CHAPTER I
INTRODUCTION

Life on earth depends on the uptake of energy. While photosynthetic organisms can directly use the sunlight, all other forms of organism obtain their energy from nutrients they take up from the environment. One of the most energy-rich sources of food is starch, a carbohydrate food reserve in plants. Starch is the major source of energy for non-photosynthetic organisms, and composed of amylose and amylopectin. Amylose is a linear polymer consisting of up to 6,000 glucose units with α-1,4-glycosidic bonds and amylopectin consists of short α-1,4-linked linear chains of 10-60 glucose units and α-1,6-linked side chains with 15-45 glucose units.

For humans, animals, fungi and bacteria to degrade starch, a whole range of starch-degrading enzymes with different reaction specificities has evolved in these organisms yielding a wide variety of products (Fig. 1.1). A number of these enzymes find application in the industrial processing of starch, either for modification of starch molecules or the production of specific degradation products. This large and diverse group of starch hydrolytic enzymes has been classified in glycosyl hydrolase family 13, the α-amylase family, and presently covers about twenty different enzyme specificities. These enzymes display specificity for either the α- or β-glycosidic bond, have either a retaining or an inverting mechanism depending on whether they retain or invert the anomeric structure of the substrate, and whose activities classify them as either hydrolases (EC 3.2.1) or transferases (EC 2.4.1) (Janecek,1997; van der Veen *et al.*, 2000a; del Rio *et al.*, 1997; Uitdehaag *et al.*, 2002).

An especially interesting and unique family 13 enzyme is the cyclodextrin glucanotransferase (CGTase; 1,4-α-D-glucan: 1,4-α-D-glucopyranosyltransferase, EC. 2.4.1.19), first observed to be produced by *Bacillus macerans* and has since been found by a variety of bacteria including:

- aerobic mesophilic bacteria: *Pseudomonas* spp., *B. cereus, B. megaterium, B. ohbensis, B. circulans, Paenibacillus macerans, Klebsiella oxytoca, K. pneumoniae, Micrococcus luteus, Brevibacillus brevis*;

- aerobic thermophilic bacteria: *B. stearothermophilus*;
- anaerobic thermophilic bacteria: *Thermoanaerobacterium thermosulfunigens, Thermoanaerobacter* sp.; and
- aerobic halophilic bacteria: *B. halophilus* (Tonkova, 1998).

When these bacteria excrete CGTases into the starch medium, these CGTases convert starch into CDs, which are subsequently absorbed and hydrolyzed by another enzyme cyclodextrinase (CDase; cyclomaltodextrin dextrin-hydrolase (decyclizing), EC 3.2.1.54).

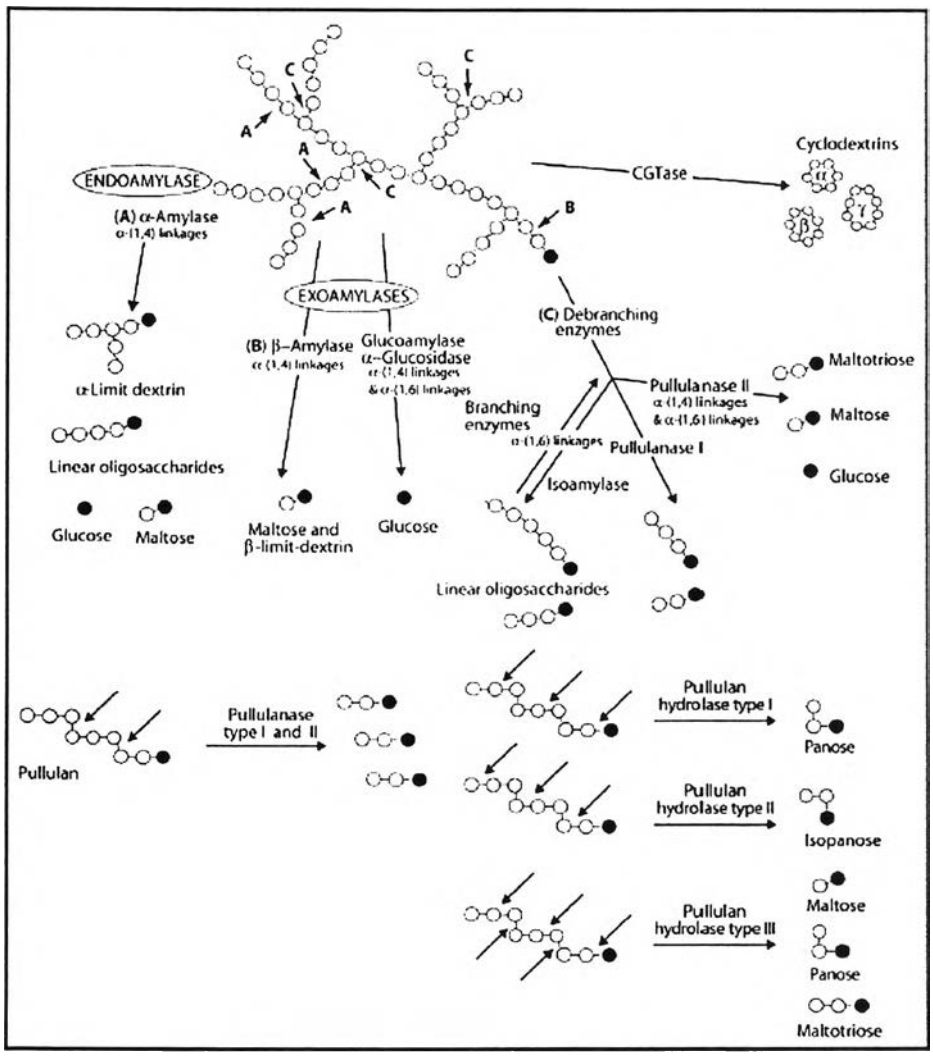

Fig. 1.1. Schematic representation of the action of starch-processing enzymes. (•) Glucose molecule with a reducing end; (o) glucose molecule without a reducing end. (Hobel, 2004).

