CHAPTER I

INTRODUCTION

Plants produce starch, a high molecular weight polymer of glucose, for storage as a carbon and energy source. These starch molecules are mostly found in seeds (e.g., wheat) or roots (e.g., potato) in the form of granules consisting of two types of glucan polymers, highly branched amylopectin and linear amylose, in which the glucose units are linked by either $\alpha(1,6)$ and $\alpha(1,4)$ glycosidic bonds (van der Veen *et al.*, 2000c).

To obtain glucose from starch, different enzymes are used, most of which are grouped in the α-amylase family (Jespersen, *et al.*,1993). This group includes enzymes that display specificity for either of the two types of 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 4.2.1) (del Rio *et al.*, 1997).

A whole range of starch-degrading enzymes with different reaction mechanisms gives rise to a wide variety of products (Fig.1) (van der Veen *et al.*, 2000c). Enzymatic degradation of starch by α-glucosidase or α-amylase normally produces molecules as glucose, maltose or maltotriose, besides a long series of straight or branched-chain oligomers, known as dextrins. The degradation of starch that produces dextrins is more accurately described as hydrolysis, as the primary product of the broken glycoside bond with one molecule of water.

If the starch is treated instead by the enzyme cyclodextrin gluconotransferase (CGTase; 1,4-α-D-glucan: 1,4-α-D-glucopyranosyltransferase, EC. 2.4.1.19), the primary reaction product undergoes an intramolecular reaction without water, to produce the α-1,4-cyclic compounds known as cycledextrins. CGTase is produced by a variety of bacteria including

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

- anaerobic thermophilic bacteria: *Thermoanaerobacterium thermosulfunigens, Thermoanaerobacter* sp.
- aerobic halophilic bacteria: *B. halophilus*

When these bacteria excrete CGTases into the starch medium, these CGTases convert starch into cyclodextrins, which are subsequently transported and degraded by the action of another enzyme cyclodextrinase located at the cytosolic side (Fig. 2). There are two possible explanations for the existence of this complicated system:

i) The organism may build up an external storage form of glucose, not accessible to most other organisms because they are not able to metabolize cyclodextrins (Penninga, 1996b).

ii) Cyclodextrins are used to form inclusion complexes with toxic compounds in the environment or with compounds needed for growth (Aeckersberg, 1991).
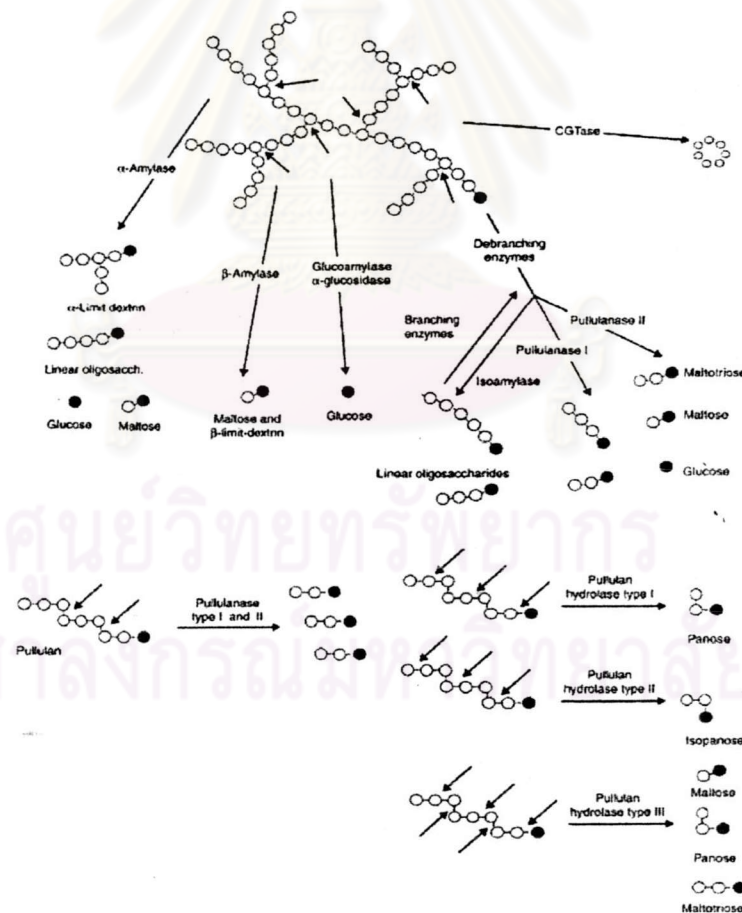

Figure 1. Schematic representation of the action of starch-processing enzymes. Black circles indicate the reducing sugars (Source: Bertodo and Antrnikian, 2002).
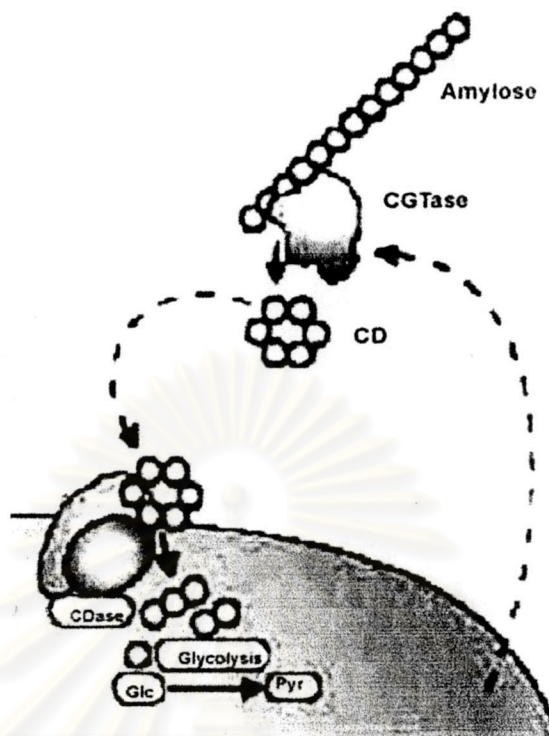
Figure 2.   Schematic representation of the location and action of CGTase and cyclodextrinase (CDase) (Saha and Zeikus, 1990).

1. Application of cyclodextrins

CGTases are able to convert starch to cyclodextrins via the cyclization reaction. The cyclodextrins are cyclic $\alpha(1,4)$ linked oligosaccharides mainly consisting of 6, 7 or 8 glucose residues, designated as α-, β- and γ-cyclodextrins (α-, β- and γ-CDs), respectively (Fig. 3a). The glucose residues in the cyclodextrin 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 (Fig. 3b). 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 cyclodextrins to form inclusion complexes with a wide variety of guest molecules (Fig. 3c). Their three dimensional forms and sizes provide unlimited spaces for forming complexes with various hydrophobic compounds or functional groups (Table 1) (van der Veen *et al.*, 2000c).
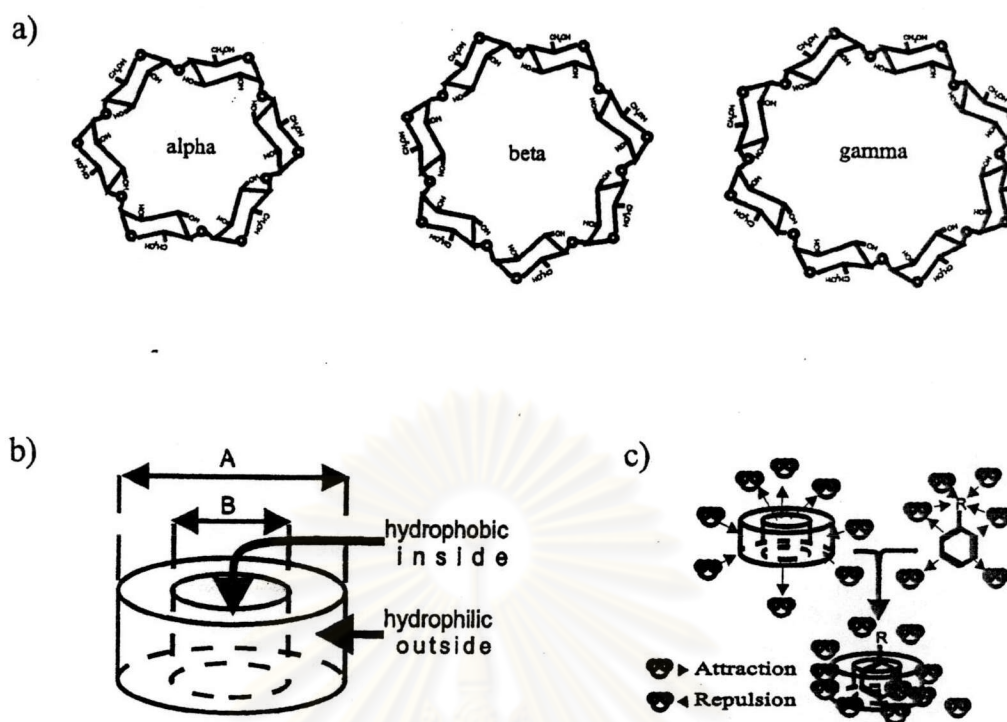
Figure 3. Structures and properties of cyclodextrins. (a) α-, β- and γ-cyclodextrins; (b) three-dimensional model and property of cyclodextrin; (c) formation of the inclusion complex of a cyclodextrin with a guest molecule (Source: van der Veen *et al.*, 2000c).

