Characterization and expression of a cellobiase Cba2 of *Cellulomonas biazotea*: the first bacterial β-glucosidase demonstrated to possess a genuine signal peptide

Siddique Faiza[1], Lam Chui Chi[1], Ng Ka Lun[1], Chan Anthony Kwun Nam[1,2] and Wong Wan Keung Raymond[1*]
1. Division of Life Science, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, CHINA
2. Department of Systems Biology, Beckman Research Institute, City of Hope, 1500 E Duarte Rd, Duarte, CA 91010, USA
*bcwkrw@ust.hk

Abstract

Cba2, a 100 kDa cellobiase found to occupy more than 40% of the β-glucosidase activity present in the culture medium of the gram-positive bacterium, Cellulomonas biazotea, was previously purified and characterized. In this study, we describe the cloning and characterization of the 2.74-kb genetic determinant of Cba2. Furthermore, we furnish supporting evidence to demonstrate that secretion of Cba2 is mediated via a genuine signal peptide. Comparing the translated Cba2 sequence with that of a hexapeptide derived from sequencing of the native Cba2 product helped localize the signal peptide. With assistance of signal peptide prediction, the exact location of the 32 aa long leader peptide was defined.

Moreover, Cba2 was found to utilize a less frequently employed start codon, GTG, to initiate its translation. Cba2 is the first bacterial cellobiase/β-glucosidase demonstrated to possess a genuine cleavable signal peptide. Due to the large size, Cba2 is prone to form insoluble aggregates during expression in E. coli. Chemical chaperones such as sorbitol were employed successfully to restore Cba2 activity. Through structural modeling and sequence comparison, Cba2 was shown to be a new member of glycoside hydrolase family 3. The cba2 gene is the 5th component of the cellobiase gene complex of C. biazotea, of which four other members have been previously characterized and reported. The availability of recombinant Cba2 will facilitate our studies of the catalytic and mechanistic properties of the enzyme involved in cellulose hydrolysis.

Keywords: *Cellulomonas biazotea*, *cba2* gene, Expression, β-glucosidase, Cellobiase, Cba2, Glycoside hydrolase family 3 (GH3), Secretion, Signal peptide.

Introduction

Cellulosic residues, which are easily and massively accessible from municipal, forestry and agricultural wastes, exist potentially as a continual source of raw materials and energy.

**** Author for Correspondence***

At present, if cellulose substrate is desired to be processed to yield glucose on a large scale, the only practical approach which is both feasible and environmentally friendly is through enzymatic hydrolysis using cellulases. However, complete hydrolysis of native cellulose requires not only the presence of a complex of three different types of cellulases: endoglucanases (Eng) (EC 3.2.1.4), exoglucanases (Exg) (EC 3.2.1.91) and β-glucosidases (Bgl) or cellobiases (Cel), but also the occurrence of synergistic action formed among these enzymes.

Cellulases from natural sources are low in specific activities and expensive to produce in large quantities.[35] Random mutagenesis offers a workable approach to the generation of cellulolytic mutants that are resistant to catabolite repression[19] or end-product inhibition.[3,33] Despite the availability of these mutants, due to inadequate understanding of their genetics, it might take a long time to comprehend the mechanistic operations at both the levels of transcriptional regulation and enzymatic hydrolysis in these mutants. Moreover, the requirement of synergistic cooperation among different types of cellulases for the degradation of native cellulose might create additional challenges for the application of the mentioned mutants to the development of large-scale hydrolytic processes.

Recombinant DNA technology offers a practical approach to the production of heterologous proteins. Employing molecular engineering to precisely fuse the desired genetic determinant with selected regulatory controls, expression of the target protein in an appropriate host strain may be achieved through a controllable and scalable format. In addition, recombinant proteins may even be co-expressed in the same host to attain multi-functional and co-operative activities.[18,46] However, the difficulties in maintaining precise regulatory controls for the co-expressed genes and operating the process cost-effectively would exert a strong inhibitory effect on applying this approach on a large-scale basis.

Our laboratory has been involved in the engineering of versatile *E. coli* and *B. subtilis* platforms for efficient expression of recombinant proteins including various cellulases,[15,20-25,45] which may be employed for studies of cooperative interactions. Among the cellulase genes, the genetic determinants encoding an Eng and an Exg, designated *cenA* and *cex* respectively have been cloned from

a soil bacterium, *Cellulomonas fimi* and expressed for years.[25,44,46] Although extracellular expression of *CenA* and *Cex* has been successfully achieved using various host systems,[24,25,42,43,46] it remains to be a challenging task to obtain an appropriate Cel enzyme(s), which may not only work synergistically with the former two cellulases but is also preferably secretory in nature. Expression of all three types of cellulase in the culture media could facilitate the formulation of a feasible process for cellulose hydrolysis.

Being a close relative of *C. fimi*, *C. biazotea* was found to produce a high level of extracellular Cel activity in the culture medium.[26] Cloning and purification endeavors have led to the identification of genetic determinants encoding 5 structurally distinct Cel components in *C. biazotea*.[4,5,26,45] Among these genes, the primary structures of 4 of them have been determined.[4,5,45] The results revealed that two of the Cel products are members of glycoside hydrolase family 1 (GH1) and the other two belong to glycoside hydrolase family 3 (GH3).[4,5]

Although all the four characterized recombinant Cel components were shown to be devoid of a conventional signal peptide, interestingly, they were all detectable in the culture media of *E. coli*,[4,5,45] supporting the notion that they are secretory in *C. biazotea*, possibly by means of an unconventional mechanism which enables protein passage through the cell membrane. Thus, they might only be minor contributors to the Cel activity detected in the growth medium of *C. biazotea*.[26]

The major component found to contribute over 40% of the extracellular Cel activity detected in *C. biazotea* was designated Cba2.[26] Another noteworthy point is that due to its notable representation of Cel activity in the culture medium, there has been a firm belief that the translocation of Cba2 is mediated by a bona fide signal peptide. In this study, we report our research findings which provide evidence to support that Cba2 secretion is indeed mediated using a genuine signal peptide. To attain this conclusion, we first cloned the 2.74-kb *cba2* gene and studied the sequence of its translated product Cba2.

Results from signal peptide prediction and sequence matching unambiguously revealed the presence of a 32 aa leader peptide at the N-terminus of Cba2. To our knowledge, Cba2 represents the first example among the known bacterial Bgl or Cel enzymes demonstrated to possess a typical signal peptide for secretion. Moreover, we provide findings to support that Cba2 is a member of GH3 and describe how Cba2, which is prone to form protein aggregates during expression, was successfully produced as a bioactive product in *E. coli*.

Material and methods

Bacterial strains: *C. biazotea* ATCC 486 (ATCC 486)[4,5,26,45] employed for the isolation of the *cba2* gene was obtained from ATCC (Rockville, MD, USA). *E. coli* JM101 and DH5α strains were employed for DNA cloning and maintenance of plasmids.[36] *E. coli* BL21-DE3 was used as the host for Cba2 expression.