There are three major forms of cyclodextrin, α-, β-, γ-CDs, in the market. In 1970, the price of β-CD was around 2,000 US$/kg, and it was available only as a rare fine chemical, 25 years later, more than half a dozen companies worldwide are producing cyclodextrins. Their total output is in excess of 1,000 tons/year, global consumption was around 6,000 metric tons in 1998 with an annual growth rate of 15-20% and the price of β-CD is lower depending on quality and delivered quantity. Also α-and γ-CDs, as well as several derivatives, (hydroxypropyl-β-CD and γ-CD, randomly methylated α- and β-CD, maltosyl-β-CD, acetylated CDs, etc.) are produced industrially. The prices of α- and γ-CD always remain higher than that of the β-CD, partly because of their lower yield (higher solubility), and partly because of the lower volume of their production (Szejtli, 1998 and McCoy, 1999). The increasing industrial demand for CDs is reflected in an increased demand for CGTase preparations. These enzymes are therefore studied extensively, leading to the knowledge of their reaction mechanisms, the factors that determine substrate and product specificity, and the stability of the enzyme.

1.1. Application of cyclodextrins

CGTases are capable of converting starch and related substrates to cyclodextrins (CDs). The closed-ring oligosaccharides of the CDs are joined by α-1,4 glycosidic bond, and mainly consisting of 6, 7 or 8 glucose residues, named as α-, β- and γ-cyclodextrins (α-, β- and γ-CDs), respectively (Fig. 1.2a). The glucose residues in the CD ring are arranged in a manner that the secondary hydroxyl groups at C2 and C3 are located on one edge of the ring and the primary hydroxyl groups (C6) on the other edge. The apolar C3 and C5 hydrogens and the ether-like oxygens are at the inside and the hydroxyl groups at the outside of these molecules. This results in a torus-shaped molecule with a hydrophilic outside, which can dissolve in water, and a hydrophobic cavity, which enables cyclodextrin to form an inclusion with a wide variety of poorly water soluble compounds, called guest molecules (Fig. 1.2b and 1.2c). The characteristics of the CDs are summarized in Table 1.1, which provide important parameters for complex formation with hydrophobic compounds or their functional groups (Szejtli, 1998 and van der Veen *et al.*, 2000a).

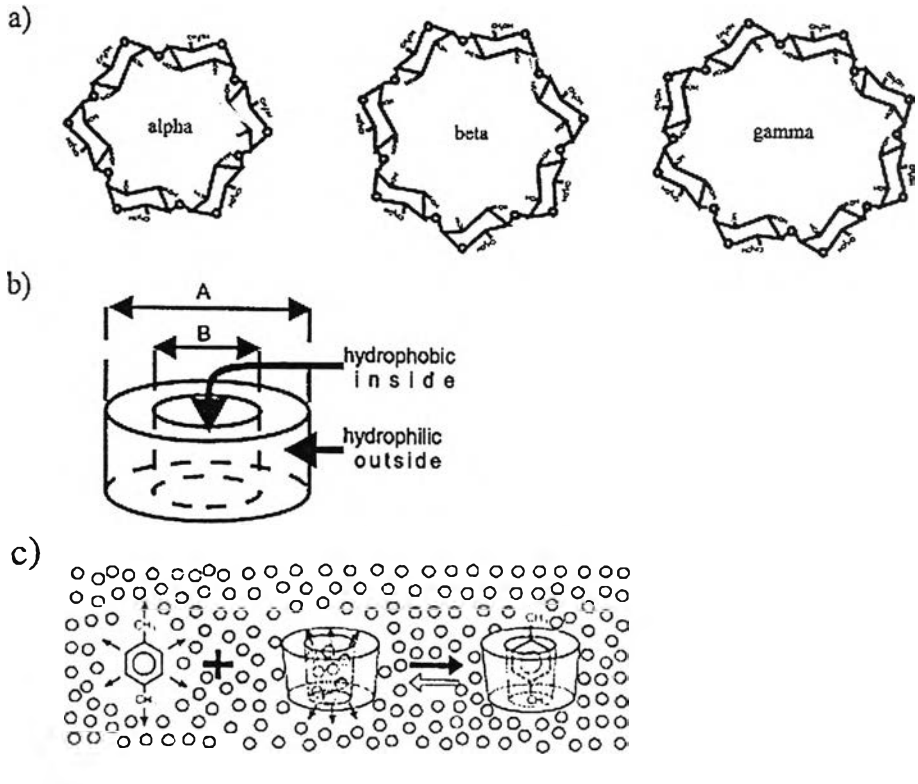

Fig. 1.2. Structures and properties of cyclodextrins. (a) α-, β- and γ-cyclodextrins; (b) three-dimensional model and property of cyclodextrin; (c) formation of the inclusion complex of a cyclodextrin with a guest molecule. p-Xylene is the guest molecule; the small circles represent the water molecules (Szejtli, 1998 and van der Veen et al., 2000a).

Table 1.1. Characteristics of α-, β-, and γ-CDs. (Szejtli, 1998)

Cyclodextrin	α	β	γ
no. of glucose units	6	7	8
Molecular weight (g/mol)	972	1135	1297
solubility in water, g 100 mL^{-1} at room temp.	14.5	1.85	23.2
cavity diameter, Å	4.7-5.3	3 6.0-6.5	7.5-8.3
height of torus, Å	7.9 ± 0.1	7.9 ± 0.1	7.9 ± 0.1
approx volume of cavity, Å	174	262	427
crystal forms (from water)	hexagonal plates	monoclinic parallelograms	quadratic prisms

The formation of inclusion complexes leads to changes in the chemical and physical properties of the encapsulated compounds. This has led to the various

applications of cyclodextrins in analytical chemistry, agriculture, pharmaceutical, food, cosmetics, biotechnology and toiletry.

In analytical chemistry, cyclodextrins are used for the separation of enantiomers on HPLC or GC. In nuclear magnetic resonance (NMR) studies, they can act as chiral shift agents and in circular dichroism as selective agents altering the spectra. In electrochemical chemistry, they can be used to mask contaminating compounds, allowing more accurate determinations. In agriculture, cyclodextrins can be applied to delay seed germination, or control of plant growth and to be additives in pesticides. In food industry, the cyclodextrins have found several applications such as texture-improvement of pastry and meat products, reduction of bitterness, unwanted tastes and odor, stabilization of flavors and emulsions like mayonnaise, and depletion of cholesterol from milk. In pharmaceutical industry, cyclodextrins increase the water solubility of several poorly water-soluble substances, improve of bioavailability and reduction of side effects, facilitate the handling of volatile products, and reduce the dose of the drug administered. Cyclodextrins also improve the stability of substances, *i.e.* increase their resistance to hydrolysis, oxidation, heat, light and metal salts. The inclusions of irritating medicines in cyclodextrins can also protect the gastric mucosa for the oral route, and reduce skin damage for the dermal uses (Penninga, 1996a).

