International Scholars Journals

Advanced Journal of Microbiology Research ISSN 2241-9837 Vol. 13 (5), pp. 001-008, May, 2019. Available online at www.internationalscholarsjournals.org © International Scholars Journals

Full Length Research Paper

Molecular cloning and extracellular expression of cyclodextrin glycosyltransferase gene from *Bacillus* sp. NR5 UPM

Norhayati Ramli[1], Suraini Abd-Aziz[1]*, Mohd Ali Hassan[1], Noorjahan Banu Alitheen[2], Kamarulzaman Kamaruddin[3] and Zoolhilmi Ibrahim[3]

[1]Department of Bioprocess Technology, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.

[2]Department of Cell and Molecular Biology, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia.

[3]Environment and Bioprocess Technology Centre, SIRIM Berhad, No. 1, Persiaran Dato' Menteri, Section 2, 40911 Shah Alam, Malaysia.

Accepted 16 March, 2019

The cloning of a polymerase chain reaction (PCR) gene fragment from *Bacillus* sp. NR5 UPM isolated from the soil in Malaysia into an *Escherichia coli* expression vector was successfully carried out. Analysis of the nucleotide sequences revealed the presence of an open reading frame of 2112 bp which encoded a protein containing 704 amino acids with a putative molecular weight of 78.6 kDa. The deduced amino acids sequence showed about 98% homology with the CGTase from *Bacillus* sp. KC201. Compared to the wild type, the CGTase that was produced in *E. coli* cells only required one-fourth of culture time and neutral pH to produce CGTase. After 12 h of cultivation, the CGTase activity in the culture medium reached 29.6 U/ml, which was approximately 2.5-fold higher than the CGTase from the parental strain. The CGTase was produced extracellularly by *E. coli* (94%) indicating the signal peptide was functional in *E. coli*.

Key words: Molecular cloning, nucleotide sequence, cyclodextrin glycosyltransferase, *Bacillus* sp. NR5 UPM.

INTRODUCTION

Cyclodextrin glycosyltransferase (CGTases, 1,4-α-D-glucopyranosyltransferase (cyclizing), EC 2.4.1.19) is an important enzyme that catalyzes the formation of α-CD, β-CD and γ-CD, containing 6, 7 and 8 glucose residues linked with α-1,4-glucosidic bonds, respectively. Due to their unique abilities to form inclusion complexes with a variety of hydrophobic materials and to entrap volatile compounds, these CDs have found extensive applications in food, pharmaceuticals, agricultural chemicals, cosmetics, industrial chemicals and others (Hashimoto, 2002).

Recently, many researchers have studied the molecular cloning of CGTase genes and analysed the genetic information in order to provide a better CGTase production method. The over expression of CGTase genes could enhance the enzyme activity, reduce cultivation time and produce less contaminating proteins compared to wild type (Charoensakdi et al., 2007). In this study, we have succeeded in isolating the CGTase gene from *Bacillus* sp. NR5 UPM. The isolated CGTase gene was cloned into an *E. coli* expression vector and over expressed to study the improved properties of the enzyme.

*Corresponding author. E-mail: suraini@biotech.upm.edu.my. Tel: +6-03-89467516/7479. Fax: +6-03-89467510.

MATERIALS AND METHODS

Bacterial strain, plasmids and media

Bacillus sp. NR5 UPM was grown in Horikoshi medium II, containing 1% (w/v) soluble starch, 0.5% (w/v) yeast extract, 0.5% (w/v) peptone, 0.1% (w/v) K_2HPO_4, 0.02% (w/v) $MgSO_4.7H_2O$ and

Table 1. Primers used for amplification of upstream and downstream adjacent sequences based on known sequences of the CGTase gene.

Primers	Sequence (5' to 3' direction)
DW2-ACP2[a]	ACPCTCGA
DW2-ACP3[a]	ACPCTACG
DW2-ACPN[a]	ACPN
Universal Primer, UniP2	GAGTTTAGGTCCAGCGTCCGT
TSP1R[b]	GCGAAGTGGTGGTTTTGG
TSP2R[b]	GGTTTGCCACCCTTCTGACATA
TSP3R[b]	GCACCATTTTCAACATAGTTAGGG
TSP1F[b]	AAAGGGATTGATGGTATTCG
TSP2F[b]	CGAAATCTACTCGCATAAACCTG
TSP3F[b]	CCCCAAAACCACCACTTCG
nTSP3F[c]	GGCTATGGCACTACAACTGAAC
n2TSP3F[c]	ACAATGCCACAACGAGCCCAG
n3TSP3F[c]	AGCGTAGCGGTTTGGCAAGTAA

[a]Annealing control primer, provided by the manufacturer; [b] Target specific primer, designated on the basis of known sequence of CGTase gene; [c] Target specific primer, designated on the basis of the newly determined sequence in the third PCR operation.

1% (w/v) Na_2CO_3 at 37°C (Park et al., 1989). *Escherichia coli* JM 109, used as the cloning host, was grown in Luria-Bertani (LB) and terrific broth (TB). For solid medium, 1.5% agar was added to LB broth. Plasmid pUC19 was used as an expression vector. Ampicillin (100 µg/ml) was added to the medium to allow the growth of the plasmid-carrying strain.

