Running title:
A Novel Chitinase from *C. shinanonensis*

Title:

Heterologous Expression and Functional Characterization of a Novel Chitinase from the Chitinolytic Bacterium, *Chitiniphilus shinanonensis*

Lanxiang HUANG, Arisa SHIZUME, Masahiro NOGAWA, Goro TAGUCHI, and Makoto SHIMOSAKA†

Division of Applied Biology, Faculty of Textile Science and Technology, Shinshu University, 3-15-1 Tokida, Ueda, Nagano 386-8567, Japan

Received October 28, 2011; Accepted December 12, 2011

†To whom correspondence should be addressed.
Makoto Shimosaka
Division of Applied Biology, Faculty of Textile Science and Technology, Shinshu University, 3-15-1 Tokida, Ueda, Nagano 386-8567, Japan
Tel/fax: +81-(0)268-21-5341 Email: mashimo@shinshu-u.ac.jp

Abbreviations: CBM, carbohydrate-binding module; ChBD, chitin-binding domain; GlcNAc, *N*-acetyl-D-glucosamine; ORF, open reading frame; pNP, *p*-nitrophenyl

Abstract: *Chitiniphilus shinanonensis* strain SAY3^{T} is a chitinolytic bacterium isolated from the moat water of Ueda Castle, Japan. Fifteen genes encoding putative chitinolytic enzymes (*chiA-chiO*) have already been isolated from this bacterium. Five of these constitute a single operon (*chiCDEFG*). The open reading frames of *chiC*, *chiD*, *chiE*, and *chiG* show sequence similarity to family 18 chitinase, while the *chiF* encodes a polypeptide with two chitin-binding domains but no catalytic domain. Each of the five genes was successfully expressed in *Escherichia coli* and the resulting recombinant proteins were characterized. Four of the recombinant proteins (ChiC, ChiD, ChiE and ChiG) exhibited an endo-type chitinase activity toward chitinous substrates, while ChiF showed no chitinolytic activity. In contrast to most endo-type chitinases, which mainly produce a dimer of *N*-acetyl-D-glucosamine (GlcNAc) as the final product, ChiG could completely split the GlcNAc dimer into GlcNAc monomers, indicating that it is a novel chitinase.

Key words: *Chitiniphilus shinanonensis*; chitinase; chitin-binding domain; *N*-acetyl-D-glucosamine

Introduction

Chitin is a linear polysaccharide composed of β-1, 4-linked *N*-acetyl-D-glucosamine (GlcNAc), and is found in crustaceans, mollusks, algae, insects, and fungi. In these organisms, chitin provides an architectural reinforcement to biological structures, such as invertebrate exoskeletons and fungal cell walls.[1, 2] Chitin is the second most abundant biomass (after cellulose) on Earth, with an annual production of 100 billon tons.[3] Chitin has been receiving attention not only as an unutilized biological resource, but also as a new functional biomaterial having favorable properties, such as biodegradability and biocompatibility. Hence, chitin has been applied in various industrial fields, including textile production, food processing, and agriculture.[4, 5] Additionally, GlcNAc oligomers prepared from chitin have useful biological activities, such as immunostimulation and induction of plant defense

1 responses.[6)] Moreover, GlcNAc monomers can be also used as a sweetener and as a
2 nutritional supplement. At present, chitin degradation on an industrial scale mainly
3 relies on chemical reactions involving strong acid; however, the resulting waste fluid is
4 a problematic environmental pollutant. Hence, bacterial chitinolytic enzymes have been
5 investigated for their potential use in the enzymatic production of GlcNAc and its
6 oligomers in an environmentally friendly manner.[7)]

7 Previously, we isolated the strong chitinolytic bacterium from the moat water of
8 Ueda Castle in Nagano Prefecture, Japan.[8, 9)] It was identified as a novel species
9 belonging to a novel genus, and was registered as the type strain under the name of
10 *Chitiniphilus shinanonensis* SAY3^{T}.[10)] To understand the chitinolytic system of this
11 bacterium, we isolated 15 genes that encoded the putative chitinolytic enzymes
12 (*chiA*-*chiO*).[11)] Sequence analysis indicated that these putative enzymes could be
13 classified into 12 family 18 chitinases, a family 19 chitinase, a family 20
14 β-*N*-acetylglucosaminidase, and a polypeptide with two chitin-binding domains but no
15 catalytic domain.[12)] Among them, *chiN*, encoding a family 19 chitinase, has been
16 studied in detail, because it is very rare in microorganisms, and the origin of *chiN* was
17 of interest.[11)] Functional analysis of the other genes, which encode a variety of
18 chitinolytic enzymes, will enlarge our knowledge on the repertoire of chitinolytic
19 enzymes, and could also lead to the development of an enzymatic reaction process to
20 efficiently produce GlcNAc and its oligomers.