Table 1. Properties of cyclodextrins

Cyclodextrin	α-CD	β-CD	γ-CD
Number of glucose units	6	7	8
Molecular weight (g/mol)	972	1135	1297
Solubility in water at 25 °C (%W/V)	14.5	1.85	23.2
Outer diameter (Å)	14.6	15.4	17.5
Inner diameter (Å)	4.7 – 5.3	6.0 – 6.5	7.5 – 8.3
Height of torus (Å)	7.9	7.9	7.9
Approximate cavity volume (Å^3)	174	262	427

(source: Szejtli, 1982 and Ukema and Irie, 1987)

The driving force of inclusion complex formation is the entropic effect of displacement of water molecules from the hydrophobic environment in the cavity. This is probably combined with the fact that the water causes strain on the cyclodextrin ring, which is released after complexation, producing a more stable and lower energy state (van der Veen *et al.*, 2000c). 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 using 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. In food industry, the cyclodextrins have found several applications such as texture-improvement of pastry and meat products, reduction of bitterness, ill smell and taste, 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 bioavailability, 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, 1996b).

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).

2. Limitation of CD production by CGTase

There are several disadvantages of CD production by CGTase. First, the CGTase which is used for the commercial production of cyclodextrins (Riisgaard, 1990) is poorly active on native starch due to the well organized structure of the granules held together by internal hydrogen bonds. Liquefied by α-amylase treatment

or heating in water to weaken the hydrogen bonds away from starch molecules is needed. The α-amylase used for liquefaction produces maltodextrins, which can act as acceptor molecules in the coupling reaction of the CGTase, severely reducing the yield of cyclodextrins. However, many thermostable CGTases have been isolated and characterized from the thermophilic bacteria. These CGTases are active and stable at high temperature, and are able to solubilize starch, thereby eliminate the need for α-amylase pretreatment(Starnes *et al.*, 1991).

Second, all wild type CGTases usually produce a mixture of α-, β- and γ-cyclodextrin and are sensitive to product inhibition (van der Veen *et al.*, 2000c, d). Two different industrial approaches are used to purify the produced cyclodextrins, solvent and non-solvent processes. The solvent process is the selective crystallization of complexes of cyclodextrins with organic solvents. The toluene and cyclohexane are commercially used for the complexation and selective precipitation of β-cyclodextrin. For α-cyclodextrin, 1-decanol can be used, but this compound is difficult to remove from aqueous solution because of its high boiling point (229 °C). For γ-cyclodextrin, cyclododecanone can be used but this solvent is very expensive. Further, disadvantages of the use of organic solvents are their toxicity, which limits the applications to human consumption, their flammability and the need for a solvent recovery process (van der Veen *et al.*, 2000c).

The non-solvent process was first developed for β-cyclodextrin production. Due to its low solubility, β-cyclodextrin can be easily purified by crystallization. The purification of α- and γ-cyclodextrin is achieved via complex and expensive chromatography with low yields and a wide range of by-products. Compared to the solvent process, a lower yield of the enzyme reaction, a more complex purification process, ineffective crystallization, a higher energy demand and a large number of by-products are the major disadvantages of the non-solvent process (Biwer *et al.*, 2002). Clearly, the availability of CGTase enzymes capable of producing an increased ratio of one particular type of cyclodextrin and with reduced product inhibition would help to avoid the disadvantages described above.

3. The three-dimensional structure of CGTase

The three-dimensional structure of α-amylases (Fujimoto *et al.*,1998) and CGTase (Lawson *et al.*, 1994 and Kanai *et al.*,2001) are quite similar. The α-amylases generally consist of three structural domamins, A, B and C, while CGTases show a similar domain organization with two additional domains, D and E (Fig. 4 and 5). Domain A consists of 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 so-called (α/β)$_8$ or TIM barrel (Fig.6) catalytic domain 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 found to be highly conserved in these enzymes (Janecek,1997 and 1999). The catalytic 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 CGTase 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 contains at least nine sugar-binding subsites, designated from the nonreducing end to the reducing end as -7 though +2 (Fig. 7)(van der Veen *et al.* 2000c and Strokopytov *et al.*,1996). The loop between β-strand 3 and helix 3 of the catalytic domain is relatively large and is regarded as a separate structural domain. This domain B consists of 44-133 amino acid residues and contributes to substrate binding. The domain C is approximately 100 amino acids long and has an antiparallel β-sandwich fold. Domain C of the CGTase from *B. circulans* strain 251 contains one of the three maltose binding sites (Lawson *et al.*, 1994) observed from maltose dependent crystals. This maltose binding site was found to be involved in raw starch binding (Penninga *et al.*, 1996a), suggesting a role of C-domain in substrate binding. Some authors suggest that this domain is involved in bond specificity, since in enzymes hydrolyzing or forming the α(1,6) bonds (e.g. pullulanase, isoamylase and branching enzymes), the A-domain is followed by a different domain (Fig.4)(Jesperson *et al.*, 1991). The domain D, consisting of approximately 90 amino acids with an immunoglobulin fold, is almost exclusively found in CGTases and its function is unknown. Domain E, consists of approximately 110 amino acids and is found to be responsible for substrate binding, This is where the other two maltose binding sites resided (Penninga *et al.*, 1996a).