DNA manipulation and amplification of known sequences of CGTase gene

Chromosomal DNA was isolated from *Bacillus* sp. NR5 UPM by the method of Ish-Horowicz and Burke (1981). Other general molecular experiments involving PCR reaction and transformation were performed according to the methods described by Sambrook and Fritsch (1989). Genomic DNA was extracted and used as the template for the preparation of the CGTase gene by PCR technique. A pair of degenerate primers; forward primer C1, 5'-GGN GGN GAY TGG CAR GGN-3' and reverse primer C2, 5'-CAT RTC RTG RTT RTC DAT RAA-3' designed based on the conserved sequences of the CGTase gene was used as PCR primers (Ong et al., 2008). The reaction contained 1 µg of *Bacillus* sp. NR5 UPM DNA, 100 pmol of each forward and reverse primer, 0.01 U/µl *Taq* Polymerase in 1X reaction buffer and 0.2 mM of each dNTPs. The DNA was initially denatured at 94°C for 5 min, followed by denaturation at 94°C for 20 s. Then, the annealing step was carried out at 50°C for 20 s and extension at 72°C for 5 s. The total number of cycles was 30.

Sequence analysis of the adjacent region of the known sequences

The adjacent region of known sequences of the CGTase gene was amplified by PCR using DNA walking *SpeedUp*™Premix Kit II (Seegene, Inc). The PCR reaction was performed with the following primers (Table 1) and *Bacillus* sp. NR5 UPM DNA as the template. The primers used were designed to amplify the upstream and downstream adjacent sequences based on known sequences of the CGTase gene. For the first PCR reaction, the primers used were a combination of either DW2-ACP2 or DW2-ACP3 together with TSP1R and TSP1F to amplify the upstream and downstream regions, respectively. Each of the first PCR products was used as a template for the second PCR. The second PCR was performed with DW2-ACPN together with TSP2R for amplification of the upstream region and TSP2F for amplification of the downstream region.

The same operation was repeated with the other primers to determine the nucleotide sequence of the extending region. The primers used were TSP3R for the amplification of the upstream region and TSP3F for amplification of the downstream region together with Universal Primer, UniP2 for each of them. Other primers were designed on the basis of the newly determined sequence in the third PCR reaction, namely nTSP3F, n2TSP3F and n3TSP3F for amplification of the downstream region. DNA sequencing of the resulting third PCR was performed by First Base Laboratories, Biosyntech Malaysia on the ABI PRISM 377 DNA sequencer.

Cloning of CGTase gene

The expression plasmid for the new ORF was constructed as follows. The new ORF of the CGTase gene was amplified by PCR. The chromosomal DNA of *Bacillus* sp. NR5 UPM was used as the template with 5'-AGCGGATCCTTGTTTTATTTTATATACGTT-3' (*BamH*I site underlined) as the forward primer and 5'-GTCAAGCTTTTACCAATTAATCATAACCGT-3' (*Hind*III site underlined) as the reverse primer. The amplified fragment was digested with both *BamH*I and *Hind*III, and then ligated into the corresponding sites on the vector pUC 19. The resultant plasmid was designated as pCGT3D. The ligation products were used for transformation into *E. coli* JM109 cells. All the white colonies were tested by the agar plate assay method which contained 1% soluble starch and 100 µg/ml ampicillin in LB agar. After growth at 37°C for 24 h, the ability of the possible white colony to form halo zone after exposure to $KI-I_2$ was observed.

Expression of CGTase gene

A single colony of *E. coli* JM109 cells harbouring plasmid pCGT3D was inoculated into 10 ml of LB medium containing 100 µg/ml ampicillin and grown at 37°C overnight. 10% of overnight culture was then diluted into 100 ml of terrific broth (TB) containing 100 µg/ml ampicillin in a 250 ml flask. The culture was incubated at 37°C on a rotary shaker (200 rpm). Samples were withdrawn hourly and analyzed for enzyme activities.

Nucleotide and protein sequence analysis

The blastn and blastx programs provided by the National Center for Biotechnology Information (NCBI) were used to search for existing sequences that were similar to the sequences we obtained. The amino acid and nucleotide sequences we obtained were compared with other sequences using the BioEdit 7.01 program.

Cell fractionation

Cell fractionation was performed as described by Ding et al. (2010). The culture broth was centrifuged at 12,000 rpm for 5 min at 4°C to obtain the extracellular fraction. Then, 1 ml culture was harvested and washed twice with 1 ml of 30 mM Tris-HCl buffer (pH 7.0). The cells were resuspended in the same buffer containing 25% (w/v) sucrose and 1 mM EDTA to separate the periplasmic fraction. The cell suspension was incubated on ice for 2 h and centrifuged at 12,000 rpm for 5 min at 4°C. The supernatant obtained was collected as the periplasmic fraction.

Assay of ß-CGTase

The phenolphthalein assay (Kaneko et al., 1987) was used to determine the CGTase activity. The reaction mixture which contained 1 ml of 40 mg of soluble starch in 100 mM phosphate buffer (pH 6.0) and 0.1 ml enzyme solution was incubated at 60°C for 10 min in a water bath. To stop the reaction, 3.5 ml of 30 mM NaOH was added to the mixture. Subsequently, 0.5 ml of 0.02% (w/v) phenolphthalein in 5 mM Na_2CO_3 solution was added to the reaction mixture and mixed well.

Then, the reduction in colour intensity was measured at 550 nm after leaving at room temperature for 15 min. Control lacking CGTase was analysed simultaneously with each batch of samples. The soluble starch and enzyme were replaced by 0.5 mg of β-CD and 0.1 ml of water, respectively as a standard. One unit of enzyme activity was defined as the amount of the enzyme that formed 1 µmol β-CD per min under the conditions aforestated.

