Biosci. Biotechnol. Biochem., **71** (7), 1703–1716, 2007

Molecular Cloning of cDNAs and Genes for Three α-Glucosidases from European Honeybees, *Apis mellifera* L., and Heterologous Production of Recombinant Enzymes in *Pichia pastoris*

Mamoru NISHIMOTO,[1,*] Haruhide MORI,[1,*] Tsuneharu MOTEKI,[1] Yukiko TAKAMURA,[1] Gaku IWAI,[1] Yu MIYAGUCHI,[1] Masayuki OKUYAMA,[1] Jintanart WONGCHAWALIT,[1] Rudee SURARIT,[2] Jisnuson SVASTI,[3] Atsuo KIMURA,[1,†] and Seiya CHIBA[1]

[1]*Division of Applied Bioscience, Research Faculty of Agriculture, Hokkaido University, Sapporo 060-8589, Japan*
[2]*Faculty of Dentistry, Mahidol University, Bangkok 10400, Thailand*
[3]*Faculty of Science, Mahidol University, Bangkok 10400, Thailand*

Received March 2, 2007; Accepted April 17, 2007; Online Publication, July 7, 2007
[doi:10.1271/bbb.70125]

cDNAs encoding three α-glucosidases (HBGases I, II, and III) from European honeybees, *Apis mellifera*, were cloned and sequenced, two of which were expressed in *Pichia pastoris*. The cDNAs for HBGases I, II, and III were 1,986, 1,910, and 1,915 bp in length, and included ORFs of 1,767, 1,743, and 1,704 bp encoding polypeptides comprised of 588, 580, and 567 amino acid residues, respectively. The deduced proteins of HBGases I, II, and III contained 18, 14, and 8 putative *N*-linked glycosylation sites, respectively, but at least 2 sites in HBGase II were unmodified by *N*-linked oligosaccharide. In spite of remarkable differences in the substrate specificities of the three HBGases, high homologies (38–44% identity) were found in the deduced amino acid sequences. In addition, three genomic DNAs, of 13,325, 2,759, and 27,643 bp, encoding HBGases I, II, and III, respectively, were isolated from honeybees, and the sequences were analyzed. The gene of HBGase I was found to be composed of 8 exons and 7 introns. The gene of HBGase II was not divided by intron. The gene of HBGase III was confirmed to be made up of 9 exons and 8 introns, and to be located in the region upstream the gene of HBGase I.

Key words: honeybee α-glucosidase; *Apis mellifera*; cloning of α-glucosidase; *Pichia pastoris*

α-Glucosidase (EC 3.2.1.20, α-D-glucoside glucohydrolase) is a group of typical exo-type carbohydrolase that catalyzes the hydrolysis of α-glucosidic linkage on the non-reducing terminal side of substrate to produce α-glucose. α-Glucosidases have been classified into the two families (I and II) based on the primary structure.[1)] Family I enzymes, for instance, α-glucosidases derived from *Saccharomyces cerevisiae*,[2)] *Candida albicans*,[3)] *Bacillus stearothermophilus*,[4)] *Drosophila melanogaster*,[5)] and *Aedes aegypti*,[6)] have four conserved regions important in the catalytic action. The regions are shared by glycoside hydrolase family 13 (GH 13) enzymes,[7,8)] such as α-amylase, α-glucosidase, dextrin 6-α-D-glucanohydrolase, pullulanase, neopullulanase, cyclomaltodextrin glucanotransferase, and branching enzyme. Family II α-glucosidases, for instance, enzymes derived from *Aspergillus niger*,[9,10)] *Schizosaccharomyces pombe*,[11,12)] human lysosome,[13)] and barley,[14)] do not have four such regions, and belong to glycoside hydrolase family 31 (GH 31),[8)] consisting mainly of α-glucosidases. It has been ascertained that some members in GH 13 possess the $(\beta/\alpha)_8$ barrel structure as the catalytic domain, and that the C-terminal domain is comprised of β-sheets. *S. pombe* α-glucosidase[11,12)] belonging to family II has four regions, but the two of the regions are not involved in the catalytic action. Of the α-glucosidases in GH 13 and GH 31, the three-dimensional structures of *Bacillus cereus* and *Sulfolobus solfataricus* α-glucosidases[15,16)] were solely solved by X-ray crystallographic analysis. The tertiary structures of both the enzymes were shaped into $(\beta/\alpha)_8$ barrel as the catalytic domain, but the features of the active sites involved in the catalytic reactions appears to be different between GH 13 and GH 31.

We have reported that the imagoes of European honeybees (*Apis mellifera*) have three kinds of α-glucosidases (HBGases I, II, and III), which are different in their substrate specificities.[17–20)] HBGases I, II, and III exhibited high activity towards maltose, and towards heterogeneous substrates such as aryl α-glucosides and sucrose. Although the three α-glucosidases were mono-

* These two authors equally contributed to this paper.
† To whom correspondence should be addressed. Fax: +81-11-706-2808; E-mail: kimura@abs.agr.hokudai.ac.jp

meric enzymes, allosteric properties were observed in the reaction of HBGases I and II. HBGase I, which was present in the ventriculus,[21] showed positive- and negative-cooperative reactions, depending on the kinds of substrates.[18] Sucrose was hydrolyzed in a mode of negative cooperativity, meaning that the catalytic reaction was unsuitable for the normal Michaelis–Menten equation due to acceleration with an increase in the concentration of sucrose.[17,18] HBGase I showed little or no activity towards nigerose, isomaltose, and soluble starch.[17] HBGase II, which was present in the ventriculus and the haemolymph, showed the positive-cooperativity in the reaction on sucrose, and the broadest substrate specificity of the three enzymes, that is, HBGase II was capable of attacking maltooligosaccharides, glucobioses having α-1,2-, α-1,3-, and α-1,6-α-glucosidic linkages, phenyl α-glucoside, sucrose, soluble starch, and also α-glucosyl-phosphate.[19] HBGase III followed the typical Michaelis–Menten type of hydrolytic reaction, and showed no cooperativity.[20] The substrate specificity was relatively narrow to be limited in maltose, maltotriose, sucrose, and aryl α-glucosides. HBGase III was clarified to be localized in the hypopharyngeal gland,[21,22] and to be secreted into nectar gathered from flowers by honeybees to make up honey through the hydrolysis of sucrose in nectar.[21,22]

Three HBGases were characterized by differences in their substrate specificities, in particular, the individual reactivity toward sucrose.[18–20] The enzymes were discriminated also by their own antibodies, since the antiserum prepared from each α-glucosidase showed no cross-reaction against the two other α-glucosidases.[19–21] The finding suggests that HBGases I, II, and III are isoenzymes having different amino acid sequences. The overall genome sequencing of European honeybee, *Apis mellifera*, was recently completed.[23] In the case of eukaryote, however, even if the genome sequence is disclosed, the amino acid sequence of the desired protein is not always deduced readily from the nucleotide sequence which contains thousands of introns. Therefore, we could not clarify the amino acid sequences of three α-glucosidases of the honeybee using only the data of GenBank. As for the amino acid sequence of honeybee α-glucosidase, the sequence[22] of α-glucosidase from the hypopharyngeal gland was reported on the basis of the nucleotide sequence analysis of cDNA encoding the enzyme, corresponding to HBGase III.

This paper describes the molecular cloning of cDNAs encoding HBGases I, II, and III, the genomic DNA structures of the three enzymes, and their heterologous expression in *Pichia pastoris* cells giving the active recombinant HBGases (rHBGases) II and III, which have almost the same properties as those of native enzymes. This is the first report that three HBGases are proven to be isoenzymes encoded by the individual gene, and that the insect α-glucosidase isoenzymes are produced in heterologous host cells.

Materials and Methods

Honeybees and enzymes. Honeybees were purchased from a beekeeper in Ishikari, Japan, and kept in hives on the Hokkaido University campus. In this study, we used honeybees kept at $-80\,^{\circ}$C. Three α-glucosidases were purified from adult honeybees, as described previously.[17,20]

Strains. The strains of *Escherichia coli* were NM522 and JM109 for DNA manipulations, and XL1-Blue MRA (P2) and C600 *hfl* for plaque preparation. The *Pichia pastoris* GS115 (*his4*) strain (Invitrogen, Carlsbad, CA, USA) was used as a host for the enzyme production.

Analysis of amino acid sequence. To analyze the N-terminal and internal amino acid sequences of the α-glucosidases, *S*-pyridylethylated enzymes were blotted onto PVDF membrane using ProSorb (Applied Biosystems, Foster City, CA, USA) or digested with lysyl endopeptidase (Wako Pure Chemical, Osaka, Japan) in 100 mM Tris–HCl buffer (pH 9.0) containing 4 M urea at $37\,^{\circ}$C for 48 h. The peptides were isolated from the digest by reverse-phase high performance liquid chromatography using a Capcell Pak C_{18} UG120 ($\phi 4.6 \times 15$ mm, Shiseido Fine Chemicals, Tokyo, Japan). The peptides were transferred on the blotted membranes and then subjected to automated Edman degradation using protein–peptide sequencer model 477A with an on-line phenylthiohydantoin analyzer (model 120A; Applied Biosystems).