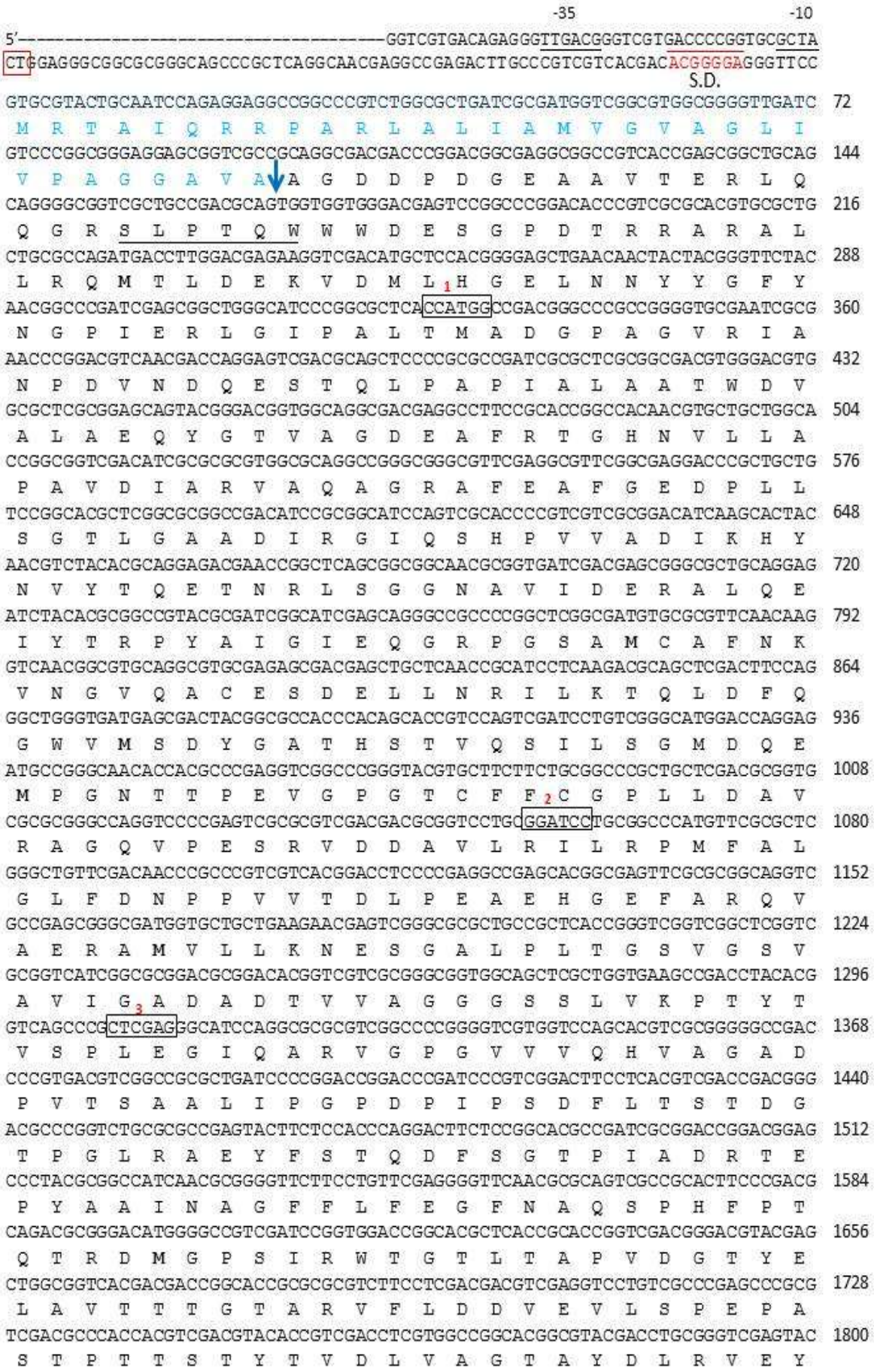
Media, chemicals and growth conditions: The 2 × YT and LB media employed for the cultivation of both *C. biazotea* and *E. coli* cultures at 30 °C were previously described.[4,5,45] When solid media were required, Bacto agar was added to the media to a final concentration of 1.5% (w/v). *E. coli* transformants were selected on culture media supplemented with 100 μg ml^{-1} of ampicillin. Induction of *E. coli* cultures was achieved using 0.1 mM IPTG. Tryptone, yeast extract and Bacto agar were purchased from Oxoid Limited (Hampshire, UK). IPTG was purchased from USB Corporation (Cleveland, USA). Phusion® High-Fidelity DNA polymerase, restriction and DNA modifying enzymes were purchased from New England Biolabs (MA, USA). PCR and sequencing primers were purchased from Invitrogen (CA, USA). All other chemicals were purchased from Sigma-Aldrich Corporation (MO, USA) unless otherwise specified.

Preparation of genomic DNA from *C. biazotea*: The isolation of genomic DNA from ATCC 486 cells was described previously.[4,45]

Cloning of the *cba2* gene encoding mature Cba2 from *C. biazotea*: The *cba2* gene encoding mature Cba2 (mCba2) was cloned in the expression vector pM[45] through separate rounds of PCR using specifically designed oligos as PCR primers (Table 1) and ATCC 486 genomic DNA as the template.

Two rounds of PCR were initially performed to clone the *cba2* sequence encoding the N-terminal segment (33rd to 436th aa) of mCba2. Initially, two pairs of primers: (i) ktac and RBS-r (Table 1) which primed at the *tac* promoter and the *lac* operator-RBS region respectively of the Tac cassette in the pM vector; (ii) RBS-cba2-f and kB2X-r, which primed at the RBS-ATG-*cba2* (at nt 97 of *cba2*) fusion region and the *Bgl* II site at nt 719 (Fig. 1) respectively were employed to attain a *cba2* segment extending from nt 97 to the *Bgl* II site at nt 719 (Fig. 1). The process resulted in a fusion product containing the *tac* promoter, the *lac* operator-RBS region of pM, followed by the newly added start codon, ATG and the cba2 sequence extending from nt 97 to nt 719.

In the second round of PCR, two primers: 6Bgl2X-f and Bxx-r, which primed at a region just opposite to primer kB2X-r (comprising the *Bgl* II site at nt 719) and the *Xho* I site at nt 1,306 of *cba2* respectively were employed. The product formed was a 590-bp fragment spanning from the two sites: *Bgl* II (nt 719) to *Xho* I (nt 1,306) of the *cba2* gene. The two fragments were digested with *Bgl* II followed by ligation to form a 1,310-bp product. Subsequent to restriction with *Kpn* I and *Xba* I, the 1,310-bp sequence was inserted into vector pM to form an intermediate construct pM2X.


-35
-10
S.D.

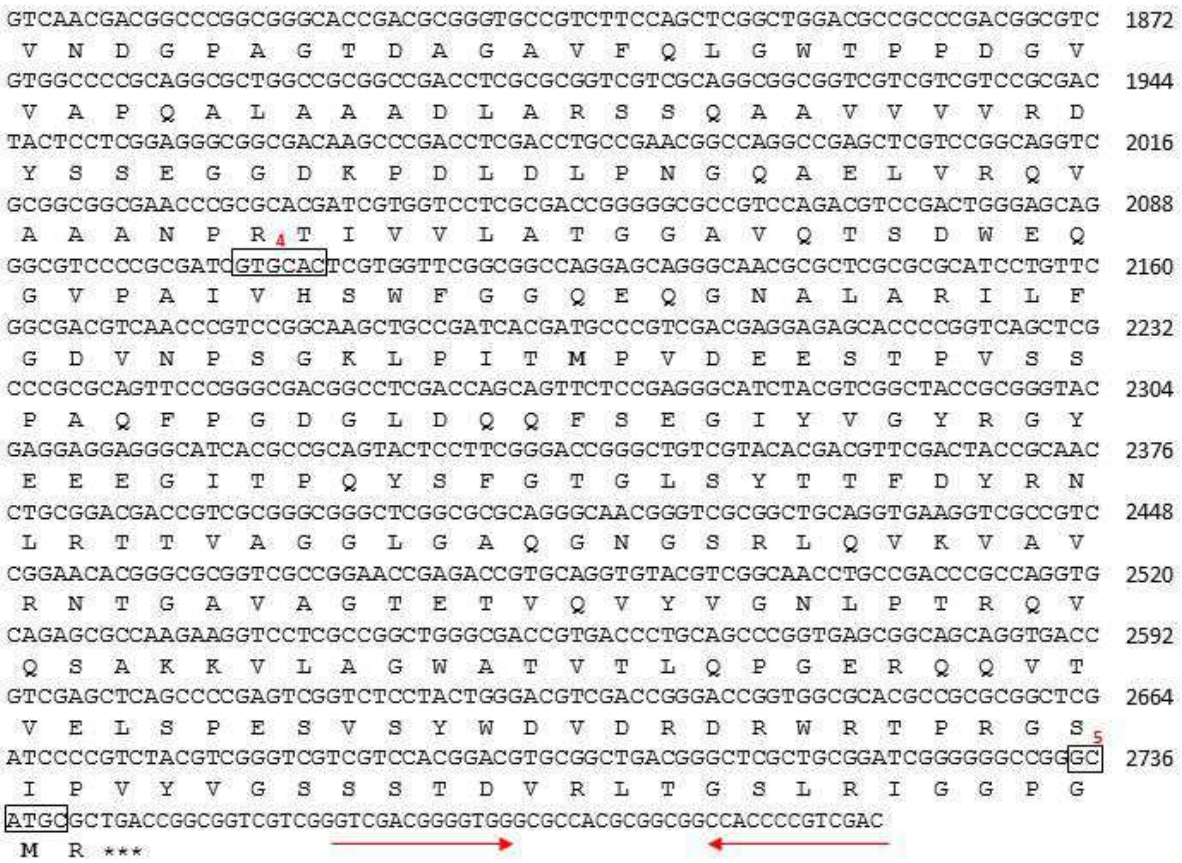


Fig. 1: The nucleotide sequence of the *cba2* gene and the deduced amino acid sequence of Cba2. The coding region starts at +1 and ends at +2742, with the translated protein sequence shown below. The -35 and -10 sites of the potential promoter are underlined. A putative Shine-Dalgarno (S.D.) sequence before the start site is overlined. The first 32 aa constitute the signal peptide and the downward arrow (↓) indicates the leader sequence processing site. The hexapeptide obtained from sequencing of native Cba2 is underlined. A 13 bp palindromic sequence representing the putative transcription terminator following the stop codon is indicated by the inverted arrows. The unique restriction sites: 1. *Nco* I (nt 326), 2. *Bam* HI (nt 1,055), 3. *Xho* I nt 1,306), 4. *Apa* L1 (nt 2,104) and 5. *Sph* I (nt 2,735) are boxed

Cloning of the intact 2.74-kb *cba2* gene: With ATCC 486 genomic DNA and 2 pairs of primers: 6XhoIA-f plus XApaLI-r and 6ApaLX-f plus XbaIA-r, a new round of PCR was performed to attain a *cba2* fragment spanning from the *Xho* I site at nt 1,306 to the 3' terminal codon comprising a *Sph* I site at nt 2,735 (Fig. 1). The 1.44-kb product was then inserted into vector pM2X to form a new intermediate construct pM2X-V.

