



RESEARCH

Open Access

# Microbial $\beta$ -glucosidases from cow rumen metagenome enhance the saccharification of lignocellulose in combination with commercial cellulase cocktail

Mercedes V Del Pozo<sup>1</sup>, Lucía Fernández-Arrojo<sup>1</sup>, Jorge Gil-Martínez<sup>2</sup>, Alejandro Montesinos<sup>3</sup>, Tatyana N Chernikova<sup>4</sup>, Taras Y Nechitaylo<sup>5</sup>, Agnes Waliszek<sup>6</sup>, Marta Tortajada<sup>3</sup>, Antonia Rojas<sup>3</sup>, Sharon A Huws<sup>7</sup>, Olga V Golyshina<sup>4</sup>, Charles J Newbold<sup>7,8</sup>, Julio Polaina<sup>9</sup>, Manuel Ferrer<sup>1\*</sup> and Peter N Golyshin<sup>4,8\*</sup>

## Abstract

**Background:** A complete saccharification of plant polymers is the critical step in the efficient production of bio-alcohols. Beta-glucosidases acting in the degradation of intermediate gluco-oligosaccharides produced by cellulases limit the yield of the final product.

**Results:** In the present work, we have identified and then successfully cloned, expressed, purified and characterised 4 highly active beta-glucosidases from fibre-adherent microbial community from the cow rumen. The enzymes were most active at temperatures 45–55°C and pH 4.0-7.0 and exhibited high affinity and activity towards synthetic substrates such as *p*-nitrophenyl-beta-D-glucopyranoside (*p*NPbetaG) and *p*NP-beta-cellobiose, as well as to natural cello-oligosaccharides ranging from cellobiose to cellopentaose. The apparent capability of the most active beta-glucosidase, herein named LAB25g2, was tested for its ability to improve, at low dosage (31.25 units g<sup>-1</sup> dry biomass, using *p*NPbetaG as substrate), the hydrolysis of pre-treated corn stover (dry matter content of 20%; 350 g glucan kg<sup>-1</sup> dry biomass) in combination with a beta-glucosidase-deficient commercial *Trichoderma reesei* cellulase cocktail (5 units g<sup>-1</sup> dry biomass in the basis of *p*NPbetaG). LAB25g2 increased the final hydrolysis yield by a factor of 20% (44.5 ± 1.7% vs. 34.5 ± 1.5% in control conditions) after 96–120 h as compared to control reactions in its absence or in the presence of other commercial beta-glucosidase preparations. The high stability (half-life higher than 5 days at 50°C and pH 5.2) and 2–38000 fold higher (as compared with reported beta-glucosidases) activity towards cello-oligosaccharides may account for its performance in supplementation assays.

**Conclusions:** The results suggest that beta-glucosidases from yet uncultured bacteria from animal digestomes may be of a potential interest for biotechnological processes related to the effective bio-ethanol production in combination with low dosage of commercial cellulases.

**Keywords:** Beta-glucosidases, Bio-ethanol, Glycosyl hydrolase, Lignocellulose, Metagenome, Rumen

\* Correspondence: mferrer@icp.csic.es; p.golyshin@bangor.ac.uk

<sup>1</sup>CSIC, Institute of Catalysis, 28049, Madrid, Spain

<sup>4</sup>School of Biological Sciences, Bangor University, LL57 2UW, Gwynedd, UK

Full list of author information is available at the end of the article

## Background

Cellulose and hemicellulose are the major components of plant cell walls and the most abundant biopolymeric materials on our Planet [1-3]. The natural breakdown of plant matter performed by hemi-cellulases [4,5], has been exploited by biotechnologists to produce bio-fuels, e.g. bio-ethanol [5-7]. Although existing technological developments for biomass pre-treatment facilitate the saccharification process [8], there are some technological barriers hindering an efficient scale-up of the technology. First, the high costs and the low specific activity of the enzymatic cocktails available and actually used; second, the limited number of companies producing the enzymes for the saccharification of plant polymers. These factors weaken the worldwide production of bio-ethanol from plant matter and increase the costs of bio-ethanol production [2]. The high costs of enzymes for biomass decomposition have been identified as a major impediment to the economic conversion of lignocellulosic feedstocks into the bio-ethanol [9,10]. In that context, in 2004 the National Laboratory of Renewable Energy of USA, in association with the two major producers of industrial enzymes (Genencor International and Novozymes A/S), have achieved a drastic reduction in the cellulase price, namely by a factor of 20–30. The new technology utilized an enzyme cocktail with three major enzymes needed to hydrolyse cellulose (endoglucanases, exoglucanases and  $\beta$ -glucosidases). The two first enzymes, working in a synergetic manner, perform the hydrolysis of cellulose chains to cellobiose, destroying the crystalline structure, while the third enzyme transforms the cellobiose to glucose [5]. Therefore, it is crucial to further reduce the enzyme costs in order to develop and maintain sustainable industrial installations that would use cellulose for the bio-ethanol production as common as petroleum in refineries [11]. For example, in industrial settings that use the plant starch-rich materials instead of cellulose as the feedstock, the cost of amylase needed for the degradation of starch is lower than 0,002-0,004 € per litre; by contrast, the most affordable enzyme cocktails that use cellulose do cost 10 times higher [12]. Furthermore, it is worth mentioning that to hydrolyse 1 g of pre-treated plant biomass *ca.* 25 mg enzyme cocktail is needed and that the actual market objective is to achieve a decrease in the enzyme cost from 250 € to 100 € per m<sup>3</sup> of produced bio-ethanol.

To deal with the objective of reducing the enzyme costs, two different approaches are complementarily used, namely, the production of new or improved transgenic yeasts/fungi capable to secrete wild-type or engineered enzymes, and the identification of new or improved enzymes in combination with low-dosage of commercial cellulases. Since the currently used commercial cellulases generally exhibit low  $\beta$ -glucosidase activity

(despite recent developments), it has been suggested that the  $\beta$ -glucosidase of the glycosyl hydrolase family 3 (GHF3;  $\beta$ -glucosidase/xylosidase) acting in the degradation of cellulosic glucans can be used for the saccharification of intermediate gluco-oligosaccharides produced by endoglucanases [13]. Certainly, the addition of  $\beta$ -glucosidases, endoxylanases and  $\alpha$ -arabinofuranosidases for the hydrolysis of steam-exploded wheat straw has been found to improve the hydrolysis from 10% to 29.5% as compared to the non-supplemented reactions [14-17].

Bacterial GHF3 contain about 2,081 known enzymes (<http://www.cazy.org>; [18]). Microbial communities operating in gastrointestinal tracts of herbivorous animals, termites and earthworms, are continuously exposed to a strong diet-driven selective pressure by chemically diverse and complex plant polymeric compounds, and thus represent a rich hotspot for diverse  $\beta$ -glucosidases. The latter can be identified and characterized using functional metagenomics approach, a promising technique to assess the genetic content of complex microbial communities without culturing individual members of microbial communities. Nevertheless, despite recent discoveries and availability of novel  $\beta$ -glucosidases from uncultured microbes (for examples see ref. [19-24]), their potential for the enzymatic conversion of lignocellulose into fermentable sugars has not been assessed as yet.

Here, four  $\beta$ -glucosidases from cow rumen microbial community were identified and characterized, and their potential utilization for supplementation of enzyme cocktails for hydrolysis of plant biomass, namely pre-treated corn stover, was tested. Our findings are pointing at the importance of rumen digestomes as rich resources for novel lignocellulases potentially useful for the effective enzymatic conversion of plant biomass into fermentable sugars.

## Results

### Library screening and general characteristics

The rumen libraries (approx. 17,000 clones harbouring *ca.* 600 Mbp of metagenomic DNA) were screened for the ability to hydrolyse *p*-nitrophenyl- $\beta$ -D-glucopyranoside (*p*NP $\beta$ G) and *p*-nitrophenyl- $\beta$ -D-cellobioside (*p*NP $\beta$ C). We identified three positives (designated SRF2, LAB20 and LAB25) as being highly active against both substrates. The fosmids with inserts SRF2 (34,226 bp; G + C content of 50.04%), LAB20 (38,523 bp; G + C 61.78%) and LAB25 (6,892 bp; G + C 52.16%) were fully sequenced [NCBI acc. nr: JX163905, JX163906 and JX163904, in the same order]. In total, 68 predicted coding sequences (CDS) were identified, which were named according to the fosmid ID and the number of the CDS in the genomic fragment sequenced (provided in Additional file 1: Table S1). Eight of them corresponded to putative glycosyl hydrolase (GH)-like polypeptides with

fosmids SRF2, LAB20 and LAB25 containing 4, 3 and 1, in the same order. Those included a GHF31  $\alpha$ -glucosidase (SRF2g9), four GHF3  $\beta$ -glucosidases (SRF2g14, SRF2g18, LAB25g2 and LAB20g4), a GHF10  $\beta$ -xylanase (LAB20g10), an  $\alpha$ -amylase/ $\beta$ -xylanase/carbohydrate esterase (SRF2g13) and a GHF43/62/32/68  $\beta$ -1,4-xylosidase/ $\alpha$ -L-arabinofuranosidase (LAB20g2). These putative proteins exhibited from 55 to 92% amino acid (AA) sequence identities and 70-96% AA sequence similarities (provided in Additional file 1: Table S1) to homologous proteins found in genomes of bacteria of the Bacteroidetes phylum (e.g. *Prevotella ruminicola*). Those microbes are known to be abundant in the ruminal environment and are thought to play key roles in the breakdown of proteins and carbohydrate polymers [25].