Construction of cDNA library. Fresh honeybee imagoes (thorax, abdomen, and head) were crashed in a mortar in liquid nitrogen, and total RNA was prepared by the acid guanidine thiocyanate–phenol–chloroform method.[24] Poly(A)$^+$RNA was isolated using Oligotex-dT30 Super (Takara Bio, Otsu, Japan). cDNA was synthesized with Superscript Choice System (Life Technologies, Rockville, MD, USA) using 10 µg of poly(A)$^+$RNA as a template, ligated to an *Eco*RI-*Not*I-*Bam*HI adapter (Takara Bio), and dephosphorylated with alkaline phosphatase (Life Technologies). After size-selection using a Sephacryl S-500 HR column (Life Technologies), double-stranded cDNA containing *Eco*RI termini was ligated into the unique *Eco*RI site of λgt10 phage DNA. The recombinant DNAs were packaged with Gigapack III Gold Packaging Extract (Stratagene Cloning System, La Jolla, CA, USA), and propagated in *E. coli.*

Screening by plaque hybridization. As a probe for screening cDNA libraries, cDNA was generated by polymerase chain reaction (PCR) of 25 cycles at $94\,^{\circ}$C for 30 s, $55\,^{\circ}$C for 30 s, and $72\,^{\circ}$C for 1 min using degenerated primers designed on the internal amino acid sequences (see the sequences in Figs. 1, 2, and 3). The

designated primers for HBGase I were GAGTGGIGIGATTTTGTIGATAATTATGCIGAGGA (sense) and GTTGTCCCATTGGAAIGGIGTICIGCAICCGTC (antisense) on Glu277 to Glu288 and Asp410 to Asn420; for HBGase II, CCITATGATGAGTATTATGTITGGTTIGATGC (sense) and TCIGCIATCAACATGAAGTTCAAIGGIAC (antisense) on Pro141 to Ala151 and Val310 to Asp319; and for HBGase III, GTIGATAATTGGATGACITATATGCCICC (sense) and TCGTCCCATTGGAAIGGIGTICIIGCIGGGTC (antisense) on Val327 to Pro336 and Asp414 to Asp424, respectively. The PCR products were labeled with digoxigenin (DIG) using DIG-High Prime (Boehringer Mannheim, Mannheim, Germany), and were used to isolate cDNA clones. Plaques were transferred onto duplicate nylon membrane (Hybond-N+ membrane, Amersham Biosciences, Buckinghamshire, UK) and were hybridized to the DIG labeled probes in rapid hyb buffer (Amersham Biosciences) at 65 °C for 15 h, followed by washing and detection (DIG Luminescent Detection Kit, Boehringer Mannheim). Isolation of the positive clones was completed by the second and third screenings, and DNA prepared from the clones using QIAGEN Lambda Midi Kit (Qiagen, Hilden, Germany) was subcloned into pBluescript II SK(+) (Stratagene), and sequenced in both strands by the dideoxy chain termination method[25] using automated DNA sequencer model 373A (Applied Biosystems).

5′ Rapid amplification of cDNA end (RACE). To the 3′ end of single-stranded cDNA, extended from random hexamer using poly $(A)^+$RNA from honeybees as template and Superscript II (Life Technologies), poly (dA) tail was added with terminal deoxynucleotidyl transferase (Life Technologies). 5′ RACE was performed using an anchor primer (AP, GGCCACGCGTCGACTAGTACTTTTTTTTTTTTTTT) and gene-specific primers (GSP, CGCTGTTATTCCAGATTCGATGAAG for HBGase I, and GTACCAGTTGGCGTCCACGC for HBGase II) with 1 cycle at 98 °C for 5 min, 50 °C for 2 min, and 72 °C for 40 min, followed by 28 cycles of 94 °C for 1 min, 54 °C for 1 min, and 72 °C for 3 min after the addition of universal amplification primer (UAP, GGCCACGCGTCGACTAGTAC). Triplet PCRs were done for each cDNA, and the products were sequenced as mentioned above.

Construction of expression vectors. For production of the enzymes secreted using their own signal sequences, the cDNA part including the 5′ untranslated region (UTR) was amplified by PCR to generate a *Bam*HI site using the following primers: CCGTA**GGATCC**AATCATGAAGAGC (SP1) and CGCTGTTATTCCAGATTCGATGAAG (ASP1) for HBGase I; GTCAAT**GGATCC**GCAAAATGTTTC (SP2) and TCIGCIATGAACATGAAGTTGAAIGGIAC (ASP2) for HBGase II; and ACTTCT**GGATCC**TAGCATGAAGGC (SP3) and CTTCGATATCACCAATACCA (ASP3) for HBGase III, where the underlined and bold-faced nucleotides the indicate initial Met codon and *Bam*HI site introduced, respectively. The *Bam*HI-*Hin*dIII DNA fragment of the PCR products and the *Hin*dIII-*Not*I fragment of cDNA were ligated into pPIC3.5 vector (Invitrogen) at the *Bam*HI-*Not*I sites, generating pPIC3.5on, pPIC3.5tw, and pPIC3.5th for HBGase I, II, and III, respectively.

For production of the proteins equipped with N-terminal signal peptide derived from *Saccharomyces cerevisiae* α-factor, cDNA fragments encoding the mature enzymes were introduced into pPIC9 (Invitrogen) as above, but with *Eco*RI instead of *Bam*HI, and the following primers: CGCGGTCGGCCTTGGC**GAATTC**CAAAAC (SP4) and ASP1 for HBGase I, CTCCGATC**GAATTC**GTGGACGCC (SP5) and ASP2 for HBGase II, and CCATTGTG**GAATTC**GCATGGAAGC (SP6) and ASP3 for HBGase III, where the underlined and bold-faced nucleotides indicate the codon for the N-terminal amino acid and the *Eco*RI site, respectively. The resulting expression plasmids were named pPIC9on, pPIC9tw, and pPIC9th for HBGase I, II, and III, respectively. The whole cDNAs and their flanking regions in the constructed plasmids were sequenced and confirmed to have no unexpected mutations.

Transformation of P. pastoris. Transformation of *P. pastoris* GS115 with *Sac*I-linearized expression plasmids was done by electroporation using Gene Pulser (Bio-Rad, Richmond, CA, USA) according to the recommendations of Invitrogen, a manual of methods for the expression of recombinant proteins in *P. pastoris*. After incubation for 4 d at 30 °C on RDB agar plates consisting of 1 M sorbitol, 2% glucose, 1.34% Yeast Nitrogen Base (w/o amino acid; Becton Dickinson, Sparks, MD, USA), 4 μg/ml D-biotin, and 50 μg/ml of each amino acid (Glu, Leu, Ile, Lys, and Met), the His-autotrophic transformants (His^+) were selected and retained on an RDB agar plate for further study.

Induction and purification of recombinant proteins. The transformants were inoculated into 10 ml of BMGY medium (1% yeast extract, 2% peptone, 1.34% Yeast Nitrogen Base, 4 μg/ml D-biotin, 1% glycerol, 0.1 M potassium phosphate buffer, pH 6.0) in a 50-ml flask, and then incubated at 30 °C for 18 h with vigorous shaking. Induction of protein expression was achieved by incubation with 0.5% methanol. The cells, collected by centrifugation at 3,000 × *g* for 5 min at room temperature, were transferred into 20 ml of BMMY medium exchanging 1% glycerol for 0.5% methanol in BMGY medium, and cultivated at 30 °C for 50 h. At the indicated times, the supernatant was collected by centrifugation at 10,000 × *g* for 5 min at 4 °C, and the enzyme activity was measured.

To purify the recombinant HBGase II and III, the transformants were cultured in 1.0 and 1.2 liters of BMGY medium at 30 °C for 30 h, followed by induction culture in 2 and 6 liters of BMMY medium for 200 h,

respectively. The values of A_{600} at the beginning of induction were 4.8 for HBGase II and 2.1 for HBGase III. The supernatant was collected by centrifugation at 10,000 × *g* for 10 min, and then, to the clarified supernatant, ammonium sulfate was added at up to 90% saturation with stirring. The resulting turbid suspension was kept at 4 °C for 12 h, and then the precipitate was collected by filtration with Celite (Wako Pure Chemicals), dissolved in 50 mM sodium acetate buffer (pH 4.7), dialyzed against the same buffer, and applied sequentially to columns of CM Sepharose CL-6B (Amersham Pharmacia Biotech, Piscataway, NJ, USA), Bio-Gel P-100 (Bio-Rad), DEAE Sepharose CL-6B (Pharmacia), and BUTYL-TOYOPEARL 650M (Tosoh, Tokyo, Japan). The homogeneity of the purified enzymes was confirmed by SDS–PAGE.

Enzyme assay and protein measurement. Measurements of enzyme activity and protein concentration were done by methods described previously.[17,20] The protein concentrations of rHBGases II and III were calculated from the values of $E^{1\%}_{1\,cm}$ at 280 nm of 11.7 and 15.1, respectively, which were obtained from the relationship between the dry weight and A_{280} of each purified enzyme preparation.

SDS–PAGE, western blotting, deglycosylation and sugar contents. SDS–PAGE was carried out by the method of Laemmli[26] in 8% polyacrylamide gel, and proteins were stained by Rapid CBB KANTO (Kanto Chemical, Tokyo, Japan). Western blotting after SDS–PAGE was performed by transfer onto PVDF membrane (Immobilon; Millipore, Bedford, MA, USA) in a semi-dry apparatus and detection using Blotting detection kit (Amersham Pharmacia Biotech) and antisera for each HBGase as the first antibody. Deglycosylation was performed with endoglycosidase F (Boehringer Mannheim) and endoglycosidase H (Roche Molecular Biochemical, Basel, Switzerland). Native and recombinant HBGases II and III (2 μg) in 30 μl of 0.1 M sodium acetate buffer (pH 5.5) were incubated with endoglycosidase F (0.1 U) at 30 °C for 36 h. HBGase III (10 μg) was denatured by treatment at 96 °C for 10 min in 0.15% SDS, followed by incubation at 37 °C for 20 h with endoglycosidase H (5 mU) in 0.05 M sodium acetate buffer (pH 5.5) containing 0.1% SDS. The sugar contents of the enzymes were measured as mannose by the phenol-sulfate method.[27]

Properties of the recombinant enzyme. The optimum pH for hydrolysis in 5.8 mM maltose was measured at 37 °C in Britton-Robinson buffer (pH 3–10) prepared with 0.2 M NaOH and an acid mixture (40 mM acetic acid, 40 mM phosphoric acid, and 40 mM boric acid). The pH and thermal stabilities were determined by measuring the residual activity of the enzymes after they were maintained at 4 °C for 24 h in Britton-Robinson buffer (pH 2.5–12), and at various temperatures for 15 min in 70 mM sodium acetate buffer (pH 5.5) containing 0.05% Triton X-100, respectively. The kinetic parameters, K_m and k_0, for the hydrolysis of substrates were estimated by $1/s$ *versus* $1/v$ plots, in which the initial velocities (v) were measured under various substrate concentrations (s). The ranges of s for recombinant HBGase II and III were 0.5–8.0 mM and 4.0–20 mM for maltose, 1.0–6.5 mM and 4.0–18 mM for maltotriose, 2.0–40 mM and 10–60 mM for sucrose, and 0.2–20 mM and 10–60 mM for phenyl α-glucoside, respectively. The enzyme reactions were carried out under standard assay conditions.