21 In this article, we report the functional analysis of five genes (*chiCDEFG*)
22 constituting a single operon. These apparently diverse gene products are assumed to
23 play a major role in chitin degradation by acting in a cooperative manner. In particular,
24 *chiG* was studied in more detail because it encodes a novel endo-type chitinase that
25 produces GlcNAc monomers as the final product.
26
27 **Materials and Methods**

28 *Strains and culture. Escherichia coli* DH5α was used for sub-cloning of chitinase genes
29 from *C. shinanonensis*. *E. coli* BL21(DE3) and Rosetta-gami™ 2 (DE3)pLysS

1 (Novagen, Darmstadt, Germany) were used as hosts to express the chitinase genes. *E.*
2 *coli* strains were routinely cultivated in Luria-Bertani (LB) medium.
3
4 *Expression of* chiC, chiD, chiE, chiF *and* chiG *in E. coli*. Each of the five genes
5 constituting the single operon (*chiCDEFG*) was PCR-amplified using fosmid DNA
6 prepared from clone No.4, which was selected from a *C. shinanonensis* SAY3^T genome
7 library constructed using the fosmid vector pCC1FOS as the template.[11)] The set of
8 primers designed to amplify each gene are listed in Table 1. The primers were designed Table 1
9 to amplify DNA fragments corresponding to the whole ORF but without any putative
10 signal peptide that was predicted by SignalP software
11 (http://www.cbs.dtu.dk/services/SignalP). The primers also contained a cleavage site for
12 a suitable restriction endonuclease at their 5' termini to allow the amplified DNA
13 fragment to be inserted, in frame, into the expression vector, pCold I (Takara, Kyoto,
14 Japan). PCR was performed using Phusion DNA polymerase (Finnzymes, Espoo,
15 Finland), and the resulting amplified products were purified using the Sephaglas™
16 Bandprep kit (GE Healthcare, Tokyo, Japan). The DNA fragments were then digested
17 with the appropriate restriction endonucleases and separately ligated into pCold I, which
18 had been digested with the same restriction endonucleases. The resulting recombinant
19 plasmids were introduced into *E. coli* BL21 (DE3) or Rosetta-gami™ 2 (DE3)pLysS.
20 Expression of recombinant proteins was induced by addition of
21 isopropyl-β-thiogalactopyranoside, as previously described.[11)] The recombinant proteins
22 were purified by affinity chromatography using His GraviTrap Columns (GE
23 Healthcare, Tokyo, Japan) that bound the poly-histidine tag fused to the N-termini of
24 the recombinant proteins.
25
26 *Chitinase enzyme assay*. The reaction mixture (0.75 mL) contained 50 mM sodium
27 acetate buffer (pH 5.0), appropriate amount of enzyme, and either 0.025% colloidal
28 chitin (from crab shells, Nacalai Tesque, Kyoto, Japan), or 0.0625% chitosan
29 (deacetylation degree of 80%, Katokichi, Tokyo, Japan). After incubation at 37°C for

1 appropriate intervals, the reaction was terminated by boiling for 3 min. The amount of
2 reducing sugar liberated during the reaction was measured by the modified Schale's
3 method.[13)] When chromogenic substrates, *p*-nitrophenyl (pNP)-$(GlcNAc)_n$ (n=1, 2 or 3)
4 were used as the substrates, the standard reaction mixture (0.1mL) contained 50mM
5 sodium acetate buffer (pH 5.0), 1mM chromogenic substrate, and an appropriate
6 amount of the recombinant proteins. The mixture was incubated at 37°C for 30 min, and
7 then the reaction was terminated by adding 0.2 mL of 0.2 N NaOH. The amount of pNP
8 released during the reaction was measured by reading the optical density at 405 nm.
9
10 *Analytical methods.* Reaction products were analyzed by thin-layer chromatography
11 (TLC). The reaction mixture (25 μL) contained 5 or 10 mM
12 *N*-acetylchitooligosaccharides, $(GlcNAc)_n$ (n=2-6), 50 mM sodium acetate buffer (pH
13 5.0), and an appropriate amount of the purified recombinant enzyme. The mixture was
14 incubated at 37°C for appropriate time intervals, and boiled for 5 min to terminate the
15 reaction. After centrifugation at 15,000 rpm for 5 min, 6μL of the supernatant was
16 spotted onto TLC plates (Silica gel $60F_{254}$; Merck, Germany) and developed with
17 *n*-propanol: ethyl acetate: 25% ammonia aq: water (60:30:12:28, v/v). The developed
18 products were visualized by spraying with 10g/L aniline: 10g/L diphenylamine (in
19 acetone): 85% phosphoric acid (5:5:1, v/v). Nucleotide sequences were determined
20 using an ABI PRISM 3100 Genetic Analyzer and BigDye Terminator v3.1 Cycle
21 Sequencing Kit (Applied Biosystems, Tokyo, Japan). The amount of protein was
22 determined with a protein assay kit (Bio-Rad, Tokyo, Japan) using bovine serum
23 albumin as a standard.
24
25 *Nucleotide sequences.* The nucleotide sequence of *chiCDEFG* has been deposited in the
26 DDBJ/EMBL/GenBank databases under accession number AB649129.
27
28 **Results**
29 *Expression of* chiC, chiD, chiE, chiF *and* chiG *genes in E. coli.* The relative locations

and domain structures of the five genes, *chiCDEFG* constituting a single operon, are shown in Fig. 1. All five genes encode a polypeptide having a common structure, that is, a putative signal peptide at the N-terminus followed by two tandem-repeated chitin-binding domains (ChBDs) of the carbohydrate-binding module (CBM) family 5.[14)] In all cases, these two ChBDs are connected by Pro-, Thr-, and Val-rich linkers. Four ORFs (ChiC, ChiD, ChiE and ChiG) contain a C-terminal domain showing sequence similarity to the catalytic domain of family 18 chitinases, while the ORF of *chiF* contains a C-terminal domain comprised of a novel amino acid sequence with no significant similarity to any of the known proteins deposited in the public databases.