#### Catalytic parameters of GHF3 $\beta$ -glucosidases from cow rumen

The four putative GHF3  $\beta$ -glucosidases (SRF2g14, SRF2g18, LAB20g4 and LAB25g2) were cloned, expressed in *Escherichia coli* BL21 (DE3) and purified. Furthermore, their activities were tested with a battery of substrates and the half-saturation (Michaelis) coefficient ( $K_m$ ), the catalytic rate constant ( $k_{cat}$ ), the catalytic efficiency ( $k_{cat}/K_m$ ) values and/or specific activities, were determined.

As shown in Table 1, in terms of catalytic efficiencies, *p*NP $\beta$ G and *p*NP $\beta$ C were the preferred substrates for 3 and 1  $\beta$ -glucosidases, respectively:  $[(k_{cat}/K_m)]_{pNP\beta G}/[(k_{cat}/K_m)]_{pNP\beta C}$  factor of  $\sim 26/1$  (for SRF2g14),  $0.7/1$  (for SRF2g18),  $356/1$  (for LAB20g4) and  $2/1$  (for LAB25g2). All enzymes showed higher  $k_{cat}$  values for *p*NP $\beta$ G (from 2.0 to 7.3-fold), and therefore, the divergence in catalytic efficiencies were mainly due to differences in substrate affinity: 2 enzymes (SRF2g14 and LAB20g4) showed a significantly (from 14 to 49-fold) lower  $K_m$  value for *p*NP $\beta$ G substrate whereas LAB25g2 showed similar affinities and SRF2g18 circa 4 times higher affinity for

*p*NP $\beta$ C. The catalytic efficiency ( $k_{cat}/K_m$ ) while using the non activated substrate cellobiose was from 44 to 1026-fold lower than that found for the best *p*NP substrate, mainly due to a significant increase of  $K_m$  values (from 7 to 260-fold, depending on the enzyme) for the disaccharide. Enzymes SRF2g14 and SRF2g18 were the most- and less-efficient, respectively, for the hydrolysis of the disaccharide ( $k_{cat}/K_m$  ratio for SRF2g14:SRF2g18 of 14:1).

The purified recombinant  $\beta$ -glucosidases were also assayed for their activities toward different polymeric substrates. By meaning of specific activity determination,  $\beta$ -glucosidases hydrolysed all short cello-oligosaccharides tested (degree of polymerisation [DP] from 2 to 5), with longer substrates being preferred (Table 2). According to the activity level the following order could be established: LAB25g2 > SRF2g18 > SRF2g14 > LAB20g4, with LAB25g2 being 9-fold more active against cellopentaose as compared to LAB20g4. Two out of four  $\beta$ -glucosidases exhibited activity against lichenan, namely, SRF2g14 and LAB20g4 (Table 2), suggesting that they are able to hydrolyze substrates with mixed  $\beta$ -1,3/4 linkages; a  $[(mU\ mg^{-1})]_{cellobiose}/[(mU\ mg^{-1})]_{lichenan}$  factor of  $\sim 6/1$  (for SRF2g14) and  $13/1$  (for LAB20g4) was observed. No activity was detected using avicel, filter paper, as well as towards substrates without  $\beta$ -1,4 linkages such as  $\beta$ -1,3 glucan, or mixed  $\beta$ -1,3/6 linkages like laminarin. Accordingly, the enzymes showed a clear preference for short cello-oligosaccharide substrates, which may likely be produced in natural settings from the cellulose components of plant cell walls due to the action of glucanases.

#### Effect of temperature and pH on specific activity and stability

The optimum activity for GHF3  $\beta$ -glucosidases was observed within a narrow range of temperatures, between 45 and 55°C, and within a narrow pH range (4.0-7.0) being most active at pH close to 5.0 (Figure 1). Stability assays (measured at 50°C and pH 5.2 in conditions reassembling the supplementation assays – see details below) revealed LAB25g2 to be substantially stable, with only 18% loss of activity after 5 days of incubation; under similar conditions, the half life for SRF2g14, SRF2g18 and LAB20g4 was  $18.06 \pm 1.9$  h,  $37.5 \pm 3.9$  h and  $136.1 \pm 9.4$  h, respectively (Figure 1C). The stability features, namely for LAB25g2 enzyme, are better than those of many other microbial  $\beta$ -glucosidases (provided in Additional file 2: Table S2).

#### Enzymatic conversion of lignocellulose in supplementation assays

Saccharification of the cellulose in biomass results in sugar-rich liquid streams useful for the production of bio-ethanol [5]. To estimate the capability of herein discovered  $\beta$ -glucosidases from cow rumen in such process, crude extracts of *E. coli* BL21 (DE3) cells expressing  $\beta$ -

**Table 1 Kinetic parameters of the purified  $\beta$ -glucosidases**

Enzyme	Substrate	$K_m$ (mM) <sup>a</sup>	$k_{cat}$ (s <sup>-1</sup> ) <sup>a</sup>	$k_{cat}/K_m$ (s <sup>-1</sup> M <sup>-1</sup> ) <sup>a</sup>
SRF2g14	<i>p</i> NP $\beta$ Glu	0.034 $\pm$ 0.003	0.37 $\pm$ 1.34	10.76 $\times$ 10 <sup>3</sup>
	<i>p</i> NP $\beta$ C	0.48 $\pm$ 0.054	0.19 $\pm$ 0.36	4.16 $\times$ 10 <sup>2</sup>
	Cellobiose	8.02 $\pm$ 0.939	0.38 $\pm$ 0.55	47.64
SRF2g18	<i>p</i> NP $\beta$ Glu	1.73 $\pm$ 0.004	0.27 $\pm$ 0.32	1.59 $\times$ 10 <sup>2</sup>
	<i>p</i> NP $\beta$ C	0.48 $\pm$ 0.072	0.11 $\pm$ 0.22	2.32 $\times$ 10 <sup>2</sup>
	Cellobiose	25.51 $\pm$ 2.46	0.09 $\pm$ 0.31	3.35
LAB20g4	<i>p</i> NP $\beta$ Glu	0.03 $\pm$ 0.003	0.88 $\pm$ 4.05	29.49 $\times$ 10 <sup>3</sup>
	<i>p</i> NP $\beta$ C	1.47 $\pm$ 0.13	0.12 $\pm$ 0.20	82.73
	Cellobiose	7.80 $\pm$ 0.67	0.22 $\pm$ 0.35	28.75
LAB25g2	<i>p</i> NP $\beta$ Glu	0.45 $\pm$ 0.06	0.41 $\pm$ 0.68	9.20 $\times$ 10 <sup>2</sup>
	<i>p</i> NP $\beta$ C	0.39 $\pm$ 0.03	0.17 $\pm$ 0.36	4.38 $\times$ 10 <sup>2</sup>
	Cellobiose	4.88 $\pm$ 0.59	0.10 $\pm$ 0.23	20.84

<sup>a</sup>Parameters were determined at 40°C and pH 5.6 (50 mM sodium acetate) using substrate and enzyme concentrations as described in Methods.

**Table 2 Specific activity of the purified  $\beta$ -glucosidases towards various cello-oligosaccharides (degree of polymerisation [DP] from 2 to 5) and carbohydrate polymers**

Substrate	Specific activity (U mg <sup>-1</sup> )			
	Pure $\beta$ -glucosidases			
	SRF2g14	SRF2g18	LAB20g4	LAB25g2
Cellulose <sup>a</sup>	369.9 ± 0.04	852.8 ± 0.03	305.9 ± 0.05	1537.2 ± 0.01
Cellotriose <sup>a</sup>	317.6 ± 0.04	789.4 ± 0.03	60.2 ± 0.01	2170.5 ± 0.02
Cellotetraose <sup>a</sup>	343.0 ± 0.01	944.4 ± 0.01	283.9 ± 0.01	2154.3 ± 0.01
Cellopentaose <sup>a</sup>	397.4 ± 0.001	1154.8 ± 0.03	313.5 ± 0.02	2682.6 ± 0.01
Lichenan <sup>a</sup>	60.8 ± 0.13	n.d	22.8 ± 0.16	n.d
<i>E. coli</i> BL21 (DE3) cell extracts expressing $\beta$ -glucosidases <sup>b</sup>				
pNP $\beta$ Glu <sup>c</sup>	0.39	5.1 × 10 <sup>-3</sup>	0.40	1.5
Cellulose <sup>d</sup>	115.2	16.2	105.9	231.5

<sup>a</sup>Reaction assays contain 2  $\mu$ g pure enzyme, 200  $\mu$ L 50 mM sodium acetate buffer (pH 5.6) and 100  $\mu$ L of a 100 mM cellobiose stock solution (final concentration of cellobiose of 20 mM). For lichenan a concentration of 10 mg ml<sup>-1</sup> was used. Samples were incubated at 40°C during 5 h, after which the glucose concentration was measured using the glucose oxidase kit (Sigma Chemical Co., St. Louis, MO, USA). All of the measurements were analysed in triplicate; n.d. no activity detected.

<sup>b</sup>Quantification of proteins in *E. coli* BL21 (DE3) cell extracts expressing  $\beta$ -glucosidases was performed according to BCA Protein Assay kit (Thermoscientific), and are as follows: 552, 426, 458 and 508  $\mu$ g ml<sup>-1</sup> for cell extracts of SRF2g14, SRF2g18, LAB20g4 and LAB25g2, respectively.