Analysis of genomic DNA. Genomic DNA (0.2 mg) was prepared from honeybees (1 g) by the CTAB method,[28] and subjected to partial *Sau*3AI digestion. The resulting DNA fragments were fractionated by ultracentrifugation (5–20% NaCl; 200,000 × *g*, for 270 min at 20 °C), and from DNA ranging in length from 9,000 to 23,000 nt, a library was constructed on λEMBL3 at the *Bam*HI site. The other libraries were prepared from restriction fragments with expected lengths. The restriction fragments cleaved by *Hin*dIII and *Bam*HI or *Sal*I and *Bam*HI were isolated by agarose gel electrophoresis, ligated into Bluescript II SK(+) (Stratagene), and used for transformation of *E. coli* DH5α (library efficiency grade, Life Technologies). Hybridization for the screening of positive clones was performed as mentioned above. Inverse PCR[29] was performed in 25 cycles at 94 °C for 1 min and at 68 °C for 3 min using cDNA sequence-based primers and self-ligated restriction DNA fragments as template DNA. Polymerases with high fidelity, KOD dash polymerase (Toyobo, Osaka, Japan), LA Taq (Takara Bio), and PrimeStar (Takara Bio), were employed in all PCR procedures.

The sequence of *hbg1* was determined from inverse PCR products (1–6,844 at the *Eco*RI site and 9,314–13,325 at the *Hin*dIII site) and PCR products (2,764–5,918, 5,871–8,243, 8,195–9,421, 9,314–10,885). The *hbg2* sequence was from the limited library (1–1,812, *Sal*I-*Bam*HI) and PCR (877–2,759). The *hbg3* sequence was from inverse PCR (1–1,854 at *Hin*cII, 1,194–8,403 at *Bgl*II), the λ library (6,234–20,965), the limited library (19,012–25,493, *Hin*dIII-*Bam*HI), and PCR assisted by the genome project information in the antisense primer (25,459–27,643).

Results

Isolation of α-glucosidase cDNAs

Prior to cDNA cloning, many peptide fragments were isolated from lysyl endopeptidase digests of HBGases I, II, and III to analyze the N-terminal and internal amino acid sequences, and to synthesize oligonucleotide primers based on the sequences. Of 29, 22, and 12 peptide samples isolated for HBGases I, II, and III, 22, 20, and 8 peptides were successfully sequenced, respectively (shown in bold letters in Figs. 1, 2, and 3). The N-

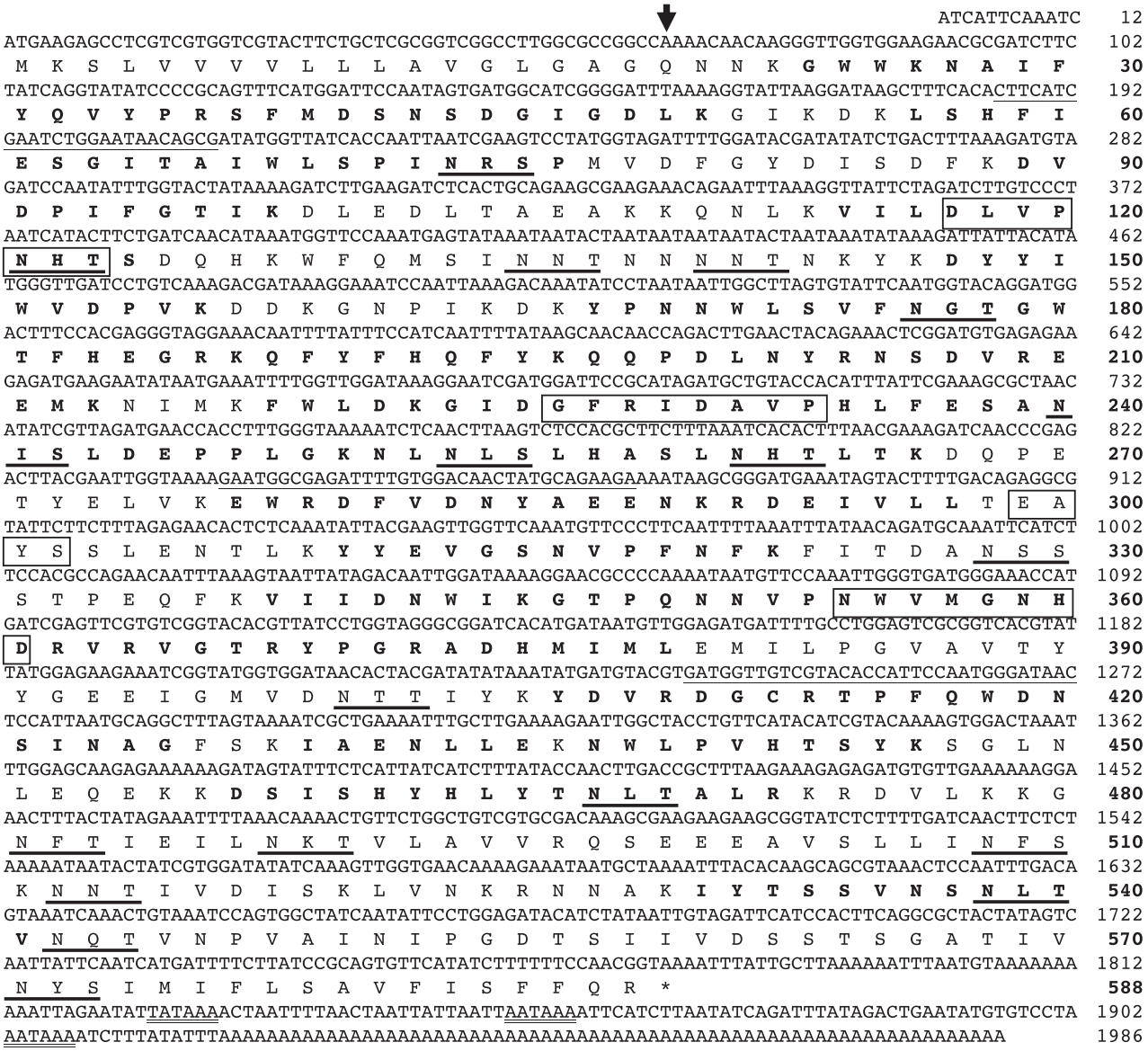

Fig. 1. cDNA and Deduced Amino Acid Sequences of HBGase I.

The predicted N-terminal amino acid of the mature protein after cleavage of the signal peptide and the stop codon are marked by an arrowhead and an asterisk, respectively. The amino acid sequences in bold underlines indicate the putative motifs (N-Xa.a.-S/T) for *N*-glycosylation, but N121 was not *N*-glycosylated because of being capable of analyzing by automated protein sequencer. Amino acids in boldface indicate sequences of internal peptides isolated from lysyl endopeptidase digest and analyzed with an automated protein sequencer. The underlined nucleotide sequences are correlated with the sense and antisense oligonucleotide primers designed on the internal peptides. Double underlines show three putative polyadenylational signals. The regions highly conserved in α-amylase family are boxed.

terminal sequence of HBGase I was not determined by the sequencer, perhaps because of blocking by a modification of the terminus. By plaque hybridization, positive clones of 27 for HBGase I and 18 for HBGase II were selected from 50,000 cDNA clones in the library representing mRNA expressed in the thorax and abdomen of adult honeybees, and 7 for HBGase III were obtained from 700 clones of the other cDNA library prepared from the head. The clones harboring the longest insert DNA, designated on211, tw522 and th611 for HBGases I, II, and III, respectively, were picked out and sequenced. On211 and tw522, lacking the 5′ terminal regions of the full-length cDNAs, had no complete open reading frames (ORFs). 5′ RACE gave 195 bp and 78 bp DNA fragments composed of the 180 bp and 58 bp known sequences already found in on211 and tw522, and 15 bp and 20 bp flanking sequences on the 5′-terminal side for HBGases I and II, respectively. The extended short sequences included the initiation codon "ATG" and 5′ untranslated region (UTR).