Fig. 1

Each of the five genes corresponding to the whole ORF, but without the putative signal peptide, was PCR-amplified and ligated into an expression vector, pCold I. In the case of *chiD*, the DNA fragment corresponding to the whole ORF was not amplified possibly because of the long repeated sequences corresponding to the Pro-, Thr-, and Val-rich linker. The upstream primer (*chiD*-Fw) was re-designed to amplify the gene fragment by removing the region corresponding to the putative signal peptide, the N-terminal first ChBD, and the linker. The resulting truncated *chiD* was amplified and successfully ligated into pCold I (encoding polypeptide is denoted by tChiD). All of the inserted DNAs in the five recombinant plasmids were sequenced to confirm that no PCR errors had occurred. Then each of recombinant plasmids was introduced into *E. coli* and expressed. As shown in Fig. 2, the recombinant proteins were successfully expressed in a soluble form and were purified by affinity chromatography targeting the poly-histidine tag fused to the N-terminus. The molecular weight of ChiC (80.8kDa), ChiE (74.4kDa), and ChiG (63.0kDa) estimated from the mobility on SDS-PAGE gel was somewhat higher than that calculated for recombinant ChiC (67.6kDa), ChiE (59.8 kDa), and ChiG (54.4kDa). This discrepancy may be caused by some distinctive amino acid sequences contained in the recombinant proteins, such as a long linker. The purified recombinant proteins were used for a characterization of chitin-degrading activity.

Fig. 2

1 *Enzymatic characteristics of recombinant Chi proteins.* The five purified recombinant
2 proteins were examined for chitin-degrading activity using colloidal chitin, soluble
3 chitosan (deacetylation degree of 80%), and chromogenic substrates, pNP-$(GlcNAc)_{1-3}$
4 (Table 2). The ORF of *chiF* contains a C-terminal domain whose amino acid sequence Table 2
5 shows no sequence similarity to any of glycosyl hydrolases deposited in the databases.
6 As predicted, the recombinant ChiF could not degrade any of the substrates tested to
7 any extent. The other four recombinant enzymes (rChiC, rtChiD, rChiE, and rChiG)
8 exhibited degrading activity toward colloidal chitin and soluble chitosan, and also
9 released pNP from pNP-$(GlcNAc)_{2 \text{ and } 3}$, but not from pNP-GlcNAc. Using
10 pNP-$(GlcNAc)_3$ as the substrate, rChiC, rtChiD, rChiE, and rChiG had optimal reaction
11 temperatures of 43°C, 45°C, 40°C, and 55°C, respectively, and showed an optimal
12 reaction pH of around 5.0.
13 The final products released from the hydrolysis of GlcNAc pentamer by rChiC,
14 rtChiD, rChiE, and rChiG were analyzed by TLC after a prolonged reaction (Fig. 3). Fig. 3
15 rChiC, rtChiD, and rChiE digested the substrate into a mixture of GlcNAc and GlcNAc
16 dimer, as is usually observed for endo-type chitinases that are able to cut internal
17 glycosydic bonds of oligomers comprising three or more GlcNAc moieties.
18 Interestingly, rChiG completely converted the GlcNAc pentamer into GlcNAc
19 monomers. Therefore, we examined this unique hydrolysis pattern of rChiG in more
20 detail.
21 First, the reactivity of rChiG toward a range of GlcNAc moieties (from dimers to
22 pentamers) was examined (Fig. 4A). All of the GlcNAc oligomers tested were Fig. 4
23 completely converted into GlcNAc monomers after a prolonged reaction. Similarly,
24 rChiG was able to produce GlcNAc as the main product from insoluble chitin polymer,
25 although it needed a longer reaction time to achieve full conversion into GlcNAc (Fig
26 4A). Second, a time course of the reaction products was examined using GlcNAc
27 hexamers as the substrate (Fig. 4B). At the initial stage of the reaction, GlcNAc
28 hexamers were digested into two trimers, and a dimer plus a tetramer, at the same time,
29 but the production of GlcNAc monomer was not detected. These results clearly

demonstrated an endo-type cleavage pattern of ChiG toward GlcNAc oligomers, and also a unique activity to split GlcNAc dimers into monomers, which has not been reported among previously-characterized endo-type chitinases belonging to both family 18 and family19 glycosyl hydrolases.[7, 15)]

Discussion

Among the 15 *chi* genes encoding putative chitin-degrading enzymes, the five genes, *chiCDEFG*, constitute a long and single operon. This was deduced from the short distance between any of the adjacent two genes and the presence of a promoter-like sequence in the upstream region of *chiC*, but not in the regions between any other two adjacent genes.[11)] We analyzed the function of the products of these genes using the recombinant proteins expressed in *E. coli*. As a result, four recombinant proteins (rChiC, rtChiD, rChiE, and rChiG) exhibited a degrading activity toward colloidal chitin, soluble chitosan, pNP-$(GlcNAc)_2$, and pNP-$(GlcNAc)_3$, although the activity of rtChiD was extremely low in comparison with the other three proteins. The rtChiD used in this work was a truncated form, lacking the N-terminal ChBD and linker. The weak activity of rtChiD might be ascribed to an improper folding of the resulting protein. The degradation rate of chitosan was much higher than that of colloidal chitin for all of the four recombinant proteins. This high degradability of chitosan can be explained by its solubility in weak acid, which permits easy access of the recombinant proteins, even though the cleavable bonds in the chitosan polymer are restricted to those located adjacent to GlcNAc moieties, which constitute only 20% of the total moieties. The reactivity toward pNP-$(GlcNAc)_2$ and pNP-$(GlcNAc)_3$, but not pNP-GlcNAc, is usually observed for endo-type chitinases. rChiC, rtChiD, and rChiG released pNP from pNP-$(GlcNAc)_3$ at a higher rate than from pNP-$(GlcNAc)_2$, while this relationship was reversed in the case of rChiE. This indicates that a subsite structure in the catalytic center of ChiE may be different from those of the other three proteins with respect to their binding affinities towards GlcNAc and pNP moieties.