To avoid mutations that might arise during PCR, a majority portion of the *cba2* sequence cloned in pM2X-V was replaced by its native genomic equivalent. In doing so, an *Nco* I - *Sph* I segment obtained from the genomic DNA was used to replace its corresponding sequence carried on pM2X-V to result in the final construct pMcba2 (Fig. 2) which was employed for Cba2 expression in *E. coli*.

To verify the authenticity of the *cba2* sequence extending from nt 97 to the *Nco* I site (Fig. 1), a *Bam* HI restricted genomic DNA covering the 5' portion of *cba*2 was cloned and sequenced. It was then confirmed that the full-length *cba*2 gene cloned in pMcba2 and its native counterpart shared the same composition.

Fusion of the *ompA* leader sequence with the *cba2* gene encoding mCba2: The placement of the *ompA* leader sequence precisely at the 5' terminus of the *cba2* gene sequence encoding mCba2 in pMcba2 was achieved employing two rounds of PCR. On one hand, the *ompA* leader sequence was precisely fused to the *tac* promoter-*lac* operator region in pMcba2 (Fig. 2) to form intermediate product I, with the help of pM as the template and a pair of PCR primers: Ktac and ompA-b (Table 1).

On the other hand, employing pMcba2 as the template and another pair of primers: ompAcba2-f and cba2NcoI-r (Table 1), PCR was undertaken to obtain intermediate product II. Products I and II were joined by overlap extension PCR (OE-PCR) and the extended fragment was then cleaved with *Kpn* I plus *Nco* I followed by its insertion into pMcba2 treated with the same two enzymes to attain construct pMOmpAscba2 for secretory expression of mCba2.

Table 1
Primers employed in the construction of plasmids harbouring the *cba2* gene and its derivatives.

S.N.	Primer	Orientation	Sequence (5' to 3')[d]
1	ktac	Forward	AAggtaccTTACTCCCCATCCCC
2	RBS-r	Reverse	AATTTTTTCCTCCTGTGTGAAATTGTTATC
3	RBS-cba2-f	Forward	GATAACAATTTCACACAGGAGGAAAAAATATGGCAGGCGACGA CCCGG
4	kB2X-r	Reverse	AAATCTAGAGCGTGTagatctCCTGCAGC
5	6Bgl2X-f	Forward	TGCAGGagatctACACGCG
6	Bxx-r	Reverse	AAAtctagaGATGCCctcgagCGGGC
7	6XhoIA-f	Forward	AGCCCGctcgagGGCA
8	XApaLI-r	Reverse	CCACGAgtgcacGATCGC
9	6ApaLX-f	Forward	GCGATCgtgcacTCGTGG
10	XbaIA-r	Reverse	AAAtctagaTCAGCgcatgcCCGGCCC
11	ompA-b	Reverse	TGCGGCCTGCGCTACG
12	ompAcba2-f	Forward	GTAGCGCAGGCCGCAGGCGACGAC
13	cba2NcoI-r	Reverse	TTTTTTccatggTGAGCGCCGGGAT
14	SphIH6-f	Forward	CCGGgcatgcGCCATCATCATCATCATCATTGAgcatgcGGCC
15	SphIH6-r	Reverse	GGCCgcatgcTCAATGATGATGATGATGATGGGCgcatgcCCGG
16	vegG-f	Forward	AAggtaccTAATTTAAATTTTATTTGACAAAAATGGGCTCGTGTTGT GCAATAAATGTAGTGAGGTGG
17	RBSlacO-r	Reverse	AATTTTTTCCTCCTGTGTGAAATTGTTATCCGCTCACAATT
18	lacUV5-f	Forward	AAggtaccGCATACTTACTCCCCATCCCCCTGTTTACACTTTATGCTT CCGGCTCGTATAATGTGTG

[d]Restriction sites used in the cloning experiments are shown in lowercase.

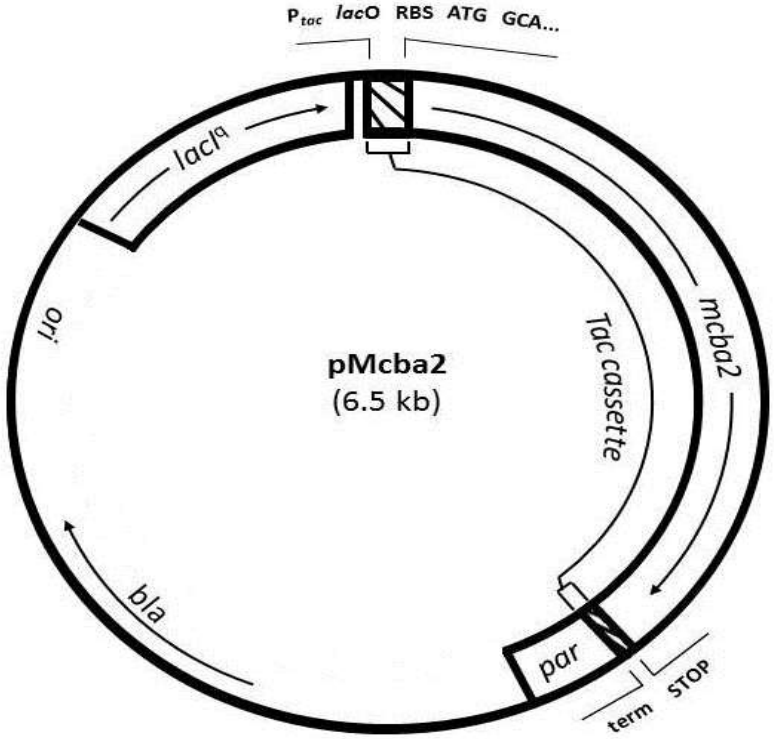


Fig. 2: Schematic representation of construct pMcba2 expressing Cba2.
Components present in the plasmid include: *mcba2*: coding sequence for mature Cba (starting from Ala 33; see Fig. 1); *par*: partition element from pSC101; *bla*: *bla* gene conferring resistance to ampicillin; *lac*I^q: *lac*I^q gene overexpressing the Lac repressor; *ori*: origin of replication. Genetic elements of the Tac cassette (hatched bars)[45] include *tac* promoter (P*tac*); *lac* operator (*lac*O); consensus ribosome binding site (RBS); start codon (ATG); the first codon (GCA) encoding mature Cba2; tandem stop codons in three reading frames (STOP); *gnt* transcriptional terminator (term). The arrows denote the direction of gene expression.

Insertion of a 6xHis-tag coding sequence at the 3' terminus of *cba2*: To facilitate the detection of Cba2 present in cellular fractions, a 6xHis-tag coding sequence was inserted at the 3' terminus of *cba2*. With the help of PCR, a pair of mutagenic primers: SphIH6-f and SphIH6-r (Table 1) and the template, pMcba2, an amplified fragment harbouring the tag sequence was obtained. The intermediate product was then restricted with *Sph* I followed by its insertion into pMcba2 restricted with the same enzyme to form pMcba2H6.

Construction of plasmids expressing *cba2* under the control of different promoters: To replace the *tac* promoter in pMcba2 with the *vegG* promoter,[42] PCR mediated mutagenesis was performed employing plasmid pM1VegGCenALacI^{q} [42] as the template and oligos vegG-f and RBSlacO-r (Table 1) as the mutagenic primers. Another PCR fragment was synthesized with pMcba2 as the template and oligos RBScba2-f and cba2NcoI-r as the primers (Table 1).

The two PCR fragments were fused by OE-PCR and the extended fragment was then restricted with *Kpn*I and *Nco*I followed by ligating the fragment with pMcba2 that had been restricted with the same two enzymes to obtain the expression construct: pMvegGcba2. Employing the same OE-PCR protocol, the *tac* promoter in pMcba2 was replaced by the *lacUV5* promoter, using pMcba2 as the template and two new pairs of oligos: lacUV5-f plus RBSlacO-r and RBScba2-f plus cba2NcoI-r (Table 1) as the primers. The resulting construct was designated pMlacUV5cba2.