Administered cyclodextrins are quite resistant to starch degrading enzymes, although they can be degraded at very low rates. α-Cyclodextrin is the slowest, and γ-cyclodextrin is the fastest degradable compound. This is due to their differences in size and flexibility. Degradation is not performed by saliva or pancreatic amylases, but by α-amylases from microorganisms in the colon flora. Absorption studies revealed that only 2-4% of cyclodextrins is absorbed in the small intestines and the remainder is degraded and taken up as glucose. This can explain the low toxicity found upon oral administration of cyclodextrins (Bar and Ulitzur, 1994; Duchene, 1988).

1.2. Limitation of commercial CD production by CGTase

In general, the CGTase from *Bacillus macerans* which is used for cyclodextrin production in an industrial scale (Riisgaard, 1990), is poorly active on native starch

because of the well-organized structure of the granules held together by internal hydrogen bonds. The process must run at two different temperatures. In the initial step, liquefied by heat-stable α-amylase treatment at 105 °C is needed to weaken the hydrogen bonds away from starch molecule. After liquefaction, the starch solution is cooled down to make it suitable for incubation at lower temperature (55 °C) required for the CGTase catalyzed production of cyclodextrins. However, this procedure is inefficient in many aspects: first, the α-amylase should be inactivated before the addition of CGTase; second, microbial contamination is possible; third, it is time consuming (Chung *et al.*, 1998) and last, the α-amylase used for liquefaction produces maltodextrins, which can act as acceptor molecules in the coupling reaction of the CGTase, severely reduces the yield of cyclodextrins (Biwer *et al.*, 2002). Thus, the use of thermostable CGTase will allow the liquefaction and cyclization to take place in one step for efficient production of cyclodextrins (Niehaus *et al.*, 1999).

Furthermore, the other disadvantage of cyclodextrin production by CGTase is that all known CGTase enzymes produce a mixture of α-, β- and γ-cyclodextrin at different ratios. The isolation of pure cyclodextrins from this mixture requires a series of additional steps, including precipitation with organic solvents that are potentially hazardous to human consumption. A CGTase, which produces only a single type of cyclodextrin, is therefore of high industrial interest (van der Veen *et al.*, 2000a).

1.3. Three-dimensional structure of CGTase

Primary and three-dimensional structural comparisons between CGTases and α-amylases have revealed both common and distinct features among the enzymes. Both CGTases and α-amylases share three structural domains: A, B and C (Rashid, *et al.*, 2002) (Fig. 1.3 and 1.4). Domain A comprises 300-400 amino acid residues and contains a highly symmetrical fold of eight parallel β-strands arranged in a barrel encircled by eight α-helices. This $(\beta/\alpha)_8$ or TIM barrel catalytic domain (Fig. 1.5) is present in all enzymes of the α-amylase family. Several proline and glycine residues flanking the loops connecting the β-strands and α-helices are highly conserved in these enzymes (Janecek, 1997 and Janecek and Sevcik, 1999). The catalytic residues, Glu257, Asp229 and Asp328 (*B. circulans* 251 CGTase numbering), and substrate

binding residues are located in the loops at the C-termini of β-strands in domain A. Moreover, on the basis of the results from many studies on X-ray crystallographic structures of CGTases with their inhibitors, substrates or products, it has been proposed that the active center of CGTase has a tandem subsite architecture in the substrate binding groove, and that it comprises at least nine sugar-binding subsites, designated from the nonreducing end to the reducing end as -7 to +2 (Fig. 1.6) (van der Veen *et al.* 2000a and Strokopytov *et al.*, 1996).

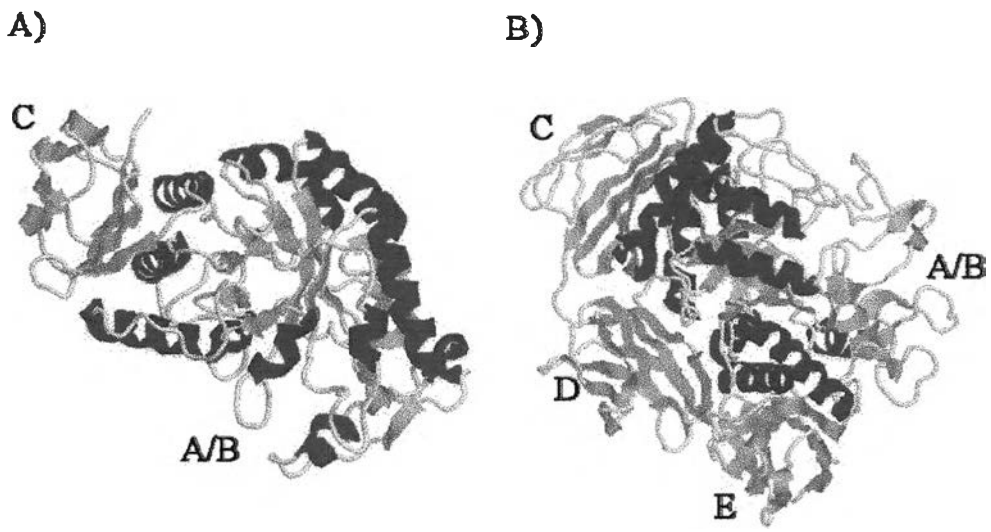

Fig. 1.3. Comparison of the three dimensional structures of α-amylase and CGTase. A) α-amylase from *Bacillus subtilis* (PDB 1BAG); B) CGTase from *Bacillus circulans* strain 251 (PDB 1CDG).

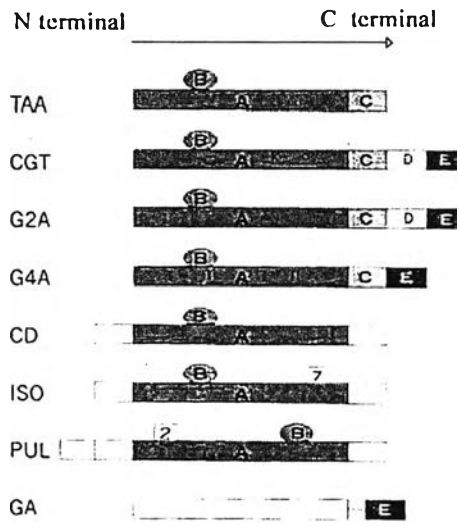

Fig. 1.4. Domain level organization of starch-degradating enzymes. CGT, CGTase from *B. circulans*; G2A, maltogenic α-amylase from *Bacillus stearothermophilus*; G4A, maltotetraose forming α-amylase from *Pseudomonas stutzeri*; TAA, α-amylase from *A. oryzae* (Taka-amylase A); CD, cyclodextrinase from *Klebsiella oxytoca*; ISO, isoamylase from *P. amyloderamosa*; PUL, pullulanase from *K. aerogenes*; GA, glucoamylase (family 15 of glycosyl hydrolases) from *Aspergillus niger*. (van der Veen *et al.*, 2000a).