Figures 1, 2, and 3 show the cDNAs and deduced amino acid sequences of HBGases I, II, and III, respectively. The cDNA for HBGase I was composed of an ORF of 1,767 bp, 5′ UTR of 12 bp, and 3′ UTR of 207 bp including 66 bp poly (A). Three putative polyadenylation signals (TATAAA or AATAAA, double

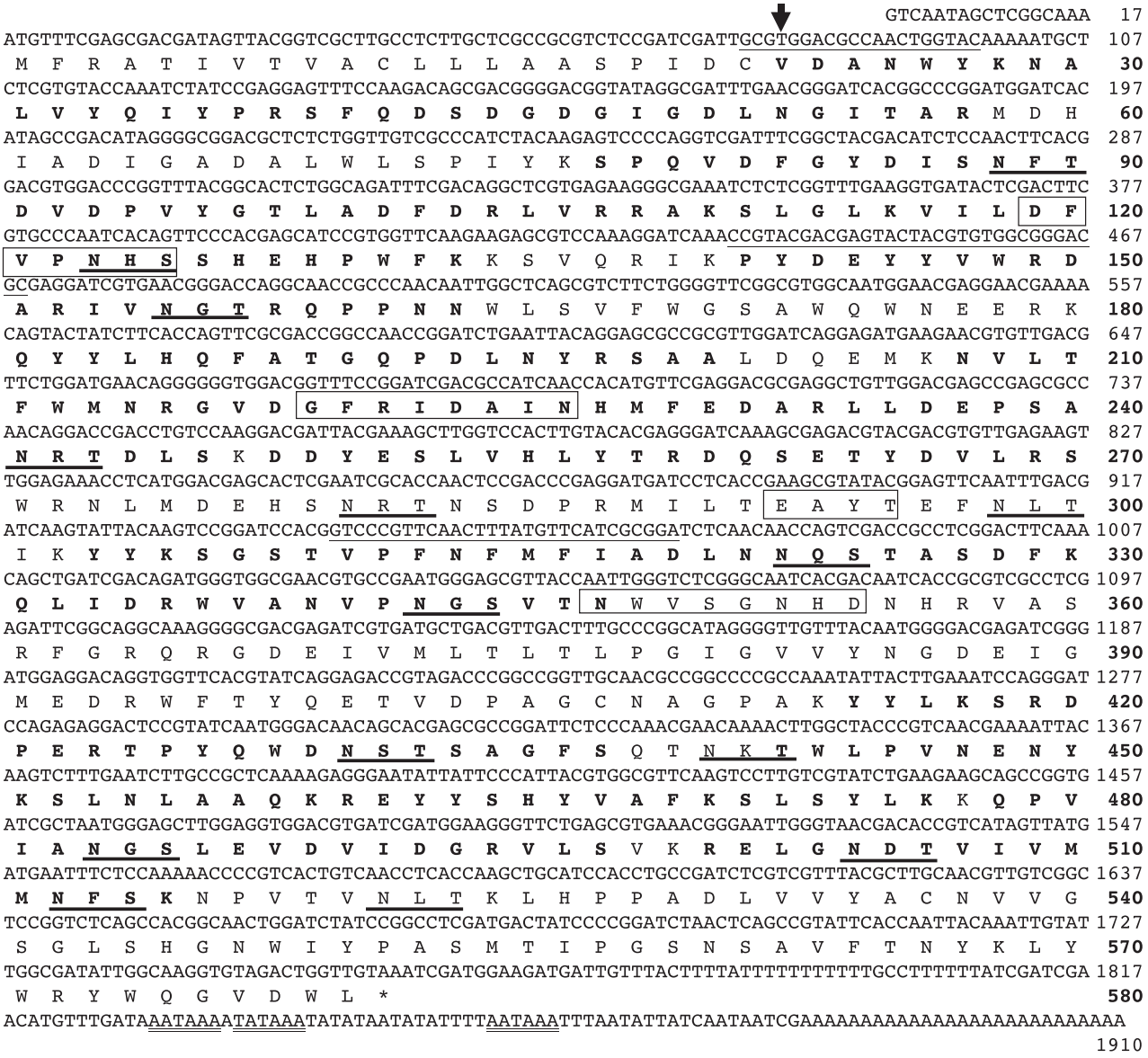

Fig. 2. cDNA and Deduced Amino Acid Sequences of HBGase II.

All indications, asterisk, single and double underlines, bold-faced characters, and boxes, are the same as in Fig. 1. The arrowhead indicates the N-terminal amino acid residue of the mature enzyme. Among Asn residues in the *N*-glycosylation motif, N123 and N512 were exclusively detected as Asn by protein sequencing, implying that N123 and N512 were not *N*-glycosylated.

underlines in Fig. 1) were located 18, 66, and 96 bp upstream the poly (A) tail. The ORF encoded a protein of 588 amino acid residues, in which the internal sequences analyzed in HBGase I were found (boldface in Fig. 1). The N-terminal residue of mature HBGase I was predicted to be Gln 19 by means of a peptide prediction program.[30] The deduced leader peptide of 18 residues had a typical signal sequence comprised of hydrophobic Val and Leu residues in the middle part and Gly and Ala residues near the cleavage point. The molecular mass of the predicted mature protein comprised of 570 amino acids was 65,636, whereas the molecular weight of HBGase I purified from honeybees, which contained 25% carbohydrate as glucose,[17] was estimated to be approximately 98,000 by SDS–PAGE.[17] The 18 putative *N*-glycosylation sites (bold underlines, N-Xa.a.-T/S in Fig. 1; Xa.a, amino acid residue except for Pro) were found in the deduced amino acid sequence.

As for HBGase II, cDNA was found to have an ORF of 1,743 bp, and 17 bp and 150 bp in the 5′ and 3′ UTR, respectively, including the continuous 27 adenosine residues at the 3′ end. The 3′ UTR of HBGase II cDNA, as in the case of HBGase I cDNA, contained three putative polyadenylation signals 26, 47, and 54 bp upstream the poly (A) tail. The ORF encoded a protein of 580 amino acid residues having 14 putative *N*-glycosylation sites. All analyzed internal sequences in the native HBGase II were included in the deduced amino acid sequence (boldface in Fig. 2). The N-terminal amino acid residue of the purified HBGase II was

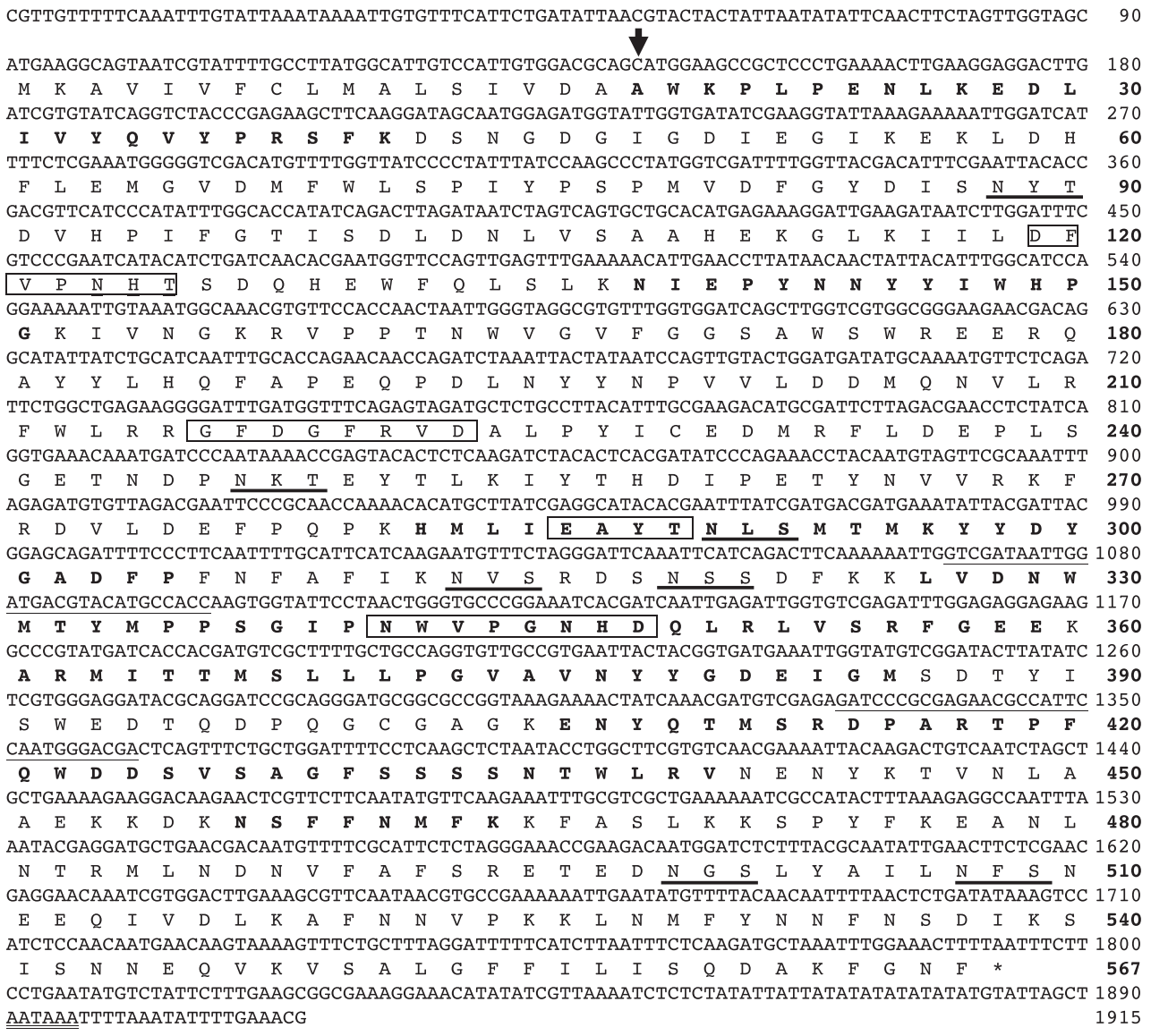

Fig. 3. cDNA and Deduced Amino Acid Sequences of HBGase III.
All indications are the same as in Figs. 1 and 2.

revealed to be Val 22. The signal peptide was a typical leader sequence, as in HBGase I, and the mature protein was composed of 559 amino acids with a calculated molecular mass of 64,377. The molecular weight of the native HBGase II was estimated to be approximately 76,000 by SDS–PAGE. This difference might have been due to sugar chain contents of 15% in the purified enzyme.[17)]