<sup>c</sup>Activity measured by using 100  $\mu$ L cell extract, 50  $\mu$ L pNP $\beta$ G solution (0.22% w/v) and 150  $\mu$ L buffer McIlvane (100 mM, pH 5). Incubation: 20 min at 40°C, after which 250  $\mu$ L Na<sub>2</sub>CO<sub>3</sub> (1 M) was added and absorbance measured at 405 nm. Specific activity was measured as follows: mU ml<sup>-1</sup> enzyme = A/KtV (A = absorbance at 405 nm; K = 0.0328 (value determined by calibration); V = volume in ml. The standard deviation was lower than 5%.

<sup>d</sup>Reaction assays contain 200  $\mu$ L enzyme extract, 200  $\mu$ L McIlvane buffer and 100  $\mu$ L of a 100 mM cellobiose stock solution (final concentration of cellobiose of 20 mM). Sample were incubated at 40°C during 4 h, after which the glucose concentration was measured using the glucose oxidase kit (Sigma Chemical Co., St. Louis, MO, USA). The standard deviation was lower than 5%.

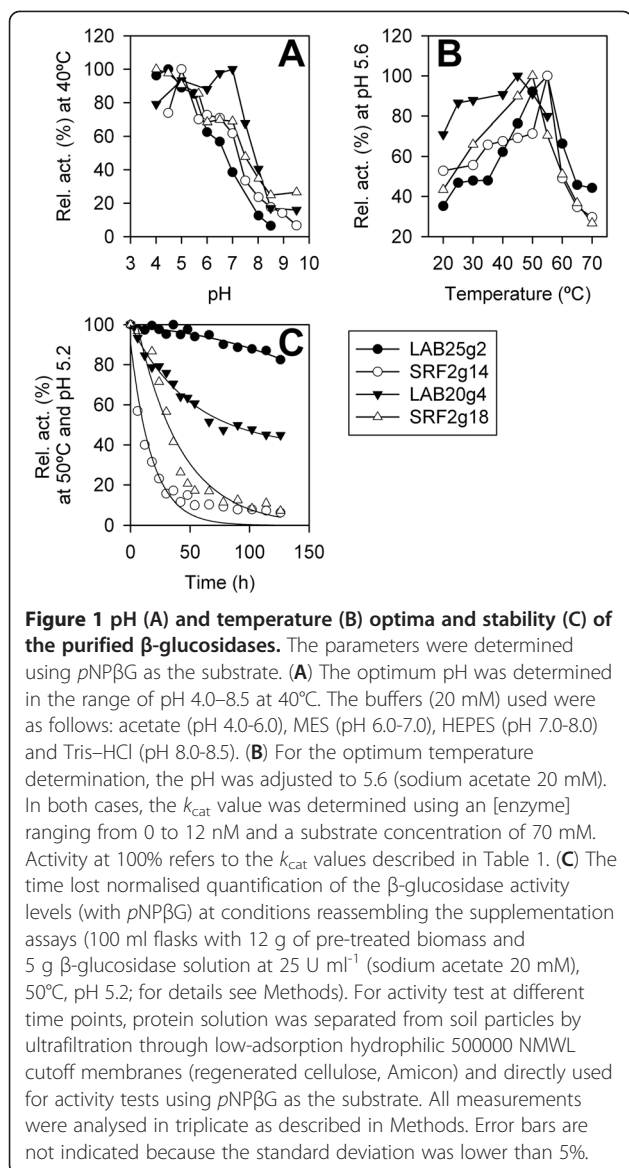
glucosidases were produced as described in Methods and used in saccharification of pre-treated corn stover biomass. First, their protein concentrations and activity levels using pNP $\beta$ G and cellobiose as substrates, were examined. As shown in Table 2 the cell extract containing  $\beta$ -glucosidase LAB25g2 showed the higher activity level toward both substrates which agrees with activity tests using pure proteins; using cellobiose (58 g l<sup>-1</sup>) as substrate,  $\beta$ -glucosidase LAB25g2 produced approx. 2.5 g glucose mg<sup>-1</sup> cell extract h<sup>-1</sup> at 40°C and pH 5.0. The differences in activity performance for the three other  $\beta$ -glucosidases when using cell extracts as compared to assays using pure proteins may be due to differences in expression level and protein solubility during production in *E. coli* BL21 (DE3), which was evidenced by SDS-PAGE analysis (not shown).

Crude extract of  $\beta$ -glucosidase LAB25g2 was then selected to investigate its effect in supplementation assays under the standard industrial conditions described in Methods: 50°C and pH 5.2. Crude cell extracts containing LAB25g2 (31.25 units g<sup>-1</sup> dry biomass, using pNP $\beta$ G as substrate, or 89.28 units g<sup>-1</sup> glucan) were supplemented to  $\beta$ -glucosidase-deficient commercial cellulase product Celluclast, derived from *Trichoderma reesei* (Novozymes A/S, Bagsvaerd, Denmark) (5 units in the basis of pNP $\beta$ G hydrolysis g<sup>-1</sup> dry biomass); control reactions without LAB25g2 as well as with two commercial  $\beta$ -glucosidase preparations [G0395 from almond (Sigma Chemical Co., St. Louis, MO, USA) and E-BGOSAG from *Agrobacterium* sp. (Megazyme; Bray, Ireland)] were also performed

under similar conditions. It should be noticed that the glucan content of the biomass used is about 35% of dry matter content (measured accordingly to the Standard Biomass Analytical Procedures; [http://www.nrel.gov/biomass/analytical\\_procedures.html](http://www.nrel.gov/biomass/analytical_procedures.html)), that is, 350 g glucan kg<sup>-1</sup> dry matter content.

Figure 2A shows the enzymatic hydrolysis profiles of control and supplementation assays. As shown, LAB25g2 addition resulted in higher production of glucose and higher cellobiose consumption, which was noticeable above 48 h of treatment, and optimal at 96 h: glucose production increased from 196.3 ± 6.0 to 234.8 ± 6.4 g glucose kg<sup>-1</sup> dry biomass, most likely as a consequence of the improved cellobiose hydrolysis by the addition of LAB25g2  $\beta$ -glucosidase. Thus, whereas control reaction contained 79.3 ± 8.0 g cellobiose kg<sup>-1</sup> dry biomass, no appreciable amount was detected in supplemented reaction. It should be noticed that other commercial  $\beta$ -glucosidase preparations such as those from almond (G0395; Sigma Chemical Co.; 42 and 2.3 units mg<sup>-1</sup> using pNP $\beta$ G and cellobiose, respectively at 50°C and pH 5.0) and *Agrobacterium* sp. (E-BGOSAG; Megazyme; (119 ± 8) × 10<sup>3</sup> and 0.29 units mg<sup>-1</sup> using pNP $\beta$ G and cellobiose, respectively at 50°C and pH 5.0), did not produce hydrolysis improvements under similar conditions as those used with LAB25g2 (not shown). To calculate the saccharification yield at an indicated time (SYt) we applied the following formula: SYt (%) = [Glucose at t (g kg<sup>-1</sup>) - Glucose initial (g kg<sup>-1</sup>)] / [Glucose maximal (g kg<sup>-1</sup>)] × 100. According to this equation,





LAB25g2 increased the final hydrolysis yield by a factor of 20% ( $44.5 \pm 1.7\%$  vs.  $34.5 \pm 1.5\%$  in control conditions) after 96–120 h treatment (Figure 2B). This suggests the potential of accessory  $\beta$ -glucosidase activity from cow rumen to improve the enzymatic hydrolysis of pre-treated corn stover.

### 3D structural analysis of biochemically characterised GHF3

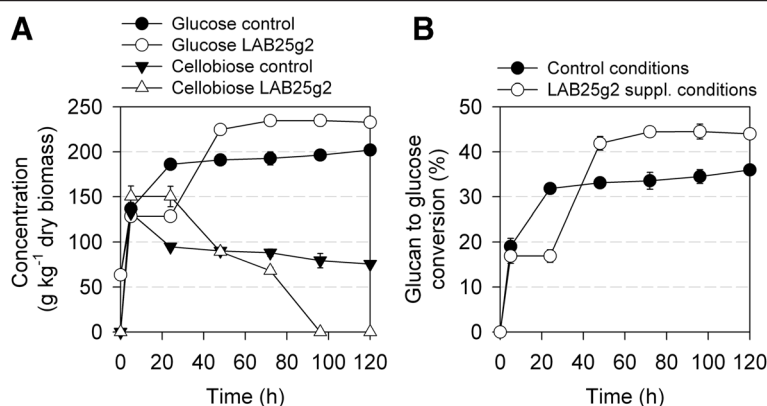
Enzymes belonging to family GHF3 are multi-domain proteins whose domains may differ in number and arrangement [26–28]. The four enzymes analysed in this work show good end-to-end sequence alignment (ca. 50% identity) and consequently high structural resemblance with the  $\beta$ -glucosidase of *Thermotoga napolitana* [26]. As shown in Figure 3, they are composed by two

domains, an  $\alpha/\beta$  sandwich and an  $\alpha/\beta$  barrel that contribute each one of the two catalytic residues, and a carboxy-terminal fibronectin-like domain, present in some but not all enzymes of this family. As expected by the high similarity of the four protein sequences, the modelling yielded nearly superimposable structures with no obvious trait (eg. amino acid insertions, loops) that could account for the observed difference in catalytic efficiency. Therefore, the structural basis of the superior performance of LAB25g2 may rely on subtle amino acid differences, possibly of residues located in the vicinity of the catalytic site, involved in substrate binding. The models should be helpful in pinpointing these residues whose role can be tested by site directed mutagenesis.