Chang *et al.*, (1998) constructed several mutants of the E-domain of *Paenibacillus macerans* CGTase. Removing the entire E-domain resulted in an inactive enzyme. Adding six amino acid residues between domain D and E caused a decrease in activity and thermostability. Replacing the domain E with the similar starch binding domain from *Aspergillus awamori* glucoamylase I caused a drastic decrease in activity, indicating the necessity of having the correct alignment of bound substrate. Substituting the Tyr634 in domain E with phenylalanine had very little effect on starch-hydrolyzing activities compared with that of wild type enzyme. The result of this study indicates that domain E is also important for the stability and integrity of *P. macerans* CGTase in addition to raw starch binding activity.
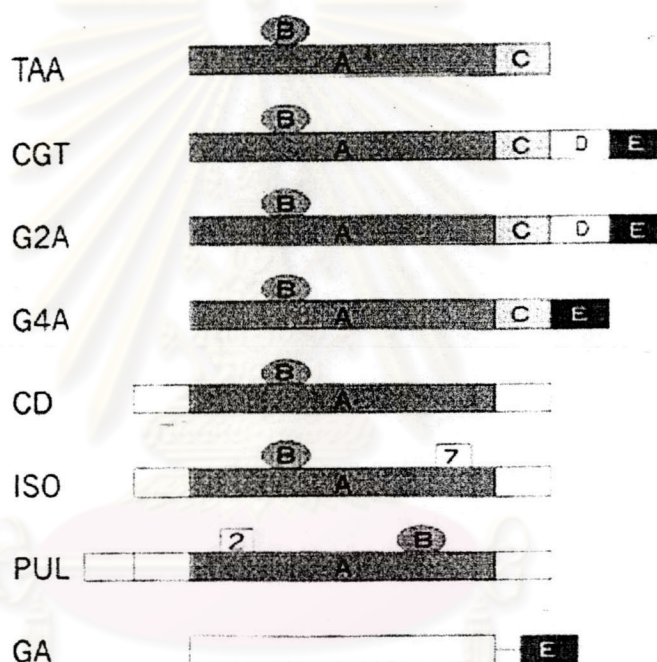
Figure.4 Domain level organization of starch-degradating enzymes. CGT, CGTase from *B. circulans*; G2A, maltogenic α-amylase from *Bacillus stearothermophilus*; G4A maltotetraose forming α-amylase from *Pseudomonas stutzeri*; TAA, α-amylase from *A oryzae* (Taka-amylase A); CD, cyclodextrinase from *Klebsiella 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.*, 2000c)
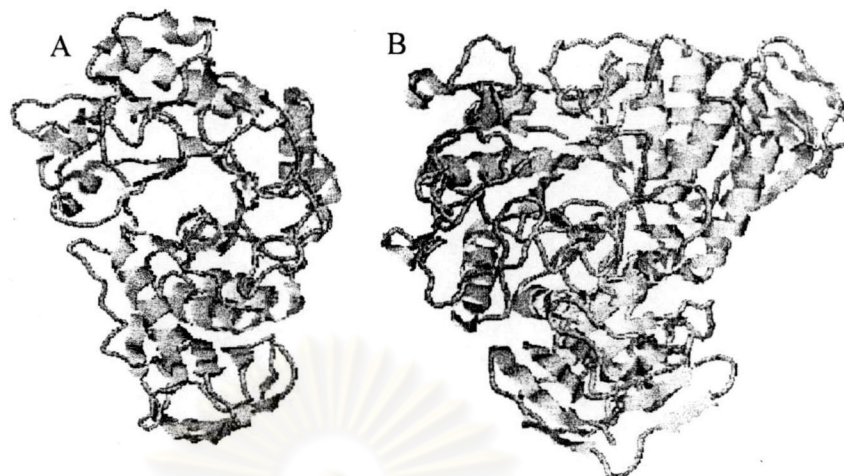
Figure.5　　Comparison of the three dimensional structure;A: α-amylase from *Bacillus subtilis* (PDB I.D. 1BAG);B: CGTase from *Bacillus circulans* strain 251(PDB I.D. 1CDG)
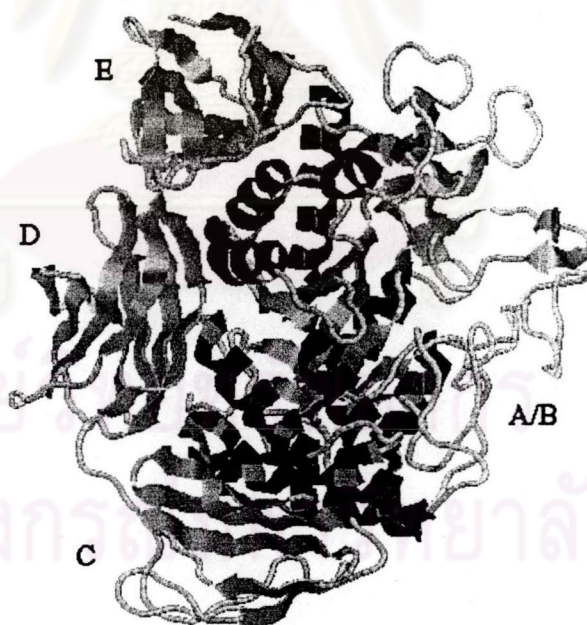
Figure 6.　　A common three dimensional structure of CGTase (Source: www.rcsb.org).
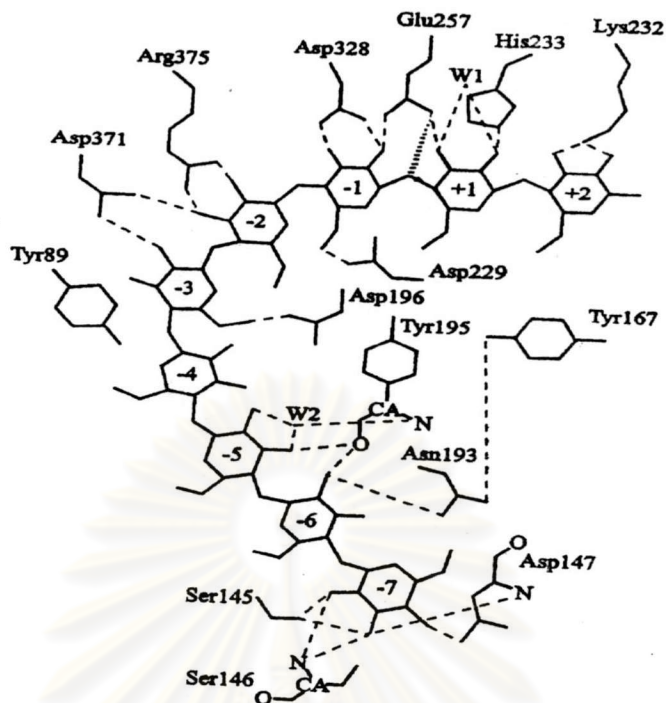
Figure7. Schematic representation of the hydrogen bonds between the *B. circulans* strain 251 CGTase and a maltononaose inhibitor bound at each subsite of the active site. W1 and W2 indtcate water and CA indicates alpha carbon. (Source: van der Veen *et al.*, 2000c).

4. CGTase reaction and catalytic mechanism of CGTase

The CGTases can catalyze 2 types of transglycosylation reaction and a weak hydrolysis reaction, though it mainly catalyses the former (Fig. 8 and 9).

i) Intramolecular transglycosylation reaction is sometimes called cyclization. The reaction proceeds by cleaving a linear oligosaccharide chain and transferring the reducing end sugar to the non-reducing end of the same oligosaccharide chain, which acts as the acceptor.

ii) Intermolecular transglycosylation reaction can be divided into coupling and disproportionation reactions. Coupling reaction is the reverse of cyclization, in which a cyclodextrin ring is cleaved and transferred to a linear acceptor oligosaccharide. The kinetic mechanism of coupling proceeds via random ternary complex mechanism (van der Veen *et al.*, 2000a). For the disproportionation reaction, a linear oligosaccharide is cleaved and transferred to a linear acceptor oligosaccharide. The kinetic

mechanism proceeds via ping-pong mechanism (Nakamura *et al.*, 1994a, van der Veen *et al.*, 2000a).

iii) Hydrolysis is a weak reaction of CGTase in which a water molecule acts as an acceptor instead of a linear oligosaccharide.
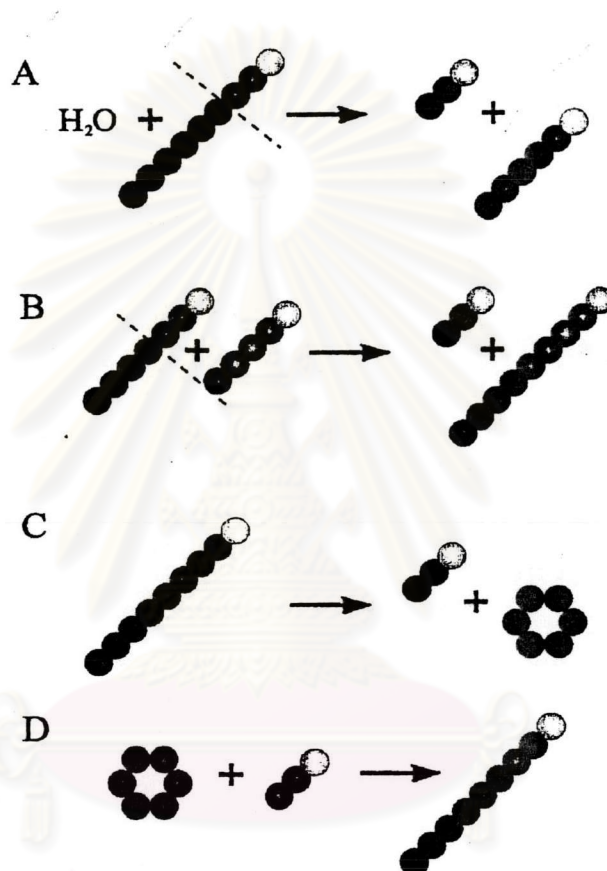
Figure 8.   Schematic representation of the CGTase-catalysed reactions. The circles represent glucose residues; the white circles indicate the reducing end sugars. (A) hydrolysis; (B) disproportionation; (C) cyclization; (D) coupling (Source: van der Veen *et al.*, 2000c)
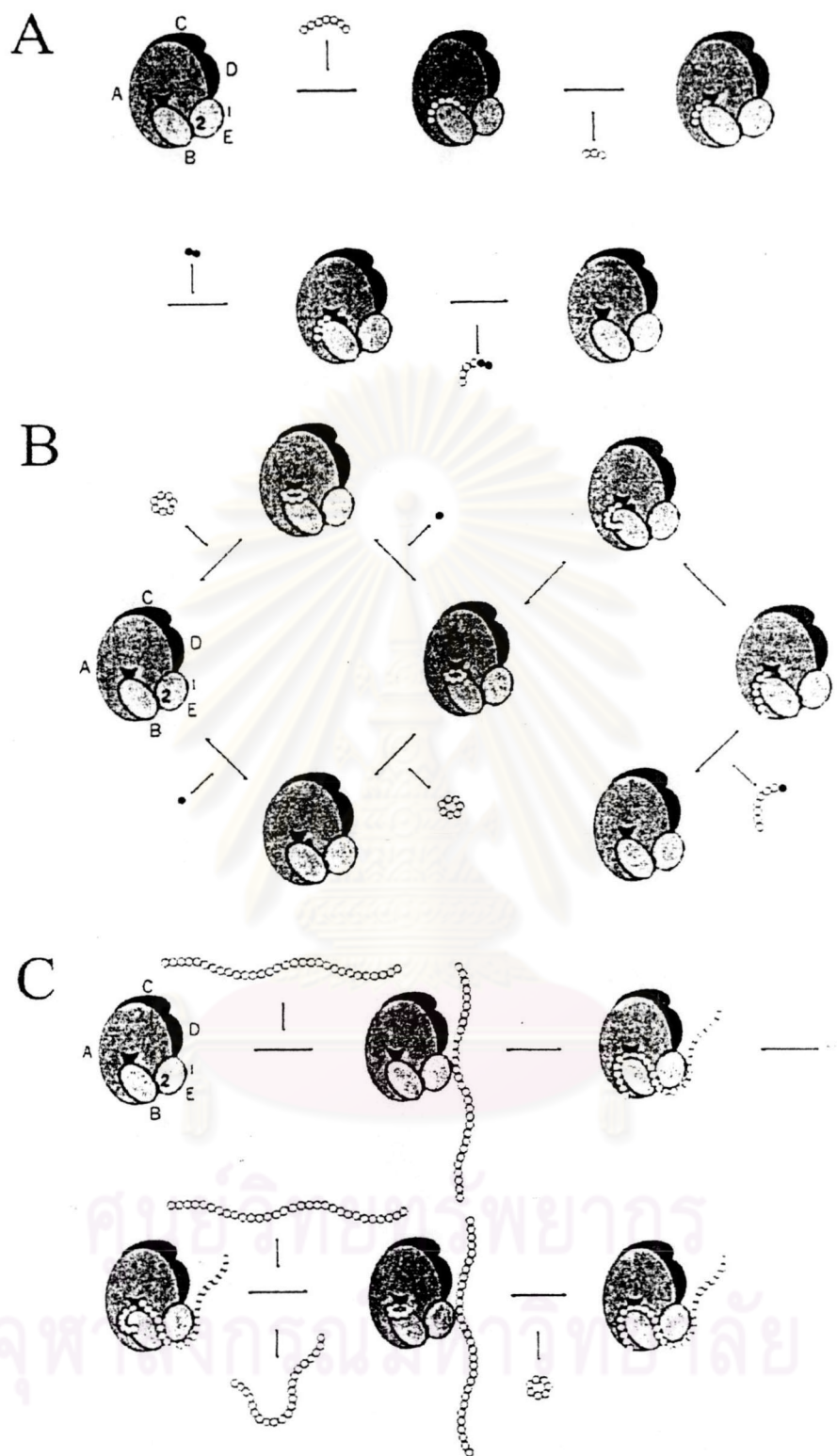
Figure 9. Proposed model of the events taking place in the CGTase-catalysed reactions. (A) disproportionation; (B) coupling; (C) cyclization. The different CGTase domains are indicated A, B, C, D and E. The circles represent glucose residues; acceptor residues are represented in black (Source: van der Veen *et al.*, 2000a).