Fractionation of *E. coli* cell cultures: To obtain an *E. coli* culture for the quantification of Cba2 activities, a colony of fresh transformants was used to inoculate 50 ml of LB medium. When the culture reached an OD_{550} value of 1.0, it was induced with 0.1 mM IPTG and cell samples were withdrawn at specified time intervals. The cell samples were fractionated into three portions: culture supernatant (SN), soluble cell lysate (CL) and insoluble cell debris (CD) of which the last two fractions were prepared by Triton-X treatment as described previously.[15]

Quantitative assay of Cba2: The presence of Cba2 activity in SN, CL and CD samples was quantified using a *p*NPGase assay as described previously.[4.26,45] One unit of *p*NPGase activity is defined as the amount of enzyme capable of releasing 1 µmol of *p*-nitrophenol per min at 37 °C.

Western blot analysis of His×6-tagged Cba2: Expression of His-tagged Cba2 was evaluated by Western blot analysis[7] using a monoclonal anti-polyhistidine antibody purchased from Santa Cruz Biotechnology.

Results

Strategy for cloning the *cba2* gene from *C. biazotea* ATCC 486 strain: The mature Cba2 Cel was previously purified by our group from the culture medium of ATCC 486 and shown to be a large protein of 109 kDa.[26] Thus, the coding sequence for Cba2 was expected to reach a size of 3,000 bp or so. Despite our success in employing a shotgun protocol together with direct expression to clone four other *cel* genes from ATCC 486,[4,5,45] the large size of *cba2* appeared to create difficulties for the gene to be readily cloned and expressed using a shotgun approach.

Cel enzymes isolated from various microbial species were shown to share high levels of identity and similarity with respect to amino acid (aa) composition.[4,5] In analyzing various Cel enzymes/genes published previously, a 2.7-kb gene sequence encoding a β-glucosidase (Bgl; GenBank: DM144958.1; designated here as Gls) was reported by a Japanese group cloned from *C. biazotea* strain N1228.[31] Interestingly, a comparison of the translated Gls sequence with that of a hexapeptide: NH_2-Ser-Leu-Pro-Thr-Gln-Trp-COOH obtained by our group from sequencing Cba2 previously[26] showed an exact match between the two protein entities. Although Cba2 and Gls were derived from two different host strains, the results suggested that they were isoforms which were expected to share a high degree of similarity at both the DNA and protein levels. PCR primers were then designed with reference to the *gls* nucleotide (nt) sequence and applied to the cloning of the *cba2* gene.

Cloning of the 5' terminal sequence encoding mature Cba2: Being a major Bgl component present in the culture medium of ATCC 486, Cba2 was postulated to be a secretory protein possessing a signal peptide.[26] The assumption gained support from the results of signal peptide prediction[32] in which a putative peptidase cleavage site was revealed to exist between two Ala residues: Ala32 and Ala33 of Cba2 (Fig. 3).

It was decided that the *cba2* sequence extending from nt 97, which was the first nt of the 33rd codon encoding Ala33 (the 1st aa of mature Cba2), to an *Xho* I site at nt 1,306 of *cba2* was initially cloned into vector pM[45] employing PCR operations and ATCC 486 genomic DNA as the template. The resulting 720 bp and 590 bp fragments were ligated to form a 1.31-kb insert, which was then cloned into vector pM to form an intermediate construct pM2X.

Cloning of the intact 2.74-kb *cba2* gene: Extension of the *cba2* sequence was achieved by PCR to obtain a 1.44-kb fragment spanning from the *Xho* I site at nt 1,306 to the 3' terminal codon comprising a *Sph* I site at nt 2,735 (Fig. 1). The entire *cba2* gene contained in the final expression construct, pMcba2 (Fig. 2), was confirmed to share the same sequence with its native counterpart from ATCC 486.

Characterization of Cba2 and its coding sequence: Despite the availability of *C. biazotea* genomic sequencing results from an online database (GenBankID: WP_130782007.1), wherefrom the *cba2* gene sequence could be located, a detailed delineation of both the *cba2* gene and its Cba2 product was not covered.

In particular, the exact location of a signal peptide, which is rarely identified among bacterial Bgl/Cel and postulated to be present on Cba2,[26] was not specified. Nevertheless, prediction results reported here lend support to the conclusion that residues Val1 to Ala32 of Cba2 comprise the putative signal peptide (Fig. 3).

In addition, the presence of: (i) a potential Shine-Dalgarno sequence (RBS)[11,12] detected 7 nt upstream from the putative start, the 1st GTG codon (Fig. 1), (ii) positively charged aa residues, which are commonly found at the N-terminal regions of signal peptides,[30] spotted closely after the 1st Val residue (encoded by GTG1) and (iii) a typical signal peptide length shown to be composed of 32 aa residues (Fig. 3), further support the conclusions that residues Val1 to Ala32 constitute the signal peptide of Cba2 and that GTG1 instead of ATG17 serves as the start codon for translational initiation of Cba2.

Sequencing results revealed that the *cba2* gene was composed of 2,742 bp encoding a 914 aa protein, which could be split into two functional domains: (i) the first 32 aa sequence comprising the signal peptide (described above); (ii) the remaining 882 aa portion representing the mature Cba2 (mCba2). Similar to other *Cellulomonas* genes which commonly possess high GC contents, *cba2* possesses also a high GC content of 72%.

Problems encountered during expression of Cba2 in *E. coli*: Studies have shown that secretory expression of foreign proteins in *E. coli* could impose a detrimental effect on the cell host[7] and that the bioactivity of a secreted protein might critically rely on the cleavage of the signal peptide from its precursor protein. Thus, it was decided that an intracellular approach was adopted to fuse the *cba2* gene encoding mCba2, which stretched from nt 97 to nt 2,742 (Fig. 1) precisely to the ATG start codon located in the Tac cassette of vector pM to form the expression construct, pMcba2 (Fig. 2).

Notwithstanding the precautions mentioned above, probably resulting from a combined effect of the large size and structural complexity of Cba2, the enzyme was expressed as inactive inclusion bodies in the cytoplasm of either BL21-DE3 (pMcba2) or BL21-DE3 (pMcba2H6) transformants (data not shown), of which the latter carried the coding sequence for a 6xHis tag at the 3' terminus of *cba2*.

Attempts to employ either weaker promoters including the *lac*UV5 and *vegG* promoters[8,25,42,53] to provide less efficient *cba2* expression or signal peptides including those of OmpA and SPA[14,24,47] to facilitate secretory expression of Cba2, unfortunately, failed also to yield bioactive Cba2.

Recovery of Cba2 bioactivity from inactive aggregates formed in *E. coli*: Chemical additives including glycerol (GLB), sucrose (SUB) and sorbitol (SOLB) were employed previously to help reduce the formation of protein aggregates.[10,34,49,50] BL21-DE3 (pMcba2) cultures grown in the presence of these chemical supplements were found to result in various levels Cba2 bioactivity. With the application of GLB, cell lysate samples were found to possess only low levels of activity. However, lysate samples obtained from cultures grown in media supplemented with SUB or SOLB were detected with significantly higher levels of Cba2 activity (Fig. 4). Despite these improvements, a major proportion of Cba2 remained to be expressed as aggregates (Fig. 5).

Cba2, a new member of GH3: Sequences of Cba2 and its homologous counterparts: Cba, Cba3, Cba4 and Cba5 of *C. biazotea*[4,5,45] were shown to share only low levels of homology (Table 2). Interestingly, however, high levels of identity (over 51%) were detected between sequences of Cba2 and some of the heterologous isoenzymes (Table 3) belonging to GH3.[5,39,45] The results support the conclusion that Cba2 is a member of GH3.

In attempting to identify the putative residues on Cba2 responsible for nucleophilic substitution and acid/base catalysis in substrate hydrolysis, the 3-D model of mCba2 was predicted and superimposed with that of a homologous GH3 member, Cba5 (data not shown), as well as with the crystal structure of a GH3 protein model, ExoI of barley.[38,39]

Despite their differences in size and composition, the results of superimposition revealed that residues Asp294 and Glu652 on mCba2 acted likely as the putative nucleophile and acid/base catalyst in substrate hydrolysis respectively (Fig. 6).

Table 2
Homologies between the primary structures of Cba2 and the 4 isozymes: Cba, Cba3, Cba4 and Cba5 of *C. biazotea* previously characterized by our group[a].

S.N.	Isozymes(GenBankID)	Identity (%)	Similarity (%)	Gaps (%)
1	Cba (AAC38196.1)	30	41.5	21.1
2	Cba3 (AEM45802.1)	11.2	15.6	68.5
3	Cba4 (ASU50143.1)	15	21.6	60.6
4	Cba5 (ASU50144.1)	25	35.2	31.5

[a]The Cba2 aa sequence was aligned individually with each of the 4 listed isozyme sequences, followed by using program EMBOSS Needle (http://www.ebi.ac.uk/Tools/psa)[27] to help obtain the optimum alignment between the two sequences in comparison.

