ПРИКЛАДНАЯ БИОХИМИЯ И МИКРОБИОЛОГИЯ, 2012, том 48, № 5, с. 522–530

UDC 577.152

CLONING AND BIOINFORMATIC ANALYSIS OF AN ACIDOPHILIC β-MANNANASE GENE, *Anman5A*, FROM *Aspergillus niger* LW-1

© 2012 S. G. Zhao, M. C. Wu, C. D. Tang, S. J. Gao, H. M. Zhang, J. F. Li

Jiangnan University, 214122 China

e-mail: biowmc@126.com; lijf@163.com

Received September 30, 2011

Using 3' and 5' rapid amplification of cDNA ends (RACE) techniques, the full-length cDNA sequence of the *Anman5A*, a gene that encodes an acidophilic β-mannanase of *Aspergillus niger* LW-1 (abbreviated to AnMan5A), was identified from the total RNA. The cDNA sequence was 1417 bp in length, harboring 5'- and 3'-untranslated regions, as well as an open reading frame (ORF) which encodes a 21-aa signal peptide, a 17-aa propeptide and a 345-aa mature peptide. Based on the topology of the phylogenetic tree of β-mannanases from glycoside hydrolase (GH) family 5, the AnMan5A belongs to the subfamily 7 of the GH family 5. Its 3-D structure was modeled by the bitemplate-based method using both MODELLER 9.9 and SALIGN programs, based on the known β-mannanase crystal structures of *Trichoderma reesei* (1QNO) and *Lycopersicon esculentum* (1RH9) from the GH family 5. In addition, the complete DNA sequence of the *Anman5A* was amplified from the genomic DNA using the pUCm-T vector-mediated PCR and conventional PCR methods. The DNA sequence was 1825 bp in length, containing a 5'-flanking regulatory region, 2 introns and 3 exons when compared with the full-length cDNA.

Mannan, the second most abundant hemicellulose after xylan, can be classified into 4 subfamilies: linear mannan, glucomannan, galactomannan and galactoglucomannan. Each of these polysaccharides presents a β-1,4-*D*-linked backbone, containing only mannose residues or a combination of mannose and glucose residues. For complete biodegradation of mannans, besides β-mannanase, β-mannosidase and β-glucosidase, additional enzymes such as acetyl mannan esterase (EC 3.1.1.6) and α-galactosidase (EC 3.2.1.22) are required to remove side-chain substituents [1]. β-Mannanase (β-1,4-*D*-mannan mannohydrolase, EC 3.2.1.78) can catalyze the random cleavage of internal β-1,4-*D*-mannosidic linkages of the mannan backbone [2]. To date, almost all known β-mannanases have been grouped into glycoside hydrolase (GH) families 5, 26 and 113 based on their primary sequence alignment and hydrophobic cluster analysis (http://www.cazy.org/). Recently, β-mannanases have attracted much attention owing to their potential applications in diverse industrial processes [1]. Like the other GH members, β-mannanases played important roles in simplifying the industrial processes improving the quality of products, and reducing the environmental pollution caused by using the chemicals [3].

So far, β-mannanases have been isolated and characterized from microorganisms, plants and metazoans, among which filamentous fungi were considered to possess great potential for their production [4]. Studies have been performed on enhancing the β-mannanase activity by mutating enzyme-producing strains and optimizing fermentation conditions, and on producing enzyme on an industrial scale. However, the commercialization and broad applications of β-mannanases are still limited by their low activities and expensive production costs [5]. Therefore, increasing the enzyme activities and improving its properties, by means of genetic engineering, are highly desirable. Up to date, some β-mannanase genes from filamentous fungi, such as *Aspergillus usamii* [4], *Aspergillus sulphureus* [6], *Aspergillus aculeatus* [7], *Biopora* sp. MEY-1 [3] and *Trichoderma reesei* [8], were cloned, characterized and rationally modified, and several recombinant β-mannanases were expressed in heterologous cells [3, 7].

In our previous research, the bimutation breeding and fermentation conditions optimizing of *Aspergillus niger* LW-1 for enhancing the β-mannanase (AnMan5A) activity were performed [5, 9]. The enzymatic properties of the AnMan5A were also characterized, which possessed some crucial properties including the superior V_M and good chemical tolerance [10]. The aim of this study was the cloning of both full-length cDNA and complete DNA sequences of the *Anman5A*, a gene that encodes an acidophilic β-mannanase of *A. niger* LW-1. Moreover, the bioinformatic analysis of the AnMan5A and *Anman5A* sequences was also described.

MATERIALS AND METHODS

Reagents and kits. Restriction enzymes, *Taq* DNA polymerase, T4 DNA ligase, X-gal, isopropyl β-D-1-thiogalactopyranoside (IPTG), DNA marker,

RNA PCR Kit and 5'-Full RACE Kit were purchased from TaKaRa (China). RNA extraction Kit, BCA-200 protein assay Kit and protein markers were obtained from Sangon (China). Phenyl-Sepharose CL-4B, DEAE-Sepharose FF and Sephadex G-75 were products of Pharmacia (Sweden). EZ-10 Spin column DNA gel extraction Kit was purchased from BBI (Canada). Locust bean gum and Coomassie brilliant blue R-250 were products of Sigma (USA). All other chemicals were of analytical grade.

Strains, media, and vector. *A. niger* LW-1, isolated from the soil in China as reported previously [9], was used to extract the total RNA and genomic DNA. The strain was cultured in a liquid medium containing (%): of tryptone – 1.0, yeast extract – 0.5, dextrose – 1.0 and locust bean gum – 0.5 (pH 6.0). *E. coli* JM109 (TaKaRa, China), used as a host strain for gene cloning and DNA sequencing, was cultured in a Luria-Bertani medium [11]. The pUCm-T vector (Sangon, China) was used both for the vector-mediated PCR method originally developed in our laboratory [4] and for the directly cloning of PCR products.

Total RNA and genomic DNA extraction. *A. niger* LW-1 was cultured at 32°C and 220 rpm for 36 h. Mycelia were collected, and washed with deionized water. Total RNA was extracted by using one-step method according to the instruction of RNA extraction Kit. Analytical results of the extracted total RNA showed that the ratio of OD_{260} to OD_{280} was 1.96, and the 18S and 28S rRNA bands, characterized for eukaryotes, on formaldehyde denatured agarose gel electrophoresis [4] were specific, indicating that the total RNA had high purity and was not decomposed. Extraction of the genomic DNA from *A. niger* LW-1 was performed according to the method reported previously [12].

