Chapter 1

Introduction

Cyclodextrin glycosyltransferase, a biocatalyst with unique properties

I. Starch, a versatile raw material
   I.1. Starch as carbon and energy storage
   I.2. Starch as carbon and energy source
   I.3. Starch as the versatile raw material

II. The α-amylase family (family 13) of glycosyl hydrolases
   II.1. Sequence similarities in the α-amylase family
   II.2. Catalytic mechanism of the α-amylase family
   II.3. Three-dimensional structure similarities in the α-amylase family
   II.4. Substrate binding of (α-)amylases and CGTases
   II.5. Product specificity of α-amylases and CGTases

III. Cyclodextrin glycosyltransferases
   III.1. CGTase catalyzed reactions
   III.2. CGTase versus α-amylase action
   III.3. Sequence similarities in CGTases
   III.4. CGTase three-dimensional structures
   III.5. Cyclodextrin product specificity of CGTases
   III.6. Cyclodextrin product inhibition

IV. Aims and outline of this thesis

Accepted for publication in Biochimica et Biophysica Acta: Protein Structure and Molecular Enzymology: Engineered Enzymes
I. Starch, a versatile raw material

I.1. Starch as carbon and energy storage

Many 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. wheats) or roots (e.g. potato) in the form of granules which consist of two types of these glucan polymers: amylopectin and amylose.

The structure of starch

Amylopectin is an α(1-4)-glucan polymer highly branched with α(1-6)-linkages (approximately one α(1-6)-linkage per 20 glucose residues) and can consist of over 200,000 glucose residues. It is arranged in organized structures, which form the basis of the starch granule (Fig. 1). The amylopectin molecules are arranged radially, with their single reducing end sugar pointed towards the center of the granule. They consist of semicrystalline regions and amorphous regions, forming the so-called growth rings of the granules (Fig. 1a), which result from fluctuations in biosynthesis depending on the growth conditions, for example, the day-night regime. The semicrystalline regions consist of crystalline and amorphous lamellae (Fig. 1b). The linear parts of the amylopectin molecules form regular packings of double helices, creating a crystalline matrix (Fig. 1c). At regular intervals (12-16 glucose residues) these crystalline lamellae alternate with clusters of branching points (Fig. 1d), resulting in more open (amorphous) lamellae. In the amorphous regions of the starch granule the organization of the amylopectin molecules is not understood (Smith et al. 1997).

Amylose molecules are linear chains of approximately 1000 glucose residues linked with α(1-4) glycosidic bonds and have a helical conformation due to the formation of internal hydrogen bonds. Amylose is found throughout the starch granule; in the amorphous regions it is interspersed with amylopectin, in the semicrystalline regions its location in the ordered amylopectin matrix is unclear.

The synthesis of starch

Starch synthesis starts with the formation of linear glucose chains from ADP-glucose by the action of starch synthase, which catalyzes the formation of an α(1-4) glycosidic linkage between the C1 of the ADP-glucose and the C4 of the non-reducing end glucose of a growing chain. Starch synthases are generally divided in two classes; soluble starch synthases (SS) and granule bound starch synthases (GBSS). SS is responsible for the elongation of amylopectin chains on the edge of the growing starch granule. When these chains are of sufficient length to assume the double helical conformation, they provide good substrates for starch branching enzyme (SBE). This SBE cleaves an α(1-4) linkage in the backbone and links the C1 of the new reducing end glucose to the C6 of another glucose in the backbone, resulting in an α(1-6) linked branching point. The concerted action of SS and SBE results in the formation of an unordered glucan at the surface of the granule. A third enzyme, the debranching enzyme (DBE), is responsible for the formation of the clusters of branches. DBE removes the branches
of the glucan, but has no access to the branch points close to the double helical regions, leaving clusters of branches with short chains at the edge of the granule, which are again elongated by SS. Thus the organized structure of amylopectin in the starch granule depends on three distinct activities, of which the specific affinity of SBE for the double helical conformation of the starch chains is responsible for the periodicity of the crystalline and amorphous lamellae in the semicrystalline regions. The synthesis of amylose depends on the other class of starch synthases, GBSS. GBSS is located and active inside the granule, producing amylose chains within the amylopectin matrix. This amylose remains largely unbranched, because it is inaccessible for SBE. The formation of amylose within the starch granule accounts for the strong variation in the ratios of amylose and amylopectin (from 11 to 51 % amylose), depending on the origin (plant species), variety within plants (plant organ, age of organ), and growth conditions. These variations may be caused by, for example, the availability of ADP-glucose inside the granule, or the available space within the amylopectin matrix (Smith et al. 1997).

**Figure 1. Structure of the starch granule.** a) Slice of a starch granule, showing the growth rings, consisting of semi-crystalline regions and amorphous regions. b) Detail of the semi-crystalline region, consisting of crystalline and amorphous lamellae. c) Part of a crystalline lamella, consisting of the linear parts of the amylopectin molecules, which form regular packings of double helices, creating a crystalline matrix. d) Part of an amorphous lamella, which contains the α(1-6) linked branch point in the amylopectin (reproduced from Smith et al. (1997) with modifications).
1.2. Starch as carbon and energy source

Many (micro)organisms capable of using starch as carbon and energy source are found in nature. Starch degradation, like starch synthesis, requires a whole range of enzymes.

Starch degrading enzymes

In order for organisms to use the glucose of the starch granule as a growth substrate, the starch molecules need to be converted extracellularly into molecules suitable for uptake and further conversion by the cells. A whole range of starch degrading enzymes with specific activities has evolved in these organisms (Fig. 2). Most enzymes are hydrolytic, cleaving the linkages in the starch molecule followed by the reaction of the cleavage product with water, resulting in a new reducing end. They can be roughly divided into amylases, hydrolyzing α(1-4) linkages, and debranching enzymes, hydrolyzing α(1-6) linkages.