### Discussion

In the present work, an expression metagenomic library was used to identify  $\beta$ -glucosidases in the DNA fragments from uncultured microorganisms populating the cow rumen. We detected four  $\beta$ -glucosidases that have been subcloned, expressed, purified and characterised; these enzymes were likely derived from representatives of the genus *Prevotella ruminicola* (Bacteroidetes phylum) known to be abundant in the ruminal environment [25].

$\beta$ -Glucosidases could be employed in combination with low dosages of commercial cellulase enzymes, in applications such as bio-ethanol production [14,16], which requires efficient enzyme cocktails to achieve a high degree of conversion of lignocellulose [9,10,29–31]. Bioconversion of lignocellulose into bio-ethanol consists of four major steps: pre-treatment, hydrolysis, fermentation, and product separation/distillation [32]. Cellulases and hemicellulases act synergistically in the second step of enzymatic hydrolysis. In this study, LAB25g2  $\beta$ -glucosidase showed synergistic effects with Celluclast, one of the most common enzymatic cocktails actually produced (Novozymes A/S, Bagsvaerd, Denmark), which can release fermentable sugars from oligosaccharides obtained during the breakdown of lignocellulose by endo-hydrolases. Such synergistic effect is of a special interest as it is known that cellulose is difficult to be hydrolysed even after pre-treatment [2,3]. Although the enzyme LAB25g2 did not show appreciable activity towards filter paper where the cellulose is mostly crystalline, it would promote the hydrolysis of short- to medium-sized cello-oligosaccharides (Table 2) that are likely produced by endo- and exo-glucanases present in the commercial cocktail. The final concentration of glucose after 96 h of treatment was  $234.8 \pm 6.4$  g glucose kg<sup>-1</sup> dry biomass which is approximately equal to 20% increase in the glucan hydrolysis yield. More interestingly, the enzyme was also able to contribute to the complete hydrolysis of cellobiose in the slurry (containing 20% dry biomass) after 96 h hydrolysis. This result

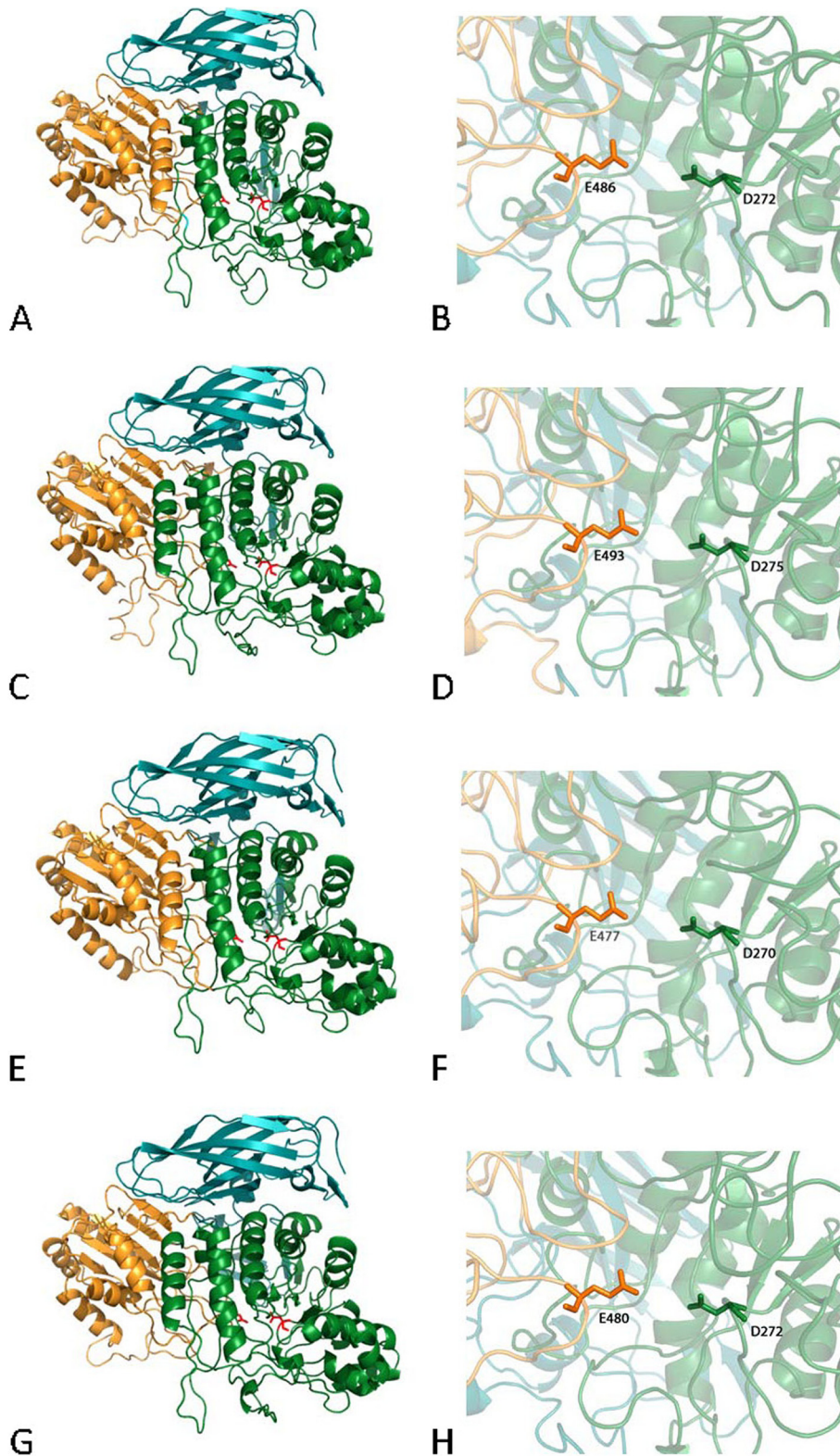


**Figure 2 Supplementation of commercial cellulases with  $\beta$ -glucosidase LAB25g2.** The dosage of  $\beta$ -glucosidases was performed on the basis of same activity towards pNP $\beta$ G, namely, 31.25 U g<sup>-1</sup> dry biomass (minimum dose of LAB25g2 to achieve saccharification of 20% w/w pre-treated corn stover under assay conditions). Saccharification of lignocellulose is shown by meaning of glucose and cellobiose concentration in assay tests (A) or glucan to glucose conversion (B). Hydrolysis reactions: 100 ml flasks, pH 5.2, 12 g of pre-treated corn stover, 3 g of enzymatic mixture (Celluclast) and 5 g LAB25g2  $\beta$ -glucosidase solution at 25 U ml<sup>-1</sup> (in the basis of pNP $\beta$ G assay). Control reactions in the absence of LAB25g2 were performed. Commercial  $\beta$ -glucosidases G0395 (from almond) and E-BGOSAG (from *Agrobacterium* sp.) were used as control test as for LAB25g2. For that protein solutions at 25 U ml<sup>-1</sup> were prepared and added to reaction mixtures to achieve a final activity of 31.25 U g<sup>-1</sup> dry biomass (pNP $\beta$ G), as for LAB25g2. Specific activity of those preparations is shown in Additional file 2: Table S2. In both cases, results were similar to those found in the control experiment without  $\beta$ -glucosidase and are not shown. Determination of glucose and cellobiose concentrations was followed by RID-HPLC. Glucose concentration was also measured by the glucose oxidase-peroxidase D-Glucose Assay Kit (Megazyme). All measurements were analysed in triplicate as described in Methods. Error bars are indicated. Note: under similar dosage conditions  $\beta$ -glucosidase Novo-188 (Novozymes A/S) consumed all cellobiose in the assay after 24 h (not shown).

seems to be highly valuable, since *Saccharomyces* strains are commonly used for fermentation of ethanol and they utilize glucose more efficiently than cellobiose, because they have no  $\beta$ -glucosidases [7]. Also, LAB25g2 displayed a high stability under experimental conditions ( $t_{1/2} > 5$  days at 50°C and pH 5.2; Figure 1C) likely similar to those required for the industrial bio-ethanol production.

To understand the reasons of the high performance of LAB25g2 in supplementation assays, a detailed comparison of the properties and kinetic studies of about hundred  $\beta$ -glucosidases from various sources, in which affinity constants ( $K_m$ ), maximal specific activity ( $V_{max}$ ), catalytic rate constant ( $k_{cat}$ ), catalytic efficiency ( $k_{cat}/K_m$ ) and/or specific activities (units mg<sup>-1</sup>) reported for assays using pNP $\beta$ G, pNP $\beta$ C and cello-oligosaccharides ranging from cellobiose to cellohexaose, was undertaken (data provided in Additional file 2: Table S2). They include those from known bacteria, archaea, yeasts, fungi, plants and insects, as well as from unknown and uncultured microbial resources. The  $\beta$ -glucosidase properties from *Aspergillus niger*, generally used to complement the cellulolytic cocktail of *Trichoderma reesei* [13] actually used in hydrolysis tests and commercial preparations, were also included. Compared with reported or commercially available  $\beta$ -glucosidases ( $K_m$ : 0.004-68 mM for pNP $\beta$ G; 4.8-17.6 mM for pNP $\beta$ C; and 0.31-66 mM for cellobiose) our enzymes exhibit similar affinity towards synthetic pNP $\beta$ G (from 0.034 to 1.73 mM) and pNP $\beta$ C (from 0.48 to 1.47) as well as for cellobiose (from 4.88 to