As a member of α-amylase family, CGTase uses the double displacement mechanism, the α-retaining mechanism (Fig. 10 and 12). Their maltose binding sites (MBS) are at the protein surface. MBS1 and MBS2 are both located on the E-domain, suggesting a role in the raw starch binding function of this domain (Lawson *et al.*,1994). Indeed, mutational studies reveal that maltose binding site 1 is important for starch binding, while maltose binding site 2 assists in guiding the linear starch chains into the active site via a groove at the surface of the CGTase protein (Penninga *et al.*, 1996a). These maltose binding sites are also found to interact strongly with CDs and oligosaccharides (Knegtel *et al.*, 1995).

The CGTase cleaves an α(1,4) glycosidic bond in its substrate between subsite −1 and +1, and forms a covalent β(1,4) linked glycosyl-enzyme intermediate. The α(1,4) glycosidic bond is reformed with an acceptor, which can be water or the C4 hydroxyl group of another sugar residue (van der Veen *et al.*, 2000a, c and d). CGTase is unique in its ability to use the free hydroxyl at the non-reducing end of the intermediate sugar chain as an acceptor, thus forming a cyclic oligosaccharides, cyclodextrins. Three amino acid residues in the active site play distinct roles in catalysis. Glu257 acts as the acid/base catalyst, which protonates the glycosidic oxygen of the scissile bond in the first step, and then deprotonates the attacking OH group in the second step. Asp229 acts as the catalytic nucleophile, which attacks the sugar, forming the covalent intermediate. Asp328 stabilizes the substrate binding and elevates the pK_a of Glu257 (Fig. 11) (Uitdehaag *et al.*, 1999).

The essential amino acid residues that are needed for cyclization are four aromatic amino acids, Phe183, Phe/Tyr195, Phe259 and Phe283, commonly located in the active site of CGTase (Fig. 13), but not in that of α-amylase. Phe259 and Phe183 at subsite +2 play a critical role in the binding of non-reducing end of starch moving into the acceptor site. Leemhuis *et al.* (2002b) and van der Veen *et al.*, (2001) has shown that both amino acids are essential for cyclization, coupling and disproportionation. Moreover, the hydrophobicity of Phe183 and Phe259 limits the hydrolyzing activity of the enzyme. When Phe183 and Phe259 are replaced by polar amino acids, the mutants have lower transglycosylation activity and higher hydrolysis activity (Nakamura *et al.*, 1994a, van der Veen *et al.*, 2001, Leemhuis *et al.*, 2002b). Phe283 is important for the hydrophobic environment around Glu257, and hence participates in raising the pK_a value of the catalytic residue, Glu257 (Penninga .,

1996b). Phe/Tyr195 is in a dominant position in the center of the active site cleft of CGTase. It was believed that Tyr195 influences the preferred cyclodextrin size. Substitution of this central amino acid by the other amino acid, Phe, however does not support the above notion. Furthermore, natural CGTases all have Tyr or Phe at this position, indicating that this residue is not involved in the differences in product specificity. Therefore, the function of this amino acid residue remains to be elucidated.