Figure 2. Action of enzymes involved in the degradation of starch. (©) Glucose molecule with a reducing end; (o) glucose molecule without a reducing end. Arrows indicate preferred cleaving points in the starch molecule (reproduced from Wind (1997), with modifications).
Examples of debranching enzymes are isoamylase (EC 3.2.1.68) and pullulanase (EC 3.2.1.41). Isoamylase specifically cleaves α(1-6) linkages in amylopectin and branched dextrins. Pullulanase hydrolyzes α(1-6) linkages in pullulan, which is a linear α-glucan consisting of maltotriose units joined by α(1-6) glycosidic linkages, but is also capable of cleaving the α(1-6) linkages in amylopectin. Amylopullulanase (pullulanase type II) hydrolyzes both α(1-4) and α(1-6) linkages. These different activities can be catalyzed by one active site, as has been shown for the amylopullulanase from *Thermoanaerobacter ethanolicus* 39E, where individual replacements of two catalytic carboxylic amino acids by their amide forms resulted in loss of both activities (Mathupala et al. 1993). The amylopullulanase from alkalophilic *Bacillus* sp. KSM-1378, however, contains two independent active sites (Ara et al. 1995; Hatada et al. 1996) and the specific hydrolytic activities can be separated by limited proteolysis with papain, yielding two protein fragments of which one has the amylase and the other the pullulanase characteristics (including pH and temperature profiles) of the parental enzyme (Ara et al. 1996).

Amylases can be further subdivided into endo- and exo-acting enzymes. A typical endo-acting enzyme is α-amylase (EC 3.2.1.1), cleaving α(1-4) bonds randomly in the starch molecule, producing (branched) oligosaccharides of various lengths. Exo-acting amylases such as β-amylase (EC 3.2.1.2) cleave α(1-4) bonds at the non-reducing end of the starch molecule and hence produce only low molecular weight products from starch (mostly glucose or maltose). Most of these enzymes are incapacable of bypassing α(1-6) linkages; degradation of branched substrates therefore remains incomplete, leaving high molecular weight compounds (limit dextrins). Some, however, are also able to cleave α(1-6) linkages, for instance glucoamylase (EC 3.2.1.3) and α-glucosidase (EC 3.2.1.20), but this reaction is slow compared to the hydrolysis of α(1-4) bonds.

Cyclodextrin glycosyltransferase (CGTase) (EC 2.4.1.19) is a unique member of the α-amylase family of glycosylases with a low hydrolytic activity. It is usually considered to be an exo-acting enzyme and unable to bypass branching points. Its main products when acting on starch are cyclic oligosaccharides consisting of 6, 7, or 8 glucose residues (named α-, β-, and γ-cyclodextrin, respectively) and highly branched high molecular weight dextrins (CGTase limit dextrins). The cyclodextrins are produced via an intramolecular transglycosylation reaction (cyclization) in which CGTase cleaves an α(1-4) bond in the starch molecule, concomitantly linking the reducing and non-reducing ends. The enzyme also catalyzes two intermolecular transglycosylation reactions: coupling, in which the ring of a cyclodextrin (donor) is opened and transferred to a linear oligosaccharide (acceptor), and disproportionation, in which part of a linear oligosaccharide chain (donor) is transferred to another linear oligosaccharide (acceptor) (Nakamura et al. 1993; van der Veen et al. 2000c).

**Metabolism of the products of starch degradation**

Degradation of starch by micro-organisms usually proceeds by the concerted action of a debranching enzyme and an amylase, resulting in the production of short linear oligosaccharides, or, in the case of CGTase, cyclodextrins. The further metabolism of these degradation products has been studied extensively in especially Gram-negative bacteria.
Chapter 1

The maltose regulon

No starch degrading enzymes have been reported in *Escherichia coli*, but the uptake and degradation of linear oligosaccharides, such as resulting from degradation of starch by the enzymes described above, have been studied in detail (Fig. 3a). The genes encoding the proteins involved in these processes are organized in the maltose regulon, consisting of at least three operons. In the mal A region the genes *malP*, *malQ*, and *malT* are located, encoding a maltodextrin phosphorylase, an amylomaltase, and a transcriptional activator protein, respectively. The mal B region contains the genes for the complete import system, consisting of a maltopin (*lamB*), located in the outer membrane; a maltose binding protein (*malE*), located in the periplasm; and the components of an ATP dependent transporter (*malFGK*), located in the cytoplasmic membrane. The third operon only contains a periplasmic α-amylase (*malS*), which probably enables the organism to grow on long chain oligosaccharides that can pass through the maltoporin, but are too big to be imported by the ATP dependent transporter. The operons described above are strongly regulated by catabolite repression (glucose effect). In the absence of an easily metabolizable compound they are transcribed constitutively and can be further induced by maltose, which binds to the transcriptional activator, resulting in its active conformation. Oligosaccharides can then enter the periplasm through the maltoporin and bind to the maltose binding protein, which targets them to the transporter in the cytoplasmic membrane. In the cytoplasm oligosaccharides consisting of 5 or more glucose residues (Gn) are converted by maltodextrin phosphorylase to glucose-1-phosphate (G1P) and G(n-1). Oligosaccharides smaller than 5 glucose units are converted to longer chains by the action of amylomaltase, catalyzing a disproportionation reaction resulting in the formation of glucose and a longer oligosaccharide, which can again be converted by maltodextrin phosphorylase (Schwartz, 1987).