rChiF did not show any chitin-degrading activity (Table 2). However, the location

1 of *chiF* within the operon (*chiCDEFG*) suggests that its gene product should have some
2 function associated with chitin degradation. Polypeptides with a CBM that can bind to
3 chitin, but have no chitinolytic activity, have been found in some chitinolytic bacteria.
4 For example, *Serratia marcescens* CBP21 is a non-catalytic chitin-binding protein
5 belonging to CBM family 33, and is able to enhance chitinolytic activity of other
6 chitin-degrading enzymes by inducing structural changes in insoluble chitin.[16] Similarly,
7 *Bacillus thuringiensis* protein Cbp50 has a modular structure composed of two
8 chitin-binding domains belonging to the CBM5 and CBM33 families.[17] Cbp50 has the
9 ability to inhibit fungal growth by binding to chitin in the fungal cell walls, although its
10 ability to accelerate chitin digestion catalyzed by other chitinolytic enzymes is unclear.
11 *C. shinanonensis* ChiF is expected to be a non-catalytic chitin-binding protein, because
12 it has two ChBDs belonging to CBM family 5. However, preliminary experiments on
13 recombinant ChiF did not show any effect in promoting the chitin-degradation rate of
14 other chitin-degrading enzymes. Further experiments are required to clarify the actual
15 function of ChiF.
16 rChiG completely degraded GlcNAc oligomers, ranging from dimers to hexamers,
17 into GlcNAc monomers, and also digested polymer chitin into GlcNAc as a main
18 product (Fig. 2). This is an unexpected result according to the catalytic abilities of
19 family 18 and family 19 chitinases reported to date.[7, 15] Some of the family 18
20 chitinases exhibited a highly processive mode of action (an exo-type cleavage fashion)
21 toward polymer chitin.[18] They released GlcNAc dimer preferentially as a final product,
22 which is clearly different from the reaction products of rChiG. The time course
23 analysis of rChiG's ability to degrade GlcNAc hexamers indicated that rChiG cut the
24 internal bonds of the hexamer to produce two trimers, or a dimer plus a tetramer, but did
25 not cut the bonds at either end to produce a monomer plus a pentamer, at least at the
26 initial stage of the reaction (Fig. 4B). These results imply that rChiG catalyzes an
27 endo-type cleavage reaction, in which the internal bonds of GlcNAc oligomers are cut
28 at random. Additionally, rChiG was able to cut GlcNAc dimers into GlcNAc monomers,
29 although the reaction rate seemed to be slower than the reaction rate in which rChiG cut

1 the internal bonds of GlcNAc oligomers with three residues or more. rChiG is not likely
2 to catalyze an exo-type cleavage reaction in which GlcNAc is released from either end
3 of GlcNAc oligomers with four residues or more. To the best of our knowledge, this is
4 the first report of an endo-type chitinase that is able to degrade GlcNAc dimers.

5 The *chiG* ORF consists of 508 amino acid residues with a calculated molecular
6 mass of 53.7 kDa (Fig. 1). The signalP software predicted the presence of a signal
7 peptide composed of 26 amino acid residues with high fidelity. The *chiG* ORF contains
8 a C-terminal domain whose sequence is similar to those of catalytic domains of family
9 18 chitinases, although the similarity is fairly low (less than 50%) (Fig. 5). Most of
10 these related genes originate from the whole genome sequence of various bacteria, and
11 the enzymatic characteristics of the gene products have not been clarified. These genes
12 have been annotated as family 18 chitinases based on comparatively low sequence
13 similarities to the well-known family 18 chitinases such as *Bacillus circulans* ChiA1[19]
14 and *Serratia marcescens* ChiB.[20] In fact, the sequence similarity of the C-terminal
15 domain of ChiG to the catalytic domain of these two representative chitinases is too low
16 to be identified by a pfam search (http://pfam.sanger.ac.uk/). In spite of the low overall
17 sequence similarities, the sequence similar to the catalytic motif, DXDXE, of family 18
18 chitinases[21] are conserved among them, including in ChiG (Fig. 5). From these
19 sequence analyses, *C. shinanonensis* ChiG would have a different three-dimensional
20 structure of the catalytic cleft from those of typical family 18 chitinases. This difference
21 may allow ChiG subsite to accommodate the GlcNAc dimer and split it into GlcNAc
22 monomers.

Fig. 5

23 The cooperative action of two enzymes, endo-type chitinase and exo-type
24 β-N-acetylglucosaminidase, is necessary for the entire conversion of chitin into GlcNAc,
25 and a production process using the two enzymes has been proposed.[22] We confirmed
26 that the crude proteins in the culture fluid of the strain SAY3 grown in chitin medium
27 converted polymer chitin into GlcNAc, although the involvement of ChiG was not clear
28 (data not shown). *C. shinanonensis* ChiG is potentially valuable in the development of a
29 single-step conversion of polymer chitin into the commercially valuable

monosaccharide, GlcNAc. Further studies, including determination of catalytic residues, the precise analysis of the reaction kinetics, and determination of the 3D structure, are now being undertaken to clarify the reaction mechanism of the novel chitinase, ChiG.