25.51 mM). The  $k_{cat}$  values ( $\leq 0.88$  s<sup>-1</sup>) were in the same range or lower than those reported (from 0.022 to 87400 s<sup>-1</sup> for pNP $\beta$ G, 18.2-17500 s<sup>-1</sup> for pNP $\beta$ C and 4.5-699 s<sup>-1</sup> for cellobiose). However, the enzyme kinetic parameters from previous studies demonstrated significant differences in their hydrolytic capacity against natural cello-oligosaccharides, with notably better values for the  $\beta$ -glucosidases reported in present study. Thus, for the enzyme LAB25g2, the specific activity for cellobiose (1537.2 units mg<sup>-1</sup> when used as pure protein or 231.5 units mg<sup>-1</sup> as crude *E. coli* BL21 (DE3) cell extracts expressing  $\beta$ -glucosidase) was up to 38000 fold higher, as compared with the reported or commercially available  $\beta$ -glucosidases (see Additional file 2: Table S2); only the  $\beta$ -glucosidase from the fungus *Penicillium funiculosum* NCL1 (1796 units mg<sup>-1</sup>) had a close activity, followed at much lower extent, by that of *Aspergillus oryzae* (938 units mg<sup>-1</sup>). Additionally, it is noteworthy that LAB25g2 did show 27-1166 fold higher specific activity towards cellotri-ose-cellopentaose (up to 2682.6  $\pm$  0.01 units mg<sup>-1</sup>), as compared to known  $\beta$ -glucosidase, with that of *Aspergillus oryzae* (512 units mg<sup>-1</sup>), being the most-active among enzymes and commercial preparations reported so far. Compared with Novozymes 188, a  $\beta$ -glucosidase derived from *Aspergillus niger* and one of the most commonly used  $\beta$ -glucosidases in lignocellulose conversion [15], we found that crude extract and pure preparations of LAB25g2 were 7 and 46 fold, respectively, more active towards



**Figure 3** Structural models of  $\beta$ -glucosidases characterised in this work: LAB25g2 (A and B,) SRFg14 (C and D), SRFg18 (E and F) and LAB20g4 (G and H). Panels A, C, E and G show overall representation of the enzyme structures. The three protein domains,  $\alpha/\beta$  barrel,  $\alpha/\beta$  sandwich and fibronectin-like, are coloured green, yellow and blue, respectively. Panels B, D, F and H show closeup of the catalytic centres, indicating the position of the catalytic nucleophile (Asp) and proton donor (Glu).



cellobiose (see Additional file 2: Table S2). Accordingly, to the best of our knowledge, LAB25g2 has the highest activity reported for short-chain cello-oligosaccharides. This finding is of a special significance, as the activity towards cello-oligosaccharides has only been reported for a limited number for  $\beta$ -glucosidases.

Results further indicated that LAB25g2 crude extract preparation was more effective than two commercial  $\beta$ -glucosidases (selected on the basis of their high  $\beta$ -glucosidase activity as compared to other commercially available preparations) for the saccharification of pre-treated corn stover, as they both produced similar results as the control tests without supplement. This may be described to the activity characteristics differences. Thus, using *pNP* $\beta$ G as substrates both enzymes were 28 (from almond) and 79000 (from *Agrobacterium* sp.) times more active than crude extract preparation of LAB25g2; however, using cellobiose, almond and *Agrobacterium* sp. preparations were 100–798 less active than LAB25g2 preparation. Additionally, whereas almond  $\beta$ -glucosidase did show low activity at 50°C and pH 5.0 that from *Agrobacterium* sp. was quite stable (Additional file 3: Figure S1), indicating that both activity and stability features accounted for the LAB25g2 performance. It has been also reported that Novozymes 188 decreased the remaining cellulose concentration in supplementation assays using *T. reesei* cellulase by 10.1% (dry matter content of 10%; 37°C, 120 h) at low dosage (3.76 units g<sup>-1</sup> dry biomass, using *pNP* $\beta$ G as substrate) [15]; this factor is about half of that produced by LAB25g2, although it should be taken into consideration that in our experiments 8.3-fold higher  $\beta$ -glucosidase supplemented activity was used. Under similar dosage conditions  $\beta$ -glucosidase Novo-188 (Novozymes A/S) was found to consume all cellobiose in our experimental conditions after 24 h (not shown), whereas LAB25g2 required several days. Other characteristics of LAB25g2  $\beta$ -glucosidase (i.e., optimum temperature, 50°C; pH 4.5–5.5, and high stability under supplementation assay conditions (50°C, pH 5.2)) suggest this enzyme has an outstanding position among the reported  $\beta$ -glucosidases isolated from different sources. Most likely, the high performance of LAB25g2 towards oligosaccharides occurring during the breakdown of lignocellulose by exo- and endo-hydrolases, together with its thermostability, seem to be major factors contributing to the biomass hydrolysis improvement when used in combination with the  $\beta$ -glucosidase-deficient commercial cocktail Celluclast (Novozymes A/S).

The discovery of a high performance novel  $\beta$ -glucosidase for the saccharification of lignocellulose under industrial operational conditions is a clear example of the utility of function-centred enzyme discovery in complex microbial communities. The natural selection by the great

polymeric substrate diversity imposed on a complex microbial community is likely a key factor driving the evolution of family 3  $\beta$ -glucosidases, which has been recognized to represent about 19% of the total carbohydrate hydrolases in the bovine rumen [25]. Taking into account that many proteins of this family share a significant degree of homology with enzyme LAB25g2 reported here, we suggest that the enzymatic potential of the microorganisms populating animal gastrointestinal tracts remains underestimated and underexploited and that animal digestomes may be a potent bio-resource for novel lignocellulases for bio-ethanol production. Although, the yield improvements reported here are far from being scaled up at the industrial level (and are lower than that reported for Novo-188 [15], one of the most commonly used  $\beta$ -glucosidases in lignocellulose conversion), the present study highlights the need for more extensive experimental work to accurately identify enzymes potentially applicable for bio-fuel production [33]. That might also result in the production of efficient setups or *a-la-carte* cocktails mixed in-house by own technological facilities upon demand and upon the choice of the feedstock.

## Conclusions

The catalytic performance of four  $\beta$ -glucosidases from cow rumen was examined for their applicability for the saccharification of lignocellulose in combination with a  $\beta$ -glucosidase-deficient commercial cellulase cocktail. One of the  $\beta$ -glucosidases was demonstrated to be effective and thermo-stable for the enzymatic breakdown of pre-treated corn stover, with a substantial role in cellobiose hydrolysis and the concomitantly increase production of glucose. Having said that, recent developments for the production of commercial cocktails with improved  $\beta$ -glucosidases have been achieved and many factors account for the cost-effective production of bio-ethanol; whatever the case, to the best of our knowledge, this study provides the first clear experimental evidences that  $\beta$ -glucosidases from ruminal bacterial metagenome are of a great potential interest for the efficient hydrolysis of lignocellulosic biomass, and confirms the necessity of isolating and characterizing new  $\beta$ -glucosidases from this bio-resource.

## Methods

### Materials and strains

Chemicals, biochemicals and solvents were purchased from Sigma Chemical Co. (St. Louis, MO, USA) and were of p.a. (pro-analysis) quality. The oligonucleotides used for DNA amplification, mutagenesis and sequencing were synthesised by Sigma Genosys Ltd. (Pampisford, Cambs, UK). The restriction and modifying enzymes were from New England Biolabs (Beverly, MA, USA). The Ni-NTA His-Bind chromatographic media were from QIAGEN (Hilden, Germany). The *E. coli*



EPI300-T1<sup>R</sup> strain (Epicentre Biotechnologies; Madison, WI, USA) used for the fosmid library construction and screening and *E. coli* GigaSingles for the cloning and BL21 (DE3) for the expression using the pET-41 Ek/LIC vector (Novagen, Darmstadt, Germany) were cultured and maintained according to the recommendations of the suppliers. *pNP-β-D-glucopyranoside (pNPβG)*, *pNP-β-D-cellobioside (pNPβC)*, the cello-oligosaccharides (ranging from cellobiose to cellopentaose), the D-glucose assay kit and the almond β-glucosidase G0395, were all provided by Sigma Chemical Co. (St. Louis, MO, USA). *Agrobacterium* sp. β-glucosidase E-BGOSAG was provided by Megazyme (Bray, Ireland).

#### Metagenomic library construction and enzyme screening

Rumen contents were collected from four rumen-fistulated, non-lactating Holstein cows (average weight of 731 kg) housed at Trawsgoed experimental farm (Aberystwyth, Ceredigion, Wales). Samples were retrieved under the authorities of the UK Animal (Scientific Procedures) Act (1986). The animals were fed a diet composed of a mixture of grass silage and straw (75:25) *ad libitum* and ~1 kg of sugar beet nuts at 07:00 am; the cows had constant access to fresh water. Sampling was done 2 h after concentrate feeding. Approximately 1 liter of mixed liquid and solid ruminal content was sampled from each cow pooled together to give a composite sample from all cows. The samples were then processed to produce two fractions: strained ruminal fluid (SRF) and liquid-attached bacteria (LAB). For SRF retrieval, total ruminal content was strained through four layers of muslin in order to remove large particles, and SRF was then frozen at -80°C until use. For LAB retrieval, approximately 1 liter of mixed total rumen content was hand squeezed to get rumen liquor, and the solid fraction was put in a large foil tray. The liquid fraction was spun at 2,000 × g, 10 min, 4°C (MSE Europa 24 M, Berthold Hermle KG, Weisbaden, Germany); the supernatant was then strained through a 1 mm<sup>2</sup> pore-sized nylon mesh to remove feed particles, and spun again at 13,000 × g, 25 min, 4°C. The pellet was washed in a saline solution (made from 180 g NaCl dissolved in 20 L distilled water), and subsequently centrifuged at 13,000 × g, 25 min, 4°C. The pellet was re-washed with distilled water, and spun down at 13,000 × g, 15 min, 4°C. The pellet, containing LAB, was then transferred into a sterile jar and kept at -80°C until DNA extraction.