The Cym operon

The components of the maltose regulon are also present in *Klebsiella oxytoca* (Fiedler et al. 1996). This organism is able to utilize starch as a sole carbon and energy source, making use of a cell surface associated pullulanase and an extracellular CGTase for the degradation of this high molecular weight substrate (Bender, 1977). The metabolism of cyclodextrins, resulting from the action of CGTase on starch, has been studied in detail and shows high homology to the metabolism of maltodextrins in *E. coli* (Fig. 3b) (Fiedler et al. 1996; Pajatsch et al. 1998). The genes encoding the proteins involved in cyclodextrin import and degradation are located in one region in the genome: the Cym operon. No distinct roles have been determined so far for 4 genes in this operon (*cymB,C,I* and *J*). One of the genes in this operon encodes the extracellular CGTase enzyme (*cgt*). CymA encodes a “cyclopin”, enabling cyclodextrins to penetrate the outer membrane. The gene product of *cymE* transports cyclodextrins across the periplasm and is homologous to the periplasmic maltose-binding protein (encoded by *malE*). The gene products of *cymF,G* and *D* are homologous to those of *malF,G* and *K*, respectively, and therefore likely to be components of an ATP dependent transporter for cyclodextrins. Finally, *cymH* is a gene encoding a cytoplasmic cyclomaltodextrinase (CDase), which converts cyclodextrins into maltose and maltotriose (Feederle et al. 1996). These oligosaccharides are further degraded to glucose and glucose-1-
CGTase, a biocatalyst with unique properties

phosphate by the action of maltodextrin phosphorylase and amylomaltase as described above. These two enzymes are the only gene products of the maltose regulon required for cyclodextrin metabolism (Fig. 3).

Figure 3. Uptake and processing of starch degradation products. a) maltodextrin metabolism by the gene products of the Mal operon; b) cyclodextrin metabolism by the gene products of the Cym operon (reproduced from Fiedler et al. (1996)).

A possible explanation for the existence of this complicated system is that, by producing cyclodextrins, the organism builds up an external storage form of glucose, not accessible for most other organisms because they are not able to metabolize cyclodextrins. Alternatively, cyclodextrins may protect bacterial cells against toxic compounds in the environment by forming inclusion complexes as in biological waste water treatment, where the addition of small amounts of β-cyclodextrin in activated sludges increases the tolerance level to toxic chemicals (Allegre and Deratani, 1994). Also the availability of compounds needed for growth may improve when present in an inclusion complex with cyclodextrins (Aeckersberg et al. 1991). These alternatives, however, do not explain the presence of the specific uptake and degradation routes for cyclodextrins.
Chapter 1

I.3. Starch as the versatile raw material

Starch can be used as a highly pure industrial raw material which is available in sufficiently large amounts. In Western Europe alone approximately 6.5 million tonnes of starch from maize, wheat, and potato are produced each year and used in the food (55 %) and non-food (45 %) sectors. It can be used as native starch, as modified starch, or as starch hydrolysates. Increasingly, the versatility of starch is shown by for instance the development of starch-derived fat replacers, carbohydrate-based detergent components, and starch copolymers as bioplastics. It has been suggested that the new millennium will carry us towards a plant-based economy. Conceivably, there is a lot of interest in starch degrading and modifying enzymes and the products resulting from their activities on starch.

The industrial importance of starch degrading enzymes

Increasingly, starch degrading enzymes are used for the production of starch hydrolysates, replacing the chemical methods used historically. These starch hydrolysates can be used for the production of a wide variety of glucose-based compounds or in fermentations for the production of for instance ethanol or lactic acid. Furthermore, microorganisms able to use starch as carbon and energy source can be used for the direct fermentation of starch, a method used for instance in the production of acetone and butanol by *Clostridium acetobutylicum* (Nigam and Singh, 1995).

Not only the versatility of starch, but also that of starch degrading enzymes and of the activities and products of individual enzymes contribute to the wide range of applications for starch. This will be further illustrated by the diversity of products from the action of CGTase on starch or starch derived substrates.

The industrial importance of CGTase

CGTase enzymes are applied for the production of cyclodextrins. Recent developments, however, also concentrate on the use of the CGTase catalyzed coupling and disproportionation reactions for the synthesis of modified oligosaccharides by using alternative acceptor substrates. Furthermore, applications of CGTase limit dextrins are being explored.

Applications of cyclodextrins

The glucose residues in the cyclodextrin rings (Fig. 4a) are arranged in such manner that the secondary hydroxyl-groups (C2 and C3) are located on one edge of the ring and the primary hydroxyl-groups (C6) on the other edge, resulting in torus shaped molecules (Fig. 4b). The apolar C3 and C5 hydrogens and ether-like oxygens are at the inside and the hydroxyl-groups at the outside of these molecules. This results in a molecule with a hydrophilic outside, which can dissolve in water, and an apolar cavity, which provides a hydrophobic matrix, enabling cyclodextrins to form inclusion complexes with a wide variety of hydrophobic guest molecules (Fig. 4c). Their three-dimensional form and size provide an important parameter for complex formation with hydrophobic compounds or functional groups (Table 1). Thus specific (α-, β-, or γ-)cyclodextrins are required for complexation of specific guest molecules. The driving force of inclusion complex formation is the entropic effect of displacement of water
molecules from the hydrophobic environment of the cavity, probably combined with the fact that this water causes strain on the cyclodextrin ring, which is released after complexation, producing a more stable, lower energy state (Saenger, 1980; Saenger, 1984). Because the inclusion complexes are quite stable they can be separated from the medium by crystallization (Starnes, 1990).

Figure 4. Structure and properties of cyclodextrins. a) α-, β-, and γ-cyclodextrins; b) 3-dimensional form and properties of cyclodextrins (for sizes of A (outer diameter) and B (inner diameter), see Table 1); c) formation of inclusion complex of a cyclodextrin with a hydrophobic molecule (reproduced from Penninga (1996)).