Terada *et al.* (2001) investigated the cyclization reactions of three CGTases from alkalophilic *Bacillus* sp. strain AZ-5a, *Paenibacillus macerans* and *Bacillus stearothermophilus*. They found that all three enzymes produced larger cyclodextrins at the early stage of the reaction but these were subsequently converted into smaller cyclodextrins, and the rate of this conversion differed among the three enzymes. For this conversion, large cyclodextrin is first linearized by either a coupling or hydrolytic reaction. The β-cyclodextrin and γ-cyclodextrin are extremely poor substrates for both reactions compared to the mixtures of larger cyclodextrins. Thus, among the initial cyclization products, larger cyclodextrins are selectively subjected to the linearization reactions, and the linear products are cyclized again into the smaller cyclodextrins. Repetition of the cyclization reaction and the linearization reaction may be the principle mechanism whereby large cyclodextrins are converted into smaller cyclodextrins, resulting in the final equilibrium composition of cyclodextrins. The factors that directly determine the composition of cyclodextrins are the product specificity of the cyclization reaction and the substrate specificity of the linearization reaction, including the coupling and hydrolytic reactions of larger cyclodextrins.
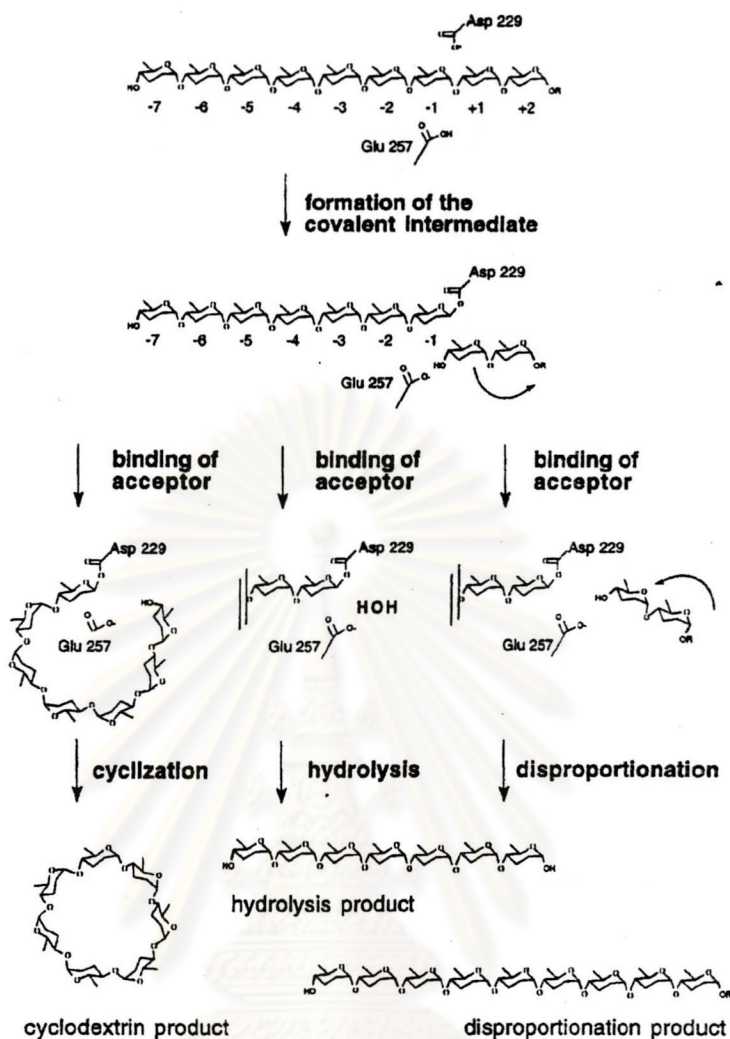
Figure 10. The catalytic reaction of CGTase, which involved cyclization, disproportionation and hydrolysis (Source: Uitdehaag *et al.*, 2000).
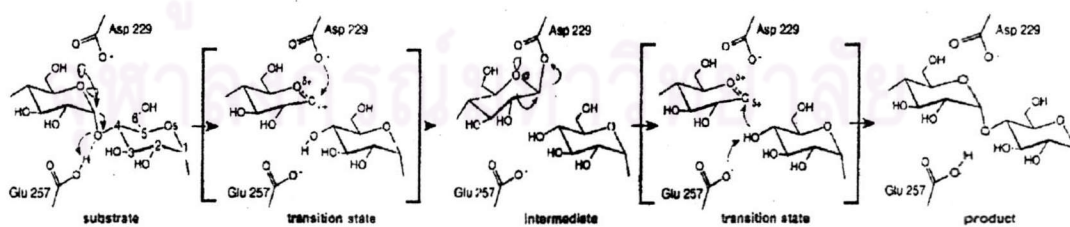
Figure 11. Scheme of the CGTase reaction mechanism (Source: Uitdehaag *et al.*, 1999a).
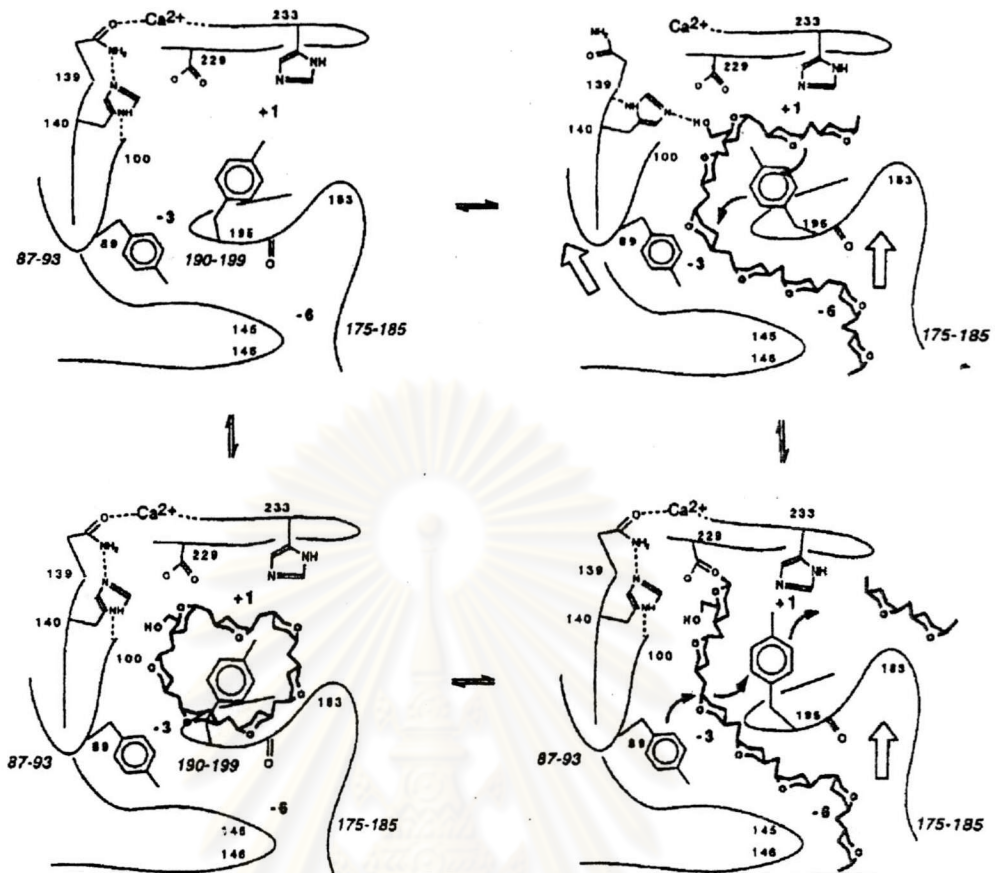
Figure 12. Possible scenario for the structural rearrangements during the CGTase reaction cycle. The pictures in clockwise order show a cyclization reaction cycle, and in counterclockwise order a coupling reaction cycle. The hydrolysis follows top left to top right and to bottom right. The disproportionation follows bottom right to top right (Source: Uitdehaag *et al.*, 2000).
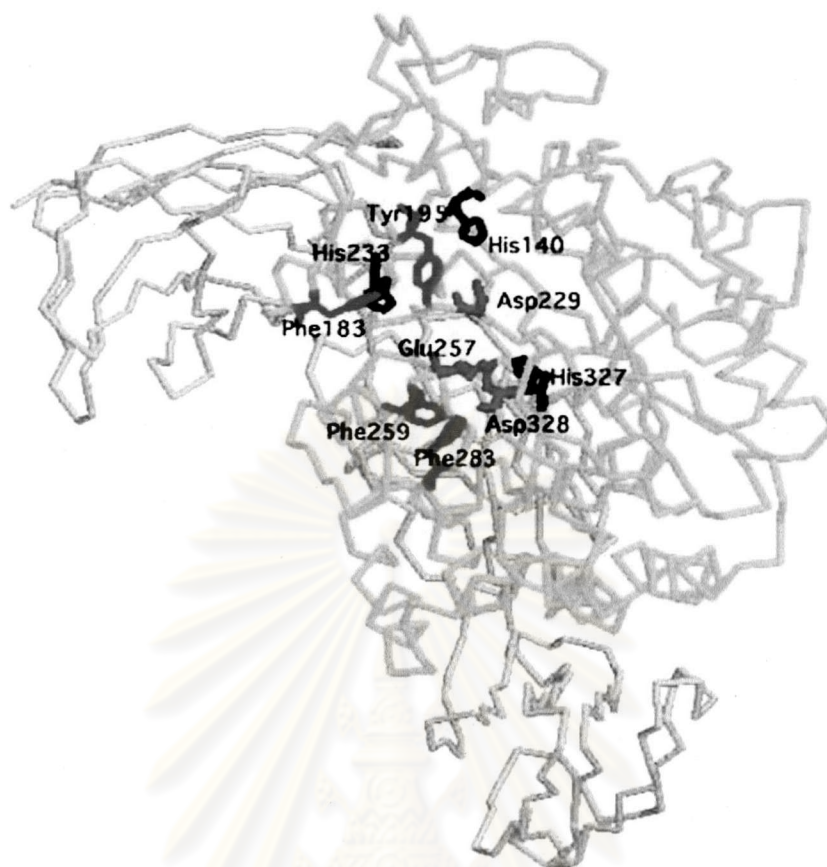
Figure 13. Essential amino acid residues for cyclization reaction of CGTase (Source: www.glycoforum.gr.jp/science/word/saccharide/SA-B03E.html).

5. Comparison of the CGTase gene sequences and the CD products

In general, CGTases from different microorganisms produce all three types of cyclodextrins with different ratios. The CGTases thus can be classified into α-CGTase, α/β-CGTase, β-CGTase, β/γ-CGTase and γ-CGTase corresponding to their major cyclodextrin products. All five types of CGTases show clear similarity in their amino acid sequences, ranging from 47% to 99%. Figure 14 and Table 2 (Takada *et al.,* 2003) show the alignment of amino acid sequences of α-, β-, β/γ- and γ-CGTases from different sources. The first specific CGTase residues, that may be involved in product specificity, are found in the region between β-strand 1 and α-helix 1, which consists of a stretch of amino acid residues, 27-53 and is involved in calcium binding. Residue 47 is involved in binding (semi)cyclic oligosaccharides and is typically an Arg in β-CGTase, Lys in α-CGTase, Thr in β/γ- and γ-CGTases and His in the CGTase

producing virtually no α-cyclodextrin (Fig. 14) (van der Veen *et al.*, 2000c and Takada *et al.*, 2003).