Total DNA was extracted from LAB and SRF microbial communities as described previously [19], using the G'NOME<sup>®</sup> DNA Isolation Kit (Qbiogene, Heidelberg, Germany). Purified and size-fractionated DNA was ligated into the pCCFOS fosmid vector and further cloned in *E. coli* EPI300-T1<sup>R</sup> according to the instructions of Epicentre Biotechnologies (WI, USA) and a procedure

described earlier [22]. Fosmid clones (16,896) harbouring approximately 600 Mbp of community genomes were arrayed using the QPix2 colony picker (Genetix Co., UK) and grown in 384-microtitre plates containing Luria Bertani (LB) medium with chloramphenicol (12.5 μg ml<sup>-1</sup>) and 15% (v/v) glycerol and stored at -80°C.

To screen for GH activity, the clones were plated onto large (22.5 × 22.5 cm) Petri plates with LB agar containing chloramphenicol (12.5 μg ml<sup>-1</sup>) and the induction solution (Epicentre Biotechnologies; WI, USA) as recommended by the supplier to induce a high fosmid copy number. An array of 2,304 clones per plate was set. After overnight incubation, each library was screened for the ability to hydrolyse *pNPβG* and *pNPβC*. For screens, the plates (22 × 22 cm) were covered with an agar buffer substrate solution (40 ml of 50 mM sodium acetate, pH 5.6, 0.4% agarose and 5 mg ml<sup>-1</sup> of *pNPβG* and *pNPβC* as substrates). Positive clones appeared due to the formation of a yellow colour. The positive clones were selected and their DNA inserts fully sequenced with a Roche 454 GS FLX Ti sequencer (454 Life Sciences, Branford, CT, USA) at Life Sequencing S.L (Valencia, Spain).

#### Cloning, expression, purification and characterisation of plant polymeric substance hydrolases

The gene cloning was performed by PCR using goTaq<sup>®</sup> DNA polymerase and custom oligonucleotide primers. To amplify the hydrolase genes, the corresponding fosmid was used as the template; the vectors and the pairs of primers are described as follows: SRF2g14 Fwd 5'-gacgacgaagATGAAGAAGACTCTGTTTTTCGCCTTTGGC-3' and SRF2g14 Rev 5'-gaggagaagcccggTTATCGTTTCAGGAGGTTTCATCTGAACCTGTGG-3' (total gene length 2361 bp); SRF2g18 Fwd 5'-gacgacgacaagATGAGAAAATCGATTTCATCAGATTAGTTTTGG-3' and SRF2g18 Rev 5'-gaggagaagcccggTTATCGTTTCAGTAGGTTGAGTTTCAATTTG-3' (total gene length 2343 bp); LAB20g4 Fwd 5'-gacgacgacaagATGAAGAAAATCATGCTCCTCTCCGC-CACC-3' and LAB20g4 Rev 5'-gaggagaagcccggTTATTT-CATTAAGAGCACGCGGTTGGCGGGC-3' (total gene length 2295 bp); LAB25g2 Fwd 5'-gacgacgacaagATGAAAATACTAACAATTTGCTTCGTAGC-3' and LAB25g2 Rev 5'-gaggagaagcccggTTACTGTTTCAGAAGATTGAGTTTCTGTTTTG-3' (total gene length 2340 bp). The PCR conditions were as follows: 95°C for 120 s, followed by 30 cycles of 95°C for 30s, 55°C for 45 s and 72°C for 120 s, with a final annealing at 72°C for 500 s. The PCR products were analysed from agarose gel-purified using the Mini Elute Gel Purification Kit (Qiagen, Hilden, Germany) and cloned into pET-41-Ek/LIC (Novagen, Darmstadt, Germany) according to the manufacturer's instructions. The resulting plasmids were introduced into the non-expressing *E. coli* GigaSingles host and then,

after plasmid extraction, into *E. coli* BL21 (DE3) for protein expression; the clones were selected on LB agar supplemented with kanamycin ( $30 \mu\text{g ml}^{-1}$ ). For the enzyme expression and purification, clones were grown overnight at  $37^\circ\text{C}$  with shaking at 200 rpm in 100 ml of LB medium containing kanamycin. Afterwards, 25 ml of this culture was used to inoculate 1 l of LB medium, which was then incubated for 4 hours to an  $\text{OD}_{600\text{nm}}$  of  $\sim 0.6$  at  $37^\circ\text{C}$ . Protein expression was induced by 1 mM isopropyl- $\beta$ -D-galactopyranoside (IPTG) followed by incubation for 16 h at  $16^\circ\text{C}$ . The cells were harvested by centrifugation at  $5000 \times g$  for 15 min to yield 2–3 g  $\text{l}^{-1}$  of pellet (wet weight). The cell pellet was frozen at  $-80^\circ\text{C}$  overnight, thawed and resuspended in 10 ml 20 mM 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES) [pH 7.0]  $\text{g}^{-1}$  wet cells; Lysonase Bioprocessing Reagent (Novagen, Darmstadt, Germany) was then added ( $4 \mu\text{l g}^{-1}$  wet cells) and incubated for 30 min on ice with rotating mixing. The cell suspension was then sonicated for a total of 1.2 min and centrifuged at  $15000 \times g$  for 15 min at  $4^\circ\text{C}$ ; the supernatant was retained. The His<sub>6</sub>-tagged enzymes were purified at  $25^\circ\text{C}$  after binding to a Ni-NTA His-Bind resin (Novagen, Darmstadt, Germany). The columns were pre-washed with 20 mM HEPES (pH 7.0), 0.15 M NaCl and 50 mM imidazole, and the enzymes were eluted with 20 mM HEPES (pH 7.0), 0.15 M NaCl and 250 mM imidazole. The monitoring of the enzyme elution was performed by SDS-PAGE and/or activity measurements, using standard assays. Pure enzyme thus obtained was treated with enterokinase, as recommended by the supplier (Novagen, Darmstadt, Germany) to remove the His tags. The purity was assessed as  $>95\%$  using SDS-PAGE, which was performed with 12% (v/v) polyacrylamide gels, using a Bio-Rad Mini Protean system.

To produce *E. coli* BL21 (DE3) cell extracts expressing  $\beta$ -glucosidases, clones were grown overnight at  $37^\circ\text{C}$  with shaking at 200 rpm in LB medium containing kanamycin ( $30 \mu\text{g ml}^{-1}$ ). Afterwards, 0.3 ml of this culture was used to inoculate 30 ml of LB medium in a 250 ml flask, which was then incubated to an  $\text{OD}_{600\text{nm}}$  of  $\sim 0.8$ . Protein expression was induced by 1 mM IPTG followed by incubation for 4 h. The cells were harvested by centrifugation at  $5000 \times g$  for 15 min and re-suspended in 10 ml McIlvaine buffer (100 mM pH 5.0) containing lysozyme ( $0.1 \text{ mg ml}^{-1}$ ). The cell suspensions were then sonicated until a completed cell lysis was observed by microscope and centrifuged at  $15000 \times g$  for 15 min; the supernatants were retained and their protein concentrations and activity levels using *pNP* $\beta$ G and cellobiose as substrates, examined.

For the enzyme characterisation, the absorbance was measured using a BioTek Synergy HT spectrophotometer under the following conditions: [Enzyme] ranging from 0 to 300 nM, [substrate] ranging from 0 to 50 mM

in 100 mM buffer,  $T = 40^\circ\text{C}$ . For the hydrolysis of the *pNP* derivatives, the corresponding volume of a *pNP* derivative stock solution (150 mM) was incubated for 2–20 min in appropriate buffer and measured at 410 nm in 96-well microtiter plates [22]. The substrates tested included *pNP* $\beta$ G and *pNP* $\beta$ C. For cello-oligosaccharides the level of released glucose was determined, under the same conditions, using a glucose oxidase kit (Sigma Chemical Co., St. Louis, MO, USA). The initial rates were fitted to the Michaelis–Menten kinetic equation using non-linear regression to determine the apparent  $K_m$  and  $k_{\text{cat}}$ ; kinetic parameter calculations were performed based on the molecular masses listed in Additional file 1: Table S1. Substrate specificity was investigated also using carboxymethylcellulose (CMC), lichenan, barley glucan, laminarin and avicel (all from Sigma Chemical Co., [St. Louis, MO, USA]) and filter paper (Whatman, England). Enzymatic activity was quantified by measuring release of reducing sugars according to Miller [34] using [enzyme] of 20 nM, [substrate] of  $10 \text{ mg ml}^{-1}$  in 50 mM sodium acetate buffer pH 5.6,  $T = 40^\circ\text{C}$ . In all cases, one unit (U) of enzyme activity was defined as the amount of enzyme producing  $1 \mu\text{mol}$  of reducing sugars in 1 min under the assay conditions. Unless otherwise stated, standard assays were performed using [enzyme] of 20 nM, [*pNP* $\beta$ G] of  $10 \text{ mg ml}^{-1}$  in 50 mM sodium acetate buffer pH 5.6,  $T = 40^\circ\text{C}$ .