RESULTS AND DISCUSSION

Amplification of known sequence of CGTase gene

Approximately, 800 bp of polymerase PCR product was successfully amplified from the genomic DNA of *Bacillus* sp. NR5 UPM. The nucleotide sequence analysis for the amplification of known sequences of CGTase gene resulted in a 752 bp size product. The sequence was subjected to a search against all known sequences in databases using blastn and blastx searches. It showed 95% maximum identity with *Bacillus* sp. KC201 CGTase (GenBank accession no. D13068.1).

Sequence analysis for amplification of CGTase gene

The DNA walking strategy was used for the amplification of the CGTase gene with genomic DNA of *Bacillus* sp. NR5 UPM as the template. The nucleotide sequence analysis of the upstream and downstream adjacent sequences based on known sequences of CGTase gene revealed a new ORF of CGTase gene from *Bacillus* sp. NR5 UPM. The nucleotide sequence has been deposited in the GenBank database under the accession number HQ876173. The new ORF of the CGTase gene consisted of 2112 bp and encoded 704 amino acids.

The open reading frame shows unique properties with TTG, rather than ATG as a start codon. The same result has also been reported in CGTase from *Bacillus* sp. G1 (Ong et al., 2008), *Bacillus* sp. TS1-1 (Rahman et al., 2006), and *Bacillus ohbensis* (Sin et al., 1991). Besides TTG and ATG, GTG and CTG are also frequently used by bacteria as start codons. Initially, TTG at nucleotide 1 was thought to be a start codon. However, the identification of Shine-Dalgarno (SD) sequence showed it was located at nucleotide 40. The identification of SD sequence is important as it aided in the recognition of the initiation codon. The SD sequence, which is also known as ribosome binding site, must be located at 6 to 13 bases upstream from the start codon and the sequence should be totally complementary to the 3' end of the 16S ribosomal RNA of *Bacillus* sp. NR5 UPM. The sequence of GGAGGA is believed to be the SD sequence which was located at 6 bp upstream from initiation the codon. Thus, TTG at nucleotide 52 is most likely to be the true initation codon (Figure 1). The presence of the ribosome binding site located in the upstream region of the initiation codon is important as it may contribute to the excellent expression of the CGTase gene in *E. coli*.

According to the amino acid sequence deduced from the DNA sequence, the first 29 amino acid residues maybe a signal peptide which is involved in the secretion of the protein. Indeed, there is no strong sequence homology among the cleavage sites and it is known to have highly variable and rapidly evolving structures. However, the comparison of signal peptide shows consistent characteristics among Gram-positive bacteria (von Heijne, 1986). According to the (-3,-1)-rule of signal peptide, there is a common pattern for a probable cleavage site and this pattern shows in agreement with the amino acid residues of signal peptide deduced from the DNA sequence. The presence of Gly at position -3 is compatible with the (-3,-1)-rule where the residue in that position must not be amino acids with electrically charged side chain, aromatic or large and polar. Besides, the presence of Gly, Ile, Glu and Leu at position -3 through +1, is in accordance with the rule that requires the absence of non-polar amino acid with hydrophobic side chain at that position.