The formation of inclusion complexes leads to changes in the chemical and physical properties of the guest molecules (Table 2). These altered characteristics of encapsulated compounds have led to various applications of cyclodextrins (or their derivatives) in analytical chemistry (Armstrong, 1988; Luong et al. 1995), agriculture (Saenger, 1980; Oakes et al. 1991), biotechnology (Allegre and Deratani, 1994; Szejtli, 1994), pharmacy (Albers and Muller, 1995; Thompson, 1997), food (Allegre and Deratani, 1994; Bicchi et al. 1999), and cosmetics (Allegre and Deratani, 1994).
Chapter 1

Table 1. Cyclodextrin properties.

|                          | α-cyclodextrin | β-cyclodextrin | γ-cyclodextrin |
|--------------------------|----------------|----------------|----------------|
| number of glucopyranose units | 6              | 7              | 8              |
| molecular weight (g/mole)  | 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            |
| approx. cavity volume (Å³) | 174           | 262            | 427            |

Values from Uekama and Irie (1987) and Szejtli (1982). Outer (A) and inner (B) diameter are indicated in Figure 4.

Other CGTase applications

In addition to production of cyclodextrins through the cyclization reaction, CGTase can be used for its coupling and disproportionation reactions for the transfer of donor substrates such as cyclodextrins or starch to acceptor molecules. Increasingly, the use of alternative acceptors is reported, resulting in novel glycosylated compounds (Kometani et al. 1994; Kometani et al. 1996a; Kometani et al. 1996b). A commercial application of this method is found in glycosylation of the intense sweetener, stevioside. This bitter compound is isolated from the leaves of the plant Stevia rebaudiana and has a low solubility. Glycosylation decreases bitterness and increases solubility (Pedersen et al. 1995).

Other applications are found for the CGTase limit dextrins. Due to the inability of CGTase to bypass α(1-6) bonds in gelatinized starches, degradation of these substrates leads to a reduction in viscosity without a corresponding decrease of the high-molecular character of starch. This CGTase limit dextrin is applied in processes for surface sizing or coating of paper, to improve the writing quality of the paper and to obtain a glossy and well printable surface. The coating or sizing liquid contains converted starch which has been obtained by treating gelatinized starch with a starch-converting enzyme selected from the group of the CGTases and the branching enzymes (Bruinenberg et al. 1996). CGTases can also be used in the preparation of doughs for baked products which comprises incorporation of the CGTase into the dough to increase the volume of the baked product (van Eijk and Mutsaers, 1995).

Table 2. Possible effects of the formation of inclusion complexes on properties of the guest molecules.

- Stabilization of light- or oxygen-sensitive compounds
- Stabilization of volatile compounds
- Alteration of chemical reactivity
- Improvement of solubility
- Improvement of smell and taste
- Modification of liquid compounds to powders
Shortcomings of industrial applications of CGTase

Like most starch degrading enzymes, the CGTase from *B. macerans*, 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. Heating in water (jet cooking) weakens these hydrogen bonds and causes swelling and gelatinization (Nigam and Singh, 1995), resulting in a very viscous starch solution when performed at starch concentrations of industrial interest. Therefore, in this initial processing step, operating at temperatures up to 105-110 °C, an α-amylase is added in order to liquefy the starch to make it suitable for incubation at the lower temperatures (55 °C) required for the CGTase catalyzed production of cyclodextrins. Unfortunately, the α-amylase used for liquefaction produces maltodextrins, which will act as acceptor molecules in the coupling reaction catalyzed by CGTase, severely reducing cyclodextrin yields (Pedersen et al. 1995).

More recently, very thermostable CGTases have been characterized from thermophilic anaerobic bacteria belonging to the genera *Thermoanaerobacter* (Starnes, 1990; Norman and Jorgensen, 1992) and *Thermoanaerobacterium* (Wind et al. 1995). These CGTases are active and stable at high temperatures and low pH values, and are able to solubilize starch, thereby eliminating the need for α-amylase pretreatment, without any traces of low molecular weight oligosaccharides produced in the initial stages of the reaction (Starnes et al. 1991). The use of these thermostable CGTases has the added advantage that the total cyclodextrin production time can be shortened (Pedersen et al. 1995). The *Thermoanaerobacter* CGTase (maximal activities at 90 °C and pH 5.8) has found commercial application since 1996.

A major disadvantage of cyclodextrin production by CGTases is that all known wild type CGTase enzymes produce a mixture of α-, β-, and γ-cyclodextrin and are subject to inhibition by these cyclic products. The *Thermoanaerobacter* CGTase for instance produces an approximately equal mixture of α- and β-cyclodextrins with a small amount of γ-cyclodextrin. Two different industrial approaches are used to purify the produced cyclodextrins: selective crystallization of β-cyclodextrin (which is relatively poorly water-soluble) and selective complexation with organic solvents. These processes not only serve to purify the cyclodextrins, but also result in decreased product inhibition, enhancing the total conversion of starch from 40% to 60% (Bergsma et al. 1988). 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 solutions because of its high boiling point (229 °C). Cyclododecanone can be used for complexation and selective precipitation of γ-cyclodextrin, but this solvent is too expensive for commercial use. Further disadvantages of the use of organic solvents are their toxicity, their flammability, and the need for a solvent recovery process (Pedersen et al. 1995). The availability of α- and γ-cyclodextrins is thus rather limited at present; consequently, there is a great demand for a process that could produce these cyclodextrins economically. Also the processes used for β-cyclodextrin production are not ideal.