The pH and temperature optima were determined in the range of pH 4.0–8.5 and  $20$ – $70^\circ\text{C}$  in assays containing [enzyme] of 12 nM and [*pNP* $\beta$ G]<sub>0</sub> of  $10 \text{ mg ml}^{-1}$ . Optimal pH was measured at  $40^\circ\text{C}$ ; the following buffers (20 mM) were used: acetate (pH 4.0–6.0), 2-(*N*-morpholino)ethanesulfonic acid (MES) (pH 6.0–7.0), HEPES (pH 7.0–8.0), Tris–HCl (pH 8.0–8.5). pH was always adjusted at  $25^\circ\text{C}$ . Optimal temperature was determined in 50 mM sodium acetate buffer pH 5.6. All of the values were determined in triplicate and were corrected for the spontaneous hydrolysis of the substrate. The results shown are the averages of three independent assays  $\pm$  the standard deviation.

pH and thermal stability assays were performed in 100 ml flasks with 12 g of pre-treated biomass (see below) and 5 g  $\beta$ -glucosidase solution at  $25 \text{ U ml}^{-1}$  (sodium acetate 20 mM),  $50^\circ\text{C}$ , pH 5.2). For activity test at different time points, protein solution was separated from soil particles by ultrafiltration through low-adsorption hydrophilic 500000 NMWL cutoff membranes (regenerated cellulose, Amicon) and directly used for activity tests using *pNP* $\beta$ G as the substrate, using the standard protocol. All measurements were analysed in triplicate.

#### ***In silico* analysis of proteins**

The MetaGeneMark tool with refined heuristic models for metagenomes (<http://exon.gatech.edu/GeneMark/>

metagenome/index.cgi; [35]) was used to predict genes in the cloned DNA fragments. The deduced proteins were analysed using blastp and psi-blast [36] against the non-redundant databases. The translation products were further analysed for protein domains using the Pfam-A database [37]. Models of rumen  $\beta$ -glucosidases were obtained from the SWISS-MODEL server (<http://swissmodel.expasy.org/>) using the coordinates of *Thermotoga napolitana*  $\beta$ -glucosidase structure [26] (PDB identity: 2x41) as the template. Figures were created with PyMOL (<http://www.pymol.org/>).

### Supplementation of commercial cellulases with $\beta$ -glucosidase LAB25g2

The activity of LAB25g2 on the hydrolysis of pre-treated corn stover biomass was tested by the supplementation of enzymatic hydrolysis reactions using a  $\beta$ -glucosidase-deficient commercial cocktail (Celluclast, Novozymes A/S, Denmark). Pre-treatment of the biomass was performed by diluted acid soaking followed by steam explosion as described previously [38]. Hydrolysis reactions were performed at 50°C in 100 ml flasks with 12 g of pre-treated biomass, 3 g of enzymatic mixture and 5 g LAB25g2  $\beta$ -glucosidase solution at 25 U ml<sup>-1</sup> (for details see Results), to reach 20 g final reaction mass. Pre-treated corn stover was adjusted to an initial pH of 5.2 by the addition of 2 N NaOH, reaching a dry matter content of 20% after addition of enzymes. Concentration of  $\beta$ -glucosidase activity during enzymatic hydrolysis was calculated in the basis of *p*NP $\beta$ G assay (the standard substrate used in supplementation assays), and corresponds to 5 U g<sup>-1</sup> dry biomass in the absence of LAB25g2 enzyme, and a supplementation of 31.25 U g<sup>-1</sup> dry biomass in the presence of the enzyme. This was the minimum dose of LAB25g2 enzyme to achieve saccharification of pre-treated corn stover at a dry matter content of 20%, under conditions reassembling industrial operations. Three control reactions were considered: (i) assay in the absence of LAB25g2 enzyme; (ii) assay in the presence of commercial  $\beta$ -glucosidases G0395 (from almond; Sigma Chemical Co.) with specific activity, measured using *p*NP $\beta$ G at 50°C and pH 5.0, of 0.042  $\pm$  0.008 mU mg<sup>-1</sup>; (iii) assay in the presence of commercial  $\beta$ -glucosidases E-BGOSAG (from *Agrobacterium* sp.; Megazyme) with specific activity, measured using *p*NP $\beta$ G at 50°C and pH 5.0, of 119  $\pm$  8 mU mg<sup>-1</sup>. For control tests protein solutions at 25 U ml<sup>-1</sup> were prepared and added to reaction mixtures as for the LAB25g2, to achieve a final activity value of 31.25 units g<sup>-1</sup> dry biomass (in the basis of *p*NP $\beta$ G assay). Enzymatic hydrolysis reactions were followed by determining glucose and cellobiose concentrations by HPLC (Agilent Technologies,

1200 Series) using a Refractive Index Detector (RID) and a Aminex HPX-87 H column. The system was operated at 65°C and a flow rate of 6 ml min<sup>-1</sup> of 5 mM H<sub>2</sub>SO<sub>4</sub> for product elution as described elsewhere [39]. Glucose concentration was also measured by the glucose oxidase-peroxidase D-Glucose Assay Kit (Megazyme; Bray, Ireland).

Note: the synthetic substrate *p*NP $\beta$ G is specific for  $\beta$ -glucosidases, while not being hydrolyzed by cellobiohydrolases or endoglucanases. The numerous assays comparing activity towards *p*NP $\beta$ G and cellobiose have demonstrated that both activities are, overall, linear functions one to another ( $\beta$ -glucosidases with high activity towards *p*NP $\beta$ G also have high activity for cellobiose; for examples see Additional file 2: Table S2). Thus, supplementation per units of activity towards synthetic *p*NP $\beta$ G for all enzymes used in saccharification ensured the same activity towards cellobiose in all cases, and for this reason, this substrate was used in our supplementation assays.

### Additional files

**Additional file 1: Table S1.** Annotation of genes predicted in fosmid clones from cow rumen. Selected fosmids were sequenced and predicted proteins ORFs were annotated by homology using BLAST alignment tool. The theoretical molecular weight (MW) and isoelectric point (pI) were calculated for the gene product using ExPASy ProtParam online tool ([http://www.expasy.org/tools/pi\\_tool.html](http://www.expasy.org/tools/pi_tool.html)). Nucleotide and amino acid sequences are available at NCBI under acc. nr: JX163905, JX163906 and JX163904.

**Additional file 2: Table S2.** Biochemical information of reported  $\beta$ -glucosidases. Data are based on bibliographic records that are specifically cited.

**Additional file 3: Figure S1.** Temperature optima of commercial  $\beta$ -glucosidases from almond (G0395; Sigma Chemical Co.) and *Agrobacterium* sp. (E-BGOSAG; Megazyme). The parameters were determined using *p*NP $\beta$ G (0.1 mg ml<sup>-1</sup>) as the substrate at pH 5.0 (sodium acetate). All measurements were analysed in triplicate and error bars are indicated. Left panel shows the specific activity and right panel the relative activity referred to that at 30°C.

### Abbreviations

*p*NP $\beta$ G: *p*-nitrophenyl- $\beta$ -D-glucopyranoside; *p*NP $\beta$ C: *p*-nitrophenyl- $\beta$ -D-cellobioside; LAB: liquid-attached bacteria; SRF: strained ruminal fluid.

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

MF and PNG designed and coordinated the study and write the manuscript. MVP, LF-A, JG-M, AM, OVG, TNG, TYN, AW, MT, AR, SAH, CJN, JP and MF carried out the experiments and analysed the results. OVG and CJN contributed preparing the manuscript. All authors read and approved the final manuscript. MF and PNG contributed equally to this work.

### Acknowledgements

The authors gratefully acknowledge the financial support provided by the Spanish CDTI (research projects CENIT 2007-1031 and Cenit BioSos, I+DEA). This work has been funded by the Ministry of Economy and Competitiveness «Fondo de inversión local para el empleo-Gobierno de España». TYN, PNG, OVG acknowledge the support of grant 0313751 K from



the Federal Ministry for Science and Education (BMBF) within the GenoMikPlus initiative, and JP grant BIO2010-20508-C04-02 from Spanish MCI. SAH, OVG, TCN, PNG and CJN acknowledge the support of Bangor-Aberystwyth Strategic Alliance partnership in frames of Centre for Integrated Research in the Rural Environment (CIRRE) and Biosciences, Environment and Agriculture Alliance BEAA.

#### Author details

<sup>1</sup>CSIC, Institute of Catalysis, 28049, Madrid, Spain. <sup>2</sup>Abengoa Bioenergía Nuevas Tecnologías S.A., 41012, Sevilla, Spain. <sup>3</sup>Biopolis S.L, 48980, Paterna, Valencia, Spain. <sup>4</sup>School of Biological Sciences, Bangor University, LL57 2UW, Gwynedd, UK. <sup>5</sup>Insect Symbiosis Research Group, Max Planck Institute for Chemical Ecology, 07745, Jena, Germany. <sup>6</sup>Environmental Microbiology Department, HZI-Helmholtz Centre for Infection Research, D-38124, Braunschweig, Germany. <sup>7</sup>Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Aberystwyth SY23 3DA, UK. <sup>8</sup>Centre for Integrated Research in the Rural Environment, Aberystwyth University-Bangor University Partnership (CIRRE), Aberystwyth SY23 3DA, UK. <sup>9</sup>CSIC, Instituto de Agroquímica y Tecnología de Alimentos, 46980, Paterna, Valencia, Spain.