Acknowledgments

This work was supported by a Grant-in-Aid from the Global COE Program by the Ministry of Education, Culture, Sports, Science, and Technology of Japan, and by The Institute for Fermentation, Osaka (IFO) Research Grant. We are indebted to the Division of Gene Research of the Research Center for Human and Environmental Sciences, Shinshu University, for providing facilities.

References

1) Tharanathan RN and Kittur FS, *Crit. Rev. Food Sci. Nutr.*, **43**, 61–87 (2003).
2) Synowiecki J and Al-Khateeb NA, *Crit. Rev. Food Sci. Nutr.*, **43**, 145–171 (2003).
3) Yu C, Bassler BL, and Roseman S, *J. Biol. Chem.*, **268**, 9405-9409(1993).
4) Ravi Kumar MNV, *React. Funct. Polym.*, **46**, 1-27 (2000).
5) Dutta PK, Dutta J, and Tripathi VS, *J. Sci. Ind. Res.*, **63**, 20-31 (2004).
6) Ebel J, *BioEssays*, **20**, 569-576 (1998).
7) Cohen-Kupiec R and Chet I, *Curr. Opin. Biotechnol.*, **9**, 270–277 (1998).
8) Sato K, Azama Y, Nogawa M, Taguchi G, and Shimosaka M, *J. Biosci. Bioeng.*, **109**, 472-478 (2010).
9) Sato K, Kato Y, Fukamachi A, Nogawa M, Taguchi G, and Shimosaka M, *Biosci. Biotechnol. Biochem.*, **74**, 636-640 (2010).
10) Sato K, Kato Y, Taguchi G, Nogawa M, Yokota A, and Shimosaka M, *J. Gen. Appl. Microbiol.*, **55**, 147-153 (2009).
11) Huang L, Garbulewska E, Sato K, Kato Y, Nogawa M, Taguchi G, and Shimosaka M, *J. Biosci. Bioeng.,* in press.
12) Henrissat B and Bairoch A, *Biochem. J.*, **293**, 781-788 (1993).

1 13) Imoto T and Yagishita K, *Agr. Biol. Chem.*, **35**, 1154-1156 (1971).
2 14) Boraston AB, Bolam DN, Gilbert HJ, and Davies GJ, *Biochem. J.*, **382**, 769-781
3 (2004).
4 15) Ubhayasekera W, *Polym. Int.*, **60**, 890-896 (2011).
5 16) Vaaje-Kolstad G, Horn SJ, van Aalten DMF, Synstad B, and Eijsink VGH, *J. Biol.*
6 *Chem.*, **280**, 28492-28497 (2005).
7 17) Mehmood MA, Xiao X, Hafeez FY, Gai Y, and Wang F, *Anton. Leeuw Int. J. G.*,
8 **100**, 445-453 (2011).
9 18) Horn SJ, Sørbotten A, Synstad B, Sikorski P, Sørlie M, Vårum KM, and Eijsink
10 VGH, *FEBS J.*, **273**, 491-503 (2006).
11 19) Watanabe T, Suzuki K, Oyanagi W, Ohnishi K, and Tanaka H, *J. Biol. Chem.*,
12 **265**, 15659-15665 (1990).
13 20) van Aalten DMF, Synstad B, Brurberg MB, Hough E, Riise BW, Eijsink VGH,
14 and Wierenga RK, *Proc. Natl. Acad. Sci. USA*, **97**, 5842-5847 (2000).
15 21) Yang J, Gan Z, Lou Z, Tao N, Mi Q, Liang L, Sun Y, Guo Y, Huang X, Zou C,
16 Rao Z, Meng Z, and Zhang KQ, *Microbiology*, **156**, 3566–3574 (2010).
17 22) Binod P, Sandhya C, Suma P, Szakacs G, and Pandey A, *Bioresour. Technol.*, **98**,
18 2742–2748 (2007).
19
20

Figure legends

Fig. 1 Deduced domain structures of ORFs encoded by *chiCDEFG*, constituting a single operon.

Domain structures deduced from the amino acid sequences are shown for each gene product, with the number of residues and the calculated molecular weight. Black and gray boxes indicate signal peptides and chitin-binding domains, respectively. White boxes indicate catalytic domains of glycosyl hydrolases with family numbers. Vertical lined bars indicate Pro-, Thr-, and Val-rich linkers.

Fig. 2 SDS-PAGE analysis of the purified recombinant proteins.

Recombinant proteins for ChiC (lane 1), tChiD (lane 2), ChiE (lane 3), ChiF (lane 4), and ChiG (lane 5) were analyzed. Lane M, molecular weight markers.

Fig. 3 TLC analysis of reaction products released from GlcNAc pentamers. GlcNAc pentamers were incubated with the purified rChiC (lane 2), rtChiD (lane 3), rChiE (lane 4), and rChiG (lane 5) overnight and analyzed by TLC. The concentration of each enzyme was in the range of 0.5-1.0 μM. The control without enzyme (lane 1) and a mixture of standard GlcNAc oligomers (lane M) are also included. G1, GlcNAc; G2, GlcNAc dimer; G3, trimer; G4, tetramer; G5, pentamer.

Fig. 4 TLC analysis of the reaction products catalyzed by rChiG.