The second region is found at subsite -3. Amino acid residues 87-94 at subsites -3 are looped. These amino acid residues show hydrophobic interactions with glucose unit bound in this region, and are remarkably different among the four types of CGTases. Both α- and β-CGTases have the sequence INYSGVN but the sequence HP-GGF- is found in β/γ- and γ-CGTases (Fig. 14 and Table 2) (Takada *et al.*, 2003). The fact that the loops in γ- and β/γ-CGTase are shorter than those in other CGTases indicates that more space for the bound glucosyl chain in this region is needed for a higher level of γ-cyclization activity as shown in Fig.15

Another specific region consists of residues 145-152 of α- and β-CGTases at subsite -7. The amino acids in this region are located in a loop at the start of the B-domain, and are SSTDPSFA and SSDQPSFA for α- and β-CGTases, respectively. The β/γ- and γ-CGTases completely lack the six amino acid residues in this region (D-I) (see Fig. 16). This indicates that the absence of these amino acid residues may also be necessary for a higher level of γ-cyclization activity as it provides more space for the bound glycosyl chain in this region as shown in Fig.15.

```
Bmace   1:-SPDTSVNNKLNFSTDTVYQIVTDRFVDGNSANNPTGAAFSSDHSNLKLYFGGDWQGITNKINDGYLTGMGITALWISQP 79
Bc251   1:AP-DTSVSNKQNFSTDVIYQIFTDRFSDGNPANNPTGAAFDGTCTNLRLYCGGDWQGIINKINDGYLTGMGVTAIWISQP 79
Bf290   1:----NENLDNV-NYAQEIIYQIVTDRFYDGDPTNNPEGTLFSPGCLDLTKYCGGDWQGVIEKIEDGYLPDMGITAIWISPP 76
Bcla    1:SNATNDLSNV-NYAEEVIYHIVTDRFKDGDPDNNPQGQLFSNGCSDLTKYCGGDWQGIIDEIESGYLPDMGITALWISPP 79
              *    *.*.****  **.  .*** .*   *.*.*******.  .*.*** **.** ***.*
              β1           Ca²⁺ binding site               α1          β2

Bmace  80:VENITAVINYSGVNNTAYHGYWPRDFKKTNAAFGSFTDFSNLIAAAHSHNIKVVNDFAPNHTNPASSTDPSFAENGALYN 159
Bc251  80:VENIYSIINYSGVNNTAYHGYWARDFKKTNPAYGTIADFQNLIAAAHAKNIKVIIDFAPNHTSPASSDQPSFAENGRLYD 159
Bf290  77:IENVMELHP-GGF--ASYHGYWGRDFKRTNPAFGSLADFSRLIETAHNHDIKVIIDFVPNHTSPVD------IENCALYD 147
Bcla   80:VENVFDLHP-EGF--SSYHGYWARDFKKTNPFFGDFDDFSRLIETAHAHDIKVVIDFVPNHTSPVD------IEDCALYD 150
             .**   *     *****  ****.**...*   *** .**  **   ***.** ****.*       *.*. **.
             G-3loop                          α2       β3         G-7loop
                                                                     II

Bmace 160:NGTLLGKYSNDTAGLFHHNGGTDFSTTESGIYKNLYDLADINQNNNTIDSYLKESIQLWLNLGVDGIRFDAVKHMPQGWQ 239
Bc251 160:NGTLLGGYTNDTQNLFHHNGGTDFSTTENGIYKNLYDLADLNHNNSTVDVYLKDAIKMWLDLGIDGIRMDAVKHMPFGWQ 239
Bf290 148:NGRLVGHYSNDSEDYFYTNGGSDFSSYEDSIYRNLYDLASLNQQNSFIDRYLKESIQMWLDLGIDGIRVDAVAHMPVGWQ 227
Bcla  151:NGTLLGHYSTDANNYFYNYGGSDFSDYENSIYRNLYDLASLNQQHSFIDKYLKESIQLWLDTGIDGIRVDAVAHMPLGWQ 230
             **. *. *.  *   *  .** ***  *  *** *.******  *...  * ***.*. **..* **** **** *** ***
                                                  Y195        α3       β4

Bmace 240:KSYVSSIYSSANPVFTFGEWFLGPDEMTQDNINFANQSGMHLLDFAFAQEIREVFRDKSETMTDLNSVISSTGSSYNYIN 319
Bc251 240:KSFMAAVNNYK-PVFTFGEWFLGVNEVSPENHKFANESGMSLLDFRFAQKVRQVFRDNTDNMYGLKAMLEGSAADYAQVD 318
Bf290 228:KNFVSSIYDY-NPVFTFGEWFTGAGGSDEYHYFINN-SGMSALDFRYAQVVQDVLRNNDGTMYDLETVLRETESVYEKPQ 305
Bcla  231:KAFISSVYDY-NPVFTFGEWFTGAQGSNHYHHFVNN-SGMSALDFRYAQVAQDVLRNQKGTMHDIYDMLASTQLDYERPQ 308
             *.....*    *********.*.  .   *    .* .*** *** ** .*      .   *. .  *  . *    .*
             α4       β5             α5    β6    α6
                     IV

Bmace 320:NMVTFIDNHDMDRFQQAGASTRPTEQALAVTLTSRGVPAIYYGTEQYMTGNGDPNNRGMMTGFDTNKTAYKVIKALAPLR 399
Bc251 319:DQVTFIDNHDMERFHASNANRRKLEQALAFTLTSRGVPAIYYGTEQYMSGGTDPDNRARIPSFSTSTTAYQVIQKLAPLR 398
Bf290 306:DQVTFIDNHDINRFSRNGHSTRTTDLGLAFLLTSRGVPTIYYGTEIYMDGDGDPDNRKMMNTFDQSTVAYQIIQQLSSLR 385
Bcla  309:DQVTFIDNHDIDRFTVEGRDTRTTDIGLAFLLTSRGVPAIYYGTENYMTGKGDPGNRKMMESFDQTTTAYQVIQKLAPLR 388
             ..********* **    .*.  . *   .** ******* ****** ...*..**    .    .* . *. *.  **
             β7          α7         β8            α8

Bmace 400:KSHNPALAYGSTTQRWVNSDVYVYERKFGSNVALVAVNRSSTTAYPISGALTALPNGTYTDVLGGLLNGNSITVNGGTVS- 478
Bc251 399:KCNPAIAYGSTQERWINNDVLIYERKFGSNVAVVAVNRNLNAPASISGLVTSLPQGSYNDVLGGLLNGNTLSVGSGGAAS 478
Bf290 386:QENRAIAYGDTTERWINEDVFIYERSFNGEYALIAVNRSLNHSYQISSLVTDMPSQLYEDELSGLLDGQSITVDQNGSIQ 465
Bcla  389:QENKAVVYGSTKERWINDDVLIYERSFNGDYLLVAINKNVNQAYTISGLLTEMPAQVYHDVLDSLLDGQSLAVKENGTVD 468
             *  *  **.**.*  **.   **  ** **   **.  .*.*.   **.. .**.  *     * *.*  **.  .

Bmace 479:NFTLAACGGTAVWQYTTTESSPI-IGNVGPTMGKPGNTITIDGRGFGTTKNKVTFGTTAVTGANIVSWEDTEIKVKVPNVA 557
Bc251 479:NFTLAACGGTAVWQYTAATATPT-IGHVGPMMAKPGVTITIDGRGFGSSKGTVYFGTTAVSGADITSWEDTQIKVKIPAVA 557
Bf290 466:PFLLAPCGEVSVWQYSNGQNVAPE-IGQIGPPIGKPGDEVRIDGSGFGSSTGDVSFAGSTMNV---LSMNDDTIIAELPEHN 542
Bcla  469:SFLLGPCEVSVWQHISESGSAPVIGQVGPPMGKPGDAVKISGSGFGSEPGTVYFRDTKIDV---LTWDDETIVITLPETL 545
             *.*.  * ***.   **.**..***  *.* ***.  . * .        .* *  *     *

Bmace 558:AGNTAVTVTNAAGTTSAAFNNFNVLTADQVTVRFKVNNATTALGQNVYLTGNVAELGNWTAANAIGPMYNQVEASYPTWY 637
Bc251 558:GGNYNIKVANAAGTASNVYDNFEVLSGDQVSVRFVVNNATTALGQNVYLTGSVSELGNWDPAKAIGPMYNQVVYQYPNWY 637
Bf290 543:GGKNSVTVTTNSGESSNGYP-FELLTGLQTSVRFVVNQAETSVGENLYVVGDVPELGSWDPDKAIGPMFNQVLYSYPTWY 621
Bcla  546:GGKAQISVTNSDGVTSNGYD-FQLLTGKQESVRFVVDNAHTNYGENVYLVGNVPELGNWNPADAIGPMFNQVVYSYPTWY 624
             .*   *.  *.   **** .*. *.  **   *. ***.*   .  *****  *** .. **. **

Bmace 638:FDVSVPANTALQFKFIKVNGSTVT-WEGGNNHTFTSPSSGVATVTVDWQN 686
Bc251 638:YDVSVPAGKTIEFKFLKKQGSTVT-WEGGSNHTFTAPSSGTATINVNWQP 686
Bf290 622:YDVSVPANQDIEYKYIMKDQNGNVSWESGNNHIYRTPENSTGIVEVNFNQ 671
Bcla  625:YDVSVPADTALEFKFIIVDGNGNVTWESGGNHNYRVTSGSTDTVRVSFRR 674
             *.****** . .*  .  .  *** *  ** .
```
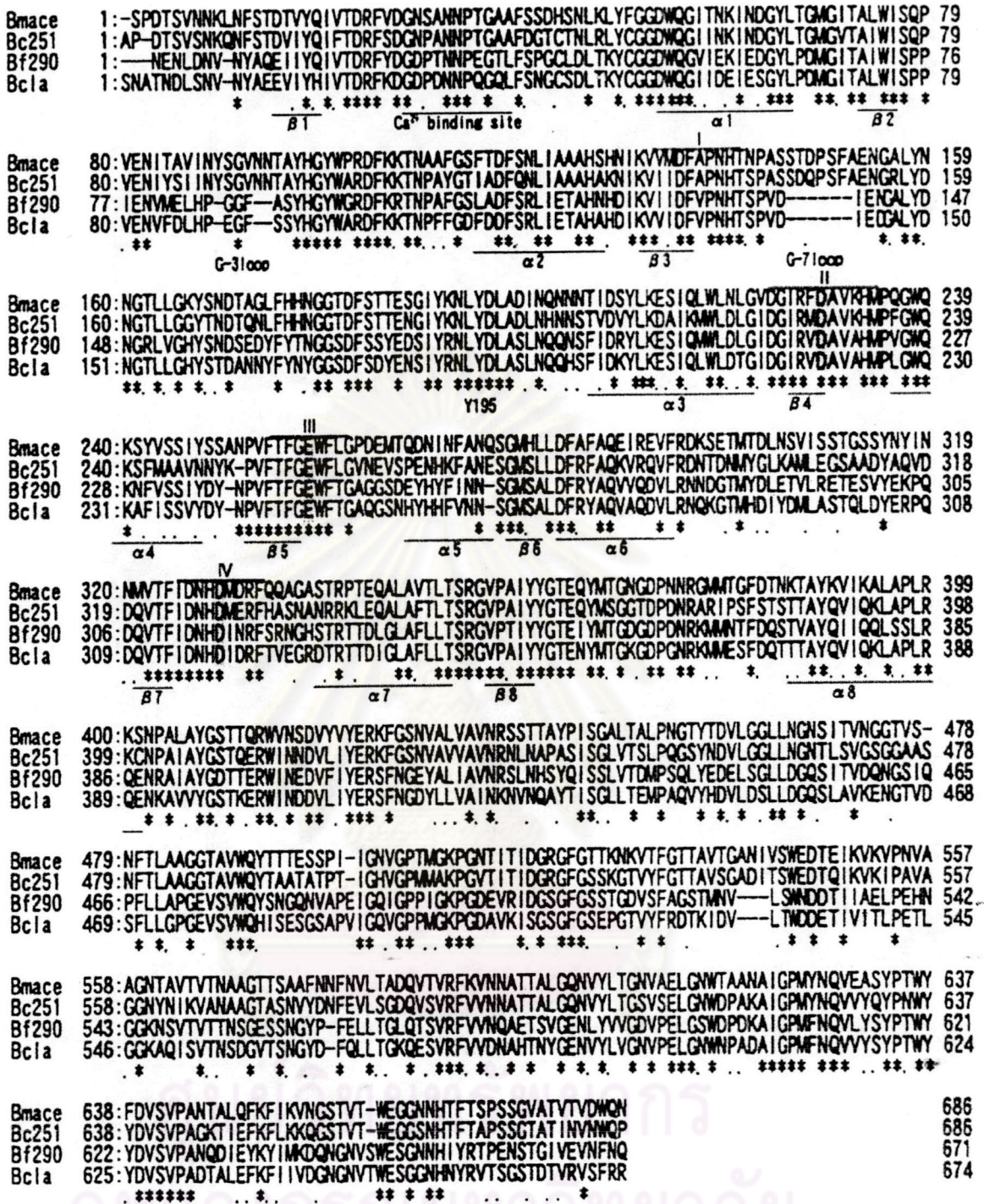
Figure 14. Alignment of amino acid sequences of typical α-, β-, β/γ-, and γ-CGTases. Bmac, Bc251, Bf290 and Bcla indicate the amino acid sequences of α-CGTase from *B (P). macerans*, β-CGTase from *B. circulans* 251, β/γ-CGTase from *Bacillus firmus* 290-3, and γ-CGTase from *B. clarkii* 7364, respectively (Source: Takada *et al.*, 2003).