The cDNA encoding HBGase III was composed of a full ORF of 1,704 bp, and 90 bp and 121 bp in the 5′ and 3′ UTR, respectively. As shown in Fig. 3, a putative polyadenylation signal (AATAAA) was found in the 3′ UTR, and there was no poly (A) region, differently from the cDNAs of the other two HBGases. The ORF encoded a protein of 567 amino acid residues, including all analyzed internal sequences and 8 putative *N*-glycosylation sites. The N-terminal amino acid residue of native HBGase III was Ala 18. The signal peptide also was a typical leader sequence of 17 amino acid residues similar to those of HBGases I and II. The mature HBGase III was a protein made up of 550 amino acids with the molecular mass of 63,758, of which the native enzyme was estimated to be approximately 68,000 by SDS–PAGE. This difference in mass might also have been due to sugar chains contents of 7.4% in the enzyme.[20)]

The deduced amino acid sequence of HBGase III was consistent with the sequence on the analysis of the cDNA encoding α-glucosidase of European honeybee hypopharyngeal gland,[22)] but there were some differences in their nucleotide sequences. The cDNA of HBGase III was 44 bp longer in 5′ UTR than that of the hypopharyngeal gland,[22)] and was lacking a poly (A) tail at the 3′ end. The replacements of 4 nucleotides were found between the hypopharyngeal gland enzyme and HBGase III cDNAs, and two of them occurred in the

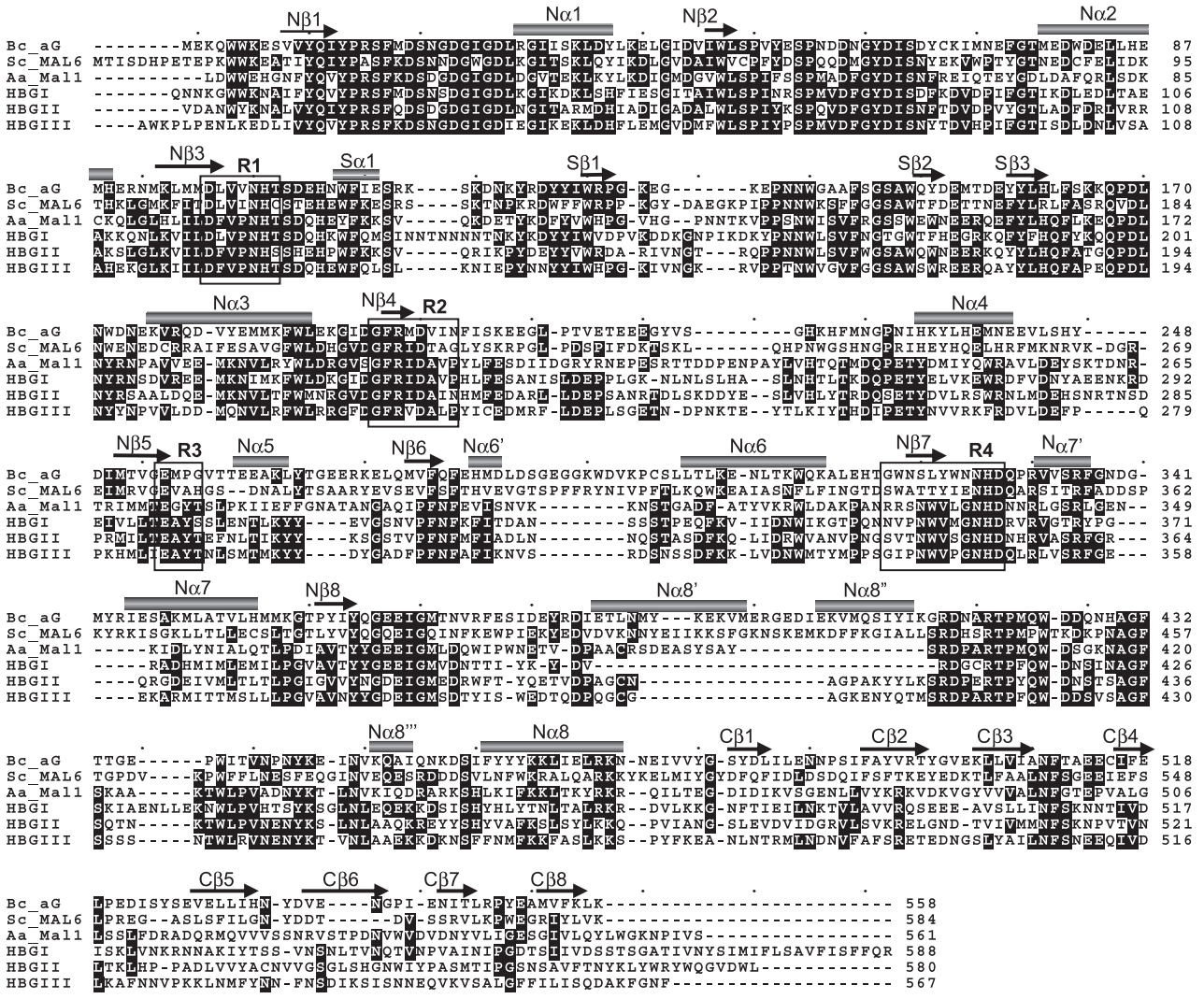

Fig. 4. Multiple Alignment of α-Glucosidases from European Honeybee *Apis mellifera* (HBGases I, II, and III), *Bacillus cereus* (Bc_aG), *Saccharomyces cerevisiae* (Sc_MAL6), and Mosquito *Aedes aegypti* (Aa_Mal1).

Six amino acid sequences are shown in one-letter code. They were aligned by the Clustal W program.[32] The N-terminal amino acids Q, V, and A for HBGases I, II, and III correspond to Q19, V22, and A18, excluding the signal peptides in Figs. 1, 2, and 3, respectively. Reversal letters represent the conserved amino acid residues in each aligned position. Arrowed lines and bars indicate β-strand and α-helix, respectively, in the three-dimensional structure of the *B. cereus* enzyme.[15] The four conserved regions are boxed with R1–R4.

ORF. However, such discrepancies brought about no change in the amino acid sequences of both enzymes.

The primary structures of HBGases I, II, and III shared the four regions highly conserved in all enzymes, designated α-amylase family enzymes,[7] as shown in Fig. 4. The multiple alignment of α-glucosidases of honeybee (*Apis mellifera*), Japanese honeybee (*A. cerana*),[31] mosquito (*Aedes aegypti*),[6] yeast (*Saccharomyces cerevisiae*),[2] and bacterium (*B. cereus*)[15] prepared using the Clustal W multiple alignment program[32] and the PHD prediction secondary structures program[33] indicated that the sequences for HBGases resembled each other, with 37 to 43% identities of the amino acid residues. Recently, an α-glucosidase (JBGase I) was purified from Japanese honeybees,[31] and the enzymatic properties were found to be highly similar to those of HBGase I. Moreover, in the amino acid sequence alignment, the enzyme showed 78% identity towards HBGase I, compared with only 38 to 42% identities towards HBGases II and III.

Selection of P. pastoris *transformants*

The expression plasmids derived from pPIC3.5, harboring the cDNA encoding full-length HBGases containing the signal sequence at the N-terminal end under the control of the alcohol oxidase (*AOX1*) promoter, were constructed, and *P. pastoris* GS115 transformants were obtained. The transformants were cultured in BMMY medium containing methanol to produce recombinant enzymes. Upon fermentation for 50 h, the recombinant HBGases II and III were secreted into the supernatant, the maltase activities were 0.08 and

Table 1. Properties of Native and Recombinant HBGases II and III

	HBGase II				HBGase III			
	Native		Recombinant		Native		Recombinant	
Optimum pH	5.1		5.3		5.5		5.7	
pH stability (4 °C, 24 h, >90%)	4.9–9.4		4.9–9.4		5.0–10.0		5.0–10.3	
Thermal stability (15 min, >90%)	<40 °C		<37 °C		<40 °C		<40 °C	
Molecular weight	76,000		110,000		68,000		70,000	
Sugar content (as mannose)	15%		21%		7.4%		10%	
	K_m (mM)	k_{cat} (s^{-1})	K_m (mM)	k_{cat} (s^{-1})	K_m (mM)	k_{cat} (s^{-1})	K_m (mM)	k_{cat} (s^{-1})
Maltose	5.4	27	2.7	33	11	37	13	39.0
Maltotriose	4.0	44	3.0	84	10	135	10	120
Sucrose	6.7[a]	110	35	80	30	195	46	200
Phenyl α-glucoside	1.6	73	1.2	83	23	148	29	150

[a]Substrate concentration giving one half of maximum velocity.[19)]

0.12 U/ml, respectively, but no production of recombinant HBGase I was found. *P. pastris* transformed with pPIC9 derivatives encoding the mature forms of HBGases II and III equipped with α-factor leader peptide at the N-terminal secreted maltase activities of 0.11 and 0.09 U/ml for the former and the latter into the same culture supernatant. The culture media showing the highest activities were selected for the production of recombinant HBGases II and III. As for HBGase I (pPIC3.5on11-9), maltase activity was at the control level even with the use of α-factor signal sequence or its own signal sequence. As for HBGase II transformant, the activity of pPIC3.5tw21 was 0.08 U/ml. And as for HBGase III transformant, the activity of pPIC3.5th13 was 0.12 U/ml, which increased up to 0.4 U/ml under long induction for 300 h.