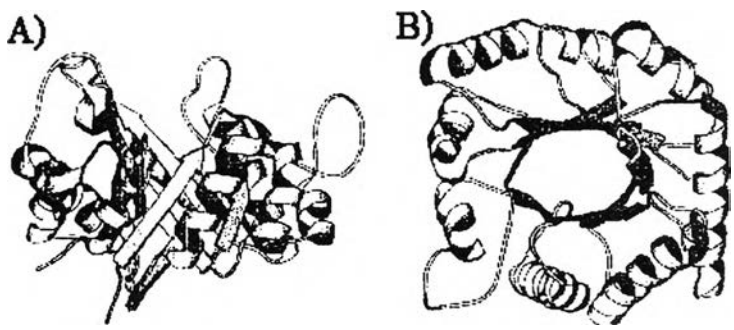

Fig. 1.5. The parallel $(\beta/\alpha)_8$ or TIM barrel. A) Side view, B) End view (Janecek, 1997).

Domain B is an extended loop region inserted between β-strand 3 and α-helix strand 3 of domain A. It is rather large and is regarded as a separate structural domain. This domain B consists of 44-133 amino acid residues and contributes to substrate binding by providing several amino acid side chain alongside a long groove on the surface of the enzyme that interact with the substrate (Uitdehaag *et al.*, 2002). Domain

C is approximately 100 amino acids long and has an antiparallel β-sandwich fold. This domain C of the CGTase from *B. circulans* strain 251 contains one of the three-maltose binding sites (Lawson *et al.*, 1994) observed from the maltose dependent crystals. This maltose binding site is found to be involved in raw starch binding (Penninga *et al.*, 1996a), suggesting a role of the domain C in substrate binding. Others suggest that this domain is involved in bond specificity, particularly in enzymes hydrolyzing or forming the α-(1,6) bonds, e.g. pullulanase, isoamylase and branching enzymes.

CGTase have two additional domains not found in α-amylases, domains D and E, which have the β-sheet structures (Fig. 1.7). The domain D, consisting of approximately 90 amino acids with an immunoglobulin fold, is almost exclusively found in CGTases and its function is unknown. The domain E consists of approximately 110 amino acids and is found to be responsible for substrate binding. Two maltose-binding sites have been identified in domain E of CGTase from *B. circulans* 251, and evidences that this domain contributes to raw-starch binding has also been obtained from other CGTases (Svensson *et al.*, 1989).

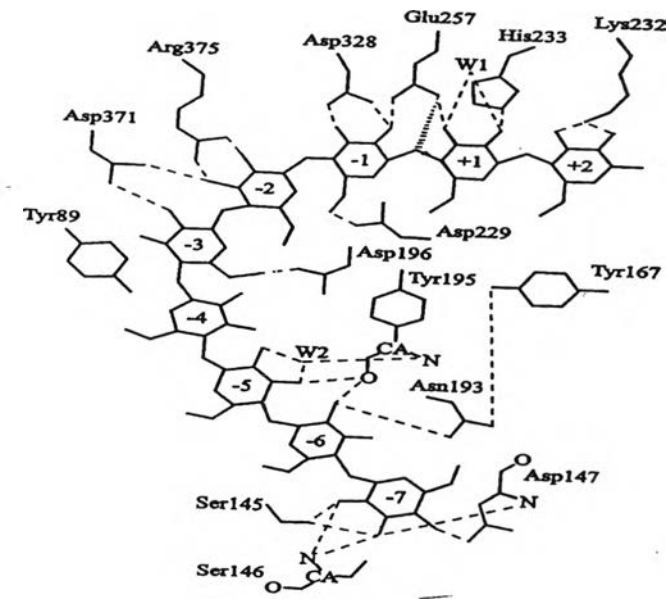

Fig. 1.6. Schematic representation of the hydrogen bonds between the *B. circulans* strain 251 CGTase and a maltononaose inhibitor bound at each subsite of the active site. W1 and W2 indtcate water and CA indicates alpha carbon (Source: van der Veen *et al.*, 2000a).

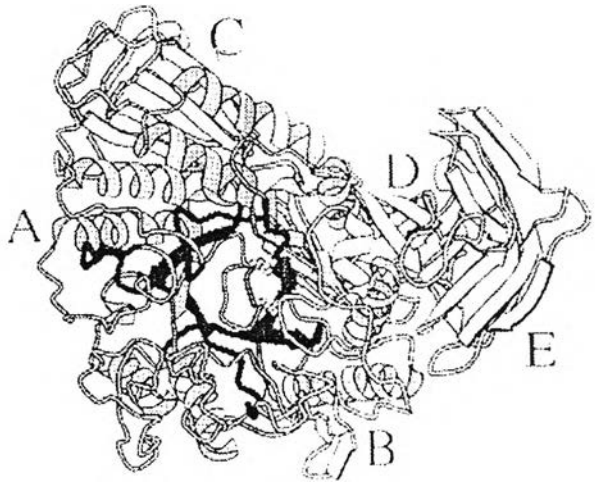

Fig. 1.7. The structure of a representive CGTase from *Bacillus circulans* 251 (Uitdehaag *et al.*, 2002).

1.4. CGTase reaction and its catalytic mechanism

The CGTase catalyzes four distinctly different reactions, depending on the nature of the acceptor molecule: three transglycosylation reactions (disproportionation, cyclization and coupling) and a hydrolysis reaction (Fig. 1.8).

i) Disproportionation is an intermolecular transglycosylation reaction. In this reaction a linear malto-oligosaccharide is cleaved and one of the products is transferred to another linear acceptor substrate. This reaction yields a mixture of smaller and longer oligosaccharides.

ii) Cyclization is an intramolecular transglycosylation reaction, in which the non-reducing end of the same oligosaccharide is transferred to the reducing end of the same oligosaccharide chain. The products of this reaction are cyclodextrins.

iii) Coupling is the reverse of the cyclization reaction in which the opening of a cyclodextrin ring is followed by the transferring of the oligosaccharide to a linear saccharide acceptor. This reaction produces a longer chain oligosaccharide.

iv) Hydrolysis is a reaction in which part of a linear oligosaccharide is transferred to water. This reaction is much weaker than the transglycosylation reactions.