Enzyme activity assay. β-Mannanase activity was assayed by measuring the amount of released reducing sugar from locust bean gum, using the 3,5-dinitrosalicylic acid (DNS) method as described previously [9]. One unit (U) of β-mannanase activity was defined as the amount of enzyme liberating 1 μmol of reducing sugar equivalent per min under the assay conditions, using *D*-mannose as standard.

Protein and carbohydrate content assays. SDS-PAGE was performed on a 12.5% gel by the method of Laemmli [13], and isolated proteins were visualized by straining with the Coomassie brilliant blue R-250. Protein concentration was determined by using the BCA-200 Protein Assay Kit, using bovine serum albumin as a standard. Carbohydrate content of the purified native AnMan5A was assayed by the phenol-sulfuric acid method [14], using *D*-mannose as a standard.

Purification of the AnMan5A. AnMan5A synthesized by *A. niger* LW-1 was secreted into the solid-state culture medium (koji) [9]. The koji (10 g) was extracted with 10 volumes (w/v) of 20 mM Na_2HPO_4–NaH_2PO_4 buffer (pH 7.0) at 30°C for 30 min with shaking (100 rpm). The crude extract was fractionated by ammonium sulfate at concentrations ranging from 40 to 75% saturation, and the fraction containing the AnMan5A activity was retained. The resulting precipitate was dissolved in 10 ml of the same buffer, and then dialyzed against it overnight. Subsequent purification manipulations were performed by using phenyl-Sepharose CL-4B hydrophobic chromatography, DEAE-Sepharose FF ion exchange chromatography and Sephadex G-75 gel filtration according to the method reported previously [10].

Primers for PCR amplification. After aligning 4 fungal β-mannanase sequences from the GH family 5: *A. usamii* (ADZ99027), *A. sulphureus* (ABC59553), *A. aculeatus* (AAA67426) and *T. reesei* (AAA34208), we found 2 most conserved peptide fragments, GYFAGTNS(/C)YW and STINTGADGLQ, located in the N-terminal region. Therefore, 2 degenerate primers ManF1 and ManF2 were designed corresponding to GYFAGTN and STINTGA, respectively. Primers dT-PR and PR (original names, oligo dT-M_{13} primer M_4 and M_{13} primer M_4), as well as OP and IP (original names, 5'-RACE outer primer and 5'-RACE inner primer) were provided by RNA PCR Kit and 5'-full RACE Kit, respectively. By pUCm-T vector-mediated PCR, the 5'-flanking regulatory region of the *Anman5A* of *A. niger* LW-1 was amplified using primers T-5prF (identical to the 21-bp fragment upstream the T/A clone site of the pUCm-T vector) and ManR1 or ManR2. Primers CManF and CManR were used to amplify the complete DNA sequence of the *Anman5A*. As listed in Table 1, all primers (except those provided by Kits) were synthesized by Sangon (China).

Cloning of the full-length cDNA. The 3'-end fragment of AnMan5A cDNA was amplified by using RNA PCR Kit and nested PCR technique. The primer dT-PR was used for reverse transcription of the first-strand cDNA from the *A. niger* LW-1 total RNA. Using the resulting first-strand cDNA as template, the first-round PCR amplification was carried out using primers ManF1 and PR as following conditions: a denaturation at 94°C for 2 min; 30 cycles of at 94°C for 30 s, then 53°C for 30 s and 72°C for 75 s; an extra elongation at 72°C for 10 min. Then the second-round PCR amplification was performed for confirmation (nested PCR) using primers ManF2 and PR under the same conditions as the first-round PCR. Next, the 5'-end fragment of AnMan5A cDNA, originating from the transcription starting point, was amplified using 5'-full RACE Kit. The first-strand cDNA was used as template for the first-round PCR using primers OP and ManR1, and then subjected to the second-round PCR using primers IP and ManR2 for

The oligonucleotide sequences of primers for PCR amplification

Cloned fragments	Primers	Oligonucleotide sequences (5'–3')*	Size, bp
3'-end fragment of AnMan5A cDNA	ManF1	GCTACTTYGCSGGVACSAAC	20
	ManF2	TCDACVATCAACACKGGNGC	20
	dT-PR	GTTTTCCCAGTCACGAC-Oligo dT	37
	PR	GTTTTCCCAGTCACGAC	17
5'-end fragment of AnMan5A cDNA	OP	CATGGCTACATGCTGACAGCCTA	23
	IP	CGCGGATCCACAGCCTACTGATGATCAGTCGATG	34
	ManR1	ATTGTCGATTTGCCGTCCTG	20
	ManR2	GCAGTTGGTACCAGACTGTG	20
5'-flanking region of the *Anman5A*	T-5prF	ACGACTCACTATAGGGCGACA	21
	ManR1	ATTGTCGATTTGCCGTCCTG	20
	ManR2	GCAGTTGGTACCAGACTGTG	20
Complete DNA of the *Anman5A*	CManF	GGATCCCACCACTCAGAGGAT	21
	CManR	TTAAGTTTTACTTGTTTGATTT	22

* Y = C/T, S = C/G, V = A/G/C, D = A/G/T, K = G/T and N = A/T/G/C.

confirmation. Finally, the full-length cDNA sequence of the *Anman5A* was identified by assembling above cloned 3'- and 5'-end fragments of AnMan5A cDNA.

Cloning of the complete DNA. The pUCm-T vector-mediated PCR amplification, a novel method initially developed in our laboratory to amplify 5'- or 3'-flanking region of a known DNA fragment, was performed by 4 steps as flowcharted in Fig. 1 (exemplified as the cloning of a 5'-flanking regulatory region of the *Anman5A*). Firstly, the *A. niger* LW-1 genomic DNA was digested using two optimum restriction enzymes, which were selected by a series of pre-experiments, to obtain a longer unknown 5'-flanking region. In this work, *Bam*HI and *Eco*RV were selected for the digestion of the genomic DNA. Secondly, a *Bam*HI cohesive end was filled in and an adenine nucleotide (A) was added to the 3'-end of each DNA strand using *Taq* DNA polymerase at 72°C for 10 min. The third step was the ligation of the second step's products into pUCm-T vector. And finally, the recombinant vectors were first amplified using T-5prF and ManR1 to obtain the 5'-flanking regulatory region of the *Anman5A*, and then subjected to the second-round PCR using T-5prF and ManR2 for confirmation. The complete DNA sequence of the *Anman5A* was directly amplified from the *A. niger* LW-1 genomic DNA by conventional PCR using CManF and CManR.