The processes described above make the production of cyclodextrins too costly for many applications, and the use of organic solvents limits applications involving human consumption. While the potential market for cyclodextrins in the USA is 32,000 Metric tons per year (Starnes, 1990), the actual world market for cyclodextrins for 1995 has been estimated...
as 5500 Metric tons only. This major gap will be overcome only when the prize of cyclodextrins is significantly reduced from US$ 15-20 per kg (for the cheapest (β) cyclodextrin) to US$ 5 per kg (Schmid, 1989). Clearly, the development of more economical cyclodextrin production processes is needed to expand the range of commercially successful technical applications of cyclodextrins.

The high energy costs for solubilization of starch, together with viscosity problems can be overcome by enzymes active on raw starch granules (Wijbenga et al. 1991). Screening thus far has yielded two different mesophilic micro-organisms capable of growth on native potato starch granules, a Microbacterium species and a B. firmus/lentus strain, producing native starch degrading amylases (Wijbenga et al. 1991).

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 above described expensive and environmentally harmful procedures involving organic solvents. This situation has strongly stimulated studies of CGTase structure-function relationships, with the mechanisms of the CGTase catalyzed reactions and inhibition by the cyclodextrin products as important research topics. In recent years detailed knowledge has become available, allowing rational design of mutant CGTase biocatalysts with improved cyclodextrin product specificity and reduced product inhibition (see below) (Pedersen et al. 1995; Dijkhuizen et al. 1996; Dijkhuizen et al. 1999).
II. The α-amylase family (family 13) of glycosyl hydrolases

CGTase is a member of the α-amylase family of glycosylases (family 13), an important group of starch converting enzymes. Enzymes belonging to this group show a wide diversity in reaction specificities, and many of them are active on starch. Whereas amylases generally hydrolyze glycosidic bonds in the starch molecules, CGTases mainly catalyze transglycosylation reactions, with hydrolysis being a minor activity (van der Veen et al. 2000c). Structure/function relationships in the α-amylase family have been studied extensively and may help to clarify the mechanistic basis of the unique activities of CGTase.

II.1. Sequence similarities in the α-amylase family

Although the overall sequence similarity within the α-amylase family of glycosylases (family 13) is relatively low (<30 %), four highly conserved regions have been identified in α-amylases by Nakajima et al. (1986). These regions were found to be also present in other members of the α-amylase family; α-glucosidases, pullulanases, isoamylases and CGTases (Svensson, 1994). An amino acid sequence alignment showing these four conserved regions for diverse members of the α-amylase family is presented in Figure 5.

|   | 132 | 140 | 223 | 233 | 253 | 260 | 324 | 332 |
|---|-----|-----|-----|-----|-----|-----|-----|-----|
| CGT| VIIDPAPNH | IDGIRMDAVKH | FTFGEWFL | IDNHDMERF |
| TAA | LMVDVVANH | IDGLRIDLTVKH | YCIGEVLD | VENHDNPFR |
| CD | VMIDAVFNH | IDGWRLDVAHN | YILGEIWH | LESHDSRNL |
| PUL | VIMDVVYNH | IDGFRFDLGMY | YFFGEGWD | VSKHDNLTL |
| ISO | VYMDVYNH | VDGFRFLASV | DLFAEPWA | IDVHDGNTL |

Figure 5. Amino acid sequence alignment of the four conserved regions for diverse members of the α-amylase family. TAA: α-amylase from Aspergillus oryzae (Taka-amylase A) (Matsuura et al. 1984); CGT: CGTase from Bacillus circulans strain 251 (Lawson et al. 1994); CD: cyclodextrinase from Klebsiella oxytoca (Fiedler et al. 1996); PUL: pullulanase from Klebsiella aerogenes (Katsuragi et al. 1987); ISO: isoamylase from Pseudomonas amylofermentans (Amemura et al. 1988). The residues are numbered according to the CGTase from Bacillus circulans strain 251. An asterisk indicates amino acid identity, a dot indicates amino acid similarity.

All four regions contain completely invariant amino acid residues within the α-amylase family and the functions of most of these have been elucidated by X-ray crystallography, site-directed mutagenesis, and chemical modification of various members of this family. These residues are directly involved in catalysis, either through substrate binding, bond cleavage, transition state
stabilization, or as ligands of a calcium binding site present near the active site. Three
carboxylic acid groups, one glutamic acid and two aspartic acid residues, were found to be
essential for catalytic activity in α-amylases and CGTases. The amino acids are equivalent to
Asp206, Glu230, and Asp297 in α-amylase from Aspergillus oryzae (Matsuura et al. 1984) and
Asp229, Glu257 and Asp328 in CGTase from B. circulans (Klein et al. 1992; Strokopytov et
al. 1995). Two conserved histidine residues, His140 and His327 (CGTase numbering), are
involved in substrate binding and transition state stabilization (Nakamura et al. 1993;
Uitdehaag et al. 1999b). A third histidine, present only in some α-amylases and CGTase
(His233, CGTase numbering), is involved in substrate binding and acts as a calcium-ligand
with its carbonyl oxygen (Lawson et al. 1994; Strokopytov et al. 1996). Arg227 is important
for the orientation of the nucleophile (Asp229, see below) (Uitdehaag et al. 1999b). The role
of Asp135 is not clear, but it is in close proximity of the catalytic site. Asn139 again is a
calcium-ligand. The importance of the calcium binding site is illustrated by the identification
of a fifth conserved region in α-amylases (Janecek, 1992) and, more recently, in several other
members of the α-amylase family (Janecek, 1995). This region consists of the stretch 197-
LADLN in CGTase from B. circulans strain 251 (173-LPDLD in the α-amylase from A.
oryzae) and contains the calcium-ligand Asp199.

