymatic treatment, the yield of ethanol is very low and far from being industrially relevant when compared with pure glucose fermentation [23]. This could be attributed to several factors. Firstly, recombinant *G. thermoglucosidasius* BZ243 and BZ244 strains were cultured for fermentation in micro-aerobic condition with no shaking and aeration, which may lead to unequal distribution of insoluble pre-treated wheat straw, compromising feedstock utilisation. Secondly, the activity of the P_{ldh} promoter is known to be threefold lower in cultures grown under micro-aerobic conditions compared to in aerobic conditions [32]. Thirdly, the effect of fermentative growth conditions on expression levels and enzyme activities, as well as the concerted action of the enzymes used, are unpredictable and hard to measure or compare to the native GHs producing organisms.

Conclusion

In this report, we have heterologously expressed GH enzymes (CtCelA, Cel6B, CglT and CbCelA) in *G. thermoglucosidasius* both individually and synergistically with detectable activities. When grown on pre-treated wheat straw as a sole carbon source, engineered *G. thermoglucosidasius* BZ9 and BZ10 strains drastically improved growth compared to the wild-type control suggesting the combined production of these GHs facilitate lignocellulosic degradation. Further, heterologous expression of GHs in an engineered ethanologenic *G. thermoglucosidasius* LS242 strain has led to substantial increase in the production of ethanol from pre-treated wheat straw. Whilst a small step of many, this advancement is supportive of the potential of *G. thermoglucosidasius* as a cellulolytic chassis for the fermentative production of biofuel. The use of the recombinant, cellulolytic *G. thermoglucosidasius* as a process organism is not, however, limited to ethanol production but has the potential to be used to produce many other biofuels and renewable chemicals with further genetic modifications.

Materials and methods
Chemicals and reagents
All the chemicals and bacteriological media were purchased from Sigma-Aldrich Company Ltd (Poole, UK) or Fisher Scientific Ltd (Loughborough, UK). DNA polymerases, restriction enzymes, deoxyribonucleoside triphosphates (dNTPs), DNA markers and associated reaction buffers were sourced from Thermo Fischer Scientific (UK) or New England Biolabs Ltd. (Hitchin, UK), T4 DNA ligase from Promega (Southampton, UK).

EDTA-Free Protease Inhibitor Cocktail Tablets were purchased from Roche (Mannheim, Germany).

Bacterial strains, media and growth conditions
The bacterial strains and plasmids used in this study are listed in Table 1. *E. coli* Top10 was used as a host for plasmid constructions and routinely grown aerobically with agitation of 200 rpm at 37 °C or 30 °C in Luria-Bertani (LB) broth or agar supplemented with 50 μg/ml kanamycin or 100 μg/ml ampicillin. *G. thermoglucosidasius* strains were routinely grown on TSA agar or in 2SPYNG broth (16-g Soy peptone, 10-g Yeast extract and 5-g NaCl) supplemented with 12.5 μg/ml kanamycin. For the protein expression, recombinant *G. thermoglucosidasius* were pre-cultured overnight in 10 ml of 2SPYNG medium supplemented with 12.5 μg/ml kanamycin in 50-ml falcon constantly rotating at 250 rpm, 52 °C. Fresh pre-warmed 50-ml 2SPYNG medium in baffled conical flask (Sigma, UK) was inoculated with the overnight cultures at a starting OD600 of 0.2 and grown until late stationary phase for almost 8–9 h. For the growth experiments with different carbohydrates, engineered *G. thermoglucosidasius* strains were grown in modified ASYE [27] medium (per 1 l: $1\times$ M9 minimal salts (sigma), 0.24-g $MgSO_4$, 0.011-g $CaCl_2$, 0.01-g Thiamine-HCl, 0.384-g citric acid, 27.8-mg $FeSO_4.7H_2O$, 4.6-mg $NiCl_3.6H_2O$, 3.05-mg biotin, 47.6-g HEPES, 5-g yeast extract, $1000\times$ dilution of Trace metal mix A5 and 10-g cellobiose or pre-treated wheat straw) at 250 rpm, 55 °C. The pretreated wheat straw sample used was generated at the DBT-ICT 2G Ethanol Demonstration Facility and was kindly gifted by Institute of Chemical Technology, Mumbai. Characterization for the ethanol production on 1% pre-treated wheat straw in modified ASYE medium at 55 °C was carried out in sealed 50 ml falcon tube as previously described method [26]. For the selection of integrated genes in *G. thermoglucosidasius*, clostridial basal medium (CBM) agar supplemented with 1% w/v xylose was used [64]. All the solidified media contained 1.5% w/v agarose.

Design and synthesis of genes encoding CtCelA, Cel6B, CglT and CbCelA
The selected GH genes; *celA* encoding endoglucanase (EC 3.2.1.4) from *Clostridium thermocellum* (designated as Ct*celA*), *cel6B* encoding exoglucanase (EC 3.2.1.91) from *Thermobifida fusca* YX, *cglT* encoding β-glucosidase (EC 3.2.1.21) from *Thermoanaerobacter brockii* and *celA* encoding multidomain cellulase (EC 3.2.1.4 and EC 3.2.1.91) from *Caldicellulosiruptor bescii* (designated as Cb*celA*) were designed as standard BioBrick-2 assembly parts. Each BioBrick-2 part consists of DNA fragment

Bashir *et al. Biotechnol Biofuels* (2019) 12:199

Page 11 of 16

Table 1 List of bacterial strains and plasmids used in this study. *G. thermoglucosidasius* is abbreviated as *G. t.* and *E. coli* as *E. c*