Amino acid sequence analysis with other CGTases and amylase

The CGTase sequence from *Bacillus* sp. NR5 UPM was

```
                                               SD
1    TTGTTTTATTTTATATACGTTTTATCACAATTTGTTGAAGGAGGAATAGAGTTGAACTGTTTAAACGGTT 70
      L  F  Y  F  I  Y  V  L  S  Q  F  V  E  G  G  I  E  L  N  C  L  N  G

71   TTTTAAAAACGATTTCATTATGCTTTATCTTTTTCTTACTTCTTTCTTTACCTACTGTTGCGGAGGCTGA 140
     F  L  K  T  I  S  L  C  F  I  F  F  L  L  L  S  L  P  T  V  A  E  A ⇑ D

141  CGTAACAAATAAAGTCAATTACTCAAAAGATGTAATTTACCAGATTGTTACCGATCGATTCTCTGACGGG 210
       V  T  N  K  V  N  Y  S  K  D  V  I  Y  Q  I  V  T  D  R  F  S  D  G

211  AATCCTGGCAACAACCCTTCAGGCGCTATCTATAGTCAGAACTGTATAGATCTCCATAAGTATTGTGGTG 280
      N  P  G  N  N  P  S  G  A  I  Y  S  Q  N  C  I  D  L  H  K  Y  C  G

281  GGGACTGGCAAGGGATTATAGACAAAATCAATGACGGTTACTTAACTGATTTAGGCATTACGGCACTATG 350
     G  D  W  Q  G  I  I  D  K  I  N  D  G  Y  L  T  D  L  G  I  T  A  L  W

351  GATCTCTCAGCCGGTCGAAAATGTTTACGCCTTACACCCAAGCGGCTATACCTCCTACCATGGATATTGG 420
       I  S  Q  P  V  E  N  V  Y  A  L  H  P  S  G  Y  T  S  Y  H  G  Y  W

421  GCTCGAGATTACAAAAAGACAAACCCTTACTATGGGAATTTCGATGACTTTGATCGTTTAATGAGTACCG 490
      A  R  D  Y  K  K  T  N  P  Y  Y  G  N  F  D  D  F  D  R  L  M  S  T

491  CACATAACAATGGGATAAAGGTAATCATGGATTTCACGCCAAATCATTCATCACCTGCACTTGAAACAAA 560
     A  H  N  N  G  I  K  V  I  M  D  F  T  P  N  H  S  S  P  A  L  E  T  N

561  CCCTAACTATGTTGAAAATGGTGCGATATATGATAATGGCGCATTATTAGGCAACTATTCAAATGATCAA 630
       P  N  Y  V  E  N  G  A  I  Y  D  N  G  A  L  L  G  N  Y  S  N  D  Q

631  CAAAACCTCTTTCACCACAATGGCGGAACAGATTTCTCTTCATATGAAGATAGCATCTACAGAAACTTAT 700
      Q  N  L  F  H  H  N  G  G  T  D  F  S  S  Y  E  D  S  I  Y  R  N  L

701  ATGATTTGGCAGACTATGATTTAAACAACACAGTCATGGATCAGTATTTAAAGGAGTCGATTAAGTTCTG 770
     Y  D  L  A  D  Y  D  L  N  N  T  V  M  D  Q  Y  L  K  E  S  I  K  F  W

771  GTTAGATAAAGGGATTGATGGTATTCGAGTGGATGCCGTTAAGCATATGTCAGAAGGGTGGCAAACCTCT 840
       L  D  K  G  I  D  G  I  R  V  D  A  V  K  H  M  S  E  G  W  Q  T  S

841  TTAATGAGCGAAATCTACTCGCATAAACCTGTTTTCACATTTGGAGAATGGTTTTTAGGATCAGTTGAAG 910
      L  M  S  E  I  Y  S  H  K  P  V  F  T  F  G  E  W  F  L  G  S  V  E

911  TTGATCCCCAAAACCACCACTTCGCCAATGAAAGTGGTATGAGCTTATTGGATTTCCAATTTGGTCAAAC 980
     V  D  P  Q  N  H  H  F  A  N  E  S  G  M  S  L  L  D  F  Q  F  G  Q  T

981  CATTCGTAACGTCTTAAAAGATCGCACAAGCAACTGGTATGATTTTCATGACATGATAAAAAGTACTGAA 1050
       I  R  N  V  L  K  D  R  T  S  N  W  Y  D  F  H  D  M  I  K  S  T  E

1051 AAAGAGTATAACGAGGTCATTGATCAAGTAACCTTTATTGATAATCACGACATGAGTCGTTTTTCGGTAG
      1120 K E Y N E V I D Q V T F I D N H D M S R F S V

1121 GATCATCTTCAAACCGTCAGACAGATATGGCGCTTGCTGTCTTGCTTACTTCCCGTGGTGTACCGACGAT
     1190 G S S S N R Q T D M A L A V L L T S R G V P T I

1191 TTACTACGGCACAGAGCAGTATGTAACAGGTGGCAACGACCCTGAAAATCGCAAACCATTGAAAACATTT
       1260 Y Y G T E Q Y V T G G N D P E N R K P L K T F

1261 GATCGGTCTACCAACTCTTATCAAATCATCAGTAAACTTGCTTCGCTACGCCAAACAAATTCCGCCTTAG
      1330 D R S T N S Y Q I I S K L A S L R Q T N S A L

1331 GCTATGGCACTACAACTGAACGTTGGCTGAACGAAGACATTTATATTTATGAAAGAACGTTTGGCAATAG
     1400 G Y G T T T E R W L N E D I Y I Y E R T F G N S

1401 TATTGTATTAACTGCTGTCAATAGCAGTAATAGCAACCAGACGATCACTAATTTAAACACCTCTTTACCT
       1470 I V L T A V N S S N S N Q T I T N L N T S L P

1471 CAAGGGAACTATACGGATGAACTACAGCAACGTTTAGATGGAAACACGATTACTGTTAACGCCAATGGCG
      1540 Q G N Y T D E L Q Q R L D G N T I T V N A N G

1541 CCGTAAATTCCTTTCCATTACGAGCAAATAGCGTAGCGGTTTGGCAAGTAAGCAACCCCTCTACGTCTCC
     1610 A V N S F P L R A N S V A V W Q V S N P S T S P

1611 TCTAATCGGCCAAGTAGGTCCTATGATGGGTAAGGCCGGGAATACCATAACAGTAAGCGGTCAAGGATTT
       1680 L I G Q V G P M M G K A G N T I T V S G Q G F

1681 GGTGATGAGAGAGGGAGCGTACTCTTTGATTCAACCTCTTCTGAAATTATTTCTTGGTCAAATACAGAAA
      1750 G D E R G S V L F D S T S S E I I S W S N T E
```

Figure 1. Nucleotide sequence of the cyclodextrin glucanotransferase (CGTase) gene from *Bacillus* sp. NR5 UPM. The two possible start codons (TTG) are shown in the boxes with the second box most likely to be a true initiation codon. The ribosome binding site, SD, is underlined. The signal peptide is shown in italics. The possible signal peptide cleavage site is shown by a vertical arrow. This sequence was submitted to the GenBank with the accession number HQ876173.

```
1681 GGTGATGAGAGAGGGAGCGTACTCTTTGATTCAACCTCTTCTGAAATTATTTCTTGGTCAAATACAGAAA 1750
       G  D  E  R  G  S  V  L  F  D  S  T  S  S  E  I  I  S  W  S  N  T  E

1751 TAAGCGTAAAGGTGCCTAATGTAGCAGGCGGTTATTATGATTTATCCGTCGTAACTGCAGCAAACATAAA 1820
     I  S  V  K  V  P  N  V  A  G  G  Y  Y  D  L  S  V  V  T  A  A  N  I  K

1821 AAGCCCTGCTTACAAAGAGTTTGAAGTATTGTCAGGCAATCAAGTCAGTGTTCGCTTTGGTGTTAACAAT 1890
        S  P  A  Y  K  E  F  E  V  L  S  G  N  Q  V  S  V  R  F  G  V  N  N

1891 GCCACAACGAGCCCAGGAACCAATTTATATATCGTTGGGAATGTGAGCGAGCTGGGGAATTGGGATGCTG 1960
       A  T  T  S  P  G  T  N  L  Y  I  V  G  N  V  S  E  L  G  N  W  D  A

1961 ATAAAGCAATTGGACCTATGTTTAACCAAGTGATGTACCAATACCCAACGTGGTACTATGATATTAGCGT 2030
     D  K  A  I  G  P  M  F  N  Q  V  M  Y  Q  Y  P  T  W  Y  Y  D  I  S  V

2031 TCCTGCTGGAAAAAACCTTGAATACAAATACATTAAAAAAGATCAGAACGGTAACGTTGTCTGGCAAAGT 2100
        P  A  G  K  N  L  E  Y  K  Y  I  K  K  D  Q  N  G  N  V  V  W  Q  S

2101 GGCAATAATCGAACCTATACGTCACCTACTACCGGAACAGATACGGTTATGATTAATTGGTAA 2163
       G  N  N  R  T  Y  T  S  P  T  T  G  T  D  T  V  M  I  N  W  *
```