(A) Final products released from GlcNAc oligomers, from dimer to pentamer, and colloidal chitin, using the purified rChiG (at the concentration of 1.5 μM), overnight. The substrates used are GlcNAc dimer (lane 1 and 2), trimer (lane 3 and 4), tetramer (lane 5 and 6), pentamer (lane 7 and 8), and colloidal chitin (lane 9 and 10). The reaction mixture was incubated in the absence (lanes 1, 3, 5, 7 and 9) or in the presence (lanes 2, 4, 6, 8 and 10) of rChiG. (B) Reaction products from GlcNAc hexamers at different time intervals. GlcNAc hexamers were incubated with rChiG (3.0 μM) for 0 min (lane 1), 1 min (lane 2), 10 min (lane 3), 30 min (lane 4), 1h (lane 5), and 24h (lane 6). Lane M, a mixture of standard oligomers. G1, GlcNAc; G2,GlcNAc dimer; G3, trimer; G4, tetramer; G5, pentamer; G6, hexamer.

Fig. 5 Comparison of the amino acid sequences of the C-terminal domain of ChiG and other similar polypeptides.

Residues conserved in all sequences are highlighted with a black ground. Those conserved in >60% of the sequences are shown with a gray background. The dashes indicate gaps. The sequences are from *Chitiniphilus shinanonensis* ChiG (this work; top line), *Catenulispora acidiphila* family 18 glycosyl hydrolase (accession number CP001700, second line), *Streptomyces lividans* secreted chitinase (GG657756, third line), *Streptomyces coelicolor* putative secreted chitinase (AL939126, fourth line), and *Kitasatospora setae*, uncharacterized protein (AP010968, bottom line). The conserved sequences similar to the motif DXDXE, which involves the catalytic residues of family 18 chitinases are marked with asterisks. Numerals show the number of amino acid residues, starting from the initial codon.

Table 1. Oligonucleotides used for PCR amplification of *chi* genes

Oligo-nucleotides	Sequences (5' -3')	Restriction cleavage sites
chiC-Fw	GCCATATGGCGCCGGCATGGCAGGAA	*Nde* I
chiC-Rv	GCAAGCTTTTACTTCACGGCGCTCATC	*Hin*d III
chiD-Fw	GCGAATTCTACAACGGCAAGAGCTAC	*Eco*R I
chiD-Rv	GCGTCGACTTACTTGATCGACTTGAT	*Sal* I
chiE-Fw	GCGGATCCGCGCCGGCCTGGCAGGAA	*Bam* HI
chiE-Rv	GCAAGCTTTCAAGGCATGGCATTGAA	*Hin*d III
chiF-Fw	GCGGATCCGCGCCGGCCTGGCAGGAA	*Bam* HI
chiF-Rv	GCAAGCTTTCAATAGCGGATGTTCGG	*Hin*d III
chiG-Fw	GCGGATCCATGGCCGCGCATGCCGCG	*Bam* HI
chiG-Rv	GCAAGCTTTCAGAACTCGCCGAAGAT	*Hin*d III

Underlined sequences are restriction endonuclease cleavage sites.

The nucleotide sequence of the *ch*i gene cluster was deposited in the DDBJ/EMBL/GenBank database under the accession number of AB649129.

Table 2. Degrading activities of five recombinant Chi proteins for various chitinous substrates

Recombinant proteins[b]	Degrading activity (μmol/min/mg-protein)[a]				
	Colloidal chitin	Soluble chitosan (DD 80%)[c]	pNP-GlcNAc	$pNP\text{-}(GlcNAc)_2$	$pNP\text{-}(GlcNAc)_3$
ChiC	3.78	10.8	ND[d]	0.68	3.5
tChiD[e]	ND	0.07	ND	0.05	0.34
ChiE	0.5	3.57	ND	8.9	1.46
ChiF	ND	ND	ND	ND	ND
ChiG	1.64	28.2	ND	0.08	0.44

[a]The amount of reducing sugar equivalent to GlcNAc liberated was measured for collidal chitin and soluble chitosan, and the amount of pNP released was determined for pNP-chromogenic substrates

[b]The concentration of each enzyme in the reaction mixture was in the range of 20-100 nM.

[c]DD, degree of deacetylation. [d]not detectable (<0.01 μmol/min/mg-protein)

[e]Truncated protein was used in the case of ChiD

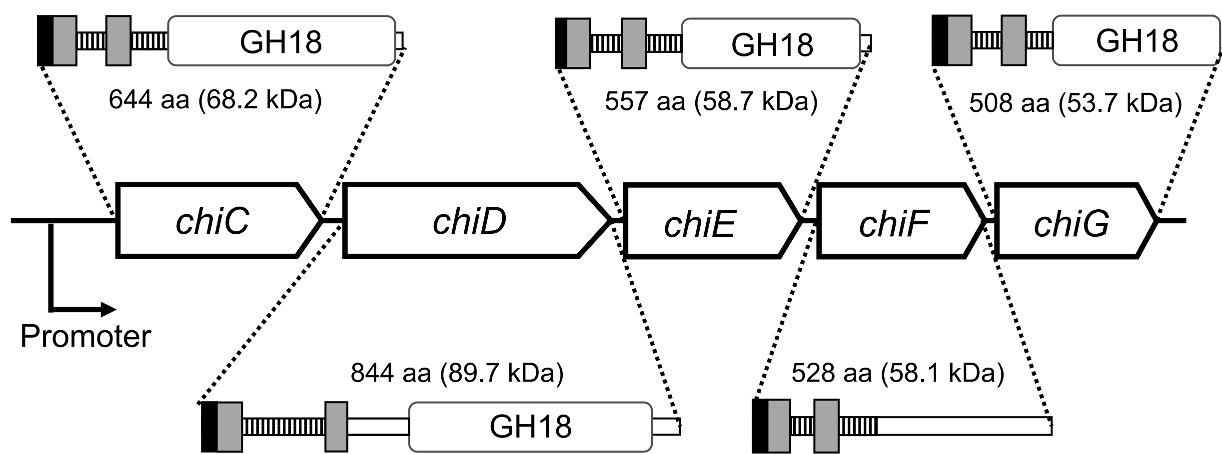

Fig. 1 Huang *et al*.