Strains/plasmids	Properties	References/source
Strains		
E. c. Top 10	*E. coli* cloning strain	NEB
G. t. NCIMB 11955	wild-type strain	TMO renewables
G. t. 11955 *ΔpyrE*	*pyrE* deletion strain of 11955	Sheng et al. [26]
G. t. LS004	*ΔpyrE Δldh Δpfl pdh*^{up}	Sheng et al. [26]
G. t. LS242	*Δldh Δpfl pdh*^{up}	Sheng et al. [26]
G. t. ZB1b	P_{ldh}::Ct*celA*; Ct*celA* expressed on plasmid constitutively; Kan^R	This study
G. t. ZB9a	P_{ldh}:: *cel6B*; *cel6B* expressed on plasmid constitutively; Kan^R	This study
G. t. ZB6d	P_{ldh}::Cb*celA*; Cb*celA* expressed on plasmid constitutively; Kan^R	This study
G. t. ZB3bInt	P_{ldh}::*cglT*; *cglT* constitutively expressed in the chromosome at *pyrE* locus	This study
G. t. BZ9	P_{ldh}::Cb*celA*; Cb*celA* expressed on autonomous plasmid; *cglT* constitutively expressed in the chromosome at *pyrE* locus; derived from *G. t.* ZB3bInt; Kan^R	This study
G. t. BZ10	P_{ldh}::Ct*celA*–*cel6B*; Ct*celA*–*cel6B* operon expressed on autonomous plasmid; *cglT* constitutively expressed in the chromosome at *pyrE* locus; derived from *G. t.* ZB3bInt; Kan^R	This study
G. t. BZ242	P_{ldh}::*cglT*; *cglT* constitutively expressed in the chromosome at *pyrE* locus; derived from *G. t.* LS242	This study
G. t. BZ243	P_{ldh}::Cb*celA*; Cb*celA* expressed on autonomous plasmid; *cglT* constitutively expressed in the chromosome at *pyrE* locus; derived from *G. t.* BZ242 strain; Kan^R	This study
G. t. BZ244	P_{ldh}::Ct*celA*–*cel6B*; Ct*celA*–*cel6B* operon expressed on autonomous plasmid; *cglT* constitutively expressed in the chromosome at *pyrE* locus; derived from *G. t.* BZ242; Kan^R	This study
Plasmids		
pMTLgSlimS	*G. thermoglucosidasius* shuttle vector. RepB, ColE1, Kan^R	Sheng et al. [26]
pBSK–*ldh*	Synthesised *ldh* promoter	Biomatik
pBSK–Ct*celA*	Synthesised Ct*celA* gene	Biomatik
pBSK–*cel6B*	Synthesised *cel6B* gene	Biomatik
pBSK–*cglT*	Synthesised *cglT* gene	Biomatik
pBSK–Cb*celA*	Synthesised Cb*celA* gene	Biomatik
pMTLgSlimS–LS3	*G. thermoglucosidasius* integration vector at *pyrE* locus	Sheng et al. [26]
pJ201–Ct*celA*	Biobrick vector harbouring Ct*celA* gene flanked by *ldh* promoter	This study
pMTLgSlimS–Ct*celA*	Expression vector for Ct*celA* with P_{ldh} promoter	This study
pJ201–*cel6B*	Biobrick vector harbouring *cel6B* gene preceded by P_{ldh} promoter	This study
pMTLgSlimS–*cel6B*	Expression vector for *cel6B* with P_{ldh} promoter	This study
pJ201–Cb*celA*	Biobrick vector harbouring Cb*celA* gene flanked by P_{ldh} promoter	This study
pMTLgSlimS–Cb*celA*	Expression vector for Cb*celA* with P_{ldh} promoter	This study
pJ201–*cglT*	Biobrick vector harbouring *cglT* gene flanked by P_{ldh} promoter	This study
pMTLgSlimS–*cglT*	Expression vector for *cglT* with P_{ldh} promoter	This study
pMTLgSlimS–LS3–*cglT*	Integration vector for Cg*lT* with P_{ldh} promoter at *pyrE* locus	This study
pMTLgSlimS–Ct*celA*–*cel6B*	Expression vector for Ct*celA*–*cel6B* operon with P_{ldh} promoter	This study

flanked by prefix (contains *Eco*RI/*Not*I/*Spe*I restriction sites) and suffix (consists of *Nhe*I/*Not*I/*Pst*I sites) sequences at 5′ and 3′ ends, respectively. In the BioBrick-2 assembly process, the combination of two parts (e.g., fusion of promoter to a gene) forms a new part which contains a scar (GCTAGT) at the junction point of the two progenitor parts while maintaining the same prefix and suffix which allows the subsequent assembly with more BioBrick-2 parts. The scar sequence represents the fusion of DNA that has been cut with *Spe*I/*Nhe*I sites

and the compatible sticky ends ligated together. At the 5′-ends of each gene, BioBrick-2 prefix and the synthetic sequence 5′-TTATATTGA**AGGAGG**ATGAATGCA-3′ in which the RBS is indicated in bold was added. FLAG-epitope before the stop codon and BioBrick-2 suffix sequence was added at the 3′-ends of the Ct*celA*, *cel6B*, *cglT* and Cb*celA* gene sequences. The P_{ldh} promoter of the lactate dehydrogenase encoding gene (*ldh*) was synthesised without its native ribosome binding site (RBS) in the BioBrick-2 format. All the four encoding genes

Bashir *et al. Biotechnol Biofuels*   (2019) 12:199

Page 12 of 16

were codon optimised based on the genome sequence of *G. thermoglucosidasius* C56-YS93 and synthesised by Biomatik (Ontario, Canada) in the standard BB2 format. The overall configuration in synthesised BioBrick-2 format is 5′-BioBrick2prefix-syntheticRBS-*gene*-FLAG tag-BioBrick2suffix-3′. The BioBrick-2 assembled nucleotide sequences of Ct*celA*, *cel6B*, *cglT* and Cb*celA* along with the P$_{ldh}$ promoter sequence used in this study are shown in Additional file 1: S1.