Figure 1. Nucleotide sequence of the cyclodextrin glucanotransferase (CGTase) gene from *Bacillus* sp. NR5 UPM. The two possible start codons (TTG) are shown in the boxes with the second box most likely to be a true initiation codon. The ribosome binding site, SD, is underlined. The signal peptide is shown in italics. The possible signal peptide cleavage site is shown by a vertical arrow. This sequence was submitted to the GenBank with the accession number HQ876173.

compared with other CGTases and amylase sequences using BLAST program. A homology search revealed that CGTase from *Bacillus* sp. NR5 UPM had highest similarity (98%) with *Bacillus* sp. KC201 CGTase (GenBank accession no. D13068.1). The other CGTases such as CGTases from *Bacillus* sp. G1-2004, *Bacillus* sp. TS1-1 and alkalophilic *Bacillus* 1-1 (GenBank accession nos. AAV38118.2, AAV38117.1 and P31746.1, respectively) also showed high homology (more than 95%) with the amino acid sequence of CGTase from *Bacillus* sp. NR5 UPM. According to Figure 2, six highly conserved regions (labelled as I to VI) were identified among CGTases and amylase. The deduced amino acid sequence of CGTases and amylase showed the homology of sites important for catalysis, which suggested a common evolutionary derivation of these two classes of enzymes (Binder et al., 1986).

Cloning and expression of CGTase gene

The sequence analysis of the open reading frame strongly suggests that it belongs to the CGTase gene. Hence, the 2.1 kb-*BamH*I-*Hind*III fragment containing the open reading frame was cloned into pUC19 under *lac* promoter. The CGTase gene was cloned into *E. coli* JM 109 as a host. The formation of clear halo zone around the colony cultured on a LB-starch plate (containing 1.5% agar, 100 µg/ml ampicillin and 1% potato soluble starch) when stained with iodine solution indicated the successful expression of starch degrading activity.

For extracellular expression of recombinant CGTase, terrific broth (TB) was used as a culture medium due to its pH-buffering capacity which gives beneficial effects on cell growth and enzyme stability. Indeed, the large amounts of yeast extract contained in the TB medium gives a critical effect on the release of the recombinant enzyme from the periplasmic space (Li et al., 2010).