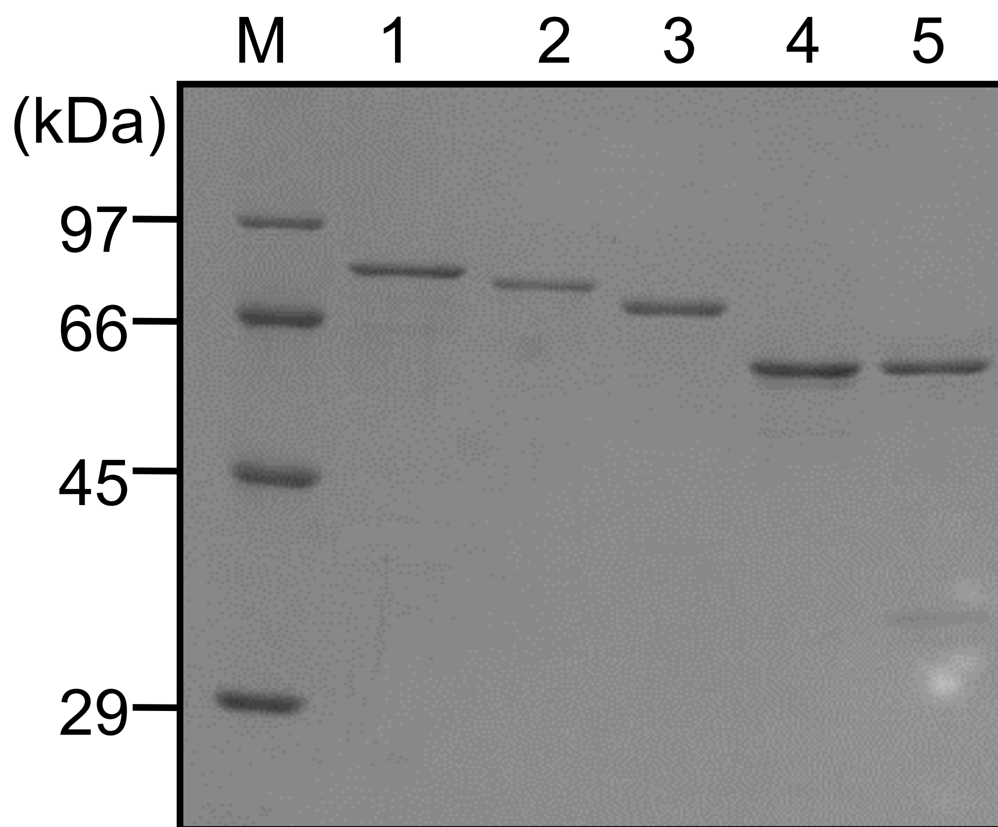

Fig. 2 Huang *et al*.

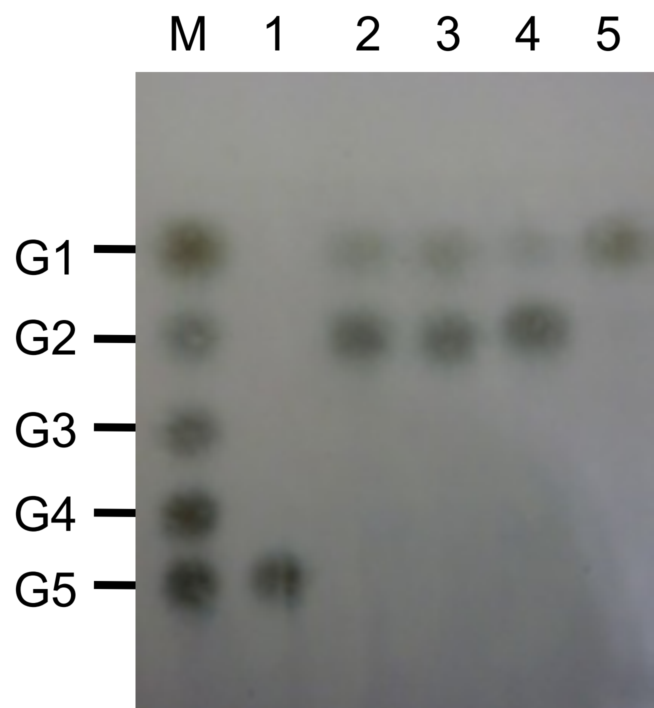

Fig. 3 Huang *et al*.