Plasmid and strain constructions

The oligonucleotide primers used are listed in Additional file 1: Table S1. NEBuilder Assembly Tool was used to design primers for NEBuilder HiFi DNA Assembly reactions and purchased from Sigma-Aldrich, UK. NEBase Changer was employed to design primers specific to the mutagenesis experiment using Q5® Site-Directed Mutagenesis Kit. For the amplification of the desired DNA product, either phusion High-Fidelity DNA polymerase or dream Taq Green PCR (Thermo Fisher Scientific, UK) master mix was used. PCR-amplification was performed in a standard 25–50 µl reaction using a thin walled PCR tubes placed into either a Biometra T$_{3000}$ thermocycler (Glasgow, UK) or Sensoquest lab cycler (Geneflow, UK). Restriction enzyme digestions and PCR products were routinely separated by agarose gel electrophoresis using 0.8–1% gels prepared with analytical grade agarose (Sigma, UK) in 1xTAE buffer (40-mM Tris, 1-mM EDTA and 0.1% (v/v) glacial acetic acid). DNA fragments were purified using Zymoclean™ Gel DNA Recovery Kit (ZymoResearch, UK) following the manufacturer's instructions. Chemically competent *E. coli* Top10 cells produced in-house were routinely used either to transform ligation reaction or complete plasmid and individual clones were selected on LB agar plates with 50 µg/ml kanamycin. Plasmids were purified using Qiagen Miniprep kit (Qiagen Ltd. Surrey, UK) according to the manufacturer's protocol and DNA was eluted in 30–50 µl nuclease-free water (Fisher Scientific, UK) followed by storage at −20 °C until further use. Sanger sequencing (Source BioSciences, Nottingham, UK) was routinely performed to confirm the plasmid constructs. Preparation of *G. thermoglucosidasius* electro-competent cells and electroporation procedure using Genepulser electroporator (BioRad, UK) was done in accordance with the method previously described [26]. *G. thermoglucosidasius* stock cultures were routinely maintained at −80 °C in 12.5% (v/v) of glycerol.

Construction of CtcelA expression vector

All the synthesised DNA fragments were delivered in pBSK vector. The Ct*celA* gene and the P$_{ldh}$ promoter synthesised in a BioBrick-2 format were assembled together in the vector pJ201 (DNA2.0) which contains the requisite prefix and suffix restriction sites for BioBrick-2 cloning. pBSK–P$_{ldh}$ and pBSK–Ct*celA* were digested with *Eco*RI/*Nhe*I and *Spe*I/*Pst*I, respectively, and ligated at *Eco*RI/*Pst*I site in pJ201 vector yielding pJ201–Ct*celA* plasmid. P$_{ldh}$–Ct*celA* cassette was subcloned from pJ201–Ct*celA* into pMTLgSlimS upon digestion with *Not*I/*Nhe*I resulting in pMTLgSlimS–Ct*celA* expression vector.

Construction of cel6B expression vector

pBSK–P$_{ldh}$ and pBSK–*cel6B* were digested with *Eco*RI/*Nhe*I and *Spe*I/*Pst*I, respectively, and ligated at *Eco*RI/*Pst*I site in pJ201 vector yielding pJ201–*cel6B* plasmid. P$_{ldh}$–*cel6B* cassette was subcloned from pJ201–*cel6B* into pMTLgSlimS upon digestion with *Not*I/*Nhe*I resulting in pMTLgSlimS–*cel6B* expression vector.

Construction of cglT expression vector

pBSK–P$_{ldh}$ and pBSK–*cglT* were digested with *Eco*RI/*Nhe*I and *Spe*I/*Pst*I, respectively, and ligated at *Eco*RI/*Pst*I site in pJ201 vector yielding pJ201–*cglT* plasmid. P$_{ldh}$–*cglT* cassette was subcloned from pJ201–*cglT* into pMTLgSlimS upon digestion with *Not*I/*Nhe*I resulting in pMTLgSlimS–*cglT* expression vector.

Construction of CbcelA expression vector

pBSK–P$_{ldh}$ and pBSK–Cb*celA* were digested with *Eco*RI/*Nhe*I and *Spe*I/*Pst*I, respectively, and ligated at *Eco*RI/*Pst*I site in pJ201 vector yielding pJ201–Cb*celA* plasmid. P$_{ldh}$–Cb*celA* cassette was subcloned from pJ201–Cb*celA* into pMTLgSlimS upon digestion with *Not*I/*Nhe*I resulting in pMTLgSlimS–Cb*celA* expression vector.