Purification and characterization of recombinant HBGases II and III

Recombinant enzymes, rHBGases II (6.7 mg) and III (42 mg), were purified from 2 and 6 liters of culture supernatant, respectively, after induction for 200 h, by salting-out chromatography with ammonium sulfate, and by CM Sepharose CL-6B, Bio-Gel P-100, DEAE Sepharose CL-6B, and BUTYL-TOYOPEARL 650 M column chromatographies. rHBGase II was separated into three glycoprotein components 1 (3.3 mg), 2 (2.5 mg), and 3 (0.92 mg) by BUTYL-TOYOPEARL column chromatography, and component 1, with the highest sugar content (21% as mannose), was used as an rHBGase II preparation in the subsequent experiments. The three components 1, 2 (sugar contents, 19%), and 3 (sugar contents, 16%) gave nearly the same K_m values in the hydrolysis of maltose, 2.7, 3.0, and 2.7 mM, respectively, though they were different in sugar contents. As shown in Fig. 5, rHBGase II showed a smear band (molecular weight, 110,000) on SDS–PAGE. The preparation of rHBGase III was separated into plural bands between molecular weights 70,000 and 71,000 on SDS–PAGE, but the proteins corresponding to those bands were cross-reacted with antiserum[21)] specific for HBGase III on western blotting (data not shown). Therefore, the preparation giving plural bands around a molecular weight of 70,000 was used as purified rHBGase III in the subsequent experiments.

The purified rHBGases II and III were compared with the native enzymes for certain properties (Table 1). rHBGases II and III exhibited optimum pHs at 5.3 and 5.7, and were stabled up to 37 and 40 °C for 15 min, and in the ranges of pH 4.9 to 9.4 and 5.0 to 10.3 at 4 °C for 24 h, respectively. These properties were almost the same as those of the native enzymes. Their N-terminal amino acid sequences were identical to those of the native enzymes. Although the molecular weights and sugar contents of rHBGases II and III were 110,000 and 21.2%, and 70,000, 10.0%, respectively, the molecular weights of the recombinant enzymes were consistent with those of the native ones, after the deglycosylation treatment with endoglycosidase F (Fig. 5). Deglycosylated native HBGase II and rHBGases II showed approximately the same molecular weight (65,000), estimated on SDS–PAGE, and the deglycosylated native HBGase III and rHBGase III with endoglycosidases F and H gave two bands of the molecular weights 67,000 and 63,000. As shown in Fig. 5B, however, the deglycosylation with endoglycosidase H gave a single molecular weight, 63,000, for both native HBGase III and rHBGase III, fully denatured by boiling in 0.15% SDS. This implies that HBGase III was unsusceptible to deglycosylation with endoglycosidase F, compared with endoglycosidase H. The molecular weight was almost consistent with the one calculated from the deduced amino acid sequence. The K_m and k_{cat} values for maltose, maltotriose, sucrose, and phenyl α-glucoside were 2.7, 3.0, 35 and 1.2 mM, 33, 84, 80, and 83 s^{-1} (rHBGase II), 13, 10, 46, and 29 mM, and 39, 120, 200, and 150 s^{-1} (rHBG III), respectively (Table 1). The recombinant enzymes expressed by *P. pastoris* appeared to be no significant differences in the enzymatic properties of their native enzymes, except for sugar contents.

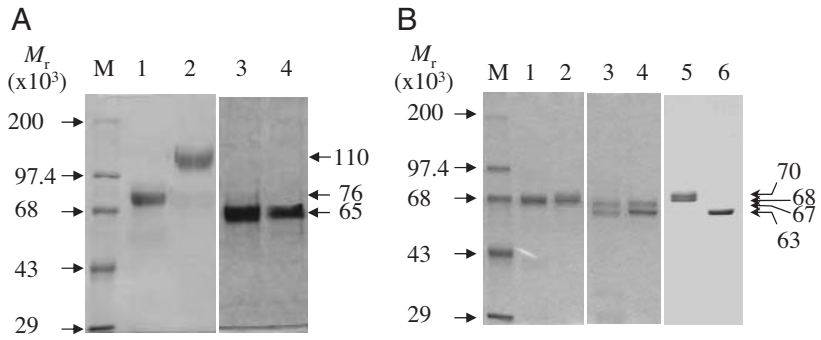

Fig. 5. SDS–PAGE Analysis of Deglycosylated Native and Recombinant HBGases II and III.

Panel-A shows native HBGase II before (lane 1) and after (lane 3) deglycosylation with endoglycosidase F, and recombinant HBGase II before (lane 2) and after (lane 4) deglycosylation with the same enzyme. Lane M indicates the size markers: myosin H-chain (M_r, 200 × 10^3), phosphorylase *b* (M_r, 97.4 × 10^3), bovine serum albumin (M_r, 68 × 10^3), ovalbumin (M_r, 43 × 10^3), and carbonic anhydrase (M_r, 29 × 10^3). Panel-B shows native HBGase III before (lane 1) and after (lane 3) deglycosylation with endoglycosidase F, and recombinant HBGase III before (lanes 2 and 5) and after (lane 4) deglycosylation with the same enzyme. Recombinant HBGases III migrated as a single band (M_r, 63,000) after deglycosylation with endoglycosidase H in the denatured state (lane 6).

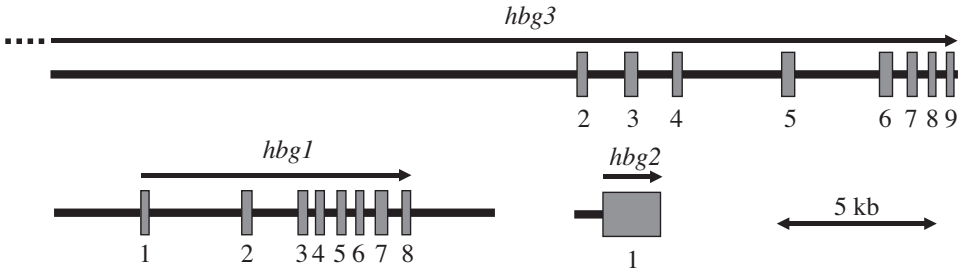

Fig. 6. Schematic Representation of Gene Structures for α-Glucosidases I, II, and III from European Honeybees.

The geomic DNAs isolated and sequenced in this study are indicated by plane lines. Gray boxes indicate the exons, and the genes for α-glucosidases are expressed by arrows. The 5′- and 3′-termini are 1GAA–CTT13,325 in *hbg1*, 1CTC–TCC2,759 in *hbg2*, and 1GAC–ATT27,643 in *hbg3*.

Genomic DNA structures of HBGases

Genomic DNAs of 13,325 bp (1GAA–CTT13,325), 2,759 bp (1CTC–TCC2,759), and 27,643 bp (1GAC–ATT27,643) were cloned and sequenced for HBGases I, II, and III (accession numbers: AB253415, AB253416, and AB253417 in DDBJ/EMBL/GenBank), respectively. Figure 6 shows the schematic diagram of the gene structures encoding the three HBGases.

The gene of HBGase I, which was located in the region of 2,767 to 10,905 nt, was composed of 8 exons and 7 introns: exon 1 was located in the range of 2,767 to 2,929 nt; exon 2, in 5,874 to 6,059 nt; exon 3, in 7,407 to 7,708 nt; exon 4, in 8,006 to 8,248 nt; exon 5, in 8,786 to 8,971 nt; exon 6, in 9,230 to 9,434 nt; exon 7, in 9,878 to 10,229 nt; and exon 8, in 10,623 to 10,905 nt, in the DNA sequence.

The gene for HBGase II was included in the sequenced 2,759 bp genomic DNA fragment in the range of 877 to 2,759 nt. It was considered to be prominent characteristic of the HBGase II gene that the gene was not divided by intron. The enzyme gene was located continuously in the region between 1G and 1,883G, as shown in Fig. 2.

The gene encoding HBGase III was divided into 9 exons. The prepared genomic DNA of 27,643 bp contained exons 2 to 9, but exon 1 was not found in the sequence, as shown in Fig. 6. Therefore, the position of exon 1 in the gene was sought from the honeybee genome project.[23] Exon 1, corresponding to the nucleotides of 1C to 88A in 5′ UTR of cDNA indicated in Fig. 3, was confirmed to be located in front of the huge intron 1 (34,514 bp) upstream of exon 2. In addition, HBGase III became apparent to be located in the upstream region (5,247 bp) of the HBGase I gene in the same genetic linkage of chromosomal LG6.

In the prepared genomic DNA of HBGase III, exon 2 was located in 16,534 to 16,692 nt; exon 3, in 17,726 to 18,060 nt; exon 4, in 19,108 to 19,233 nt; exon 5, in 22,397 to 22,807 nt; exon 6, in 25,339 to 25,594 nt; exon 7, in 26,208 to 26,403 nt; exon 8, 26,914 to 27,048 nt; and exon 9, in 27,363 to 27,571 nt. All exon/intron boundries in *hbg1* and *hbg3* conformed to the GT-AG rules for splicing.

Discussion

In α-glucosidase family I, grouped into GH 13,[8] the three-dimensional structure has been solved so far

only for *B. cereus* α-glucosidase.[15] As shown in Fig. 4, HBGases I, II, and III showed high homologies (43 to 45% similarity) in the primary structure to *B. cereus* α-glucosidase, and had four highly conserved regions (R1, R2, R3, and R4, in Fig. 4) including the catalytic sites, as well as *B. cereus* enzyme. This suggests that the three-dimensional structures of HBGases are similar to that of *B. cereus* α-glucosidase.[15] The enzyme from *B. cereus* has been elucidated to be comprised of three domains, the N-terminal, sub-, and C-terminal domains, and the catalytic amino acid residues are situated in the N-terminal domain, consisting of $(\beta/\alpha)_8$ barrel. On the basis of the structure of *B. cereus* α-glucosidase, the structures of HBGases were inferred from multiple alignment of the primary structures and the predicted secondary ones (Fig. 4). The numbers of the first amino acid residues, Gln, Val, and Ala, for HBGases indicated in Fig. 4, start off with Gln19, Val22, and Ala18, excluding the signal peptides in Figs. 1, 2, and 3, respectively. The regions of Gln19 to Thr123 and Ser206 to Gly484 in HBGase I, Val22 to Ser125 and Ala199 to Gly484 in HBGase II, and Ala18 to Thr125 and Pro199 to Ala478 in HBGase III corresponded to N-terminal domain consisting of $(\beta/\alpha)_8$ barrel. The amino acid sequences from Nβ1 to Nβ4 regions in *B. cereus* α-glucosidase constituting N-terminal domain and sub-domain were highly conserved not only in HBGases but also in yeast and mosquito enzymes. Such primary structure is assumed to be important in the formation of the active site. The catalytic residues were thought to be Asp230, Glu299, and Asp361 in HBGase I, Asp223, Glu292, and Asp354 in HBGase II, and Asp223, Glu286, and Asp348 in HBGase III.