Like the α-amylase family, CGTase has the catalytic mechanism as the α-retaining double displacement mechanism. This mechanism involves two catalytic residues, an acid/base catalyst (Glu257 in CGTase) and a nucleophile (Asp229) (Fig. 1.9). When substrate has bound, Glu 257 (as acid) donates a proton to the scissile glycosidic bond oxygen. The scissile bond between subsite +1 and -1 is then cleaved, during which reaction the -1 sugar goes through an oxocarbonium ion-like transition state. In the subsequent reaction step, this oxocarbonium ion-like transition state collapse into a stable covalent glycosyl enzyme reaction intermediate which is β-glycosidically linked to the nucleophile Asp229. The leaving group (at subsite +1) is exchanged for a new compound, the acceptor, which gets activated by Glu257 (now a base). This activated acceptor attacks the covalent intermediate, and via another oxocarbonium ion-like transition state, the final α-1,4 glycosidic product bond is formed (Uitdehaag *et al.*, 2002). However, there is Asp328, catalytic residue, stabilizes the substrate binding and elevates the pK_a of Glu257 that exists only when no substrate or inhibitor is present (Knegtel *et al.*, 1995).

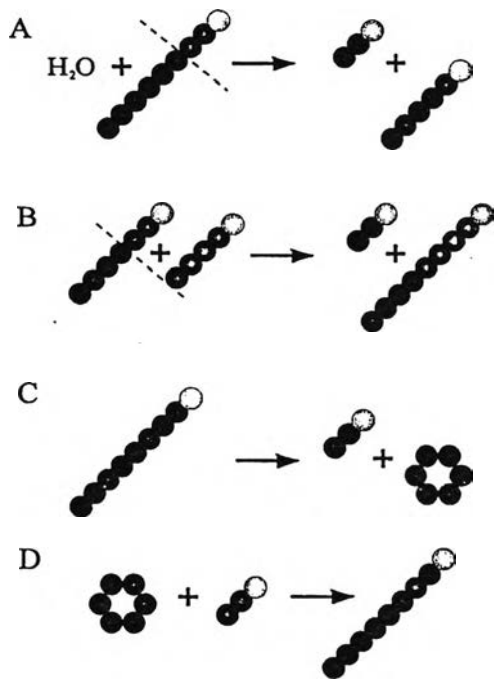

Fig. 1.8. Schematic representation of the CGTase-catalysed reactions. The circles represent glucose residues; the white circles indicate the reducing end sugars. (A) hydrolysis; (B) disproportionation; (C) cyclization; (D) coupling (van der Veen *et al.*, 2000a).

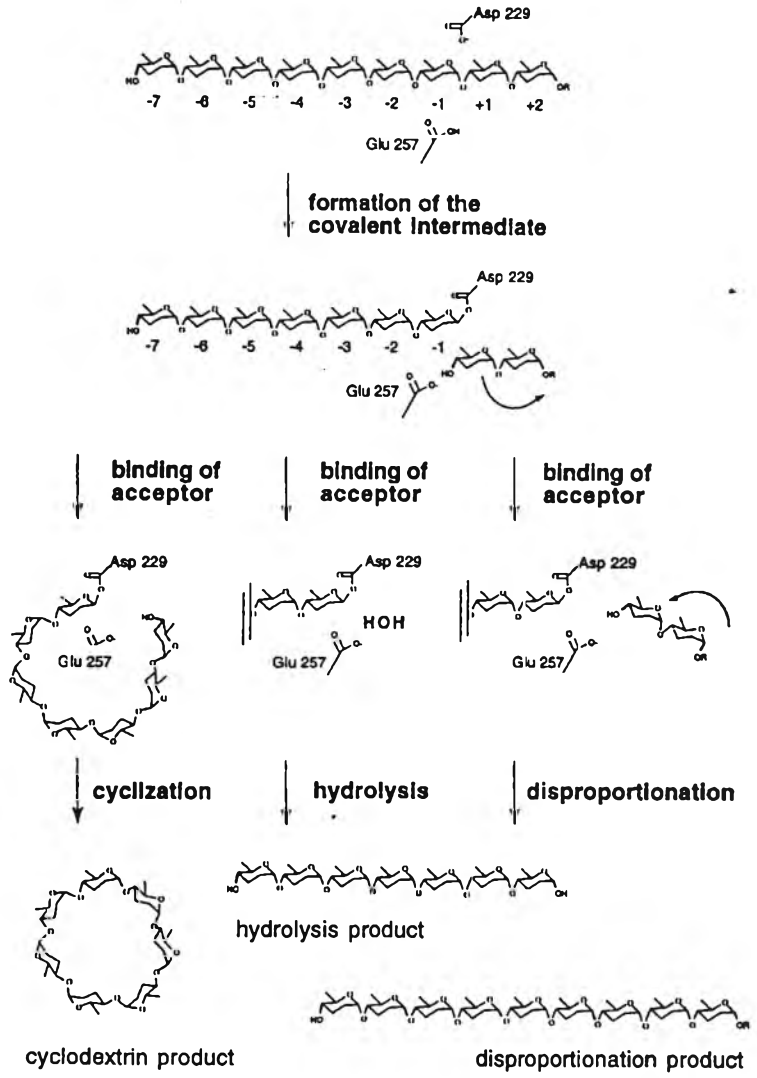

Fig. 1.9. The catalytic reaction of CGTase, which involved cyclization, disproportionation and hydrolysis (Uitdehaag *et al.*, 2000).

1.5. Thermostability of CGTases

Understanding the molecular basis of protein thermal stability is an important fundamental problem with obvious practical applications. One approach to this problem involves the comparison of structures and sequences of protein from mesophilic and thermophilic organisms (Chakravarty and Varadarajan, 2000) which could provide the important clues to themostability of proteins.

In general, CGTases from many microorganisms have optimum temperature in the range of 40-60 °C. In contrast, the thermophilic enzymes produced from some anaerobic and aerobic thermophilic bacteria were characterized with high thermostability and a temperature optimum in the range of 80-90 °C (Table 1.2) (Tonkova, 1998). The three-dimensional structures of two thermostable CGTases from *Thermoanaerobacterium thermosilfurigenes* EM1 and *Bacillus stearothermophilus* are known (Fig. 1.10). Structural studies of the CGTase revealed that the NH_2-terminal domains are similar to those of α-amylase. The COOH-terminal domains, however, are unique to the CGTase. Hence, the COOH-terminal domains are considered to be involved in cyclization of linear maltooligosaccharide. In thermostable CGTase which is functional in extremely thermal environment, the NH_2-terminal and the COOH-terminal domains seem to be suitably arranged for efficient CD production (Yamamoto *et.al.*, 1999) and only NH_2-terminal of the enzyme was important for heat stability (Kaneko *et.al*, 1989).