Secretion of extracellular CGTase from *E. coli* transformant

The CGTase of *Bacillus* sp. NR5 UPM in *E. coli* was produced extracellularly into the culture medium, indicating the signal peptide of CGTase was functional in *E. coli*. The CGTase activities localized extracellularly and in the periplasm were 29.6 and 1.9 U/ml, respectively; that is 94 and 6%, respectively. A comparative study on the characteristics of a signal peptide shows the importance of an amino-terminal positively charged region (n-region, 1-5 residues) and a central-hydrophobic region (h-region, 7-15 residues) in targeting the enzyme into the extracellular medium. The signal peptidases play a critical role in cleaving the amino terminus of the protein once its targeting function has

```
KC201    1   MADTWASEIGKLSQDKPFDLFSGKKVEKGVSGGITSFIFFLLLSLPTVAEADVTNKVNYS 60
BC       1   -MFYFIYVLSQFVEGG-IELNCLNGFLKTIS---LCFIFFLLLSLPTVAEADVTNKVNYS 55
BO       1   -------------------MKNLTVLLKTIP---LALLLFILLSLPTAAQADVTNKVNYT 38
NR5      1   -MFYFIYVLSQFVEGG-IELNCLNGFLKTIS---LCFIFFLLLSLPTVAEADVTNKVNYS 55
amylase  1   --------------------------------MMVAWWSLFLYGLQVAAPALAATPADWR 28
                    I                                   II
KC201    61  KDVIYQIVTDRFSDGNP    -NNPSGAIFSQNCIDLHKYCGGDWQGIIDKINDGYLTDLGITA 119
BC       56  KDVIYQIVTDRFSDGNPGNNPSGAIYSQNCIDLHKYCGGDWQGIIDKINDGYLTDLGITA 115
BO       39  RDVIYQIVTDRFSDGDPSNNPTGAIYSQDCSDLHKYCGGDWQGIIDKINDGYLTDLGITA 98
NR5      56  KDVIYQIVTDRFSDGNPGNNPSGAIYSQNCIDLHKYCGGDWQGIIDKINDGYLTDLGITA 115
amylase  29   SQSIYFLLTDRFARTDGSTTATCNTADQ------KYCGGTWQGIIDKLD--YIQGMGFTA 80
                                 III
KC201    120 LWISQPVENVYALHPSGYTSYHGYWARDYKKTNPYYGNFDDFDRLMSTAHSNGIKVIMDF 179
BC       116 LWISQPVENVYALHPSGYTSYHGYWARDYKKTNPYYGNFDDFDRLMSTAHNNGIKVIMDF 175
BO       99  IWISQPVENVYALHPSGYTSYHGYWARDYKRTNPFYGDFSDFDRLMDTAHSNGIKVIMDF 158
NR5      116 LWISQPVENVYALHPSGYTSYHGYWARDYKKTNPYYGNFDDFDRLMSTAHNNGIKVIMDF 175
amylase  81  IWITPVTAQLPQTTAYG-   DAYHGYWQQDIYSLNENYGTADDLKALSSALHERGMYLMVDV 139

KC201    180 TPNHSSPALETNPNYVENGAIYDNGALLGNYSNDQQNLFHHNGGTDFSSYEDSIYR--NL 237
BC       176 TPNHSSPALETNPNYVENGAIYDNGALLGNYSNDQQNLFHHNGGTDFSSYEDSIYR--NL 233
BO       159 TPNHSSPALETDPSYAENGAVYNDGVLIGNYSNDPNNLFHHNGGTDFSSYEDSIYR--NL 216
NR5      176 TPNHSSPALETNPNYVENGAIYDNGALLGNYSNDQQNLFHHNGGTDFSSYEDSIYR--NL 233
amylase  140 VANHMG--------YDGAGSSVDYSVFKPFSSQDYFHPFCLIQNYEDQTQVEDCWLGDNT 191
                                             IV
KC201    238 YDLADYDLNNTVMDQYLKESIKFWLDKG- IDGIRVDAVKHMSEGWQTSLMSEIYSHKPVF 296
BC       234 YDLADYDLNNTVMDQYLKESIKFWLDKG- IDGIRVDAVKHMSEGWQTSLMSEIYSHKPVF 292
BO       217 YDLADYDLNNTVMDQYLKESIKLWLDKG- IDGIRVDAVKHMSEGWQTSLMSDIYAHEPVF 275
NR5      234 YDLADYDLNNTVMDQYLKESIKFWLDKG- IDGIRVDAVKHMSEGWQTSLMSEIYSHKPVF 292
amylase  192 VSLPDLDTTKDVVKNEWYDWVGSLVSNYSIDGLRIDTVKHVQKDFWPGYN----KAAGVY 247

KC201    297 TFGEWFLGSGEVDPQNHHFANESGMSLLDFQFGQTIRNVLKDRTSNWYDFNEMITSTEKE 356
BC       293 TFGEWFLGSVEVDPQNHHFANESGMSLLDFQFGQTIRNVLKDRTSNWYDFHDMIKSTEKE 352
BO       276 TFGEWFLGSGEVDPQNHHFANESGMSLLDFQFGQTIRDVLMDGSSNWYDFNEMIASTEED 335
NR5      293 TFGEWFLGSVEVDPQNHHFANESGMSLLDFQFGQTIRNVLKDRTSNWYDFHDMIKSTEKE 352
amylase  248 CIGEVLDG----DPAYTCPYQNVMDGVLNYPIYYPLLNAFKSTSGSMHDLYNMINTVKSD 303
                         V                                            VI
KC201    357         YNEVIDQVTFIDNHDMSRFSVGSSSNRQTDMALAVLLTSRGVPTIYYGTEQYVTGGNDPE        416
BC       353         YNEVIDQVTFIDNHDMSRFSVGSSSNRQTDMALAVLLTSRGVPTIYYGTEQYVTGGNDPE        412
BO       336         YDEVIDQVTFIDNHDMSRFSFEQSSNRHTDIALAVLLTSRGVPTIYYGTEQYLTGGNDPE        395
NR5      353         YNEVIDQVTFIDNHDMSRFSVGSSSNRQTDMALAVLLTSRGVPTIYYGTEQYVTGGNDPE        412
amylase  304         CPDSTLLGTFVENHDNPRFASYTNDIALAKNVAAFIILNDGIPIIYAGQEQHYAGGNDPA        363

KC201    417 NRKP--LKTFDRSTNSYQIISKLASLRQTNSALGYGTTTERWLNEDIYIYERTFGNSIVL 474
BC       413 NRKP--LKTFDRSTNSYQIISKLASLRQTNSALGYGTTTERWLNEDIYIYERTFGNSIVL 470
BO       396 NRKP--MSDFDRTTNSYQIISTLASLRQNNPALGYGNTSERWINSDVYIYERSFGDSVVL 453
NR5      413 NRKP--LKTFDRSTNSYQIISKLASLRQTNSALGYGTTTERWLNEDIYIYERTFGNSIVL 470
amylase  364 NREATWASGYPTDSELYKLIASANAIRN-------------------YAISKDTG----- 399
```

Figure 2. Comparison of the deduced amino acid sequence of CGTase from *Bacillus* sp. NR5 UPM with other CGTases and amylase. The six highly conserved regions in different CGTases and amylase are boxed. The numbers within brackets are GenBank accession numbers. KC201: CGTase from *Bacillus* sp. KC201 (D13068.1), BC: CGTase from *B. circulans* (X68326.1), BO: CGTase from *B. ohbensis* (D90243.1), NR5: CGTase from *Bacillus* sp. NR5 UPM (HQ876173.1), amylase: Taka-amylase A from *Aspergillus oryzae* (M33218.1).

been carried out. However, many cases show the translocation of the proteins ended up in the periplasm due to the failure of the integral membrane protein to remove the translocation signal which instead remained anchored to the membrane by an uncleaved signal peptide.