Table 3
Homologies between the primary structures of Cba2 and examples of published GH3 proteins[b].

S.N.	Organisms	GenBankID	Identity (%)	Similarity %	Gaps (%)
1[c]	*Cellulomonas biazotea*	WP_130782007.1	98.2	98.2	1.8
2	*Cellulomonas xylanilytica*	WP_146926856.1	83.5	89.5	2.5
3	*Cellulomonas cellasea*	WP_141372097.1	81.9	88.1	2.2
4	*Cellulomonas fimi*	WP_013771207.1	78.3	86.6	1.6
5	*Leifsonia flava*	WP_135120555.1	70.7	80.6	2.9
6	*Cellulomonas hominis*	WP_154730242.1	69.7	80.3	1.5
7	*Cellulosimicrobium cellulans*	WP_158370858.1	70.3	80.3	1.5
8	*Cellulomonas massiliensis*	WP_019136739.1	66.7	76.2	2.2
9	*Cellulomonas timonensis*	WP_082812984.1	51.3	63.8	8.6

[b]The aa sequence of Cba2 was aligned individually with each of the isoenzyme sequences [available from NCBI protein BLAST search against the reference protein (refseq_protein) database] using program EMBOSS Needle (http://www.ebi.ac.uk/Tools/psa)[27] to help obtain the optimum alignment of two sequences.
[c]The *C. biazotea* genomic sequencing results were obtained from an online database (GenBankID: WP_130782007.1), wherefrom the information lacks the first 16 amino acids of Cba2.

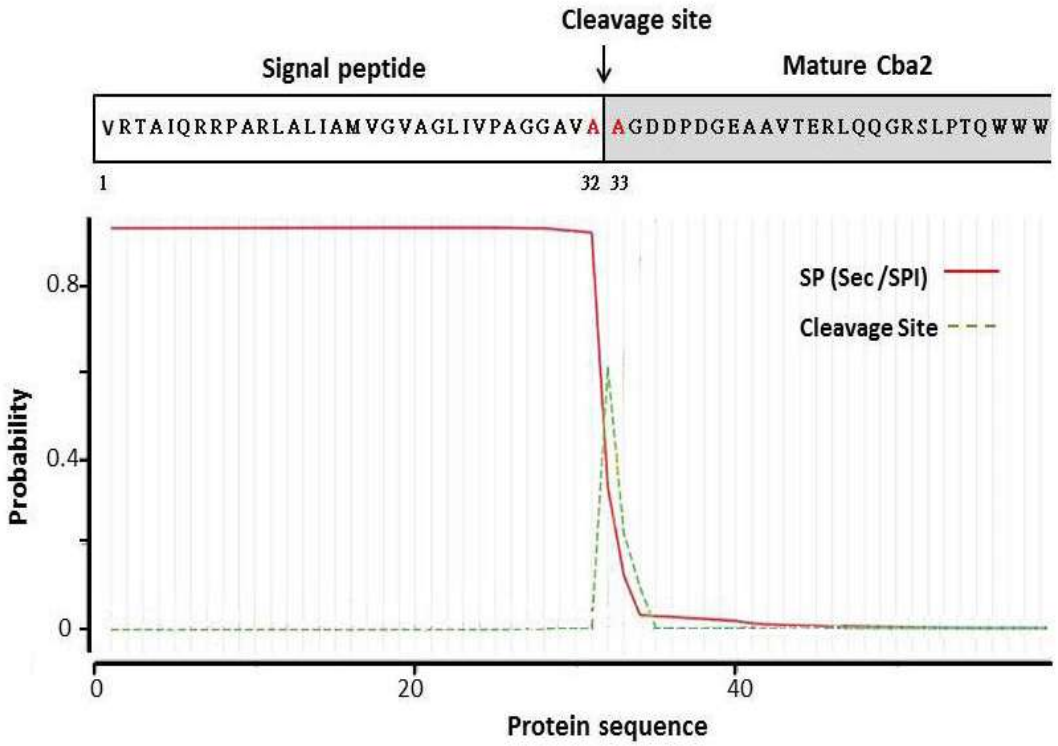


Fig. 3: The putative signal peptide in Cba2.
The presence of a signal peptide comprising the first 32 aa residues in Cba2 was predicted using the SignalP 5.0 server (for gram-positive bacteria). The arrow indicates the potential peptidase cleavage site located between Ala32 and Ala33. SP (Sec/SPI) stands for Signal Peptide (Sec/signal peptidase I).

Discussion

To effectively "forage for" and hydrolyse cellulosic residues present in their surrounding habitats e.g. forestry, agricultural and municipal wastes, cellulolytic microorganisms are normally capable of producing extracellular cellulases.[13,37,44,45,52]

It is generally accepted that fungal cellulases are more potent than their bacterial counterparts. A possible reason for this disparity is that extracellular cellulases of bacterial origins are typically devoid of Bgl/Cel activities which have been commonly shown to be intracellular or cell-associated.[51]

ATCC 486 was found to produce at least 5 Cel isoenzymes; the coding sequences for 4 of them: Cba, Cba3, Cba4 and Cba5 were cloned, characterized and expressed.[4,5,26,45] Interestingly, despite lacking an apparent signal, these 4 Cel were all shown to be secretory in *E. coli*.[4,5,45] The results lent support to the conclusion that ATCC 486 is capable of producing extracellular Cel activities.

Among the 5 Cel isoenzymes of ATCC 486, Cba2 was the only member purified and characterized in its native form. Cba2 was shown to be a major contributor to over 40% of the total Cel activity detected in the culture medium of ATCC 486.[26] Sequencing of secreted Cba2 revealed the presence of a hexapeptide comprising a structure of NH_2-Ser-Leu-Pro-Thr-Gln-Trp-COOH.[26] Since the sequence was deduced from extracellular Cba2, it was postulated that secretion of the high level of Cba2 activity was mediated through the use of a bona fide signal peptide.

Notwithstanding the identification of a perfect match between the translated sequence of a Bgl (designated as Gls in this work) of a variant strain of ATCC 486, *C. biazotea* N1228[31] and that of the hexapeptide (Fig. 1), which provided helpful hints on the cloning of the *cba2* gene from ATCC 486 (See Results), it was not straightforward to achieve an explicit delineation of Cba2.

First, the unveiled *cba2* gene sequence (Fig. 1) did not lead to a conclusive answer regarding whether codon GTG1 or ATG17 of the gene was used as the translational start codon. Nevertheless, the observations from this study including: (i) identification of a putative RBS upstream from GTG1 (Fig. 1), (ii) the results of the computational prediction (Fig. 3) and (iii) a length of 32 aa (extending from GTG1 to ALA32) being a typical size commonly shared among known signal peptides, unequivocally support that GTG1 but not ATG17 is the start codon and that the first 32 aa residues constitute the signal peptide of Cba2.

Secondly, it was unclear why the potential peptidase cleavage site determined to be present between residues Ala32 and Ala33 (Fig. 1) did not match with the predicted result in which the cleavage was expected to occur between Arg51 and Ser52 according to the hexapeptide sequence (Fig. 1).[26] The uncertainty was then resolved after realizing that tryptic cleavage could have occurred between Arg51 and Ser52 during sequencing of mCba2 (Fig. 1). A tryptically digested fragment beginning at Ser52 was then determined to yield the hexapeptide sequence.

Since Cba2 was shown to be a large (~100 kDa) secreted protein,[26] an elaborate characterization of it might not only facilitate the adoption of a feasible recombinant approach for its expression, but also the understanding of its role working cooperativity with other cellulase components in cellulose hydrolysis. On one hand, care was taken to avoid the formation of insoluble Cba2 aggregates during expression.

On the other hand, secretory expression of proteins, in particular larger ones, might cause cell death due to a detrimental phenomenon called saturated translocation.[7,25] Thus, an intracellular approach was first adopted for mCba2 expression. Despite the precautions, the expression still led to the formation of inactive mCba2 inclusion bodies in BL21-DE3 transformants harbouring construct pMcba2 (Fig. 2).

Since later attempts including reduced expression and secretory production of mCba2 failed to offer a workable solution to the aggregation problem, it was concluded that the difficulty was attributable to the structural complexity of the enzyme. It was speculated that the Cys residues present in mCba2 (totally 4; Fig. 1) might be the cause of the formation of protein aggregates during expression. Similar outcomes were reported to occur in other secreted proteins previously.[2,16] Nevertheless, cultures grown in the presence of chemical chaperones such as SUB and SOLB appeared to be effective in enabling proteins to retain their native conformations during synthesis.[10,34,49]

Indeed, cultivation of BL21-DE3 (pMcba2) cells supplemented with SUB or SOLB resulted in active Cba2 activities (Fig. 4), although a major portion of the mCba2 molecules remained to exist as inactive aggregates under the growth conditions employed (Fig. 5).