Production of recombinant HBGases was tried with yeast *P. pastoris*. As mentioned above, however, attempts to produce recombinant HBGase I were unsuccessful, and the reason was not apparent. There were no predominant differences in the properties between native and recombinant enzymes of HBGases II and III. Some recombinant enzymes were prepared bacteria and yeast α-glucosidases,[3,4,34,35] but no recombinant enzyme has been reported on insect α-glucosidase. HBGase I and II, both of which are monomeric proteins, are unique enzymes having allosteric properties, though HBGase III shows no such property. The construction of an expression system of recombinant enzymes is useful for investigations aimed at uncovering the reactivities and protein structures of HBGases.

The overall genome sequences of European honeybee, *Apis mellifera*, have recently become apparent. According to this information, the honeybee has 17 genetic linkages, chromosomal LG1–LG16, and mitochondorial MT. In the genetic linkage groups, the HBGase II gene (hereafter referred to as *hbg2*) was localized on LG8 (accession NC_007077). The genes of HBGases III (LOC406131, *hbg3*) and I (*hbg1*) were tandemly aligned on LG6 (accession NC_007075) in that order, with 5,247 bp insertion between them.

The exons of *hbg1* determined in this study were identical to the cDNA and the equivalent part presented by the genome project (*hbg1* in Amel4) in nucleotide sequences, while the introns contained many differences, 991 nt in total (Table 2). The large distinctions in intron 2 were mainly caused by 891 bp sequence not yet sequenced in Amel4.

The overall identity in the *hbg1* regions between this study and the genome project was 87%, including the 891 bp undetermined sequence.

As mentioned above, *hbg3* was continuously located in the upstream region of *hbg1* on LG6. Based on the sequences of the prepared genomic DNA for HBGase III and the latest data of the genome project, the gene of the enzyme was ascertained to be comprised of 9 exons and 8 introns. In *hbg3*, the variance was observed between cDNA, our genome sequence, and the genome project data even in the exons (Table 2). The nucleotide sequences of exons 2, 3, 4, 6, and 7 of *hbg3* determined in this study completely coincided with the equivalent part of cDNA and the data of the genome project, but in exons 1, 5, and 8, nucleotide substitutions were found (Table 2). The single substitution in exon 1 was found in the 5′UTR, and the substitutions of other 8 nucleotides were found in the ORF. Only one nucleotide replacement at G1069 (cDNA numbering) gives an amino acid substitution, from Val327 to Ile (Table 2). Introns contained greater differences in nucleotides. In particular, long extra sequences of 45 and 22 bp in introns 5 and 6, respectively, were found only in Amel4 sequence. The overall identity between our genome sequence and the Amel4 data was 99%.

The *hbg2* gene was confirmed to be located on LG8 without an intron. The cDNA and the genome data in this study coincided with each other, but 15 nucleotides different from the Amel4 sequence (Table 2). Only one substitution at A1255 (in cDNA) led to a substitution of Lys413 for Arg, and the others were all silent mutations. The identity in the *hbg2* region was 99%. The sequences of 876 bp upstream *hbg2* were identical with each other, except for 11 bp substitutions (Table 2), and showed 99% identity.

The difference in genomic sequences between this work and the genome project is thought to be due to polymorphism observed in the same species. Genetic polymorphism is also suggested in the nucleotide differences in HBGase cDNAs between our data and reference 22. Even though there were many nucleotide substitutions, most of them were in introns, and only two amino acids, one in HBGase II and one in HBGase III, were replaced by similar residues, implying that the genes produce active enzymes.

Besides the three HBGases genes mentioned above, a BLAST search suggested the possibility that another family I α-glucosidase encoded by the gene (LOC-552357) exists in LG5. The homologue exhibited 34 to 37% identity to HBGases I, II, and III in amino acid sequence. However, it is not clear as to whether such an

Table 2. Nucleotide Substitution Found in cDNAs and *hbg1*, *2*, and *3* of This Study and Genome Project (Amel4[23])

HBGase I				HBGase II				HBGase III			
Exon	cDNA[a]	*hbg1*[a]	*hbg1*[b]	Exon	cDNA[a,c]	*hbg2*[a]	*hbg2*[b,d]	Exon	cDNA[a,c]	*hbg3*[a,d]	*hbg3*[b]
1 to 8	No difference			1	C92[3] (N)	C	T	1	A74	nd[e]	G
					A557[3] (K)	A	G	5	T711[3] (N)	C	T
					G875[3] (P)	G	C		C837[3] (T)	T	C
					A893[3] (E)	A	G		C879[3] (T)	T	C
					G896[3] (A)	G	T		A978[3] (K)	G	A
					T899[3] (Y)	T	C		T1032[3] (V)	C	T
					G902[3] (T)	G	A		G1069[1] (V327)	A (I)	G
					T1211[3] (Y)	T	C	8	C1584[3] (D)	T	T
					C1250[3] (P)	C	A		C1638[3] (D)	T	T
					A1255[2] (K413)	A	G (R)				
					T1463[3] (A)	T	G				
					T1466[3] (A)	T	C				
					C1535[3] (T)	C	G				
					C1538[3] (V)	C	G				
					C1565[3] (N)	C	T				

Intron	Number of distinct nucleotide
1	75
2	905
3	2
4	3
5	3
6	2
7	1

Upstream region[f]	Number of distinct nucleotide
	11

Intron	Number of distinct nucleotide
1	72
2	5
4	12
5	103
6	25

[a]This study. [b]Genome project, Amel4.[23] [c]Superscript 1, 2, and 3 against nucleotides indicate the first, second, and third nucleotides of codons, respectively. The corresponding amino acids are shown in parentheses, in which the amino acids with numbers are replaced by the nucleotide substitutions. [d]Letters in parentheses indicate the amino acids inconsistent with those deduced from the cDNA. [e]Not determined. [f]Reagion upstream the initial Met codon (ATG).

enzyme is actually expressed in honeybee organs, due to a lack of confirmation of its presence in the honeybees.

Ten homologous genes of family I α-glucosidase gene have been analyzed in *Drosophila melanogaster*,[5,36] and two similar genes, in *Drosophila virilis*.[37] The amino acid sequences deduced from the genes showed 34 to 44% identity to those of HBGases I, II, and III, although there was no evidence of the presence of the enzymes in vinegar flies. There are also introns in the α-glucosidase genes, but their nucleotides are appreciably short, below 100 bp, compared with the introns on honeybee α-glucosidase genes. In vinegar flies, *Drosophilae*, several α-glucosidase homologous genes containing relatively short introns are continuously localized in a cluster on the genomes, but no such clustering is found in the case of honeybee genomes.

References

1) Chiba, S., Molecular mechanism in α-glucosidase and glucoamylase. *Biosci. Biotechnol. Biochem.*, **61**, 1233–1239 (1997).
2) Hong, S. H., and Marmur, J., Primary structure of the maltase gene of the *MAL6* locus of *Saccharomyces carlsbergensis*. *Gene*, **41**, 75–84 (1986).
3) Geber, A., Williamson, P. R., Rex, J. H., Sweeney, E. C., and Bennett, J. E., Cloning and characterization of a *Candida albicans* maltase gene involved in sucrose utilization. *J. Bacteriol.*, **174**, 6992–6996 (1992).
4) Takii, Y., Daimon, K., and Suzuki, Y., Cloning and expression of a thermostable exo-α-1,4-glucosidase gene from *Bacillus stearothermophilus* ATCC12016 in *Escherichia coli*. *Appl. Microbiol. Biotechnol.*, **38**, 243–247 (1992).
5) Snyder, M., and Davidson, N., Two gene families clustered in a small region of the *Drosophila* genome. *J. Mol. Biol.*, **166**, 101–118 (1983).
6) James, A. A., Blackmer, K., and Racioppi, J. V., A salivary gland-specific, maltase-like gene of the vector mosquito, *Aedes aegypti*. *Gene*, **75**, 73–83 (1989).
7) Svensson, B., Protein engineering in the α-amylase family: catalytic mechanism, substrate specificity, and stability. *Plant Mol. Biol.*, **25**, 141–157 (1994).
8) Henrissat, B., A classification of glycosyl hydrolases based on amino-acid sequence similarities. *Biochem. J.*, **280**, 309–316 (1991).
9) Kimura, A., Takata, M., Sakai, O., Matsui, H., Takai, N., Takayanagi, T., Nishimura, I., Uozumi, T., and Chiba, S., Complete amino acid sequence of crystalline α-glucosidase from *Aspergillus niger*. *Biosci. Biotechnol. Biochem.*, **56**, 1368–1370 (1992).
10) Nakamura, A., Nishimura, I., Yokoyama, A., Lee, D.-G., Hidaka, M., Masaki, H., Kimura, A., Chiba, S., and Uozumi, T., Cloning and sequencing of an α-glucosidase gene from *Aspergillus niger* and its expression in *A. nidulans*. *J. Biotechnol.*, **53**, 75–84 (1997).
11) Okuyama, M., Okuno, A., Shimizu, N., Mori, H., Kimura, A., and Chiba, S., Carboxyl group of residue Asp647 as possible proton donor in catalytic reaction of α-glucosidase from *Schizosaccharomyces pombe*. *Eur. J.*