Nucleotide sequence accession number. Both the full-length cDNA and complete DNA sequences of the *Anman5A* of *A. niger* LW-1 have been deposited in the GenBank database under the accession numbers of JN123356 and JN811092, respectively.

Analysis of the AnMan5A structures from *A. niger* LW-1. The signal peptide of the preproAnMan5A was predicted by using SignalP 3.0 (http://www.cbs.dtu.dk/services/SignalP/). The putative N-glycosylation sites were located by using the NetNGlyc program 1.0 (http://www.cbs.dtu.dk/services/NetNGlyc/). The physicochemical properties of the mature AnMan5A were assayed by using the Protparam (http://au.expasy.org/tools/protparam.html). The homology alignment of the protein primary structures between the AnMan5A from *A. niger* LW-1 and other β-mannanases from the GH family 5 was accomplished in GenBank using the BLAST program. The phylogenetic tree was constructed by Swiss-Prot protein knowledgebase (http://www.ebi.ac.uk/swissprot/) using the PC/GENE CLUSTAL method [4]. The 3-D structure of the AnMan5A was modeled by the bitemplate-based method using both MODELLER 9.9 (http://salilab.org/modeller/) and SALIGN programs (http://salilab.org/DBAli/?page=tools &action=f_salign), based on known crystal structures of the *T. reesei* β-mannanase (1QNO) and *Lycopersicon esculentum* β-mannanase (1RH9) from the GH family 5.

Analysis of the *Anman5A* sequence from *A. niger* LW-1. The open reading frame (ORF) was determined by using the program of NCBI ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). The GeneMark (http://opal.biology.gatech.edu/GeneMark/eukhmm.cgi) was applied for the exact localization of the exon/intron boundaries. The prediction of 5'-promoter region and its characterization were carried out by using both the Berkeley Drosophila Genome Project (http://www.fruitfly.org/seq_tools/promoter.html) and PLACE (http://www.dna.affrc.go.jp/ PLACE/signalscan.html).

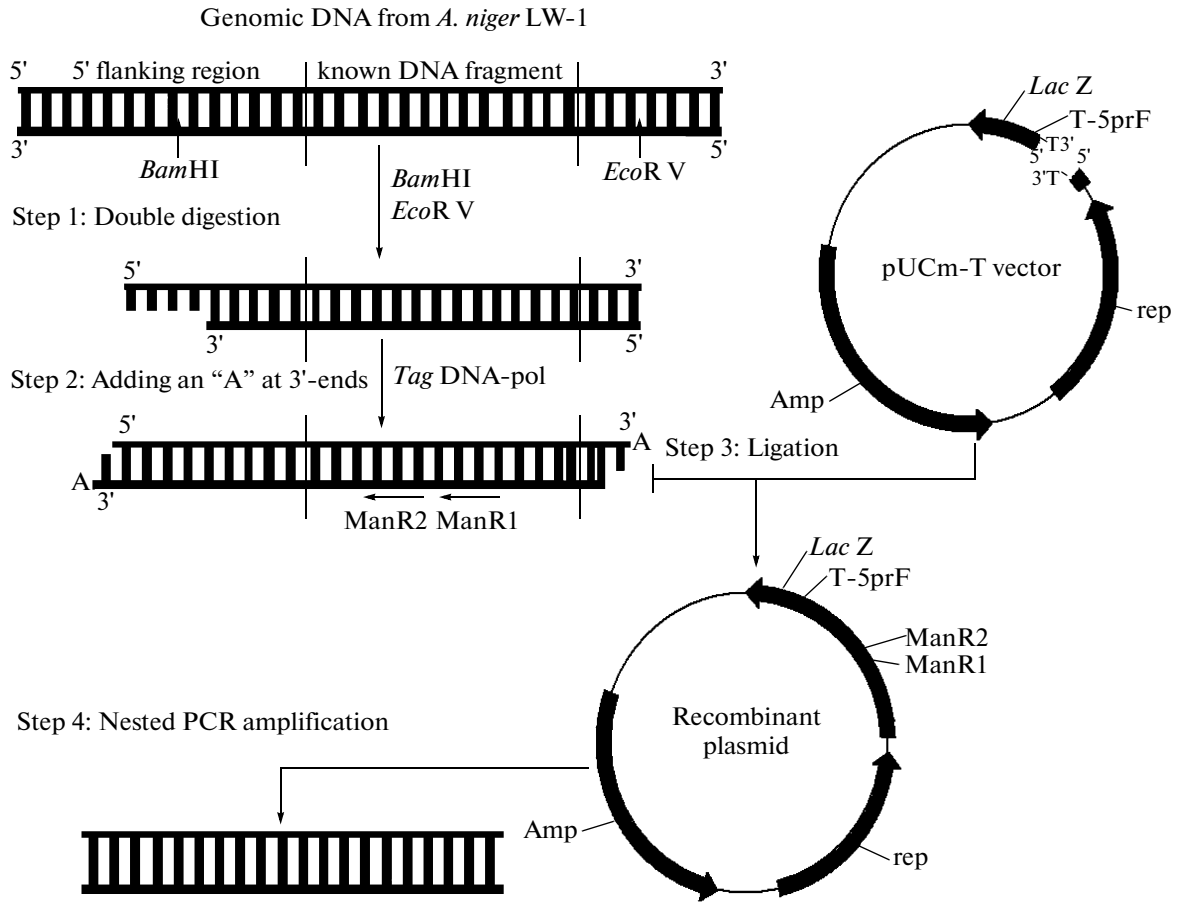


Fig. 1. Flowchart of the pUCm-T vector-mediated PCR and nested PCR amplification methods (exemplified as the cloning of a 5'-flanking region of the *Anman5A* from *A. niger* LW-1).

RESULTS AND DISCUSSION

Purification and N-terminus sequencing. The native AnMan5A from the cultivated koji of *A. niger* LW-1 was purified to homogeneity by using ammonium sulfate fractionation, phenyl-Sepharose CL-4B hydrophobic chromatography, DEAE-Sepharose FF ion exchange chromatography and Sephadex G-75 gel filtration [10]. The specific enzyme activity of the native AnMan5A towards locust bean gum was 238.1 U/mg. SDS-PAGE analysis revealed a single protein band of about 48.0 kDa, an apparent molecular weight of the native AnMan5A, which is larger than the theoretical molecular weight of 38 kDa. The carbohydrate content of the purified AnMan5A was detected to be 25.2% using the phenol-sulfuric acid method. The N-terminal amino acid sequence of the native AnMan5A was analyzed on 470A automatic sequencer obtained from Applied Biosystems (Foster, CA, USA). The sequence of N-terminal 15 amino acid residues was determined to be SFASTSGLQFTIDGE, which is identical to those of the *A. usamii* and *A. sulphureus* β-mannanase [4, 6].