II.2. Catalytic mechanism of the α-amylase family

The reactions catalyzed by the enzymes belonging to the α-amylase family proceed
with retention of the substrate’s anomeric (α-)configuration. Since each substitution at a chiral
center results in inversion of configuration, catalysis must proceed through a double
displacement reaction (Fig. 6) (Koshland, 1953). The first step involves a protonation of the
glycosidic oxygen by a general acid catalyst, creating an oxo-carbonium transition state which
subsequently collapses into an intermediate (McCarter and Withers, 1994; McCarter and
Withers, 1996). This intermediate is attacked by a water nucleophile (or the C4-OH at the non-
reducing end of another oligosaccharide in case of transglycosylases (e.g. CGTase)) in the
second step, assisted by the base form of the acid catalyst. The roles of the three carboxylic
amino acids in this mechanism have been clarified by X-ray crystallographic studies on α-
amylase (Qian et al. 1994) and CGTase (Strokopytov et al. 1995) with acarbose, a potent
pseudotetraose inhibitor, bound in the active site. Glu257 (CGTase numbering) is the general
acid catalyst, acting as proton donor; Asp229 serves as the nucleophile, stabilizing the
intermediate, and Asp328 has an important role in substrate binding. For retaining enzymes
the intermediate could either be an oxo-carbonium ion which is electrostatically stabilized by
a carboxylate, or involves formation of a covalent bond, in which one of the catalytic
aspartates is presumed to act as a nucleophile (see Fig. 6). Although initially the nature of the
intermediate was disputed, it is now generally accepted that the reaction proceeds via a
covalent intermediate. Clear evidence for a covalent glycosyl-enzyme intermediate in family
13 has been obtained from rapid trapping studies with natural substrates. Low-temperature
13C NMR experiments have provided evidence for the formation of a β-carboxylacetal ester
covalent adduct between maltotetraose and porcine pancreas α-amylase (Tao et al. 1989).
Conclusive evidence recently came from experiments involving trapping of a covalent intermediate with 4-deoxymaltotriosyl \(\alpha\)-fluoride as a substrate in the virtually inactive Glu257Gln mutant of \(B.\) circulans 251 CGTase (Mosi et al. 1997) and elucidation of the X-ray crystallographic structure of the enzyme with the covalently linked intermediate (Uitdehaag et al. 1999b).

\[ \text{Figure 6. Reaction mechanism of family 13 glycosylases as revealed by X-ray crystallographic studies of CGTase.} \]

\(\text{Retaining enzymes act via a double displacement mechanism. The first step involves a protonation of the glycosidic oxygen by a general acid catalyst (Glu257 in CGTase), creating an oxo-carbonium transition state which subsequently collapses into an intermediate covalently linked to the nucleophile (Asp229 in CGTase). This intermediate is attacked by the C4-OH at the non-reducing end of another oligosaccharide (or a water nucleophile in case of hydrolysis) in the second step, assisted by the base form of the acid catalyst (reproduced from Uitdehaag et al. (1999b)).} \]

\[ \text{II.3. Three-dimensional structure similarities in the \(\alpha\)-amylase family} \]

In contrast to a limited similarity in primary structure (< 30%), the three-dimensional structures of \(\alpha\)-amylases (Matsuura et al. 1984; Boel et al. 1990; Brady et al. 1991; Qian et al. 1993; Qian et al. 1994; Kadziola et al. 1998; Machius et al. 1995b; Machius et al. 1995a) and CGTases (Kubota et al. 1991; Klein and Schulz, 1991; Lawson et al. 1994; Knegtel et al. 1996; Harata et al. 1996) are quite similar. \(\alpha\)-Amylases generally consist of three structural domains, A, B, and C, while CGTases show a similar domain organization with two additional domains, D and E (see Fig. 7). Domain A contains a highly symmetrical fold of eight parallel \(\beta\)-strands arranged in a barrel encircled by eight \(\alpha\)-helices. This so-called (\(\beta/\alpha\)_\(\beta\))\_ or TIM-barrel catalytic domain (Janecek, 1994) of 300-400 residues is present in all enzymes of the \(\alpha\)-amylase family. The (\(\beta/\alpha\)_\(\beta\))-barrel was first found in the structure of chicken muscle triose-phosphate isomerase (TIM) (Banner et al. 1975), but it has been shown to be very wide-spread in functionally diverse enzymes (Svensson and Sogaard, 1991). Several prolines and glycines flanking loops
connecting the β-strands and α-helices have been found to be highly conserved in these enzymes (Janecek, 1996). The catalytic and substrate binding residues conserved in the α-amylase family are located in loops at the C-termini of β-strands in domain A. The loop between β-strand 3 and α-helix 3 of the catalytic domain is rather large and is regarded as a separate structural domain. This B-domain consists of 44-133 amino acid residues and contributes to substrate binding. The C-domain 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 maltose binding sites observed in the structure derived from maltose dependent crystals (Lawson et al. 1994). This maltose binding site was found to be involved in raw starch binding (Penninga et al. 1996), suggesting a role of the C-domain in substrate binding. Some authors suggest that this domain is involved in bond specificity, since in enzymes hydrolyzing or forming α-1,6-bonds (e.g. pullulanase, isoamylase, branching enzyme) the A-domain is followed by a different domain (see Fig. 7) (Jespersen et al. 1991). The D-domain, consisting of approximately 90 amino acids with an immunoglobulin fold, is almost exclusively found in CGTases and has an unknown function. The E-domain, following the D-domain in CGTases is more widespread in starch degrading enzymes. Besides in the α-amylase family, where it is found as the C-terminal domain when present, it is also found in glucoamylases (family 15 of glycosylases), where it is attached to the C- or N-terminus of the catalytic domain via a glycosylated linker (see Fig. 7). The E-domain consists of approximately 110 amino acids and was found to be responsible for the adsorption onto granular starch (see below).