The successful expression of active mCba2 represents the first demonstration of recombinant production of a bacterial Cel which employs a genuine signal peptide for secretion. Although mCba2 was expressed as a deletant lacking the first 32 aa residues (Fig. 1), it remained to be active (Fig. 4). The outcome is in line with the conclusion that the first 32 residues constitute a signal peptide which plays an important role in the secretion of mCba2 rather than in enzyme catalysis. Thus, the remaining sequence commencing from Ala33 until the C-terminus of mCba2 (Fig. 1) is concluded to be essential for both substrate binding and catalytic action in enzymatic hydrolysis.

The structural elucidation of mCba2 would facilitate our understanding of functional, organizational and phylogenetic relationships of this enzyme with other Cel/Bgl. Although mCba2 shows only low levels of identity with its four homologous isoenzymes: Cba, Cba3, Cba4 and Cba5, interestingly it shares relatively high levels (> 51%) of identity with quite a few heterologous Bgl (Table 3). The large size of Cba2, which is also a property commonly noted among Bgl of Glycoside Hydrolase Family 3 (GH3),[1,5,6] and the high levels of sequence identity shared between mCba2 and GH3 members (Table 3), supports the conclusion that Cba2 is a GH3 Cel/Bgl.

Cba2 is not the first GH3 enzyme of *C. biazotea* characterized. Making efforts to understand the organizational and functional complexities of Cel in *C. biazotea*, we have successfully cloned 5 individual gene sequences encoding different Cel components: Cba,[45] Cba2 (this work), Cba3,[4] Cba4 and Cba5[5] from the bacterium. Cba3 and Cba4, which are at least 35% smaller than the other three homologous counterparts,[5] possess two short stretches of aa residues constituting the discernible conserved NEP and ENG motifs.[4,29] The relatively smaller sizes of Cba3 and Cba4 and the presence in them of the characteristic NEP and ENG motifs unequivocally support the conclusion that these two Cel are members of GH1.[4,5]

On the other hand, GH3 enzymes e.g. Cba and Cba5 contain trackable aa residues that are believed to be involved in nucleophilic and acid/base catalysis. These residues could be identified by superimposing a 3D structure of the studied GH3 enzyme (obtained through homology modelling) with that of a GH3 protein model e.g. the crystal structure of barley β-glucan ExoI.[38,39] Using this approach, the putative nucleophilic and acid/base residues in Cba2 were identified as Asp294 and Glu652 respectively (Fig. 6). In the same manner, the nucleophilic and acid/base residues in Cba5 potentially involved in enzyme catalysis were located previously.[5] The results reaffirm the previous conclusion that Cba2 is a GH3 member.

A precise delineation of Cba2 may not only facilitate the formulation of a practical approach for enhanced production of mCba2, but also the application of mCba2 to enzymatic studies conducted together with other Cel isoenzymes and cellulase components.

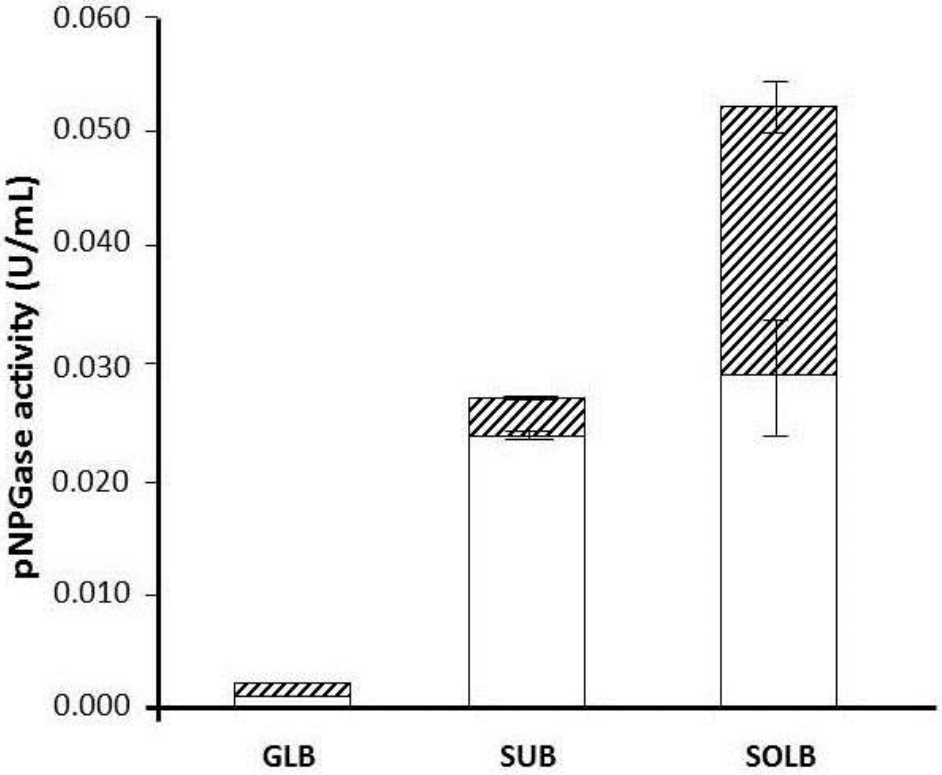


Fig. 4: Cba2 activities expressed by BL21-DE3 (pMcba2) transformants cultivated in the presence of different chemical additives.
The figure reveals the *p*NPGase activities obtained from supernatant (hatched bars) and lysate (open bars) fractions of BL21-DE3 (pMcba2) cells grown in LB medium supplemented with 1M glycerol (GLB), 1M sorbitol (SOLB) or 1M sucrose (SUB). The activities detected in supernatant samples resulted likely from cell lysis during growth. Cultures were fractionated subsequent to growth with IPTG induction for 24 h. All data shown are mean ±SEM values obtained from three independent experiments

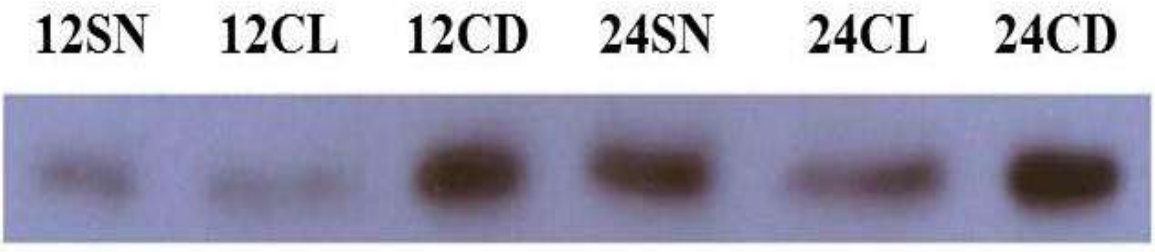


Fig. 5: Formation of Cba2 aggregates in *E. coli* BL21-DE3 (pMcba2H6) cells grown in LB medium supplemented with 1M SOLB.
Cultures were fractionated into supernatant (SN), cell lysate (CL) and cell debris (CD) samples subsequent to growth with IPTG induction for 12 h and 24 h. The samples were analysed using Western blotting with a monoclonal anti-polyhistidine antibody.