The availability of the crystal structures at 2.0 Å resolution of the CGTases of the mesophilic *B. circulans* 251 (PDB 1CGD) (Lawson *et al.*, 1994) and 8 (PDB 1CGT) (Klein and Schulz, 1991) (Fig. 1.11), at 2.5 Å resolution of the thermophilic *B. stearothermophilus* (PDB 1CYG) (Kubota *et al.*, 1991 cited in Knegtel *et al.*, 1995) and at 2.3 Å of the *Thermoanaerobacterium thermosilfurigenes* EM1 (PDB 1CIU) (Knegtel *et al.*, 1996) allows for detailed amino acid sequence and structural comparison between the mesophilic and thermostable CGTase molecules. Significant differences are observed in a few regions particularly at the loop regions near the active site (Fig. 1.12). The amino acid residues in these regions possibly contribute with hydrogen bonds and apolar contacts to the stabilization of the thermostable enzyme.

Recently, the structural studies of thermophilic enzymes have suggested that an increase in surface salt-bridge could be responsible for structural stability of proteins at elevated temperatures (Chan *et al.*, 1995 and Day *et al.*, 1992). The mesophilic CGTases from *B. circulans* 251 and *B. circulans* 8 have a total of 27 and 16 salt-bridges, respectively (using a 3.5 Å distance limit for the Arg, Lys, Asp, and Glu charged side-chain atoms), while the *T. thermosilfurigenes* EM1 CGTase has 20. By comparing the location of salt-bridges, Leemhuis *et al.* (2004) indicated that

certain salt-bridges were absent in CGTase from *B. circulans* 251 and carried out a mutagenesis experiment to create the salt-bridge. They found that one of the mutant had slightly increased thermostability. The fact that CGTase from *B. circulans* 251 has more salt-bridge than that from *T. thermosilfurigenes* EM1 and the creation of salt-bridge in CGTase has little effect on thermostability indicates that salt-bridge alone is not the major determinant for thermostability.

Table 1.2. Some mesophilic and thermophilic CGTase characteristics.

Microorganism	Optimum Temp. (°C)	Product	References
Bacillus circulans 8	ND	β-CD	Nitschke *et al.*, 1990
Bacillus circulans 251	ND	β-CD	Lawson *et al.*, 1994
Bacillus stearothermophilus No 2	80	α/β-CD	Fujiwara *et al.*, 1992
Bacillus stearothermophilus ET1	80	β-CD	Chung *et al.*, 1998
Thermoanaerobacterium thermosilfurigenes EM1	80-85	β-CD	Wind *et al.*, 1995
Thermoanaerobacter sp. ATCC 53627	90	α/β-CD	van der Veen *et al.*, 2000a

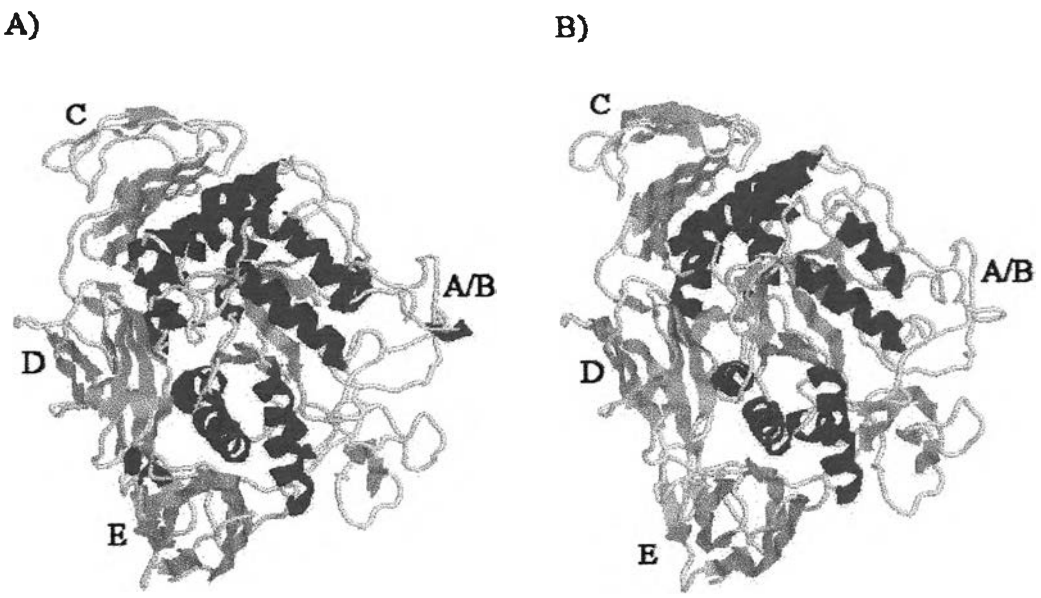

Fig. 1.10. The three dimensional structures of the two thermostable CGTases. (A), *Thermoanaerobacterium thermosilfurigenes* EM1 (PDB 1CIU) and (B), *Bacillus stearothermophilus* (PDB 1CYG).