Meanwhile, the polar carboxy-terminal domain (c-region, 3-7 residues) is needed to specify the signal peptidase cleavage site for proper removal of the signal peptide from the mature chain (von Heijne, 1990). Many studies on molecular cloning of the CGTase gene have been carried out, in which the translocation of recombinant enzymes among the construct varied from extracellular medium into periplasmic and intracellular space (Table 2). Among the recombinants constructed, the NR5 transformant showed the best translocation of the CGTase into the extracellular medium (94%) compared to other constructs which were around 68 to 74% (Kim et al., 1998; Yong et al., 1996). The ability of the signal peptide to secrete the enzyme predominantly into the culture medium is beneficial because this will aid in the downstream processing, enhance the *in vivo*

Table 2. Translocation of recombinant enzymes into extracellular, periplasmic and intracellular spaces.

Parental strain	Cloning host	Expression vector	Mature enzyme (amino acids)	Signal peptide (amino acids)	Translocation of enzyme	References
Bacillus sp. NR5 UPM	*E. coli* JM109	pUC19	675	29	Extracellular	This study
Klebsiella pneumoniae M5a1	*E. coli* RR28	pHE3	625	30	Extracellular	Binder et al. (1986)
Bacillus sp. Strain no. 8	*E. Coli* MB 406	pTZ18R and pTZ19R	684	34	Extracellular	Nitschke et al. (1990)
Bacillus sp. TS1-1	*E. coli* JM109	pUC19	666	46	Extracellular	Rahman et al. (2006)
Bacillus sp. Strain No. 38-2	*E. coli* HB101	pBR322	685	27	Periplasmic	Kaneko et al. (1988)
Bacillus sp. KC201	*E. Coli* DH5α	pUC18	674	51	Intracellular	Kitamoto et al. (1992)

Table 3. Comparison of growth and CGTase production profiles between wild type (NR5) and transformant (pNR5).

Parameter	NR5	pNR5
Growth	30 – 40°C, pH 10	37°C, pH 7
Optimum CGTase production	37°C, pH 10	37°C, pH 7
Maximum CGTase activity	11.7 U/ml	29.6 U/ml
Time of cultivation	48 h	12 h

stability of the secreted enzyme, facilitate the folding processes and enable the production of enzyme in soluble and biologically active form which can finally reduce the subsequent purification cost (Mergulhao et al., 2005). The presence of functional signal peptide presents significant advantages in terms of specific secretion of the enzyme of interest, thus minimizing the contamination of the target enzyme with other non-target proteins.

Besides, the efficient secretion of enzyme into the extracellular medium also eliminates the need for the use of other approaches for protein secretion, such as the use of fusion partners, permeabilizing proteins, nutrients or other agents that can create “leakage” and increase the permeability of the outer membrane of *E. coli* (Makrides, 1996). The study by Kim et al. (2005) showed the co-expression of folding accessory proteins facilitated the production of active CGTase of *Bacillus macerans* in recombinant *E. coli*. The soluble expression of the target protein was improved with the use of folding accessory proteins and co-expression with molecular chaperons; otherwise the proteins are mainly expressed as inclusion bodies. The formation of inclusion bodies becomes a significant obstacle in gene expression due to the challenging task in the refolding process of the targeted protein. The supplementation of medium additives also has been verified to lead to the secretion of heterologous protein which has the ability to enhance the permeability of the membrane. Aristidou et al. (1993) reported that the supplementation of glycine gave significant effect to the release of the enzyme which resulted from the extensive cell lysis caused by glycine.

The study on the appropriate time for supplying glycine is also important because the addition of glycine might harm cell growth, thus affecting the overall enzyme production (Li et al., 2009). Together with glycine, the addition of Ca^{2+}, SDS and Na^{+} also facilitated the secretion of recombinant enzyme from *E. coli* (Ding et al., 2010).

Enzymatic properties of recombinant CGTase

The recombinant CGTase was characterized and compared with native enzyme from *Bacillus* sp. NR5 UPM with respect to their cyclization activities (Table 3).

The result suggested that the CGTase from transformants required a neutral pH to carry out the cyclization activity, while the wild type required alkaline pH as optimum culture condition for CGTase production. The same finding was reported by Charoensakdi et al. (2007). The 2.5-fold increment in CGTase activities and one-fourth less in cultivation time proved the transformant CGTase as a successful recombinant enzyme. The recombinant CGTase exhibited a putative molecular weight of 78.6 kDa.

Conclusion

A CGTase gene from *Bacillus* sp. NR5 UPM was successfully isolated, cloned and expressed into pUC19 cloning vector with *E. coli* as the host. The deduced amino acid sequence showed 2.1 kb mature CGTase with putative molecular weight of 78.6 kDa and exhibited 98% maximum identity with CGTase sequence from *Bacillus* sp. KC201. The optimum temperature and pH for recombinant CGTase activity were 37°C and neutral pH, respectively. The CGTase was produced extracellularly (94%) indicating the signal peptide was functional in *E. coli*. The recombinant CGTase activity was enhanced, approximately 2.5-fold higher than the CGTase from the parent strain and it was beneficial due to it needing less culture time for CGTase production.

ACKNOWLEDGEMENTS

The authors acknowledge the financial support from the Ministry of Higher Education Malaysia (MOHE) under the Fundamental Research Grant Scheme.