Construction of CtcelA–cel6B operon

For the construction of Ct*celA* and *cel6B* operon under the transcriptional control of P$_{ldh}$ promoter, NEBuilder-HiFi DNA Assembly method was used according to the manufacturer's instructions. Briefly, pMTLgSlimS plasmid was linearized by PCR amplification employing pMTLgSLimS_HiFi_F and pMTLgSLimS_HiFi_R primers (detailed sequence in Additional file 1: Table S1). The *cel6B* nucleotide sequence was amplified from plasmid pMTLgSlimS–*cel6B* using Cel6B_F and Cel6B_R pair of primers with the incorporation of new RBS-5 (TTTTAA**AGGAGG**TATAAGCT) upstream of the start codon. P$_{ldh}$–Ct*celA* expression cassette was amplified from pMTLgSlimS–Ct*celA* plasmid without FLAG-tag at the C-terminal using P$_{ldh}$_CelA_F and P$_{ldh}$_CelA_R primers. All these modifications were done to avoid repetitions of FLAG-tag or same synthetic RBS in an operon and insertion of P$_{ldh}$–Ct*celA*–*cel6B* cassette in *lacz* region of pMTLgSlimS vector between *Xba*I and

Bashir *et al. Biotechnol Biofuels*   (2019) 12:199

Page 13 of 16

*Hin*dIII restriction recognition sites. The PCR gel-purified pMTLgSlimS, *cel6B* and P*ldh*–Ct*celA* were assembled together using master mix–HiFi assembly kit according to the manufacturer's protocol yielding pMTLgSlimS–Ct*celA*–*cel6B* expression plasmid.

Chromosomal integration of *cglT* gene via Allele-Coupled Exchange (ACE)

The *cglT* gene along with the P*ldh* promoter was subcloned from pMTLgSlims–*cglT* between the SpeI and NotI restriction sites of the integrative vector pMTLg-SlimS–LS3 [26] giving rise to pMTLgSlims–LS3–*cglT.* For the integration of *cglT* gene into the chromosome at the *pyrE* locus using ACE [26], *G. thermoglucosidasius* Δ*pyrE* cells were electroporated with 1 μg of the pMTLg-SLims–LS3–*cglT* plasmid DNA. Briefly, following the transformation procedure [23], cells were plated onto TSA plates supplemented with 12.5 μg/ml kanamycin and incubated overnight at 52 °C. Six large clones were then streaked on to CBM1X (CBM with 1% w/v xylose) plates lacking uracil and incubated at 62 °C for overnight. Further, single colonies were re-streaked twice on CBM1X plates without uracil to purify and incubated for 16–18 h at 62 °C. Single colonies were used as a template for colony-PCR to confirm the integration of *cglT* expression cassette with pyrE_C1_F and pyrE_C2_R primers (Additional file 1: Table S1).

SDS-PAGE and Western blotting

The supernatants from *G. thermoglucosidasius* strains expressing extracellular recombinant CtCelA, Cel6B and CbCelA proteins were collected by centrifuging 10-ml culture at 7000 rpm for 20 min at 4 °C. Then, 1 ml of Trichloroacetic acid (TCA, 10% w/v) was added for precipitation of proteins and the suspension was stored at −20 °C for overnight. The next day, samples were centrifuged at 10,000 rpm for 20 min at 4 °C and pellet was resuspended in 1 ml of 100 mM Tris–HCl followed by centrifugation at 13,000 rpm for 5 min at 4 °C. The washing step of pellet with 100 mM Tris–HCl was repeated three times. After the final wash, protein pellet was air dried for 60 min at room temperature and 2X NuPAGE® Lithium dodecyl suphate (LDS) Sample (Thermo Fisher, UK) buffer was added to the pellet.

Cell lysate was prepared by spinning (7000 rpm for 20 min at 4 °C) 50 ml of *G. thermoglucosidasius* cells expressing recombinant CglT protein grown in 2SPYNG medium. The cell pellet was washed with 5 ml of 50-mM Tris–HCl (pH 7.8) and resuspended in 1 ml of same buffer containing EDTA-free complete protease inhibitor cocktail tablets. Cells placed in ice bath were disrupted by sonication using Soniprep 150 Ultrasonic Disintegrator

(MSE Scientific Instruments, Wolflabs, UK) and the lysed cells were pelleted at 25,000×*g* for 30 min at 4 °C. The supernatant that corresponds to the soluble fraction of the cell extract was further used for SDS-PAGE or enzyme activity analysis.

All the protein samples were boiled at 95 °C for 5 min and separated by SDS-PAGE using precast 4–12% NuPAGE® Bis–Tris (Thermo Fischer, UK) polyacrylamide gels according to the manufacturer's instructions. To determine the size of the recombinant polypeptide, broad range precision plus protein™ Kaleidoscop™ prestained protein (Biorad, UK) marker was also loaded onto the SDS-PAGE gels. After the gel electrophoresis, proteins were electro transferred from SDS-PAGE gels onto the Trans-Blot® Turbo™ Mini PVDF membranes (Bio-Rad, UK) for the western blotting. The transfer was carried out using the Trans-Blot® Turbo™ Transfer System (Bio-Rad) at an electric current of 1.3 A, voltage of 25 V for 7 min. After the run was complete, membrane was blocked with 1.5% dried skimmed milk powder (Sigma, UK) in 30 ml Tris-Buffered Saline (TBS) (0.05-M Tris pH 8.0, 0.15-M NaCl) for 1 h at room temperature. The membrane was further incubated with fresh blocking solution, 1.5% dried skimmed milk powder in TBS containing at a dilution of 1:3000 monoclonal Anti-FLAG® M2-peroxidase (HRP) antibody (Sigma, UK) and incubated overnight in a shaker at 4 °C. The monoclonal Anti-FLAG M2-Peroxidase is a mouse IgG antibody covalently conjugated to horseradish peroxidase (HRP) and binds to FLAG fusion proteins. After overnight incubation, membrane was washed six times with TBS containing 0.1% Tween 20 for 10 min each before developing the blot using the colorimetric 3,3′,5,5′-tetramethylbenzidine (TMB) as a substrate (Sigma, UK). After 5 min of exposure with TMB, the blot was dried and acquired for the image using scanner.