Cloning of the full-length cDNA of the *Anman5A*. Using the reverse transcribed first-strand cDNA as template, one 1.2-kb band and the other faint band were amplified by the first-round PCR amplification using primers ManF1 and PR (Fig. 2, lane 1). Based on the principle of the nested PCR technique, each band was agarose gel-purified and subjected to the second-round PCR for confirmation using primers ManF2 and PR. Only the 1.2-kb band can be amplified again (Fig. 2, lane 2), and then was ligated into pUCm-T vector. DNA sequencing result verified that the cloned 3'-end cDNA fragment was exactly 1213 bp in length (except complementary sequence of dT-PR). An approximate 450-bp band of the 5'-end cDNA fragment was first-round amplified using primers OP and ManR1 as a major PCR product (Fig. 2, lane 3), and then subjected to the second-round PCR for confirmation using primers IP and ManR2 (Fig. 2, lane 4). DNA sequencing result showed that the first-round major PCR product was exactly 394 bp in length (except primers OP and IP), containing a 190 bp of sequence identical to that between ManF1 and ManR1, and a 204 bp of sequence in which a transcription

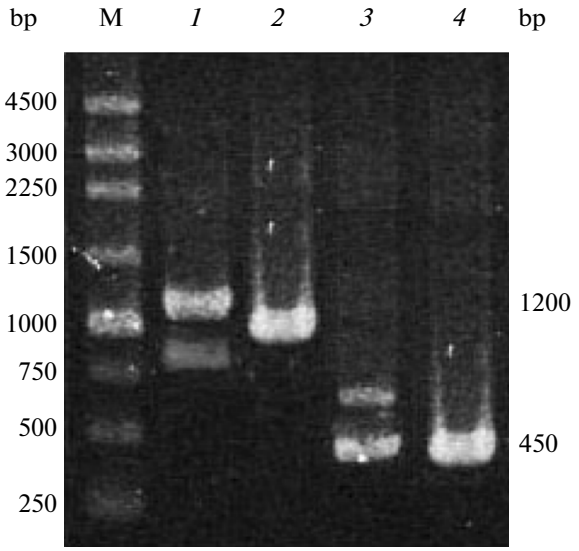


Fig. 2. Cloning of the full-length cDNA sequence of the *Anman5A* from *A. niger* LW-1. M – 250 bp DNA marker; *1* – one 1.2-kb band and the other faint band amplified by the first-round PCR using primers ManF1 and PR; *2* – an agarose gel-purified 1.2-bp band in lane 1 subjected to the second-round PCR for confirmation using primers ManF2 and PR; *3* – an approximate 450-bp band amplified by first-round PCR using primers OP and ManR1 as a major PCR product; *4* – a gel-purified 450-bp band in lane 3 subjected to the second-round PCR for confirmation using primers IP and ManR2.

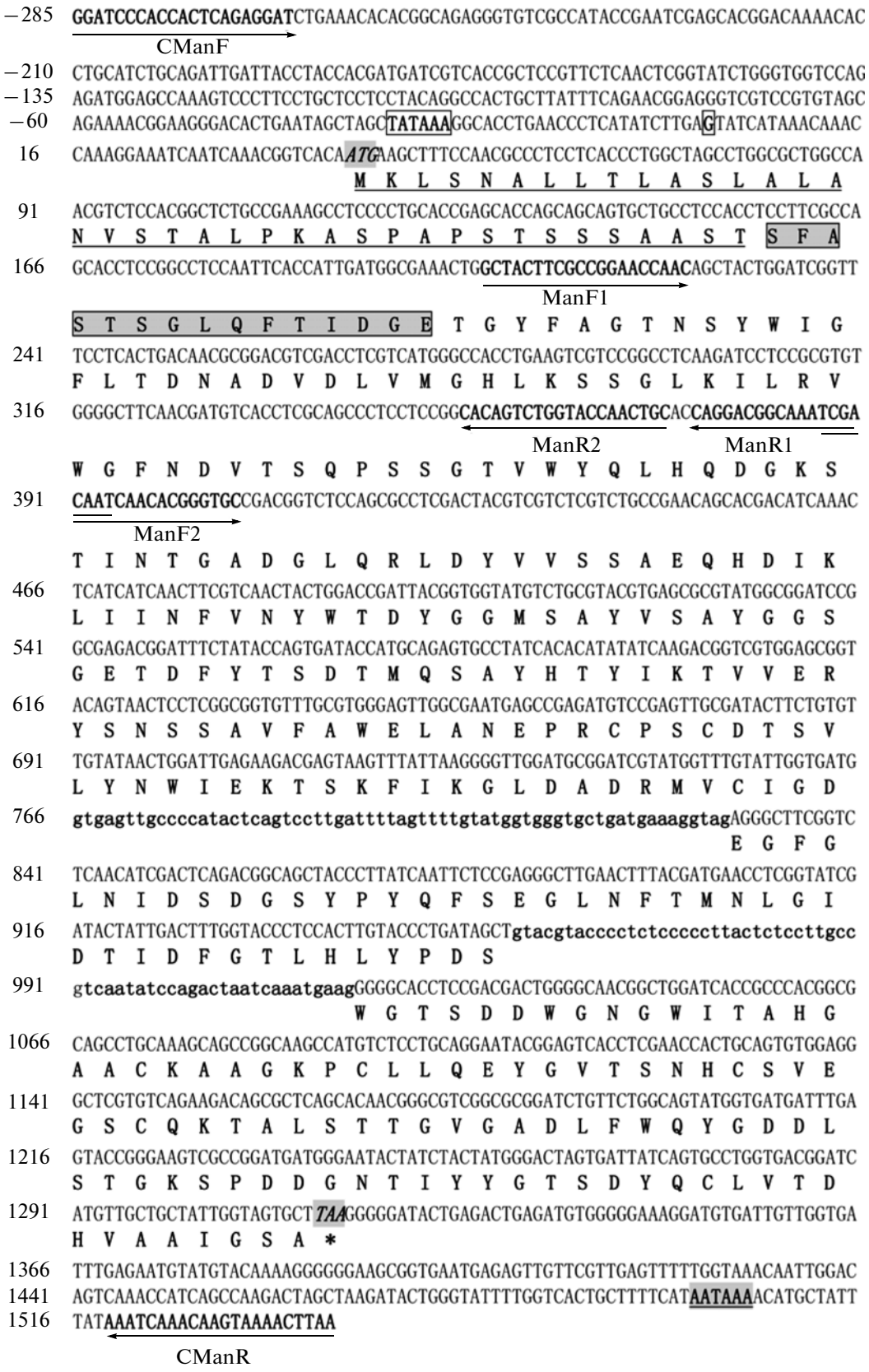
starting point (G) and a cDNA fragment encoding a 21-aa signal peptide, a 17-aa propeptide and a determined N-terminal 15 amino acid residues (SFASTSGLQFTIDGE) of the native AnMan5A were recognized (Fig. 3). The full-length cDNA sequence was obtained by assembling two cloned 3'- and 5'-end cDNA fragments. It was 1,417 bp in length (except polyA), including a 41 bp of 5'-untranslated region (5'-UTR), a 1,152 bp of ORF encoding a 383-aa preproAnMan5A, and a 224 bp of 3'-untranslated region (3'-UTR).