Biochem., **268**, 2270–2280 (2001).
12) Okuyama, M., Tanimoto, Y., Ito, T., Anzai, A., Mori, H., Kimura, A., Matsui, H., and Chiba, S., Purification and characterization of the hyper-glycosylated extracellular α-glucosidase from *Schizosaccharomyces pombe*. *Enzyme Microb. Technol.*, **37**, 472–480 (2005).
13) Hoefsloot, L. H., Hoogeveen-Westerveld, M., Kroos, M. A., van Beeumen, J., and Reuser, A. J., Primary structure and processing of lysosomal α-glucosidase; homology with the intestinal sucrase-isomaltase complex. *EMBO J.*, **7**, 1697–1704 (1988).
14) Tibbot, B. K., and Skadsen, R. W., Molecular cloning and characterization of a gibberellin-inducible, putative α-glucosidase gene from barley. *Plant Mol. Biol.*, **30**, 229–241 (1996).
15) Watanabe, K., Hata, Y., Kizaki, H., Katsube, Y., and Suzuki, Y., The refined crystal structure of *Bacillus cereus* oligo-1,6-glucosidase at 2.0 Å resolution: structural characterization of proline-substitution sites for protein thermostabilization. *J. Mol. Biol.*, **269**, 142–153 (1997).
16) Ernst, H. A., Leggio, L. L., Willemoës, M., Leonard, G., Blum, P., and Larsen, S., Structure of the *Sulfolobus solfataricus* α-glucosidase: implications for domain conservation and substrate recognition in GH31. *J. Mol. Biol.*, **358**, 1106–1124 (2006).
17) Takewaki, S., Chiba, S., Kimura, A., Matsui, H., and Koike, Y., Purification and properties of α-glucosidases of the honey bee *Apis mellifera L. Agric. Biol. Chem.*, **44**, 731–740 (1980).
18) Kimura, A., Takewaki, S., Matsui, H., Kubota, M., and Chiba, S., Allosteric properties, substrate specificity, and subsite affinities of honeybee α-glucosidase I. *J. Biochem.*, **107**, 762–768 (1990).
19) Takewaki, S., Kimura, A., Kubota, M., and Chiba, S., Substrate specificity and subsite affinities of honeybee α-glucosidase II. *Biosci. Biotechnol. Biochem.*, **57**, 1508–1513 (1993).
20) Nishimoto, M., Kubota, M., Tsuji, M., Mori, H., Kimura, A., Matsui, H., and Chiba, S., Purification and substrate specificity of honeybee, *Apis mellifera L.*, α-glucosidase III. *Biosci. Biotechnol. Biochem.*, **65**, 1610–1616 (2001).
21) Kubota, M., Tsuji, M., Nishimoto, M., Wongchawalit, J., Okuyama, M., Mori, H., Matsui, H., Surarit, R., Svasti, J., Kimura, A., and Chiba, S., Localization of α-glucosidases I, II, and III in organs of European honeybees, *Apis mellifera L.*, and the origin of α-glucosidase in honey. *Biosci. Biotechnol. Biochem.*, **68**, 2346–2352 (2004).
22) Ohashi, K., Sawata, M., Takeuchi, H., Natori, S., and Kubo, T., Molecular cloning of cDNA and analysis of expression of the gene for α-glucosidase from the hypopharyngeal gland of the honeybee *Apis mellifera L. Biochem. Biophys. Res. Commun.*, **221**, 380–385 (1996).
23) The Honeybee Genome Sequencing Consortium, Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature*, **433**, 931–949 (2006).
24) Chomczynski, P., and Sacchi, N., Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Anal. Biochem.*, **162**, 156–159 (1987).
25) Sanger, F., Nicklen, S., and Coulson, A. R., DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467 (1977).
26) Laemmli, U. K., Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685 (1970).
27) Dubois, M., Gilles, K. A., Hamilton, K., Rebers, P. A., and Smith, F., Colorimetric method for determination of sugars and related substances. *Anal. Chem.*, **28**, 350–356 (1956).
28) Murray, M. G., and Thompson, W. F., Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.*, **8**, 4321–4326 (1980).
29) Ochman, H., Gerber, A. S., and Hartl, D. L., Genetic application of an inverse polymerase chain reaction. *Genetics*, **120**, 621–623 (1988).
30) Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G., Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.*, **10**, 1–6 (1997).
31) Wongchawalit, J., Yamamoto, T., Nakai, H., Kim, Y.-M., Sato, N., Nishimoto, M., Okuyama, M., Mori, H., Saji, O., Chanchao, C., Wongsiri, S., Surarit, R., Svasti, J., Chiba, S., and Kimura, A., Purification and characterization of α-glucosidase I from Japanese honeybee (*Apis cerana japonica*) and molecular cloning of its cDNA. *Biosci. Biotechnol. Biochem.*, **70**, 2889–2898 (2006).
32) Thompson, J. D., Higgins, D. G., and Gibson, T. J., CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**, 4673–4680 (1994).
33) Rost, B., and Sander, C., Prediction of protein secondary structure at better than 70% accuracy. *J. Mol. Biol.*, **232**, 584–589 (1993).
34) Nakao, M., Nakayama, T., Kakudo, A., Inohara, M., Harada, M., Omura, F., and Shibano, Y., Structure and expression of a gene coding for thermostable α-glucosidase with a broad substrate specificity from *Bacillus* sp. SAM1606. *Eur. J. Biochem.*, **220**, 293–300 (1994).
35) Schonert, S., Burder, T., and Dahl, M. K., Identification and enzymatic characterization of the maltose-inducible α-glucosidase MalL (sucrose–isomaltase–maltase) of *Bacillus subtilis*. *J. Bacteriol.*, **180**, 2574–2578 (1998).
36) Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., Scherer, S. E., Li, P. W., Hoskins, R. A., Galle, R. F., George, R. A., Lewis, S. E., Richards, S., Ashburner, M., Henderson, S. N., Sutton, G. G., Wortman, J. R., Yandell, M. D., Zhang, Q., Chen, L. X., Brandon, R. C., Rogers, Y. H., Blazej, R. G., Champe, M., Pfeiffer, B. D., Wan, K. H., Doyle, C., Baxter, E. G., Helt, G., Nelson, C. R., Gabor, G. L., Abril, J. F., Agbayani, A., An, H. J., Andrews-Pfannkoch, C., Baldwin, D., Ballew, R. M., Basu, A., Baxendale, J., Bayraktaroglu, L., Beasley, E. M., Beeson, K. Y., Benos, P. V., Berman, B. P., Bhandari, D., Bolshakov, S., Borkova, D., Botchan, M. R., Bouck, J., Brokstein, P., Brottier, P., Burtis, K. C., Busam, D. A., Butler, H., Cadieu, E., Center, A., Chandra, I., Cherry, J. M., Cawley, S., Dahlke, C., Davenport, L. B., Davies, P., de Pablos, B., Delcher, A., Deng, Z., Mays, A. D., Dew, I., Dietz, S. M., Dodson, K., Doup, L. E., Downes, M., Dugan-Rocha, S., Dunkov, B. C., Dunn,

P., Durbin, K. J., Evangelista, C. C., Ferraz, C., Ferriera, S., Fleischmann, W., Fosler, C., Gabrielian, A. E., Garg, N. S., Gelbart, W. M., Glasser, K., Glodek, A., Gong, F., Gorrell, J. H., Gu, Z., Guan, P., Harris, M., Harris, N. L., Harvey, D., Heiman, T. J., Hernandez, J. R., Houck, J., Hostin, D., Houston, K. A., Howland, T. J., Wei, M. H., Ibegwam, C., Jalali, M., Kalush, F., Karpen, G. H., Ke, Z., Kennison, J. A., Ketchum, K. A., Kimmel, B. E., Kodira, C. D., Kraft, C., Kravitz, S., Kulp, D., Lai, Z., Lasko, P., Lei, Y., Levitsky, A. A., Li, J., Li, Z., Liang, Y., Lin, X., Liu, X., Mattei, B., McIntosh, T. C., McLeod, M. P., McPherson, D., Merkulov, G., Milshina, N. V., Mobarry, C., Morris, J., Moshrefi, A., Mount, S. M., Moy, M., Murphy, B., Murphy, L., Muzny, D. M., Nelson, D. L., Nelson, D. R., Nelson, K. A., Nixon, K., Nusskern, D. R., Pacleb, J. M., Palazzolo, M., Pittman, G. S., Pan, S., Pollard, J., Puri, V., Reese, M. G., Reinert, K., Remington, K., Saunders, R. D., Scheeler, F., Shen, H., Shue, B. C., Siden-Kiamos, I., Simpson, M., Skupski, M. P., Smith, T., Spier, E., Spradling, A. C., Stapleton, M., Strong, R., Sun, E., Svirskas, R., Tector, C., Turner, R., Venter, E., Wang, A. H., Wang, X., Wang, Z. Y., Wassarman, D. A., Weinstock, G. M., Weissenbach, J., Williams, S. M., Woodage, T., Worley, K. C., Wu, D., Yang, S., Yao, Q. A., Ye, J., Yeh, R. F., Zaveri, J. S., Zhan, M., Zhang, G., Zhao, Q., Zheng, L., Zheng, X. H., Zhong, F. N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H. O., Gibbs, R. A., Myers, E. W., Rubin, G. M., and Venter, J. C., The genome sequence of *Drosophila melanogaster*. *Science*, **287**, 2185–2195 (2000).

37) Vieira, C. P., Vieira, J., and Hartl, D. L., The evolution of small gene clusters: evidence for an independent origin of the maltase gene cluster in *Drosophila virilis* and *Drosophila melanogaster*. *Mol. Biol. Evol.*, **14**, 985–993 (1997).