Determination of protein concentration and enzyme activity assays

Total protein concentrations in the supernatant or soluble fraction of cell lysates were determined using BCA-kit (Thermo Fisher Scientific, UK) following the manufacturer's instructions and standard curve generated from BSA (200–1000 μg/ml) was used to calculate the protein concentrations. For Congo-red plate staining assay, wild type and recombinant *G. thermoglucosidasius* strains producing CtCelA were streaked on TSA agar plates containing 0.5% CMC followed by incubation at 52 °C overnight. Next day, plates were flooded with 0.1% Congo red solution for 30 min at room temperature and then washed with 1-M NaCl solution [65]. RAC as a substrate for exoglucanase activity was prepared from Avicel® PH-101 (Sigma, UK) as described by the method [66].

Bashir *et al. Biotechnol Biofuels* (2019) 12:199

Page 14 of 16

The quantitative enzyme activities of heterologous endoglucanase (CtCelA, CbCelA) and exoglucanase (Cel6B, CbCelA) were assayed with CMC (carboxy-methyl cellulose) and RAC, respectively, using the DNSA (3, 5-Dinitrosalicylic acid) method [67, 68]. Briefly, the reaction mixture containing 0.125 ml of culture supernatant (as an enzyme solution) and 0.125 ml of 1% CMC or RAC in 100 mM citrate phosphate buffer (pH 6.0) was incubated at 60 °C. Reactions were terminated after 1 and 16 h for CMC and RAC assays, respectively, by adding 0.25 ml of DNSA reagent and heating the mixture at 100 °C for 10 min. 0.1 ml of mixture was transferred into 96-well plate and the absorbance measured at 540 nm using CLARIOstar plate reader (BMG Labtech, UK). As a control, enzyme blank and substrate-independent reactions were performed simultaneously, and the absorbance determined was subtracted from the sample absorbance. One unit of enzyme activity was defined as the amount of enzyme catalysing the release of 1 μmol of reducing sugar per min from substrate under the specified assay conditions.

Synthetic chromogenic substrates p-nitrophenyl-β-D-glucopyranoside (pNPG) (Sigma, UK) and p-nitrophenyl-β-D-cellobioside (pNPC) (Sigma, UK) were used for detection of β-glucosidase (CglT) and cellobiohydrolase activity, respectively [67]. The reaction mixture containing 50-mM citrate phosphate buffer (pH 6.0), 5-mM pNPG or pNPC and 0.125-ml supernatant (used as enzyme) in a total volume of 0.55 ml was incubated at 60 °C for 10 min. After incubation 1 ml of 1% Na_2CO_3 was added in the mixtures to stop the reaction and absorbance measured at 400 nm with CLARIOstar plate reader (BMG Labtech, UK). Enzyme activity was calculated using calibration curve prepared with p-nitrophenol standard and one unit of enzyme activity is defined as the release of 1 μmol of pNP per minute under the above conditions.

Analytical techniques

Bacterial cell growth was measured at an optical density of 600 nm (OD600) using a Novaspec II spectrophotometer (Pharmacia LKB Ltd., Cambridge, UK). Fermentation samples at various time points were collected, centrifuged at 14,000×g for 10 min and cell-free supernatant was used for analysis of ethanol by HPLC. Internal standard solution (80-mM valeric acid in 0.005-M H_2SO_4) was mixed with equal volume (0.2 ml) cell-free supernatant and filtered through 0.22-μm HPLC certified syringe filter (Whatman® Spartan®, GE Healthcare Life Sciences, UK) and transferred into a HPLC vial with a 100 μl insert. Ethanol concentrations were quantified by Dionex UltiMate 3000 HPLC system equipped with a 300 mm × 7.8 mm Aminex® HPX-87H (Bio-Rad, UK)

column set at 35 °C using 5-mM H_2SO_4 as the mobile phase and monitored using a refractive index (RI) and Diode Array UV–Vis-detector. All the experiments were performed in triplicates and the statistically significant differences between the test samples were calculated in the PRISM (GraphPad Software, La Jolla, USA) software using Student's t test or ANOVA. P values ≤ 0.05 were considered statistically significant.

Additional file

Additional file 1. Figures S1–S5, Tables S1, File S1.

Authors' contributions
ZB carried out the laboratory work, data analysis and drafted the manuscript. NPM helped design the study and edited the manuscript. LS participated in the laboratory work and coordination of the final part of the study. AA helped with the biomass pre-treatment. AL helped and coordinated the biomass quality control. YZ conceived the study, oversaw its design and coordination, helped with the data analysis and revised the manuscript. All authors read and approved the final manuscript.

Availability of data and materials
All data and material used in the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate
Not applicable.

Consent for publication
All authors consented to the publication of this work.

Competing interests
The authors declare that they have no competing interests.

Bashir *et al. Biotechnol Biofuels*   (2019) 12:199

Page 15 of 16

Author details
¹ Clostridia Research Group, BBSRC/EPSRC Synthetic Biology Research Centre (SBRC), School of Life Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, UK. ² DBT-ICT Centre for Energy Biosciences, Institute of Chemical Technology, Nathalal Parikh Marg, Mumbai 400019, India.