REFERENCES

Aristidou AA, Yu P, San KY (1993). Effects of glycine supplement on protein production and release in recombinant *Escherichia coli*., Biotechnol. Lett., 15: 331-336.

Binder F, Huber O, Bock A (1986). Cyclodextrin-glycosyltransferase from *Klebsiella pneumonia* M5a1: cloning, nucleotide sequence and comparison. Gene, 47: 269-277.

Charoensakdi R, Iizuka M, Ito K, Rimphanitchayakit V, Limpaseni T (2007). A recombinant cyclodextrin glycosyltransferase cloned from *Paenibacillus* sp. strain RB01 showed improved catalytic activity in coupling reaction between cyclodextrins and disaccharides. J. Incl. Phenom. Macrocyl. Chem., 57: 53-59.

Ding R, Li Z, Chen S, Wu D, Wu J, Chen J (2010). Enhanced secretion of recombinant α-cyclodextrin glucosyltransferase from *E. coli* by medium additives. Process Biochem., 45: 880-886.

Hashimoto H (2002). Present status of industrial application of cyclodextrins in Japan. J. Incl. Phenom. Macrocycl. Chem., 44: 57-62.

Ish-Horowicz D, Burke J (1981). Rapid and efficient cosmid cloning. Nucleic Acids Res., 9: 2989-2998.

Kaneko T, Hamamoto T, Horikoshi K (1988). Molecular cloning and nucleotide sequence of the cyclomaltodextrin glucanotransferase gene from the alkalophilic *Bacillus* sp. strain no. 38-2. J. Gen. Microbiol., 134: 97-105.

Kaneko T, Kato T, Nakamura N, Horikoshi K (1987). Spectrophotometric determination of cyclization activity of β-cyclodextrin-forming cyclomaltodextrin glucanotransferase. J. Jpn. Soc. Starch Sci., 29: 45-48.

Kim MH, Sohn CB, Oh TK (1998). Cloning and sequencing of a cyclodextrin glycosyltransferase gene from *Brevibacillus brevis* CD162 and its expression in *Escherichia coli*. FEMS Microbiol. Lett., 164: 411-418.

Kim SG, Kweon DH, Lee DH, Park YC, Seo JH (2005). Coexpression of folding accessory proteins for production of active cyclodextrin glycosyltransferase of *Bacillus macerans* in recombinant *Escherichia coli*. Protein Expres. Purif., 41: 426-432.

Kitamoto N, Kimura T, Kito Y, Ohmiya K (1992). Cloning and sequencing of the gene encoding cyclodextrin glucanotransferase from *Bacillus* sp. KC201. J. Ferment. Bioeng., 74: 345-351.

Li Z, Gu Z, Miao W, Du G, Wu J, Chen J (2009). Delayed supplementation of glycine enhances extracellular secretion of the recombinant α-cyclodextrin glycosyltransferase in *Escherichia coli*, Appl. Mircobiol. Biotechnol., 85: 553-561.

Li Z, Li B, Gu Z, Du G, Wu J, Chen J (2010). Extracellular expression and biochemical characterization of α-cyclodextrin glycosyltransferase from *Paenibacillus macerans*. Carbohyd. Res., 345: 886-892.

Makrides SC (1996). Strategies for achieving high-level expression of genes in *Escherichia coli*. Microbiol. Rev., 60: 512-538.

Mergulhao FJM, Summers DK, Monteiro GA (2005). Recombinant protein secretion in *Escherichia coli*. Biotechnol. Adv., 23: 177-202.

Nitschke L, Heeger K, Bender H, Schulz GE (1990). Molecular cloning, nucleotide sequence and expression in *Escherichia coli* of the β-cyclodextrin glycosyltransferase gene from *Bacillus circulans* strain no. 8. Appl. Microbiol. Biotechnol., 33: 542-546.

Ong RM, Goh KM, Mahadi NM, Hassan O, Rahman RNZRA, Illias RM (2008). Cloning, extracellular expression and characterization of a predominant β-CGTase from *Bacillus* sp. G1 in *E. coli*., J. Ind. Microbiol. Biotechnol., 35: 1705-1714.

Park SC, Park HH, Kim SH (1989). A rapid screening method for alkaline β-cyclodextrin glucanotransferase using phenolphthalein-methyl orange containing solid media. Agric. Biol. Chem., 53: 1167-1169.

Rahman K, Illias RM, Hassan O, Mahmood NAN, Rashid NAA (2006). Molecular cloning of a cyclodextrin glucanotransferase gene from alkalophilic *Bacillus* sp. TS1-1 and characterization of the recombinant enzyme. Enzyme Microb. Tech., 39: 74-84.

Sambrook J, Fritsch E (1989). Molecular cloning: A laboratory Manual, Cold Spring Harbor Laboratory Press, New York, U.S.A.

Sin KA, Nakamura A, Kobayashi K, Masaki H, Uozumi T (1991). Cloning and sequencing of a cyclodextrin glucanotransferase gene from *Bacillus ohbensis* and its expression in *Escherichia coli*., Appl. Microbiol. Biot., 35: 600-605.

von Heijne G (1986). A new method for predicting signal sequence cleavage sites. Nucleic Acids Res., 14: 4683-4690.

von Heijne G (1990). The signal peptide. J. Membr. Biol., 115: 195-201.

Yong J, Choi JN, Park SS, Park CS, Park KH, Choi YD (1996). Secretion of heterologous cyclodextrin glycosyltransferase of *Bacillus* sp. E1 from *Escherichia coli*. Biotechnol. Lett., 18: 1223-1228.