Analysis of the AnMan5A primary structure of *A. niger* LW-1. The SignalP 3.0 predicted an unambiguous signal peptide cleavage site between Ala21 and Leu22, indicating that the AnMan5A is a secretory protein. With the information of the determined sequence of N-terminal 15 amino acids of the native AnMan5A purified from *A. niger* LW-1, it was probable that a 383-aa preproAnMan5A was predicted to contain a 21-aa signal peptide from Met1 to Ala21, a 17-aa propeptide from Leu22 to Thr38 and a 345-aa mature peptide (namely AnMan5A) (Fig. 3). Propeptides also exist in other β-mannanases or microbial enzymes [15, 16]. The molecular weight of 38 kDa and pI of 4.15 were calculated from the deduced AnMan5A. There were 2 putative N-glycosylation sites (N156-S-S158 and N225-F-T227) in the AnMan5A sequence. Amino acid homology alignment showed that identities of the AnMan5A from *A. niger* LW-1 with other 4 fungal β-mannanases of *A. usamii* (ADZ99027), *A. sulphureus* (ABC59553), *A. aculeatus* (AAA67426) and *T. reesei* (AAA34208) from GH family 5 were 98.6, 93.0, 73.9 and 56.3%, respectively. Alignment analysis also revealed that the AnMan5A sequence contained 7 functional amino acid residues that were strictly conserved among the GH family 5 members [6]: 2 catalytic residues (acid/base, Glu168 and nucleophile, Glu276) located respectively in the β4 and β7 strands, and 5 active site residues (Arg52, Asn167, His241, Tyr243, and Trp306) (Fig. 4). These features verified that the AnMan5A from *A. niger* LW-1 was a member of GH family 5.

As the amino acid sequence identity among the members of GH family 5 is less than 20%, an additional classification into more than 10 subfamilies has been introduced. As a result, GHs within any subfamily show at least 25% sequence identity and may display similar substrate specificity. On the basis of their sequence similarity and substrate specificity, β-mannanases belong to subfamilies 7, 8 and 10 [17]. In 1998 β-mannanases from the eukaryotic and prokaryotic organisms were grouped under subfamily 7 and 8, respectively [18]. Phylogenetic tree of 24 known β-mannanases from various microorganisms, plants and metazoans of the GH family 5 and the *A. niger* LW-1 AnMan5A was constructed using the PC/GENE CLUSTAL method as shown in Fig. 5. Based on the topology of a phylogenetic tree, the AnMan5A is related to β-mannanases from the subfamily 7 and is closest to those of *A. usamii* (ADZ99027), *A. sulphureus* (ABC59553) and *A. aculeatus* (AAA67426).

Analysis of the 3-D structure. Based on crystal structures of the *T. reesei* β-mannanase (1QNO) and *L. esculentum* β-mannanase (1RH9) from the GH family 5, we modeled the 3-D structure of the AnMan5A from *A. niger* LW-1 using the bitemplate-based method (Fig. 4). The 3-D structure consisted principally of the $(\alpha/\beta)_8$ TIM-barrel protein fold. The structure has been likened to a 'salad bowl', with one

Fig. 3. Nucleotide sequence of the full-length cDNA or complete DNA of the *Anman5A* from *A. niger* LW-1 and its deduced amino acid sequence of the AnMan5A. Two introns with sizes of 63 and 60 bp are shown in lowercase letters. A signal peptide from Met1 to Ala21 and a propeptide from Leu22 to Thr38 are underlined. The determined N-terminal 15 amino acid sequence of the native AnMan5A is indicated in a grayed box. The bold letters of TATAAA and G in boxes indicate a putative TATA box and a transcription starting point, respectively. The grayed italic letters of ATG and TAA represent the starting codon and stop codon, respectively. A polyadenylation signal, AATAAA, is shown as grayed underlined letters. The bold arrows below the letters represent the primers for PCR amplification.