References
1. Amin FR, Khalid H, Zhang H, Rahman SU, Zhang R, Liu G, Chen C. Pretreatment methods of lignocellulosic biomass for anaerobic digestion. AMB Express. 2017;7(1):72.
2. Socha AM, Parthasarathi R, Shi J, Pattathil S, Whyte D, Bergeron M, George A, Tran K, Stavila V, Venkatachalam S, et al. Efficient biomass pretreatment using ionic liquids derived from lignin and hemicellulose. Proc Natl Acad Sci USA. 2014;111(35):E3587–95.
3. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res. 2009;37(Database issue):D233–8.
4. Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD. Biomass recalcitrance: engineering plants and enzymes for biofuels production. Science. 2007;315(5813):804–7.
5. Mazzoli R, Lamberti C, Pessione E. Engineering new metabolic capabilities in bacteria: lessons from recombinant cellulolytic strategies. Trends Biotechnol. 2012;30(2):111–9.
6. Akinosho H, Yee K, Close D, Ragauskas A. The emergence of *Clostridium thermocellum* as a high utility candidate for consolidated bioprocessing applications. Front Chem. 2014;2:66.
7. Zhao L, Cao G-L, Wang A-J, Ren H-Y, Zhang K, Ren N-Q. Consolidated bioprocessing performance of *Thermoanaerobacterium thermosaccharolyticum* M18 on fungal pretreated cornstalk for enhanced hydrogen production. Biotechnol Biofuels. 2014;7(1):178.
8. Turner P, Mamo G, Karlsson EN. Potential and utilization of thermophiles and thermostable enzymes in biorefining. Microb Cell Fact. 2007;6:9.
9. Keller M, Loder A, Basen M, Izquierdo J, Kelly RM, Adams MWW. Production of lignofuels and electrofuels by extremely thermophilic microbes. Biofuels. 2014;5(5):499–515.
10. Chang T, Yao S. Thermophilic, lignocellulolytic bacteria for ethanol production: current state and perspectives. Appl Microbiol Biotechnol. 2011;92(1):13–27.
11. 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.
12. Egorova K, Antranikian G. Industrial relevance of thermophilic Archaea. Curr Opin Microbiol. 2005;8(6):649–55.
13. Verbeke TJ, Giannone RJ, Klingeman DM, Engle NL, Rydzak T, Guss AM, Tschaplinski TJ, Brown SD, Hettich RL, Elkins JG. Pentose sugars inhibit metabolism and increase expression of an AgrD-type cyclic pentapeptide in *Clostridium thermocellum*. Scientific Reports. 2017;7:43355.
14. Papanek B, O'Dell KB, Manga P, Giannone RJ, Klingeman DM, Hettich RL, Brown SD, Guss AM. Transcriptomic and proteomic changes from medium supplementation and strain evolution in high-yielding *Clostridium thermocellum* strains. J Ind Microbiol Biotechnol. 2018;45(11):1007–15.
15. Tian L, Cervenka ND, Low AM, Olson DG, Lynd LR. A mutation in the AdhE alcohol dehydrogenase of *Clostridium thermocellum* increases tolerance to several primary alcohols, including isobutanol, n-butanol and ethanol. Scientific Rep. 2019;9(1):1736.
16. Shaw AJ, Hogsett DA, Lynd LR. Identification of the [FeFe]-hydrogenase responsible for hydrogen generation in *Thermoanaerobacterium saccharolyticum* and demonstration of increased ethanol yield via hydrogenase knockout. J Bacteriol. 2009;191(20):6457–64.
17. Shaw AJ, Podkaminer KK, Desai SG, Bardsley JS, Rogers SR, Thorne PG, Hogsett DA, Lynd LR. Metabolic engineering of a thermophilic bacterium to produce ethanol at high yield. Proc Natl Acad Sci USA. 2008;105(37):13769–74.
18. Shaw AJ, Hogsett DA, Lynd LR. Natural competence in *Thermoanaerobacter* and *Thermoanaerobacterium* Species. Appl Environ Microbiol. 2010;76(14):4713.
19. Chung D, Cha M, Guss AM, Westpheling J. Direct conversion of plant biomass to ethanol by engineered *Caldicellulosiruptor bescii*. Proc Natl Acad Sci USA. 2014;111(24):8931–6.