Table 2. Comparison of the amino acid residues around the active center in the four types of CGTases.

Residue No.	Residue in α-CGTase	Residue in β-CGTase	Residue in β/γ-CGTase	Residue in γ-CGTase	Function
Subsite +2					
183	F	F	F	F	Cyclization
232	K	K	A	A	Disproportionation
259	F/Y	F/Y	F	F	Cyclization and disproportionation
Subsite +1					
194	L	L	L	L	Cyclization
230	A	A	A	A	Transglycosylation*
233	H	H	H	H	General activity
Subsite −1					
The catalytic residues (**D229, E257 and D328**) are completely conserved.					
Subsite −2					
98	H	H	H	H	General activity
101	W	W	W	W	-
375	R	R	R	R	-
Subsite −3					
89	Y	Y	-	-	Cyclization specificity
87-93	INYSGVN	INYSGVN	HP-GGF-	HP-GGF-	Cyclization specificity
371	D	D	D	D	Cyclization
196	D	D	D	D	Cyclization
47	K	R	T	T	Cyclization
Subsite −4 and −5					
No side chain contacts					
Subsite −6					
167	Y	Y	Y	Y	-
179-180	GG	GG	GG	GG	Cyclization
193	N	N	N	N	Cyclization
Subsite −7					
145-152	SSTDPSFA	SSDQPSFA	D.......I	D.......I	Cyclization, Central specificity
195	Y	Y/F	Y	Y	Cyclization, Ca^{2-} binding site
32-36	NNPTG	NNPTG	NNPEG	NNPQG	Ca^{2+} binding

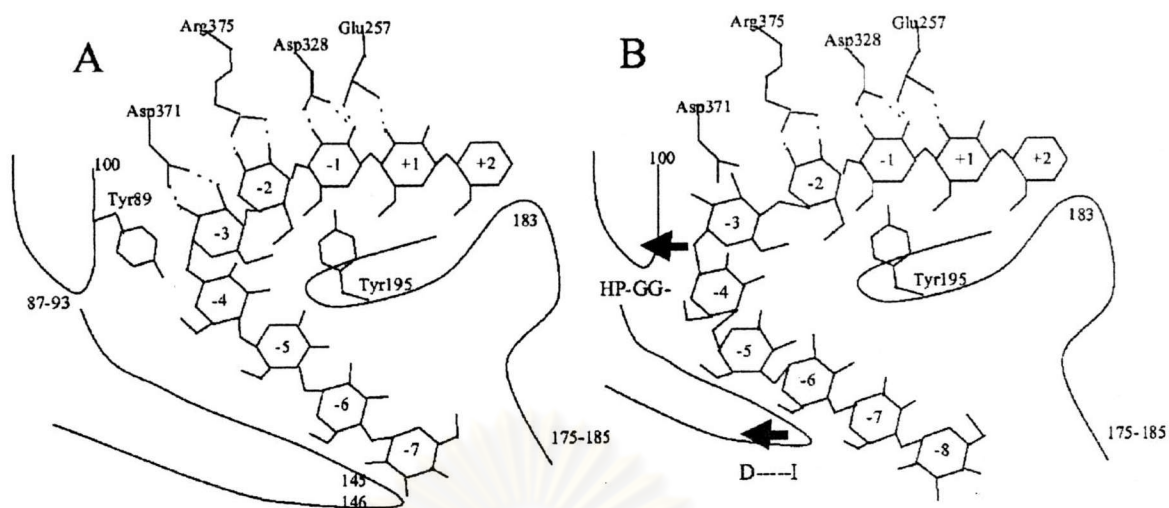(Source: Takada *et al.*, 2003, *Leemhuis *et al.*, 2003b)

Figure 15.　Schematic structure of the substrate binding sites of CGTase. (A) β-CGTase, (B) γ-CGTase (Source: Takada *et al.*, 2003).