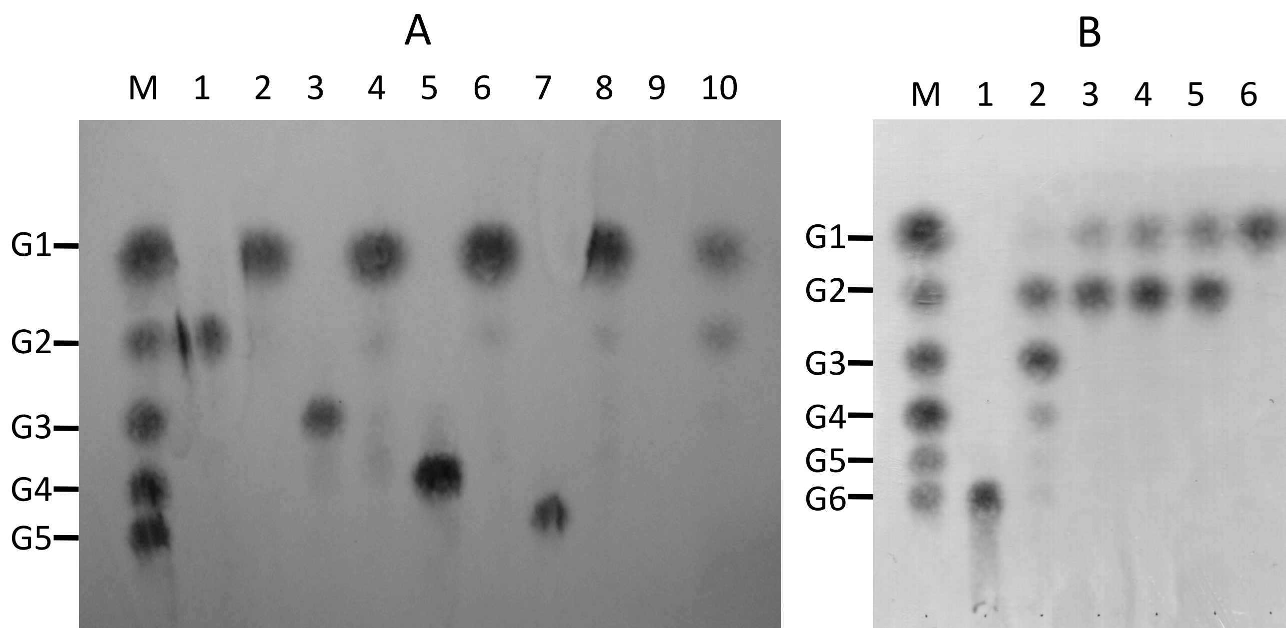

Fig. 4 Huang *et al*.

```
ChiG            214  --FAPYIDVSGSMDLSGWAQATGGRYVSLAFFNSA--GGCNGGW------PTGEAGLLSQ  263
C. acidiphila    50  HVFAPYFEAYNGDDPAALSQASGAKYLTMAFIQTASKGSCTVDWDGDTSTPISSSTYGSA  109
S. lividans      37  RVFAPYFEAWTGESPAALAAQSGAKHLTMAFLQTATPGSCTPYWNGDTSMPIAQSTFGAD   96
S. coelicolor    40  RVFAPYFEAWTGESPAALAAQSGAKHLTMAFLQTATPGSCTPYWNGDTSMPIAQSTFGAD   99
K. setae         37  HVFAPYFEAWTGESPAALAAQSGAKHLTMAFLQAATKGSCTPLWNGDTAMPVSNSIFGAD   96

                                                                 *****
ChiG            264  ANSLKAFGGNVIVSSGGWNAN----DLARSCTDASALATTYENVLDRFGTNRLDLDPENA  319
C. acidiphila   110  ISTIRAGGGDVIPSFGGYAADNGGTEIADSCTNVSSIAAAYENVITTYNVTRLDLDTEDN  169
S. lividans      97  FDTIQANGGDVIPSFGGYTADTTGTEIADSCTDVQQIAAAYEKVITTYDVSRLDMDIEID  156
S. coelicolor   100  FDTIQANGGDVIPSFGGYTADTTGTEIADSCTDVQQIAAAYEKVITTYDVSRLDMDIEID  159
K. setae         97  IATIRANGGDVIPSFGGYTADNTGTEIADSCTDVNQIAAAFESVITTYDVTRIDLDIEDN  156

ChiG            320  PGNNNLEVAVVDRRNAALKILQDRAKAKGKTVFVSYTLGVNPDGGFNSENLYVLQSAKNA  379
C. acidiphila   170  SLTNT---AGIDRRNKAIKMVEDWAAANGRTVQFTYTLPTTTS-GLAGSGLKVLQNAVTN  225
S. lividans     157  ALDNT---AGIDRRNKAIKLVQDWAAANGRDLEISYTLPTTTR-GLASNGVALLENAVKN  212
S. coelicolor   160  ALDNT---AGIDRRNKAIKLVQDWAAANGRDLEISYTLPTTTR-GLASNGVALLENAVKN  215
K. setae        157  SLTNT---AGIDRRNKAIKIVEDWAAANGRTIQFSYTLPTTTS-GLASSGLAVLKNAVTN  212

ChiG            380  GVEVSLVNPMIMDYYDGVSGNQMGARSILALQKVHAQIKNLWPGKTDAQYWGMLSATAMI  439
C. acidiphila   226  NARVDVVNIMTFDYYDG-ATHQMANDTKTAATGLEKQLATLYPTKTAAQLWGMVGVTEMP  284
S. lividans     213  GTKVDVVNLMTFDYYDN-QQHDMARDTQTATQGLHDVLARLHPGKSSADLW---------  262
S. coelicolor   216  GTKVDVVNLMTFDYYDN-QQHDMARDTQTATQGLHDVLARLHPGKSSADLW---------  265
K. setae        213  NARIDVVNMMTFDYYDN-AAHDMAADTQTSAQGLYNQLA---------------------  250
```

Fig. 5 Huang *et al*.