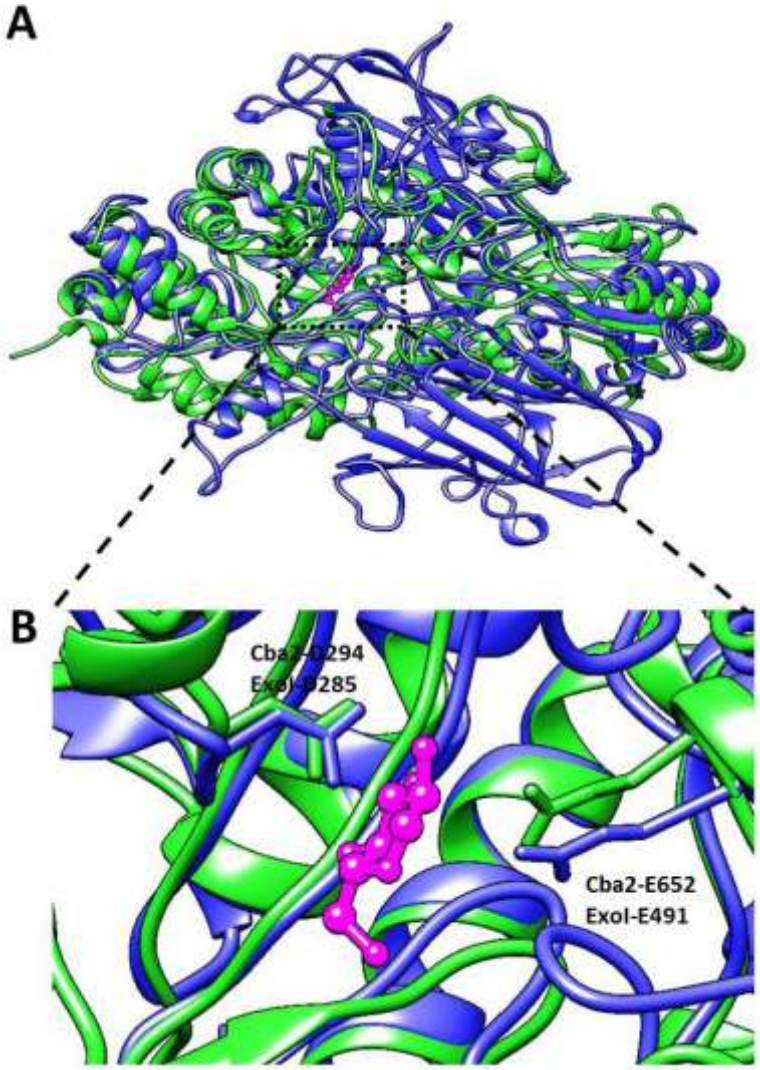


Fig. 6: Structural superimposition of Cba2 homology model (blue) with barley GH3 beta-D-glucan exohydrolase (ExoI) (PDB ID: 1EX1, green) in complex with glucose (purple).
(A) Overview. (B) Close-up view of the active site region [dotted square in (A)]. Side chains of key residues are denoted using stick notation. The co-crystallized glucose molecule in 1EX1 is presented in purple. The nucleophiles and acid/base catalysts of Cba2 (D294 and E652, respectively) and ExoI (D285 and E491, respectively) are indicated. The picture was generated with assistance from UCSF Chimera.

Fusion approaches of target proteins with appropriate tag partners e.g. inteins[22,23,48] and secretion signal peptides[8,48] have been developed to enable intracellular and/or extracellular expression of protein products possessing not only high potencies but also desired primary structures.[15,22,23,40]

On the other hand, chemical additives/chaperones e.g. β-cyclodextrin,[17] have been employed individually or cooperatively to achieve expression of soluble proteins.[41] Applying the mentioned strategies, production of bioactive mCba2 can be obtained for use in synergistic studies whereby better understanding of the mechanistic aspects of cellulose hydrolysis may be achieved.

Being a major Cel component present in the culture medium (>40% of total extracellular Bgl activity) of *C. biazotea*,[26] mCba2 is envisaged to be a potent hydrolytic enzyme. Obviously, the genuine signal peptide (its presence confirmed in this work) is also believed to work efficiently in translocating mCba2 to the culture medium.

These findings lend support to the speculation that *C. biazotea* and likely many other cellulolytic microorganisms might adopt a more dynamic approach than the multi-modular cellulosome mode of operation, which is believed to be employed by some anaerobic bacteria species,[9,28] to the exploitation of extracellular cellulases for use in the acquisition and hydrolysis of cellulose substrates.

Conclusion and Future Prospects

In this study, we provide evidence to demonstrate the existence of a bona fide signal peptide on a bacterial cellobiase, Cba2, of *C. biazotea*. Our analysis shows that the 2.74-kb *cba2* gene encodes a 100 kDa Cba2 product, of which the first 32 amino acids form the signal peptide. The

findings paved the way for the expression of recombinant Cba2 in *E. coli*. These achievements will help clarify whether Cba2 secretion plays a crucial role as a major contributor to over 40% of the extracellular BgI activity in *C. biazotea*.

Moreover, a convenient source of recombinant Cba2 will facilitate the determination of kinetic parameters that are required for the best performance of the enzyme. Cba2 may eventually prove to be a crucial component in synergistic studies conducted together with its isoenzymes and other types of cellulases in the hydrolysis of native cellulosic substrates.

Acknowledgement

This study was partially supported by RGC project: GRF16101515. We thank Kate K.Y. Ng for her technical assistance.

References

1. Ara K.Z.G., Manberger A., Gabrisko M., Linares-Pasten J.A., Jasilionis A., Friethjonsson O.H., Hreggviethsson G.O., Janecek S. and Nordberg Karlsson E., Characterization and diversity of the complete set of GH family 3 enzymes from *Rhodothermus marinus* DSM 4253, *Sci. Rep.*, **10(1),** 1-15 **(2020)**

2. Bhandari B.K., Gardner P.P. and Lim C.S., Solubility-Weighted Index: fast and accurate prediction of protein solubility, *Bioinformatics*, **36(18)**, 4691-4698 **(2020)**

3. Chand P., Aruna A., Maqsood A.M. and Rao L.V., Novel mutation method for increased cellulase production, *J. Appl. Microbiol.*, **98**, 318-23 **(2005)**

4. Chan K.N., Wang Y., Ng K.L., Fu Z. and Wong W.K.R., Cloning and characterization of a novel cellobiase gene, cba3, encoding the first known beta-glucosidase of glycoside hydrolase family 1 of *Cellulomonas biazotea*, *Gene*, **493(1)**, 52-61 **(2012)**

5. Chan K.N., Ng K.L., Ng K.K. and Wong W.K.R., Cloning and characterization of two novel beta-glucosidase genes encoding isoenzymes of the cellobiase complex from *Cellulomonas biazotea*, *Gene*, **642**,367-375 **(2018)**

6. Dodd D., Kiyonari S., Mackie R.I. and Cann I.K.O., Functional diversity of four glycoside hydrolase family 3 enzymes from the rumen bacterium *Prevotella bryantii* B14, *J. Bacteriol.*, **192(9)**, 2335-45 **(2010)**

7. Fu Z.B., Ng K.L., Lam T.L. and Wong W.K.R., Cell death caused by hyper-expression of a secretory exoglucanase in *Escherichia coli*, *Protein. Expr. Purif.*, **42,** 67-77 **(2005)**

8. Fu Z.B., Ng K.L., Lam C.C., Leung K.C., Yip W.H. and Wong W.K.R., A two-stage refinement approach for the enhancement of excretory production of an exoglucanase from *Escherichia coli*, *Protein Expr. Purif.*, **48(2)**, 205-14 **(2006)**

9. Galera-Prat A., Vera A.M., Morais S., Vazana Y., Bayer E.A. and Carrión-Vázquez M., Impact of scaffoldin mechanostability on cellulosomal activity, *Biomater. Sci.*, **8(13)**, 3601-3610 **(2020)**

10. Gheibi N., Saboury A.A., Haghbeen K. and Moosavi-Movahedi A.A., The effect of some osmolytes on the activity and stability of mushroom tyrosinase, *J. Biosci.*, **31(3)**, 355–62 **(2006)**

11. Harley C.B. and Reynolds R.P., Analysis of *E.coli* pormoter sequences, *Nucleic Acids Res.*, **15(5)**, 2343–2361 **(1987)**

12. Hawley D.K. and McClure W.R., Compilation and analysis of *Escherichia coli* promoter DNA sequences, *Nucleic Acids Res*., **11(8)**, 2237-55 **(1983)**

13. Holtzapple M., Cognata M., Shu Y. and Hendrickson C., Inhibition of Trichoderma reesei cellulase by sugars and solvents, *Biotechnol. Bioeng.*, **36(3)**, 275-87 **(1990)**

14. Huang R.C., Lam E., Chen Y.H., Hackett J., Lam T.L., Liu D., Ma M.C., Siu K.L., Sivakesava S., Xu Z.N., Wong S.C. and Wong W.K.R., Human epidermal growth factor excreted by recombinant *Escherichia coli* K-12 has the correct N-terminus and is fully bioactive, *Process Biochem.*, **35(1-2)**, 1–5 **(1999)**

15. Hu X., Lai C.Y.N., Sivakumar T., Wang H., Ng K.L., Lam C.C. and Wong W.K.R., Novel strategy for expression of authentic and bioactive human basic fibroblast growth factor in *Bacillus subtilis*, *Appl. Microbiol. Biotechnol.*, **102**, 7061-7069 **(2018)**

16. Idicula-Thomas S. and Balaji P.V., Understanding the relationship between the primary structure of proteins and its propensity to be soluble on overexpression in *Escherichia coli*, *Protein Sci.*, **14(3)**, 582–592 **(2005)**

17. Irie T. and Uekama K., Cyclodextrins in peptide and protein delivery, *Adv. Drug Deliv. Rev.*, **36(1)**, 101-123 **(1999)**

18. Kholod N. and Mustelin T., Novel vectors for co-expression of two proteins in *E. coli*, *Biotechniques*, **31(2)**, 322-328 **(2001)**

19. Kotchoni O.S., Shonukan O.O. and Gachomo W.E., *Bacillus pumilus* BpCRI 6, a promising candidate for cellulase production under conditions of catabolite repression, *African J. Biotechnol.*, **2**, 143–157 **(2003)**