6. Mutagenesis of CGTase

In the past, various site-directed mutagenesis affecting the product specificity of CGTases had been made on the aromatic amino acid residue 195 (Phe or Tyr), located dominantly in the center of the active site cleft of the enzymes, and a mechanism for the α-(1,4)-glucan chain folding around this residue had been proposed (Strokopytov *et al.*,1996). It was believed that Tyr195 played an important role in the cyclization reaction of the CGTase. Attempts had been made to change the product specificity of CD by substituting this amino acid residue with other amino acid residues. Substitution of this amino acid with Trp (Tyr188Trp) (equivalent to Tyr195 in CGTase of *B. circulans* 251) in *B. ohbensis* (Sin *et al.*, 1994), Tyr195Leu in *B. circulans* strain 251 (Penninga *et al.*, 1995) and Tyr195Trp in the *B. circulans* strain 8 (Parsiegla *et al.*, 1998) could improve product specificity to a limited extent with higher production of γ-cyclodextrin. However, several other Tyr188 and Tyr195 mutations as well as the Phe191Tyr (equivalent to Tyr195 in CGTase of *B. circulans* 251) in the CGTase of *G. stearothermophilus* NO2 did not support the importance of this central amino acid (Fujiwara *et al.*, 1992 and van der Veen *et al.*, 2000c). Furthermore, the natural α-, α/β-, β-, β/γ- and γ-CGTases all have Tyr or Phe at this position, indicating that this residue is not involved in product specificity.

Leemhuis *et al.* (2002a) found that the subsite -6 in the active site of β-CGTase from *B. circulans* strain 251 was of great importance for all three transglycosylation

reactions (cyclization, coupling and disproportionation) but not for hydrolysis reaction. They found that Gly180 and Asn193 were important for β-and γ-cyclization, while Gly179 was important for α-cyclization.

In 1996, Strokopitov *et al.* studied the X-ray structure of the CGTase from *B. circulans* strain 251 in complex with a maltononaose inhibitor and suggested that the sugar binding subsites further away from the catalytic site could be important for the enzyme-product specificity and that it might be possible to change the ratio of the produced cyclodextrins by altering the affinities for glucose residues at these sugar binding subsites. Subsites -6, -7 and -8 may be the key sites for the product specificity.

In 2000, Uitdehaag *et al.* determined the X-ray structures of CGTase in complex with maltoheptaose, an equivalent oligosaccharide for β-cyclodextrin, and maltohexaose, an equivalent oligosaccharide for α-cyclodextrin, they found that the conformations of maltoheptaose and maltohexaose were different at subsites -3, -6 and -7 (Fig. 16). The conformation differences at specific sugar binding subsites suggested that the determinants for cyclodextrin size specificity are located at subsites -3, -6 and -7. This finding was in agreement with the amino acid sequence comparison of natural CGTases with different size specificities, which showed the highest sequence variation in subsites -3 and -7.

Parsiegla *et al.* (1998) constructed β-CGTase mutant from *B. circulans* strain 8 by replacing residues 145-151 with a single Asp,Δ(145-151)D, which removed most interactions at subsite -7. The mutant showed reduced production of β-cyclodextrin and increased production of γ-cyclodextrin. The higher γ-cyclization activity of this mutant were explained as a widening of the active site cleft in the deleted portion of amino acid sequence to produce more space for the bound glucosyl chain (Parsiegla *et al.*, 1998). This Δ(145-151)D deletion mutant had a stretch of HTSPADAE similar to a stretch of HTSPVDIE in the β/γ-CGTase from *B. firmus* strain 290-3 (Englbrecht *et al.*, 1990) and γ-CGTase from *B. clarkii* 7364 (Takada *et al.*, 2003). This indicates that it is necessary for γ-cyclization activity to have more space for the bound glucosyl chain at subsite -7 (Fig. 15).
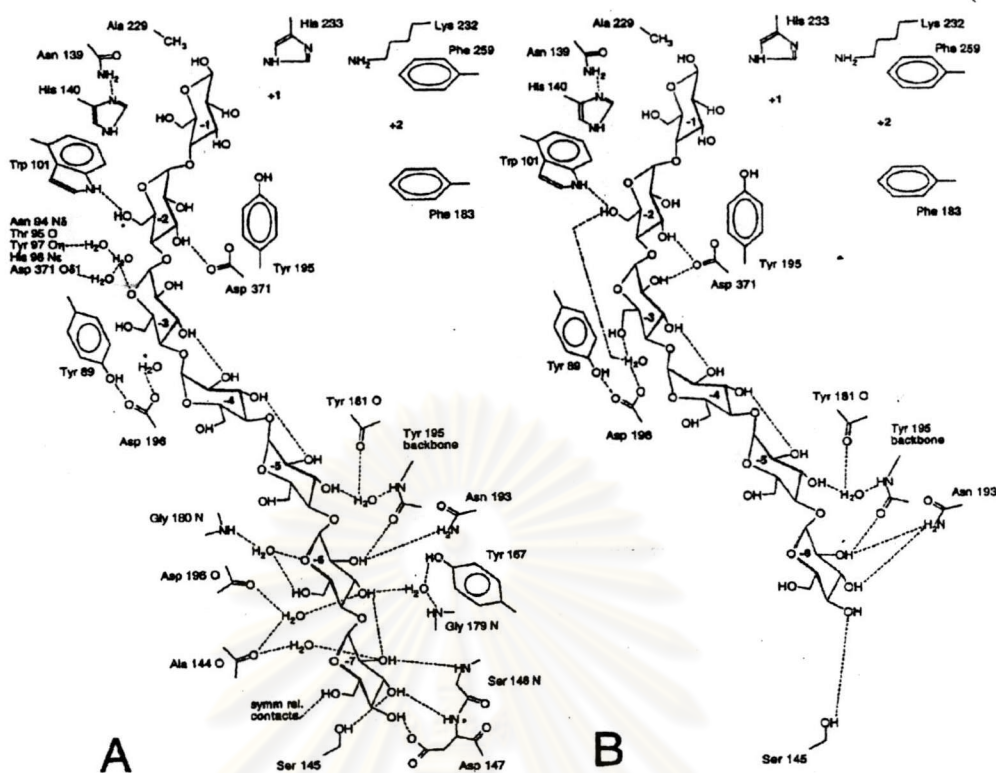
Figure 16. Schematic representation of all interaction of (A) maltoheptaose, and (B) maltohexaose with the CGTase from *B. circulans* strain 251 (Source: Uitdehaag *et al.*, 2000).

7. Scope of this study

As mentioned above, the separation of the different CDs is costly and time-consuming, CGTases that predominantly produce one type of CD are of great importance for industry (Englbrecht *et al.*, 1990 and Martins *et al.*, 2002). The CDs can be synthesized enzymatically on an industrial scale by using CGTases from various bacteria. These CGTases convert starch into a mixture of α-, β- and γ-CDs with either one of them as the main product. The CD ratio varies depending on the source of enzymes. Most CGTases are found to produce α- and β-CDs as major products. Although, there is a demand for γ-CD, which can trap larger molecules that cannot be trapped by α- or β-CDs, in various industries, the industrial production of γ-CD is not yet practical due to the low yield of this substances with currently utilized CGTase(Kato *et al.*,1986 and van der Veen *et al.*,2000d).

Several attempts have been made to enhance the yield of γ-CD production. There are two possible ways to increase the γ-CD production. The first simple way is to screening for microorganisms that produce large amounts of γ-CD. The other way

is to mutagenize the existing CGTase to change the enzyme such that it favors the γ-CD production. The research will work on β-CGTase gene from *B. circulans* A11. (Rimphanitchayakit *et al.*, 2000). The gene encodes a β-CGTase of 713 amino acid residues, including a signal peptide. The molecular weight is approximately 72 kDa.

By using information from the GenBank and Bioinformatics, the amino acid sequence alignments between β-CGTase from *B. circulans* A11 and other typical CGTases, α-CGTase from *B. mercerans*, β/γ-CGTase from *B. firmus* 290-3 and γ-CGTase from *B. clarkii* 7364 are carried out. The differences among the sequences are analyzed. The A11 CGTase is mutagenized according to the sequence analysis. The mutated enzymes are prepared from the mutants and the CD products are subjected to HPLC analysis. By doing so, the regions of CGTase amino acid sequences that govern the γ-CD production might be determined.