−285 GGATCCCACCACTCAGAGGATCTGAAACACACGGCAGAGGGTGTCGCCATACCGAATCGAGCACGGACAAAACAC
CManF
−210 CTGCATCTGCAGATTGATTACCTACCACGATGATCGTCACCGCTCCGTTCTCAACTCGGTATCTGGGTGGTCCAG
−135 AGATGGAGCCAAAGTCCCTTCCTGCTCCTCCTACAGGCCACTGCTTATTTCAGAACGGAGGGTCGTCCGTGTAGC
−60 AGAAAACGGAAGGGACACTGAATAGCTAGCTATAAAGGCACCTGAACCCTCATATCTTGAGTATCATAAACAAAC
16 CAAAGGAAATCAATCAAACGGTCACAATGAAGCTTTCCAACGCCCTCCTCACCCTGGCTAGCCTGGCGCTGGCCA
M K L S N A L L T L A S L A L A
91 ACGTCTCCACGGCTCTGCCGAAAGCCTCCCCTGCACCGAGCACCAGCAGCAGTGCTGCCTCCACCTCCTTCGCCA
N V S T A L P K A S P A P S T S S S A A S T S F A
166 GCACCTCCGGCCTCCAATTCACCATTGATGGCGAAACTGGCTACTTCGCCGGAACCAACAGCTACTGGATCGGTT
ManF1
S T S G L Q F T I D G E T G Y F A G T N S Y W I G
241 TCCTCACTGACAACGCGGACGTCGACCTCGTCATGGGCCACCTGAAGTCGTCCGGCCTCAAGATCCTCCGCGTGT
F L T D N A D V D L V M G H L K S S G L K I L R V
316 GGGGCTTCAACGATGTCACCTCGCAGCCCTCCTCCGGCACAGTCTGGTACCAACTGCACCAGGACGGCAAATCGA
ManR2
ManR1
W G F N D V T S Q P S S G T V W Y Q L H Q D G K S
391 CAATCAACACGGGTGCCGACGGTCTCCAGCGCCTCGACTACGTCGTCTCGTCTGCCGAACAGCACGACATCAAAC
ManF2
T I N T G A D G L Q R L D Y V V S S A E Q H D I K
466 TCATCATCAACTTCGTCAACTACTGGACCGATTACGGTGGTATGTCTGCGTACGTGAGCGCGTATGGCGGATCCG
L I I N F V N Y W T D Y G G M S A Y V S A Y G G S
541 GCGAGACGGATTTCTATACCAGTGATACCATGCAGAGTGCCTATCACACATATATCAAGACGGTCGTGGAGCGGT
G E T D F Y T S D T M Q S A Y H T Y I K T V V E R
616 ACAGTAACTCCTCGGCGGTGTTTGCGTGGGAGTTGGCGAATGAGCCGAGATGTCCGAGTTGCGATACTTCTGTGT
Y S N S S A V F A W E L A N E P R C P S C D T S V
691 TGTATAACTGGATTGAGAAGACGAGTAAGTTTATTAAGGGGTTGGATGCGGATCGTATGGTTTGTATTGGTGATG
L Y N W I E K T S K F I K G L D A D R M V C I G D
766 gtgagttgccccatactcagtccttgattttagttttgtatggtgggtgctgatgaaaggtagAGGGCTTCGGTC
E G F G
841 TCAACATCGACTCAGACGGCAGCTACCCTTATCAATTCTCCGAGGGCTTGAACTTTACGATGAACCTCGGTATCG
L N I D S D G S Y P Y Q F S E G L N F T M N L G I
916 ATACTATTGACTTTGGTACCCTCCACTTGTACCCTGATAGCTgtacgtacccctctccccttactctccttgcc
D T I D F G T L H L Y P D S
991 gtcaatatccagactaatcaaatgaagGGGGCACCTCCGACGACTGGGGCAACGGCTGGATCACCGCCCACGGCG
W G T S D D W G N G W I T A H G
1066 CAGCCTGCAAAGCAGCCGGCAAGCCATGTCTCCTGCAGGAATACGGAGTCACCTCGAACCACTGCAGTGTGGAGG
A A C K A A G K P C L L Q E Y G V T S N H C S V E
1141 GCTCGTGTCAGAAGACAGCGCTCAGCACAACGGGCGTCGGCGCGGATCTGTTCTGGCAGTATGGTGATGATTTGA
G S C Q K T A L S T T G V G A D L F W Q Y G D D L
1216 GTACCGGGAAGTCGCCGGATGATGGGAATACTATCTACTATGGGACTAGTGATTATCAGTGCCTGGTGACGGATC
S T G K S P D D G N T I Y Y G T S D Y Q C L V T D
1291 ATGTTGCTGCTATTGGTAGTGCTTAAGGGGGATACTGAGACTGAGATGTGGGGGAAAGGATGTGATTGTTGGTGA
H V A A I G S A *
1366 TTTGAGAATGTATGTACAAAAGGGGGGAAGCGGTGAATGAGAGTTGTTCGTTGAGTTTTTGGTAAACAATTGGAC
1441 AGTCAAACCATCAGCCAAGACTAGCTAAGATACTGGGTATTTTGGTCACTGCTTTTCATAATAAAACATGCTATT
1516 TATAAATCAAACAAGTAAAACTTAA
CManR

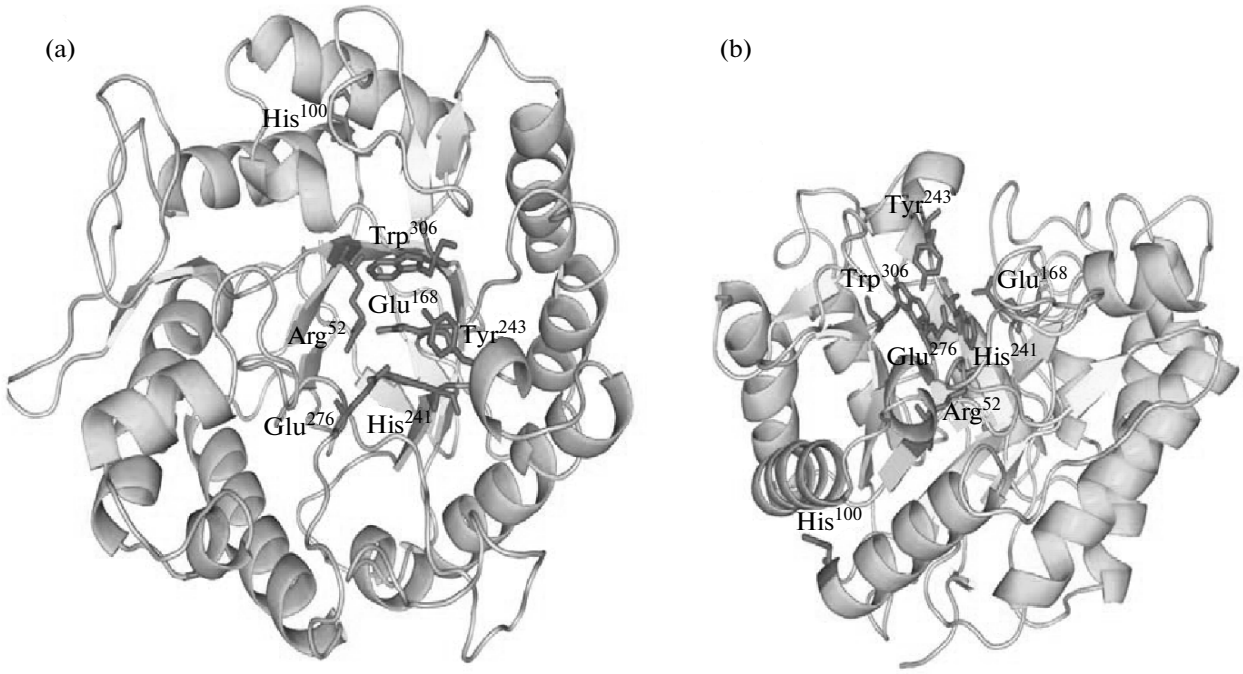


Fig. 4. The 3-D structure of the AnMan5A from *A. niger* LW-1 modeled with the bitemplate-based method using both MODELLER 9.9 and SALIGN programs, based on crystal structures of the *T. reesei* β-mannanase (1QNO) and *L. esculentum* β-mannanase (1RH9). The 3-D structure consists principally of the $(\alpha/\beta)_8$ TIM barrel fold. Two catalytic residues (Glu168 and Glu276) and 5 active site residues (Arg52, Asn167, His241, Tyr243, and Trp306) were strictly conserved among the GH family 5 members. (a) View along the barrel axis. (b) The "salad bowl" view showing the hydrophobic cleft for substrate binding and cleaving.