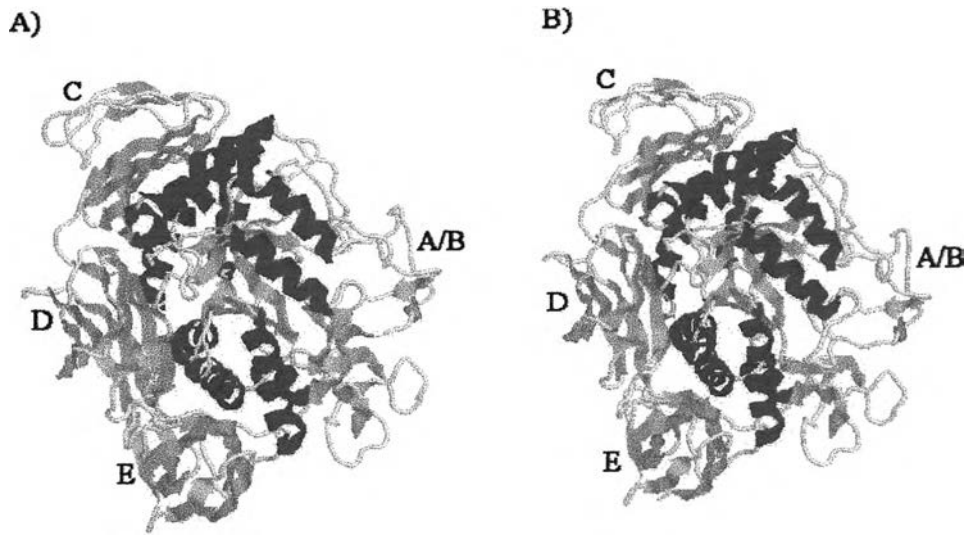

Fig. 1.11. The three dimensional structures of the two mesophilic CGTases. (A), *B. circulans* 251 (PDB 1CDG) and (B), *B. circulans* 8 (PDB 1CGT).

```
BC251           -----MKKFLKSTAALAL--GLSLTFGLFSPAQAAPDTSVSNKQNFSTDVIYQIFTDRFS 53
BC8             MFQMAKRAFLSTTLTLGLLAGSALPFLPASAVYADPDTAVTNKQSFSTDVIYQVFTDRFL 60
TthEM1          -----MKKTFKLILVLML--SLTLVFGLTAPIQAASDTAVSNVVNYSTDVIYQIVTDRFV 53
BST2            -----MRRWLSLVLSMSFVFSAIFIVSDTQKVTVEAAGNLN-KVNFTSDVVYQIVVDRFV 54
                      :  ::.   : :  .  : .       . .   :.   .:::**:**:..***

BC251           DGNPANNPTGAAFDGTCTNLRLYCGGDWQGIINKINDGYLTGMGVTAIWISQPVENIYSI 113
BC8             DGNPSNNPTGAAYDATCSNLKLYCGGDWQGLINKINDNYFSDLGVTALWISQPVENIFAT 120
TthEM1          DGNTSNNPTGDLYDPTHTSLKKYFGGDWQGIINKINDGYLTGMGVTAIWISQPVENIYAV 113
BST2            DGNTSNNPSGALFSSGCTNLRKYCGGDWQGIINKINDGYLTDMGVTAIWISQPVENVFSV 114
                ***.:***:*  :.   :.*: * ******:******.*::.:****:*********:::

BC251           INYSGVN-NTAYHGYWARDFKKTNPAYGTIADFQNLIAAAHAKNIKVIIDFAPNHTSPAS 172
BC8             INYSGVT-NTAYHGYWARDFKKTNPYFGTMADFQNLITTAHAKGIKIVIDFAPNHTSPAM 179
TthEM1          LPDSTFGGSTSYHGYWARDFKRTNPYFGSFTDFQNLINTAHAHNIKVIIDFAPNHTSPAS 173
BST2            MNDASGS--ASYHGYWARDFKKPNPFFGTLSDFQRLVDAAHAKGIKVIIDFAPNHTSPAS 172
                :  :     ::**********:.** :*:::***.*: :***:.**::**********

BC251           SDQPSFAENGRLYDNGTLLGGYTNDTQNLFHHNGGTDFSTTENGIYKNLYDLADLNHNNS 232
BC8             ETDTSFAENGRLYDNGTLVGGYTNDTNGYFHHNGGSDFSSLENGIYKNLYDLADFNHNNA 239
TthEM1          ETDPTYAENGRLYDNGTLLGGYTNDTNGYFHHYGGTDFSSYEDGIYRNLFDLADLNQQNS 233
BST2            ETNPSYMENGRLYDNGTLLGGYTNDANMYFHHNGGTTFSSLEDGIYRNLFDLADLNHQNP 232
                . :.:: **********:******::  *** **: **: *:***:**:****:*::*.

BC251           TVDVYLKDAIKMWLDLGIDGIRMDAVKHMPFGWQKSFMAAVNNYKPVFTFGEWFLGVNEV 292
BC8             TIDKYFKDAIKLWLDMGVDGIRVDAVKHMPLGWQKSWMSSIYAHKPVFTFGEWFLGSAAS 299
TthEM1          TIDSYLKSAIKVWLDMGIDGIRLDAVKHMPFGWQKNFMDSILSYRPVFTFGEWFLGTNEI 293
BST2            VIDRYLKDAVKMWIDMGIDGIRMDAVKHMPFGWQKSLMDEIDNYRPVFTFGEWFLSENEV 292
                .:* *:*.*:*:*:*:*:****:*******:****. *  :  ::**********.
```