Received: 18 June 2012 Accepted: 14 September 2012

Published: 21 September 2012

#### References

- Sánchez OJ, Cardona CA: Trends in biotechnological production of fuel ethanol from different feedstocks. *Bioresour Technol* 2008, **99**:5270–5295.
- Kovacs K, Macrelli S, Szakacs G, Zacchi G: Enzymatic hydrolysis of steam-pretreated lignocellulosic materials with *Trichoderma atroviride* enzymes produced in-house. *Biotechnol Biofuels* 2009, **2**:1–14.
- Zhang M, Su R, Qi W, He Z: Enhanced enzymatic hydrolysis of lignocellulose by optimizing enzyme complexes. *Appl Biochem Biotechnol* 2010, **160**:1407–1414.
- Bayer EA, Lamed R, White BA, Flint HJ: From cellulosome to cellulosomes. *Chem Rec* 2008, **8**:364–377.
- Rubin EM: Genomics of cellulosic biofuels. *Nature* 2008, **454**:841–845.
- Sánchez C: Lignocellulosic residues: biodegradation and bioconversion by fungi. *Biotechnol Adv* 2009, **27**:185–194.
- Gurgu L, Lafraja Á, Polaina J, Marín-Navarro J: Fermentation of cellobiose to ethanol by industrial *Saccharomyces* strains carrying the  $\beta$ -glucosidase gene (BGL1) from *Saccharomycopsis fibuligera*. *Bioresour Technol* 2011, **102**:5229–5236.
- Hendriks ATWM, Zeeman G: Pretreatments to enhance the digestibility of lignocellulosic biomass. *Bioresour Technol* 2009, **100**:10–18.
- Gnansounou E, Dauriat A: Techno-economic analysis of lignocellulosic ethanol: A review. *Bioresour Technol* 2010, **101**:4980–4991.
- Walton J, Banerjee G, Car S: GENPLAT: an automated platform for biomass enzyme discovery and cocktail optimization. *J Vis Exp* 2011, **56**:e3314.
- Weusthuis RA, Aarts JMMJG, Sanders JPM: From biofuel to bioproducts: is bioethanol a suitable fermentation feedstock for synthesis of bulk chemicals? *Biofuels Bioprod Bioref* 2011, **5**:486–494.
- Zinoviev S, Müller-Langer F, Das P, Bertero N, Fornasiero P, Kaltschmitt M, Centi G, Miertus S: Next-generation biofuels: Survey of emerging technologies and sustainability issues. *ChemSusChem* 2010, **3**:1106–1133.
- Chauve M, Mathis H, Huc D, Casanave D, Monot F, Lopes Ferreira N: Comparative kinetic analysis of two fungal  $\beta$ -glucosidases. *Biotechnol Biofuels* 2010, **3**:3.
- García-Aparicio MP, Ballesteros M, Manzanera P, Ballesteros I, González A, Negro MJ: Xylanase contribution to the efficiency of cellulose enzymatic hydrolysis of barley straw. *Appl Biochem Biotechnol* 2007, **137–140**:353–365.
- Han Y, Chen H: Characterization of  $\beta$ -glucosidase from corn stover and its application in simultaneous saccharification and fermentation. *Bioresour Technol* 2009, **99**:6081–6087.
- Alvira P, Negro MJ, Ballesteros M: Effect of endoxylanase and  $\alpha$ -L-arabinofuranosidase supplementation on the enzymatic hydrolysis of steam exploded wheat straw. *Bioresour Technol* 2011, **102**:4552–4558.
- Pallapolu VR, Lee YY, Garlock RJ, Balan V, Dale BE, Kim Y, Mosier NS, Ladisch MR, Falls M, Holtzapfel MT, Sierra-Ramirez R, Shi J, Ebrik MA, Redmond T, Yang B, Wyman CE, Donohoe BS, Vinzant TB, Elander RT, Hames B, Thomas S, Warner RE: Effects of enzyme loading and  $\beta$ -glucosidase supplementation on enzymatic hydrolysis of switchgrass processed by leading pretreatment technologies. *Bioresour Technol* 2011, **102**:11115–11120.
- 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**:D233–D238.
- Ferrer M, Golyshina OV, Chernikova TN, Khachane AN, Reyes-Duarte D, Santos VA, Strompl C, Elborough K, Jarvis G, Neef A, Yakimov MM, Timmis KN, Golyshin PN: Novel hydrolase diversity retrieved from a metagenome library of bovine rumen microflora. *Environ Microbiol* 2005, **7**:1996–2010.
- Scharf ME, Tartar A: Termite digestomes as sources for novel lignocellulases. *Biofuels Bioprod Bioref* 2008, **2**:540–552.
- Li LL, McCorkle SR, Monchy S, Taghavi S, van der Lelie D: Bioprospecting metagenomes: glycosyl hydrolases for converting biomass. *Biotechnol Biofuels* 2009, **2**:10.
- Beloqui A, Nechitaylo TY, López-Cortés N, Ghazi A, Guazzaroni ME, Polaina J, Strittmatter AW, Reva O, Waliczek A, Yakimov MM, Golyshina OV, Ferrer M, Golyshin PN: Diversity of glycosyl hydrolases from cellulose-depleting communities enriched from casts of two earthworm species. *Appl Environ Microbiol* 2010, **76**:5934–5946.
- Zhou F, Chen H, Xu Y: GASdb: a large-scale and comparative exploration database of glycosyl hydrolysis systems. *BMC Microbiol* 2010, **10**:69.
- Hess M, Sczyrba A, Egan R, Kim TW, Chokhawala H, Schroth G, Luo S, Clark DS, Chen F, Zhang T, Mackie RI, Pennacchio LA, Tringe SG, Visel A, Woyke T, Wang Z, Rubin EM: Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. *Science* 2011, **331**:463–467.
- Ferrer M, Ghazi A, Beloqui A, Vieites JM, López-Cortés N, Marín-Navarro J, Nechitaylo TY, Guazzaroni M-E, Polaina J, Waliczek A, Chernikova A, Reva ON, Golyshina OV, Golyshin PN: Functional metagenomics unveils a multifunctional glycosyl hydrolase from the family 43 catalysing the breakdown of plant polymers in the calf rumen. *PLoS One* 2012, **7**:e38134.
- Pozzo T, Pasten JL, Karlsson EN, Logan DT: Structural and functional analyses of  $\beta$ -glucosidase 3B from *Thermotoga neapolitana*: a thermostable three-domain representative of glycoside hydrolase 3. *J Mol Biol* 2010, **397**:724–739.
- Marín-Navarro J, Gurgu L, Alamar S, Polaina J: Structural and functional analysis of hybrid enzymes generated by domain shuffling between *Saccharomyces cerevisiae* (var. diastaticus) *Sta1* glucoamylase and *Saccharomycopsis fibuligera* *Bgl1*  $\beta$ -glucosidase. *Appl Microbiol Biotechnol* 2011, **89**:121–130.
- Nakatani Y, Cutfield SM, Cowieson NP, Cutfield JF: Structure and activity of exo-1,3/1,4- $\beta$ -glucanase from marine bacterium *Pseudoalteromonas* sp. BB1 showing a novel C-terminal domain. *FEBS J* 2011, **279**:464–478.
- Berlin A, Maximenko V, Gilkes N, Saddler J: Optimization of enzyme complexes for lignocellulose hydrolysis. *Biotechnol Bioeng* 2007, **97**:287–296.
- Hu J, Arantes V, Saddler JN: The enhancement of enzymatic hydrolysis of lignocellulosic substrates by the addition of accessory enzymes such as xylanase: is it an additive or synergistic effect? *Biotechnol Biofuels* 2011, **4**:36.
- Lin ZX, Zhang HM, Ji XJ, Chen JW, Huang H: Hydrolytic enzyme of cellulose for complex formulation applied research. *Appl Biochem Biotechnol* 2011, **164**:23–33.
- Balat M, Balat H: Recent trends in global production and utilization of bioethanol fuel. *Appl Energy* 2009, **86**:2273–2282.
- Horn SJ, Vaaje-Kolstad G, Westereng B, Eijsink VGH: Novel enzymes for the degradation of cellulose. *Biotechnol Biofuels* 2012, **5**:45.
- Miller GL: Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal Chem* 1959, **5**:193–219.
- Zhu W, Lomsadze A, Borodovsky M: Ab initio gene identification in metagenomic sequences. *Nucleic Acids Res* 2010, **38**:e132.
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: Gapped BLAST and PS I-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997, **25**:3389–3402.
- Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, Khanna A, Marshall M, Moxon S, Sonnhammer EL, Studholme DJ, Yeats C, Eddy SR: The Pfam protein families database. *Nucleic Acids Res* 2004, **32**:D138–D141.

38. Nguyen QA, Tucker MP, Keller FA, Eddy FP: **Two stage dilute-acid pretreatment of softwoods.** *Appl Biochem Biotechnol* 1998, **70–72**:77–87.
39. Langston JA, Shaghasi T, Abbate E, Zu F, Vlasenko E, Sweeney MD: **Oxidoreductive cellulose depolymerisation by the enzymes cellobiose dehydrogenase and glycoside hydrolase 61.** *Appl Environ Microbiol* 2001, **77**:7007–7015.

doi:10.1186/1754-6834-5-73

**Cite this article as:** Del Pozo *et al.*: Microbial  $\beta$ -glucosidases from cow rumen metagenome enhance the saccharification of lignocellulose in combination with commercial cellulase cocktail. *Biotechnology for Biofuels* 2012 **5**:73.

**Submit your next manuscript to BioMed Central  
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