face of the molecule having a large radius (approximately 45 Å) due to an elaborate loop architecture, while the opposite face, which consists of simple α/β turns, had a radius of approximately 30 Å. This was similar to the fold described for GH family 10 enzymes and both were members of GH clan-A [19]. Indeed, two GH families 5 and 10 were quite closely related and in addition to sharing a common fold they had the same type of catalytic mechanism and shared several common residues [17, 20]. Two catalytic residues, Glu168 and Glu276, located in the hydrophobic cleft of the AnMan5A, where the β-1,4-*D*-mannosidic linkages of the mannan or heteromannan backbone were inserted and got cleaved.

Cloning of the complete DNA of the *Anman5A*. An approximate 750-bp 5'-flanking regulatory region of the *Anman5A* from *A. niger* LW-1 was first amplified using primers T-5prF and ManR1, then subjected to the second-round PCR for confirmation using primers T-5prF and ManR2. DNA sequencing result showed that the second-round PCR product was exactly 657 bp in length (except the sequence from the primer T-5prF to T/A clone site of pUCm-T vector), containing a new 275-bp sequence and a 382-bp sequence identical to that from the transcription starting point (G) to ManR2. Using the *A. niger* LW-1 genomic DNA as template, an approximate 1.8-kb complete DNA sequence of the *Anman5A* was directly amplified by the conventional PCR using primers CManF and CManR, and then cloned into pUCm-T vector, followed by DNA sequencing. Result showed that the DNA sequence was exactly 1,825 bp in length, containing a 5'-flanking regulatory region, 3 exons and 2 short introns with sizes of 63 and 60 bp, respectively (Fig. 3).

Analysis of the complete DNA of the *Anman5A*. Compared with the full-length cDNA sequence of the *Anman5A*, the complete DNA sequence was composed of a 285 bp of 5'-flanking regulatory region, a 41 bp of 5'-UTR region, 2 short introns with sizes of 63 and 60 bp respectively, 1152 bp of ORF, and a 224 bp of 3'-UTR. Both of the exon/intron boundaries conform to the canonical GT-AG rule. It was predicted that the 5'-promoter region of the *Anman5A* located at the range from −40 to +10 bp, designating the transcription starting point (G) as + 1 bp. The TATAAA sequence as a classical TATA box locates at −30 bp upstream the transcription starting point, which is in agreement with the consensus distance generally found in 5'-promoter regions of eukaryote genes. In some cases, other consensus sequences, such as TTATTT, also could act as substitutes for the classical

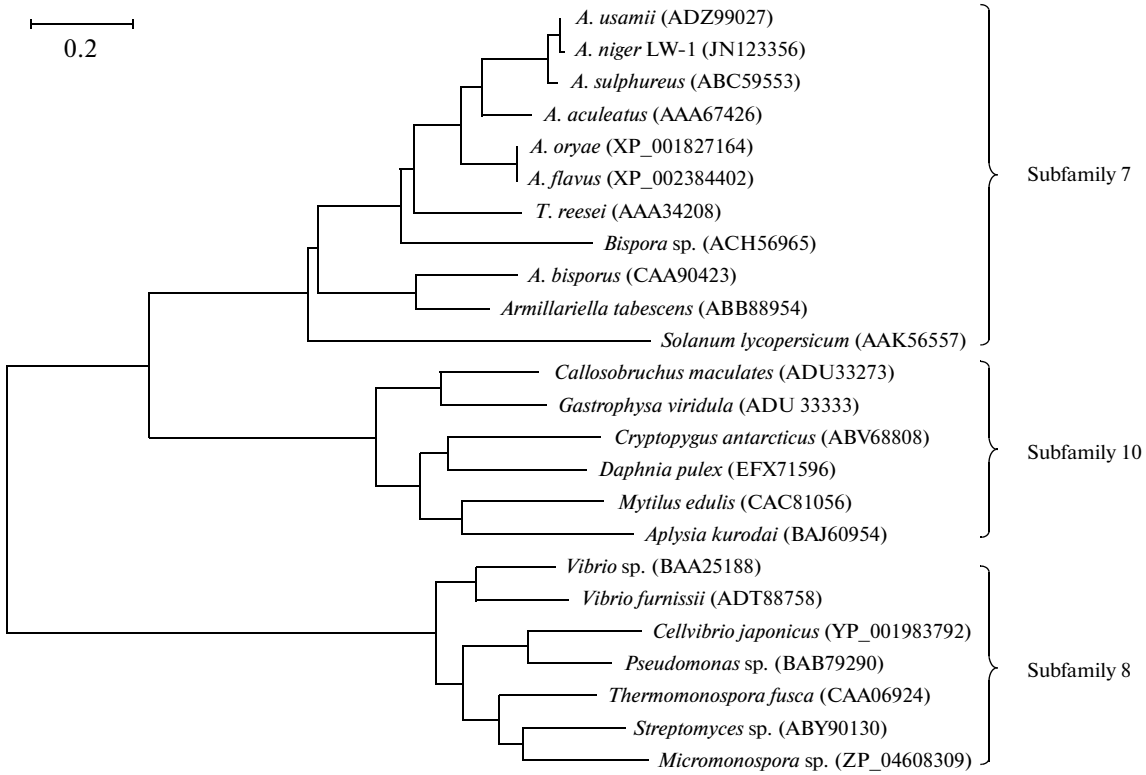


Fig. 5. The phylogenetic tree showing the evolutionary relativity and the homological degrees among the β-mannanase amino acid sequences from the GH family 5. The estimated genetic distance between sequences is proportional to the lengths of the horizontal lines connecting one sequence to another.