Fig. 1.12. Alignment of the amino acid sequences of the mesophilic and thermophilic CGTases. BC251, *B. circulans* 251 (PDB 1CDG); BC8, *B. circulans* 8 (PDB 1CGT); TthEM1, *Thermoanaerobacterium thermosilfurigenes* EM1 (PDB 1CIU) and BST2, *B. stearothermophilus* (PDB 1CYG).

```
BC251           SPENHKFANESGMSLLDFRFAQKVRQVFRDNTDNMYGLKAMLEGSAADYAQVDDQVTFID 352
BC8             DADNTDFANKSGMSLLDFRFNSAVRNVFRDNTSNMYALDSMINSTATDYNQVNDQVTFID 359
TthEM1          DVNNTYFANESGMSLLDFRFSQKVRQVFRDNTDTMYGLDSMIQSTASDYNFINDMVTFID 353
BST2            DANNHYFANESGMSLLDFRFGQKLRQVLRNNSDNWYGFNQMIQDTASAYDEVLDQVTFID 352
                . :*  ***:********** . :*:*:*:*:.. *.:. *::.:*: *  : * *****

BC251           NHDMERFHASNANRRKLEQALAFTLTSRGVPAIYYGTEQYMSGGTDPDNRARIPSFSTST 412
BC8             NHDMDRFKTSAVNNRRLEQALAFTLTSRGVPAIYYGTEQYLTGNGDPDNRAKMPSFSKST 419
TthEM1          NHDMDRFYNGGS-TRPVEQALAFTLTSRGVPAIYYGTEQYMTGNGDPYNRAMMTSFNTST 412
BST2            NHDMDRFMIDGGDPRKVDMALAVLLTSRGVPNIYYGTEQYMTGNGDPNNRKMMSSFNKNT 412
                ****:**  .    * :: ***. ******* ********::*. ** **  :.**...*

BC251           TAYQVIQKLAPLRKCNPAIAYGSTQERWINNDVLIYERKFGSNVAVVAVNRNLNAPASIS 472
BC8             TAFNVISKLAPLRKSNPAIAYGSTQQRWINNDVYVYERKFGKSVAVVAVNRNLSTSASIT 479
TthEM1          TAYNVIKKLAPLRKSNPAIAYGTTQQRWINNDVYIYERKFGNNVALVAINRNLSTSYNIT 472
BST2            RAYQVIQKLSSLRRNNPALAYGDTEQRWINGDVYVYERQFGKDVVLVAVNRSSSSNYSIT 472
                *::**.**:.**: ***:*** *::****.** :***:**..*.:**:**. .:  .*:

BC251           GLVTSLPQGSYNDVLGGLLNGNTLSVGSGGAASNFTLAAGGTAVWQYTAATATPTIGHVG 532
BC8             GLSTSLPTGSYTDVLGGVLNGNNITS-TNGSINNFTLAAGATAVWQYTTAETTPTIGHVG 538
TthEM1          GLYTALPAGTYTDVLGGLLNGNSISVASDGSVTPFTLSAGEVAVWQYVSSSNSPLIGHVG 532
BST2            GLFTALPAGTYTDQLGGLLDGNTIQVGSNGSVNAFDLGPGEVGVWAYSATESTPIIGHVG 532
                ** *:** *:*.* ***:*:**.:   :.*: . * *..* ..** * ::  :* *****

BC251           PMMAKPGVTITIDGRGFGSSKGTVYFGTTAVSGADITSWEDTQIKVKIPAVAGGNYNIKV 592
BC8             PVMGKPGNVVTIDGRGFGSTKGTVYFGTTAVTGAAITSWEDTQIKVTIPSVAAGNYAVKV 598
TthEM1          PTMTKAGQTITIDGRGFGTTSGQVLFGSTAGT---IVSWDDTEVKVKVPSVTPGKYNISL 589
BST2            PMMGQVGHQVTIDGEGFGTNTGTVKFGTTAAN---VVSWSNNQIVVAVPNVSPGKYNITV 589
                * * : *  :****.***:..* * **:** .   :.**.:.:: * :* *: *:* :.:

BC251           ANAAGTASNVYDNFEVLSGDQVSVRFVVNNATTALGQNVYLTGSVSELGNWDP-AKAIGP 651
BC8             A-ASGVNSNAYNNFTILTGDQVTVRFVVNNASTTLGQNLYLTGNVAELGNWSTGSTAIGP 657
TthEM1          KTSSGATSNTYNNINILTGNQICVRFVVNNASTVYGENVYLTGNVAELGNWDT-SKAIGP 648
BST2            QSSSGQTSAAYDNFEVLTNDQVSVRFVVNNATTNLGQNIYIVGNVYELGNWDT-SKAIGP 648
                  ::*  * .*:*: :*:.:*: ********:*  *:*:*:.*.* *****.. :.****

BC251           MYNQVVYQYPNWYYDVSVPAGKTIEFKFLKKQGS-TVTWEGGSNHTFTAPSSGTATINVN 710
BC8             AFNQVIHQYPTWYYDVSVPAGKQLEFKFFKKNGS-TITWESGSNHTFTTPASGTATVTVN 716
TthEM1          MFNQVVYQYPTWYYDVSVPAGTTIQFKFIKKNGN-TITWEGGSNHTYTVPSSSTGTVIVN 707
BST2            MFNQVVYSYPTWYIDVSVPEGKTIEFKFIKKDSQGNVTWESGSNHVYTTPTNTTGKIIVD 708
                :***::.**.** ***** *. ::***:**:.. .:***.****.:*.*:. *.:: *:

BC251           WQP 713
BC8             WQ- 718
TthEM1          WQQ 710
BST2            WQN 711
                **
```

Fig. 1.12. (continue) Alignment of amino acid sequences of mesophilic and thermophilic CGTases. BC251, *B. circulans* 251 (PDB 1CDG); BC8, *B. circulans* 8 (PDB 1CGT); TthEM1, *Thermoanaerobacterium thermosilfurigenes* EM1 (PDB 1CIU) and BST2, *B. stearothermophilus* (PDB 1CYG).

1.6. Protein engineering for CGTase thermostability

Engineering proteins for thermostability has been vigorously studied in the biotechnical research areas because of the useful application in the industrial scale. Protein from thermophilic organisms offer good model systems with which to address

the origins of thermostability (Strop and Mayo, 2000). Several reasons have been attributed to the greater stability of the thermophilic proteins. Among the most prominent ones are greater core hydrophobicity, deletion or shortening of loops, better packing, smaller and less numerous cavities, increased surface area buried upon oligomerization, amino acid substitutions within and outside the secondary structures, increased occurrence of proline residues, decreased occurrence of thermolabile residues, increased helical content, increased polar surface area, increased hydrogen bonding and salt bridges (Kumar *et al.*, 2000).

Protein engineering, usually perform through site-directed mutagenesis is the favorite mode for experimental analysis and stability enhancement. There are many site-directed mutagenesis protocols, which change nucleotide sequence by using specific primer introduced into the part of the interested gene. For CGTases, there are numerous studies to investigate the product specificity of the enzymes while only a few studies are on the thermostability of CGTase. Thus far, there has been no success in creating the thermostable CGTase from the mesophilic one.

1.7. Scope of this study

As mentioned above, thermostable CGTases are beneficial for the production of CDs. Several attempts have been made to enhance the stability of CGTases. There are two possible ways to obtain the thermostable CGTases. The first one is the screening for thermophilic microorganisms that produce the thermostable enzymes. The second is to mutagenize the existing CGTase to create an enzyme with higher thermostability. In this research, the second approach is used. The CGTase gene from *B. circulans* A11 (Rimphanitchayakit *et al.*, 2005) is used. The gene encodes a mesophilic β-CGTase of 713 amino acid residues, including a signal peptide. The molecular weight is approximately 72 kDa. The optimal temperature for CD production is 60-65 °C. The changes in product specificity of this enzyme has been studied (Rimphanitchayakit *et al.*, 2005; Kerdsin, 2003; Chotechuang, 2003). To create the thermostable version of the CGTase from *B. circulans* A11, the basic differences in the amino acid sequence from the thermostable enzymes are studied. The locations of the amino acid sequence differences on the three-dimensional

structure are taken into consideration. Mutagenesis of the CGTase is carried out using the unique site elimination (USE) mutagenesis. The activity of the mutant CGTases under various temperatures is reported.