20. Kwong K.W. and Wong W.K.R., A revolutionary approach facilitating co-expression of authentic human epidermal growth factor and basic fibroblast growth factor in both cytoplasm and culture medium of *Escherichia coli*, *Appl. Microbiol. Biotechnol.*, **97**, 9071–9080 **(2013)**

21. Kwong K.W., Ng K.L., Lam C.C., Wang Y. and Wong W.K.R., Authentic human basic fibroblast growth factor produced by secretion in *Bacillus subtilis*, *Appl. Microbiol. Biotechnol.*, **97**, 6803-11 **(2013)**

22. Kwong K.W., Ng K.L. and Wong W.K.R., Engineering versatile protein expression systems mediated by inteins in *Escherichia coli*, *Appl. Microbiol. Biotechnol.*, **100(1)**, 255-62 **(2016)**

23. Kwong K.W., Sivakumar T., Wong W.K.R., Intein mediated hyper-production of authentic human basic fibroblast growth factor in *Escherichia coli*, *Sci. Rep.*, **6(1)**, DOI:10.1038/srep33948 **(2016)**

24. Lam K.H., Chow K.C. and Wong W.K.R., Construction of an

efficient *Bacillus subtilis* system for extracellular production of heterologous proteins, *J. Biotechnol.*, **63**, 167-77 **(1998)**

25. Lam T.L., Wong S.C.R. and Wong W.K.R., Enhancement of extracellular production of a *Cellulomonas fimi* exoglucanase in *Escherichia coli* by the reduction of promoter strength, *Enzyme Microb. Technol.*, **20**, 482-8 **(1997)**

26. Lau T.Y. and Wong W.K.R., Purification and characterization of a major secretory cellobiase, Cba2, from *Cellulomonas biazotea*, *Protein Expr. Purif.*, **23(1)**, 159-66 **(2001)**

27. Li W., Cowley A., Uludag M., Gur T., McWilliam H., Squizzato S., Park Y.M., Buso N. and Lopez R., The EMBL-EBI bioinformatics web and programmatic tools framework, *Nucleic Acids Res.*, **43(W1)**, W580-4 **(2015)**

28. Matte A., Kozlov G., Trempe J.F., Currie M.A., Burk D., Jia Z., Gehring K., Ekiel I., Berghuis A.M. and Cygler M., Preparation and characterization of bacterial protein complexes for structural analysis, *Adv. Protein Chem. Struct. Biol.*, **76**, 1-42 **(2009)**

29. Moracci M., Capalbo L., Ciaramella M. and Rossi M., Identification of two glutamic acid residues essential for catalysis in the beta-glycosidase from the thermoacidophilic archaeon *Sulfolobus solfataricus*, *Protein Eng.*, **9(12)**, 1191-1195 **(1996)**

30. Nesmeyanova M.A., Karamyshev A.L., Karamysheva Z.N., Kalinin A.E., Ksenzenko V.N. and Kajava A.V., Positively charged lysine at the N-terminus of the signal peptide of the *Escherichia coli* alkaline phosphatase provides the secretion efficiency and is involved in the interaction with anionic phospholipids, *FEBS Lett.*, **403(2)**, 203–207 **(1997)**

31. Okura T., Nishimoto T., Chaen H. and Fukuda S., Beta-glucosidase, its preparation and use for producing beta-glucosidases, JP2009034028-A/1 **(2009)**

32. Petersen T.N., Brunak S., Von Heijne G. and Nielsen H., SignalP 4.0: discriminating signal peptides from transmembrane regions, *Nat. Methods*, **8(10)**, 785-6 **(2011)**

33. Peterson R. and Nevalainen H., *Trichoderma reesei* RUT-C30-thirty years of strain improvement, *Microbiology*, **158**, 58-68 **(2012)**

34. Prasad S., Khadatare P.B. and Roy I., Effect of chemical chaperones in improving the solubility of recombinant proteins in *Escherichia coli*., *Appl. Environ. Microbiol.*, **77(13)**, 4603-9 **(2011)**

35. Reese E.T. and Mandels M., Rolling with the Times: Production and Applications of *Trichoderma reesei* CELLULASE, *Annu. Rep. Ferment. Prcs.*, **7**, 1–20 **(1984)**

36. Sambrook J. and Russell D.W., Molecular Cloning: A Laboratory Manual, Third edition, CSHL Press **(2001)**

37. Stackebrandt E. and Schumann P., The family Cellulomonadaceae, The Prokaryotes: Actinobacteria, 163–184 **(2014)**

38. Thongpoo P., McKee L.S., Araujo A.C., Kongsaeree P.T. and Brumer H., Identification of the acid/base catalyst of a glycoside hydrolase family 3 (GH3) beta-glucosidase from *Aspergillus niger* ASKU28, *Biochem. Biophys. Acta*, **1830(3)**, 2739-49 **(2013)**

39. Varghese J.N., Hrmova M. and Fincher G.B., Three-dimensional structure of a barley beta-D-glucan exohydrolase, a family 3 glycosyl hydrolase, *Structure*, **7(2)**, 179-90 **(1999)**

40. Wang H., Hu X., Sivakumar T., Lai C.Y., Ng K.L., Lam C.C. and Wong W.K.R., A practical approach to unveiling auto-catalytic cleavages mediated by Mxe GyrA intein and improving the production of authentic bFGF, *J. Adv. Res. Biotech.*, **4(1)**, 1-10 **(2018)**

41. Wang L., Chen S. and Wu J., Cyclodextrin enhanced the soluble expression of *Bacillus clarkii* gamma-CGTase in *Escherichia coli*, *BMC Biotechnol.*, **18(1)**, 72 **(2018)**

42. Wang Y.Y., Ng K.L., Lam C.C., Chan A.K., Sze K.F., Fu Z.B. and Wong W.K.R., Efficient *Bacillus Subtilis* promoters for graded expression of heterologous genes in *Escherichia Coli*, *Res. J. Biotechnol.*, **5(4)**, 5–14 **(2010)**

43. Wang Y.Y., Fu Z.B., Ng K.L., Lam C.C., Chan A.N., Sze K.F. and Wong W.K.R., Enhancement of excretory production of an exoglucanase from *Escherichia coli* with phage shock protein A (PspA) overexpression, *J. Microbiol. Biotechnol.*, **21**, 637-45 **(2011)**

44. Wong W.K.R., Gerhard B., Guo Z.M., Kilburn D.G., Warren A.J. and Miller R.C., Characterization and structure of an endoglucanase gene cenA of *Cellulomonas fimi*, *Gene*, **44**, 315-24 **(1986)**

45. Wong W.K.R., Ali A., Chan W.K., Ho V. and Lee N.T., The cloning, expression and characterization of a cellobiase gene encoding a secretory enzyme from *Cellulomonas biazotea*, *Gene*, **207**, 79-86 **(1998)**

46. Wong W.K.R., Curry C., Parekh R.S., Parekh S.R., Wayman M., Davies R.W., Kilburn D.G. and Skipper N., Wood hydrolysis by *Cellulomonas fimi* endoglucanase and exoglucanase coexpressed as secreted enzymes in *Saccharomyces cerevisiae*, *Biotechnology*, **6**, 713–719 **(1988)**

47. Wong W.K.R., Fu Z., Wang Y.Y., Ng K.L. and Chan K.N., Engineering of efficient *Escherichia coli* excretion systems for the production of heterologous proteins for commercial applications, *Recent Pat. Chem. Eng.*, **5(1)**, 45–55 **(2012)**

48. Wong W.K.R., Ng K.L., Sivakumar T., Hu X., Wang H. and Lai C.Y., *Escherichia coli*: A versatile platform for recombinant protein expression. The Universe of *Escherichia coli*, IntechOpen Limited, London, UK, 22-44 **(2019)**

49. Xie G. and Timasheff S.N., Mechanism of the stabilization of ribonuclease A by sorbitol: Preferential hydration is greater for the denatured than for the native protein, *Protein Sci.*, **6(1)**, 211–221 **(1997)**

50. Yancey P.H., Water Stress, Osmolytes and Proteins, *Am. Zool.*, **41(4)**, 699–709 **(2001)**

51. Yan S. and Wu G., Secretory pathway of cellulase: a mini-review, *Biotechnol. Biofuels*, **6(1)**, 177 **(2013)**

52. Zhang L., Fu Q., Li W., Wang B., Yin X., Liu S., Xu Z. and Niu Q., Identification and characterization of a novel beta-glucosidase via metagenomic analysis of *Bursaphelenchus xylophilus* and its microbial flora, *Sci. Rep.*, **7(1)**, 1-11 **(2017)**

53. Zhinan Z., Gang L., Peilin C. and Wong W.K.R., Factors influencing excretive production of human epidermal growth factor (hEGF) with recombinant *Escherichia coli* K12 system, *Bioprocess Eng*., **23(6)**, 669–674 **(2000)**.

(Received 02nd March 2021, accepted 12th April 2021)