20. Zeldes BM, Keller MW, Loder AJ, Straub CT, Adams MWW, Kelly RM. Extremely thermophilic microorganisms as metabolic engineering platforms for production of fuels and industrial chemicals. Front Microbiol. 2015;6:1209.
21. Yan Q, Fong SS. Challenges and advances for genetic engineering of non-model bacteria and uses in consolidated bioprocessing. Front Microbiol. 2017;8:2060.
22. Nazina TN, Tourova TP, Poltaraus AB, Novikova EV, Grigoryan AA, Ivanova AE, Lysenko AM, Petrunyaka VV, Osipov GA, Belyaev SS, et al. Taxonomic study of aerobic thermophilic bacilli: descriptions of *Geobacillus subterraneus* gen. nov., sp. nov. and *Geobacillus uzenensis* sp. nov. from petroleum reservoirs and transfer of *Bacillus stearothermophilus*, *Bacillus thermocatenulatus*, *Bacillus thermoleovorans*, *Bacillus kaustophilus*, *Bacillus thermodenitrificans* to Geobacillus as the new combinations *G. stearothermophilus*, *G. th*. Int J Syst Evol Microbiol. 2001;51(Pt 2):433–46.
23. Cripps RE, Eley K, Leak DJ, Rudd B, Taylor M, Todd M. Metabolic engineering of *Geobacillus thermoglucosidasius* for high yield ethanol production. Metab Eng. 2009;11:398–408.
24. Van Zyl LJ, Taylor MP, Eley K, Tuffin M, Cowan DA. Engineering pyruvate decarboxylase-mediated ethanol production in the thermophilic host *Geobacillus thermoglucosidasius*. Appl Microbiol Biotechnol. 2014;98(3):1247–59.
25. Zhou J, Wu K, Rao C. Evolutionary engineering of *Geobacillus thermoglucosidasius* for improved ethanol production. Biotechnol Bioeng. 2016;113:2156–67.
26. Sheng L, Kovács K, Winzer K, Zhang Y, Minton NP. Development and implementation of rapid metabolic engineering tools for chemical and fuel production in *Geobacillus thermoglucosidasius* NCIMB 11955. Biotechnol Biofuels. 2017;10(1):5.
27. Lin PP, Rabe KS, Takasumi JL, Kadisch M, Arnold FH, Liao JC. Isobutanol production at elevated temperatures in thermophilic *Geobacillus thermoglucosidasius*. Metab Eng. 2014;24:1–8.
28. 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;6:430.
29. Assareh R, Shahbani Zahiri H, Akbari Noghabi K, Aminzadeh S, Bakhshi Khaniki G. Characterization of the newly isolated *Geobacillus* sp. T1, the efficient cellulase-producer on untreated barley and wheat straws. Bioresour Technol. 2012;120:99–105.
30. Bhalla A, Bischoff KM, Uppugundla N, Balan V, Sani RK. Novel thermostable endo-xylanase cloned and expressed from bacterium *Geobacillus* sp. WSUCF1. Bioresour Technol. 2014;165:314–8.
31. Ng IS, Li CW, Yeh YF, Chen PT, Chir JL, Ma CH, Yu SM, Ho TH, Tong CG. A novel endo-glucanase from the thermophilic bacterium *Geobacillus* sp. 70PC53 with high activity and stability over a broad range of temperatures. Extremophiles. 2009;13(3):425–35.
32. Bartosiak-Jentys J, Hussein AH, Lewis CJ, Leak DJ. Modular system for assessment of glycosyl hydrolase secretion in *Geobacillus thermoglucosidasius*. Microbiology. 2013;159(Pt 7):1267–75.
33. Suzuki H, Yoshida K, Ohshima T. Polysaccharide-degrading thermophiles generated by heterologous gene expression in *Geobacillus kaustophilus* HTA426. Appl Environ Microbiol. 2013;79(17):5151–8.
34. Wilson DB. Three microbial strategies for plant cell wall degradation. Ann N Y Acad Sci. 2008;1125:289–97.
35. Scott M, Gunderson CW, Mateescu EM, Zhang Z, Hwa T. Interdependence of cell growth and gene expression: origins and consequences. Science. 2010;330(6007):1099–102.
36. Chung D, Young J, Bomble YJ, Vander Wall TA, Groom J, Himmel ME, Westpheling J. Homologous expression of the *Caldicellulosiruptor bescii* CelA reveals that the extracellular protein is glycosylated. PLoS ONE. 2015;10(3):e0119508.