**Figure 7. Domain level organization of starch degrading enzymes.** TAA: α-amylase from *Aspergillus oryzae* (Taka-amyrase A); CGT: CGTase from *Bacillus circulans*; G2A: maltogenic α-amylase from *Bacillus steaoterophilus*; G4A: maltotetraose forming α-amylase from *Pseudomonas stutzeri*; CD: cyclodextrinase from *Klebsiella oxytoxica*; ISO: isoamylase from *Pseudomonas amylofera*; PUL: pullulanase from *Klebsiella aerogenes*; GA: glucoamylase (family 15 of glycosylases) from *Aspergillus niger* (reproduced from Jespersen et al. (1991) with modifications).
II.4. Substrate binding of (α-)amylases and CGTases

The first important step in enzyme catalysis is binding of the substrate. In several starch degrading enzymes a separate domain responsible for absorption unto raw starch has been found. For *Aspergillus niger* two forms of glucoamylase (GA) have been described. GAI is 114 amino acids longer than GAII and was found to contain an additional C-terminal domain required for binding raw starch (Svensson et al. 1986b). Fusion of the corresponding domain of the *Aspergillus awamori* glucoamylase to the C-terminus of β-galactosidase resulted in a protein with binding affinity for corn starch and cross-linked amylose (Dalmia et al. 1995). Sequence comparisons between the E-domain of several amylases and CGTases and the raw starch binding domain from glucoamylases revealed the presence of this domain in various starch degrading enzymes (Svensson et al. 1989). Later, evidence for a starch binding site in CGTases separate from the active site was presented (Villette et al. 1992) and fusion of the E-domain of the *Bacillus macerans* CGTase to β-galactosidase demonstrated that it can indeed function as a starch binding domain (Dalmia et al. 1995). Studies on the CGTase from *B. circulans* strain 251 have revealed the function of the E-domain in more detail. High concentrations of maltose are required for crystallization of this CGTase (Lawson et al. 1990).

Figure 8. Ribbon drawing of the structure of the *Bacillus circulans* strain 251 CGTase. The different domains (A-E) and maltose binding sites (1-3) are indicated. The catalytic residues and bound maltose molecules are indicated as black sticks.
Three maltose binding sites (MBS) were observed at the protein surface (Fig. 8), two of which (MBS1 and MBS3) contribute to intermolecular crystal contacts. 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 revealed that maltose binding site 1 is important for (raw) 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 (see below) (Lawson et al. 1994; Penninga et al. 1996). The maltose binding sites on the CGTase E-domain were found to interact strongly with cyclodextrins and oligosaccharides (Knegtel et al. 1995). Also in the raw starch binding domain of glucoamylase from *Aspergillus niger* two sites interacting with maltoheptaose and β-cyclodextrin, similar to the maltose binding sites in the E-domain of CGTase, have been identified (Sorimachi et al. 1996; Williamson et al. 1997). Further experiments showed that also the roles of these binding sites are similar to those of CGTase (Sorimachi et al. 1996; Sigurskjold et al. 1998). Recent studies revealed that the raw starch binding domain has an additional function in the disruption of the structure of granular starch (Southall et al. 1999).

As mentioned above, the A-domain contains the catalytic residues of α-amylases and CGTases, while domain B is involved in substrate binding. X-ray crystallographic studies have revealed a groove on the surface of these enzymes formed on one side by loops of the A-domain and on the other side by the B-domain. In crystal structures from pig pancreatic α-amylase (PPA) (Larson et al. 1994) and CGTase from *B. circulans* strain 251 (Lawson et al. 1994), where maltose molecules serve as contact points between the enzyme molecules in the crystals, the functionality of this groove in substrate binding has been nicely shown. In PPA the maltose is bound at one end of the groove and the contact point in the crystal is formed by interactions of this maltose with the other end of the groove of the neighboring enzyme molecule (Larson et al. 1994). Soaking of these crystals with α-cyclodextrin revealed three binding sites for this cyclic compound. The first α-cyclodextrin replaces the maltose serving as the contact point between the amylase molecules. The second binding site was found in the middle of the groove, in close proximity to the catalytic residues. The third α-cyclodextrin is further removed from the substrate binding groove and bound in a slight depression formed by an edge of the B-domain and the first turn of α-helix 3 of the A-domain. Interestingly, the depression in which the third α-cyclodextrin binds in PPA (Larson et al. 1994) corresponds to a region in CGTase which is involved in interactions between the catalytic domain (A) and the starch binding domain (E) (Harata et al. 1996). From soaking experiments with the CGTase from *B. circulans* strain 251 the structure of the enzyme with a maltononaose inhibitor was obtained (Fig. 9), revealing in more detail the mode of substrate binding in the groove (Strokopytov et al. 1996). The substrate binding sites are numbered +2 to -7 (numbering according to Davies et al. (Davies et al. 1997)), with the catalytic site between subsites +1 and -1. The non-reducing end of the oligosaccharide is bound at subsite -7, which agrees with the formation of mainly β-cyclodextrin from starch by the enzyme. The glucose residue at subsite -7 is located at the end of the substrate binding groove, interacting with amino acid residues of the B-domain. These residues correspond to the region in PPA where the first α-cyclodextrin, bound at the beginning of the groove in one enzyme molecule, interacts with the end of the groove in the neighbouring molecule. The glucose residues bound at subsites +1 and
CGTase, a biocatalyst with unique properties