TATA box [21]. In eukaryotes, the functional CAAT box is typically found about −75 bp upstream the transcription starting point. Some CAAT boxes may locate far from the starting point [22]. However, we did not found any CAAT box in the *Anman5A* until −285 bp. It was also found that the AATAAA sequence as a putative polyadenylation signal located at +1500 bp downstream the transcription starting point.

***

The native AnMan5A purified from the cultivated koji of *A. niger* LW-1 possesses some predominant properties including the superior V_M, and good chemical tolerance, pH stability and thermostability, which are very suitable for industrial applications. But, its purification processes were time-consuming and laborious. To simplify the purification manipulations and to produce the AnMan5A inexpensively, it is necessary to express the *Anman5A* in heterologous cells, such as *Pichia pastoris* GS115 that is one of the favorite hosts.

In this work, we developed a procedure to clone both the full-length cDNA and complete DNA sequences of the *Anman5A* from *A. niger* LW-1 by using 4 steps of PCR amplification based on different principles. In addition, the bioinformatic analysis of the AnMan5A and *Anman5A* sequences was also reported. Our present work provided a solid basis for further researches on the *Anman5A* expression in *P. pastoris* GS115, the relationship between AnMan5A's structure and function, and improvement of its enzymatic properties by means of genetic engineering such as the site-directed mutagenesis, directed evolution and computational design.

ACKNOWLEDGMENTS

This work was supported by a grant from the Doctoral Research Funds of Jiangnan University (No. JUDCF11011) and the Postgraduate Innovation Training Project of Jiangsu, China (CXZZ11_0480). We are grateful to Prof. Xianzhang Wu (School of Biotechnology, Jiangnan University) for providing technical assistance.

REFERENCES

1. *Moreira, L.R.S.* and Filho, E.X.F., *Appl. Microbiol. Biotechnol.*, 2008, vol. 79, no. 2, pp. 165–178.
2. *Dhawan, S.* and Kaur, J., *Crit. Rev. Biotechnol.*, 2007, vol. 27, no. 4, pp. 197–216.
3. Luo, H.Y., Wang, Y.R., Wang, H., Yang, J., Yang, Y.H., Huang, H.Q., Yang, P.L., Bai, Y.G., Shi, P.J., Fan, Y.L., and Yao, B., *Appl. Microbiol. Biotechnol.*, 2009, vol. 82, no. 3, pp. 453–461.

4. Tang, C.D., Guo, J., Wu, M.C., Zhao, S.G., Shi, H.L., Li, J.F., and Wang, J.Q., *World J. Microbiol. Biotechnol.*, 2011, vol. 27, no. 12, pp. 2921–2929.
5. Wu, M.C., Tang, C.D., Li, J.F., Zhang, H.M., and Guo, J., *Carbohydr. Res.*, 2011, vol. 346, no. 14, pp. 2149–2155.
6. Chen, X.L., Cao, Y.H., Ding, Y.H., Lu, W.Q., and Li, D.F., *J. Biotechnol.*, 2007, vol. 128, no. 3, pp. 452–461.
7. Roth, R., Moodley, V., and van Zyl, P., *Mol. Biotechnol.*, 2009, vol. 43, no. 2, pp. 112–120.
8. Stalbrand, H., Saloheimo, A., Vehmaanpera, J., Henrissat, B., and Penttila, M., *Appl. Environ. Microbiol.*, 1995, vol. 61, no. 3, pp. 1090–1097.
9. Li, J.F., Wu, M.C., and Xia, W.S., *J. Food Sci. Biotechnol.*, 2007, vol. 26, no. 4, pp. 88–92. (in Chinese)
10. Li, J.F., Wu, M.C., and Xia, W.S., *Food Ferment. Ind.*, 2008, vol. 34, no. 1, pp. 5–9.
11. Sambrook, J. and Russell, D.W., *Molecular Cloning: A Laboratory Manual* (3rd ed), Cold Spring Harbor, New York: Cold Spring Harbor Lab., 2001, pp. 1595–1604.
12. Wu, M.C., Wang, J.Q., Zhang, H.M., Tang, C.D., Gao, J.H., and Tan, Z.B., *World J. Microbiol. Biotechnol.*, 2011, vol. 27, no. 4, pp. 831–839.
13. Laemmli, U.K., *Nature*, 1970, vol. 227, no. 5259, pp. 680–685.
14. Dubois, M., Gilles, K.A., Hamilton, J.K., Rebers, P.A., and Smith, F., *Anal. Chem.*, 1956, vol. 28, no. 3, pp. 350–356.
15. Ademark, P., Varga, A., Medve, J., Harjunpaa, V., Drakenberg, T., Tjerneld, F., and Stalbrand, H., *J. Biotechnol.*, 1998, vol. 63, no. 3, pp. 199–210.
16. Wang, J.Q., Zhang, H.M., Wu, M.C., and Tang, C.D., *Biotechnol. Lett.*, 2011, vol. 33, no. 3, pp. 1029–1038.
17. Larsson, A.M., Anderson, L., Xu, B., Munoz, I.G., Uson, I., Janson, J.C., Stalbrand, H., and Stahlberg, J., *J. Mol. Biol.*, 2006, vol. 357, no. 5, pp. 1500–1510.
18. Hilge, M., Gloor, S.M., Rypniewski, W., Sauer, O., Heightman, T.D., Zimmermann, W., Winterhalter, K., and Piontek, K., *Structure*, 1998, vol. 6, no. 11, pp. 1433–1444.
19. Zhang, Y.L., Ju, J.S., Peng, H., Gao, F., Zhou, C., Zeng, Y., Xue, Y.F., Li, Y., Henrissat, B., Gao, G.F., and Ma, Y.H., *J. Biol. Chem.*, 2008, vol. 283, no. 46, pp. 31551–31558.
20. Larson, S.B, Day, J., Barba de la Rosa, A.P., Keen, N.T., and McPherson, A., *Biochem.*, 2003, vol. 42, no. 28, pp. 8411–8422.
21. Lampidonis, A.D., Stravopodis, D.J., Voutsinas, G.E., Messini-Nikolaki, N., Stefos, G.C., Margaritis, L.H., Argyrokastritis, A., Bizelis, I., and Rogdakis, E., *Gene*, 2008, vol. 427, no. 1–2, pp. 65–79.
22. Wenkel, S., Turck, F., Singer, K., Gissot, L., Le Gourrierec, J., Samach, A., and Coupland, G., *Plant Cell*, 2006, vol. 18, no. 11, pp. 2971–2984.