Bashir *et al. Biotechnol Biofuels*   (2019) 12:199

Page 16 of 16

37. Zverlov V, Mahr S, Riedel K, Bronnenmeier K. Properties and gene structure of a bifunctional cellulolytic enzyme (CelA) from the extreme thermophile '*Anaerocellum thermophilum*' with separate glycosyl hydrolase family 9 and 48 catalytic domains. Microbiology. 1998;144(Pt 2):457–65.
38. Gefen G, Anbar M, Morag E, Lamed R, Bayer EA. Enhanced cellulose degradation by targeted integration of a cohesin-fused beta-glucosidase into the *Clostridium thermocellum* cellulosome. Proc Natl Acad Sci USA. 2012;109(26):10298–303.
39. Singhania RR, Patel AK, Sukumaran RK, Larroche C, Pandey A. Role and significance of beta-glucosidases in the hydrolysis of cellulose for bioethanol production. Biores Technol. 2013;127:500–7.
40. Lali AMO, Annamma Anil; Birhade, Sachinkumar Hiraman; Victoria, Juliet Joanna and Sawant, Sneha Chandrakant.: Process for Production of Soluble Sugars from Biomass. US patent 20160333382 2016.
41. Schwarz WH, Grabnitz F, Staudenbauer WL. Properties of a *Clostridium thermocellum* endoglucanase produced in *Escherichia coli*. Appl Environ Microbiol. 1986;51(6):1293–9.
42. Adav SS, Ng CS, Arulmani M, Sze SK. Quantitative iTRAQ secretome analysis of cellulolytic *Thermobifida fusca*. J Proteome Res. 2010;9(6):3016–24.
43. Brunecky R, Alahuhta M, Xu Q, Donohoe BS, Crowley MF, Kataeva IA, Yang SJ, Resch MG, Adams MW, Lunin VV, et al. Revealing nature's cellulase diversity: the digestion mechanism of *Caldicellulosiruptor bescii* CelA. Science. 2013;342(6165):1513–6.
44. Brunecky R, Alahuhta M, Sammond DW, Xu Q, Chen M, Wilson DB, Brady JW, Himmel ME, Bomble YJ, Lunin VV. Natural diversity of glycoside hydrolase family 48 exoglucanases: insights from structure. Biotechnol Biofuels. 2017;10(1):274.
45. Breves R, Bronnenmeier K, Wild N, Lottspeich F, Staudenbauer WL, Hofemeister J. Genes encoding two different beta-glucosidases of *Thermoanaerobacter brockii* are clustered in a common operon. Appl Environ Microbiol. 1997;63(10):3902–10.
46. Yi Z, Su X, Revindran V, Mackie RI, Cann I. Molecular and biochemical analyses of CbCel9A/Cel48A, a highly secreted multi-modular cellulase by *Caldicellulosiruptor bescii* during growth on crystalline cellulose. PLoS ONE. 2013;8(12):e84172.
47. Chung D, Sarai NS, Knott BC, Hengge N, Russell JF, Yarbrough JM, Brunecky R, Young J, Supekar N, Vander Wall T, et al. Glycosylation is vital for industrial performance of hyperactive cellulases. ACS Sustain Chem Eng. 2019;7(5):4792–800.
48. Yan S, Wu G. Signal peptide of cellulase. Appl Microbiol Biotechnol. 2014;98(12):5329–62.
49. Freudl R. Signal peptides for recombinant protein secretion in bacterial expression systems. Microb Cell Fact. 2018;17(1):52.
50. Krishnappa L, Monteferrante CG, Neef J, Dreisbach A, van Dijl JM. Degradation of extracytoplasmic catalysts for protein folding in Bacillus subtilis. Appl Environ Microbiol. 2014;80(4):1463–8.
51. Pohl S, Bhavsar G, Hulme J, Bloor AE, Misirli G, Leckenby MW, Radford DS, Smith W, Wipat A, Williamson ED, et al. Proteomic analysis of *Bacillus subtilis* strains engineered for improved production of heterologous proteins. Proteomics. 2013;13(22):3298–308.
52. Wang W, Hollmann R, Deckwer WD. Comparative proteomic analysis of high cell density cultivations with two recombinant *Bacillus megaterium* strains for the production of a heterologous dextransucrase. Proteome Sci. 2006;4:19.
53. Forsberg Z, Heggset EB, Eijsink VGH, Aachmann FL, Courtade G. The carbohydrate-binding module and linker of a modular lytic polysaccharide monooxygenase promote localized cellulose oxidation. J Biol Chem. 2018;293:13006–15.
54. Pasari N, Adlakha N, Gupta M, Bashir Z, Rajacharya GH, Verma G, Munde M, Bhatnagar R, Yazdani SS. Impact of module-X2 and Carbohydrate Binding Module-3 on the catalytic activity of associated glycoside hydrolases towards plant biomass. Scientific Rep. 2017;7(1):3700.
55. Reyes-Ortiz V, Heins RA, Cheng G, Kim EY, Vernon BC, Elandt RB, Adams PD, Sale KL, Hadi MZ, Simmons BA, et al. Addition of a carbohydrate-binding module enhances cellulase penetration into cellulose substrates. Biotechnol Biofuels. 2013;6:93.
56. Lambertz C, Garvey M, Klinger J, Heesel D, Klose H, Fischer R, Commandeur U. Challenges and advances in the heterologous expression of cellulolytic enzymes: a review. Biotechnol Biofuels. 2014;7(1):135.
57. Garvey M, Klose H, Fischer R, Lambertz C, Commandeur U. Cellulases for biomass degradation: comparing recombinant cellulase expression platforms. Trends Biotechnol. 2013;31(10):581–93.
58. Bhattacharya AS, Bhattacharya A, Pletschke BI. Synergism of fungal and bacterial cellulases and hemicellulases: a novel perspective for enhanced bio-ethanol production. Biotechnol Lett. 2015;37(6):1117–29.
59. Cornet P, Millet J, Beguin P, Aubert J-P. Characterization of Two Cel (Cellulose Degradation) genes of *Clostridium thermocellum* coding for endoglucanases. Nat Biotechnol. 1983;1(7):589–94.
60. Gusakov AV, Salanovich TN, Antonov AI, Ustinov BB, Okunev ON, Burlingame R, Emalfarb M, Baez M, Sinitsyn AP. Design of highly efficient cellulase mixtures for enzymatic hydrolysis of cellulose. Biotechnol Bioeng. 2007;97(5):1028–38.
61. Amster-Choder O, Houman F, Wright A. Protein phosphorylation regulates transcription of the beta-glucoside utilization operon in *E. coli*. Cell. 1989;58(5):847–55.
62. Harwani D, Zangoui P, Mahadevan S. The beta-glucoside (bgl) operon of Escherichia coli is involved in the regulation of oppA, encoding an oligopeptide transporter. J Bacteriol. 2012;194(1):90–9.
63. Xu Y, Ye BC. GlnR and PhoP regulate beta-glucosidases involved in cellulose digestion in response to nitrogen and phosphate availability. Microbiology. 2018;164(5):779–89.
64. O'Brien RW, Morris JG. Oxygen and the growth and metabolism of *Clostridium acetobutylicum*. J Gen Microbiol. 1971;68(3):307–18.
65. Teather RM, Wood PJ. Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. Appl Environ Microbiol. 1982;43(4):777–80.
66. Zhang YH, Cui J, Lynd LR, Kuang LR. A transition from cellulose swelling to cellulose dissolution by o-phosphoric acid: evidence from enzymatic hydrolysis and supramolecular structure. Biomacromol. 2006;7(2):644–8.
67. Bashir Z, Kondapalli VK, Adlakha N, Sharma A, Bhatnagar RK, Chandel G, Yazdani SS. Diversity and functional significance of cellulolytic microbes living in termite, pill-bug and stem-borer guts. Sci Rep. 2013;3:2558.
68. Miller GL. Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal Chem. 1959;31(3):426–8.

Publisher's Note