II.5. Product specificity of α-amylases and CGTases

Individual glycosyl residues of an oligosaccharide bind at highly specific subsites in the active site cleft of the protein. At each subsite, binding energy is generated by hydrogen-bonds with the OH-groups of the carbohydrates, Van der Waals interactions with aromatic residues, or by the hydrophobic effect from displacement of bound water molecules (Johnson et al. 1988; Quiocho, 1986; Quiocho, 1989). To determine the contribution to the Gibbs free energy (∆G) of the different subsites, and to determine which amino acid residues interact with the glycosyl residues at these subsites, two methods are being applied; firstly, the hydroxyl groups of the ligands can be exchanged with hydrogen atoms to measure the individual contribution of these hydroxyl groups (Adelhorst and Bock, 1992), and secondly, site-directed mutants can be compared with the wild-type protein (Nakamura et al. 1993). Furthermore, a kinetic and product formation analysis involving substrates of different length can be used to calculate the number and positions of the subsites. For CGTase this latter method indicated a total of 9 subsites for both the CGTase of *K. pneumoniae*, forming mostly α-cyclodextrin (with 6 glucose residues in the ring), and the CGTase from *B. circulans* strain 8, forming mostly β-cyclodextrin (with 7 glucose residues in the ring) (Bender, 1990). Analysis of the preference of the products formed suggested that for the CGTase of *B. circulans* strain 8 these 9 subsites range from +2 to -7, as confirmed by the structure of the maltononaose inhibitor bound in the active site of the β-CGTase from *B. circulans* strain 251 (Fig. 9) (Strokopytov et al. 1996). For the α-CGTase of *K. pneumoniae* the 9 subsites were predicted to range from +3 to -6. Binding up to subsite +3 has also been observed in the structure of the α-CGTase from *Thermoanaerobacterium thermosulfurigenes* complexed with a maltohexaose inhibitor (Wind et al. 1998). A similar method, using 4-nitrophenyl-α-maltooligosaccharides of different lengths as substrates, revealed 5 high affinity subsites in PPA, ranging from +2 to -3 (Ajandouz and Marchismouren, 1995). From these experiments it appears likely that the number and positions of sugar binding subsites determine the differences in product specificity between individual α-amylases and CGTases. For example PPA, with an active site consisting of five subsites, produces mainly maltose and maltotriose, whereas α-amylase from barley contains at least ten subsites and yields mainly maltose, maltohexaose and maltoheptaose. For TAKA-amylase a substrate binding model was proposed involving six or seven glucose residues (Nitta et al. 1971; Matsuura et al. 1984), which has recently been confirmed by the structure of *Aspergillus oryzae* α-amylase complexed with an acarbose derived maltohexaose inhibitor binding from subsites +3 to -3 (Brzozowski and Davies, 1997). Increasingly, X-ray crystallographic studies of protein-carbohydrate complexes result in the identification of the protein-ligand interactions, also providing information about factors determining the carbohydrate substrate and product specificities of different enzymes. Amino acids on loops
in the A-domain (including the B-domain), linking the C-terminal end of a β-strand to the N-terminal end of the adjacent α-helix, form the subsites of the active site. The number of subsites can be changed by changing the length and folding of the loops of the (β/α)_{8}-barrel (MacGregor, 1993), or by changing specific amino acids using site directed mutagenesis (Matsui et al. 1992a; Matsui et al. 1992b; Matsui and Svensson, 1997; van der Veen et al. 2000b).

Figure 9. Schematic representation of the hydrogen bonds between the *B. circulans* strain 251 CGTase and a maltononaose inhibitor bound at the active site. The subsites are numbered according to the general subsite labeling scheme recently proposed for all glycosyl hydrolases (Davies et al. 1997) (reproduced from Strokopytov et al. (1996), with modifications).
III. Cyclodextrin glycosyltransferases

III.1. CGTase catalyzed reactions

Whereas α-amylases generally hydrolyze α(1-4) glucosidic bonds (Fig. 10.a), CGTases mainly catalyze transglycosylation reactions. Such reactions can be described as: G(n) + G(m) → G(n-x) + G(m+x) in which G(n) is the donor and G(m) the acceptor oligosaccharide consisting of n and m glucose residues, respectively. Disproportionation (Fig. 10.b) can be regarded as the default reaction, and is also catalyzed by several other members of the α-amylase family (e.g. 4-α-glucanotransferase, EC 2.4.1.25 (amylo maltase, disproportionating enzyme)). The specific CGTase reaction is the cyclization reaction (Fig. 10.c) in which the part of the donor that has been cleaved off also acts as the acceptor, resulting in formation of a cyclodextrin, described as: G(n) → cyclicG(x) + G(n-x). The reverse reaction is also catalyzed by the enzyme and is referred to as the coupling reaction (Fig. 10.d).

Figure 10. Schematic representation of the CGTase catalyzed reactions. The circles represent glucose residues; the white circles indicate the reducing end sugars. (a): hydrolysis, (b): disproportionation, (c): cyclization, (d): coupling.

CGTases and some α-amylases (e.g. maltotetraose forming α-amylase from *Pseudomonas stutzeri* (EC 3.2.1.60, G4α, (Robyt and Ackerman, 1971)) and maltogenic α-amylase from *B. stearothermophilus* (EC 3.2.1.133, G2α, novamyl, (Outtrop and Norman, 1984)) have been described as exo-acting enzymes degrading starch molecules from their non-reducing ends, whereas most α-amylases are endo-acting enzymes cleaving α(1→4) glucosidic bonds more randomly in the starch molecules. True exo-acting enzymes like β-amylase and glucoamylase (family 14 and 15 of glycosylases, respectively) have substrate binding clefts which are closed to one side by specific loops, forcing binding of the non-reducing end of the
substrate in the active site. In the crystal structure of G4α a similar feature has been observed (Morishita et al. 1997). However, the open groove found in CGTases and other α-amylases should allow binding of extended substrates in the active site, consequently resulting in an endo-type activity. Analysis of the action on amylose by CGTase, forming large cyclic α(1-4) glucans (Terada et al. 1997), and G2α (Christopherson et al. 1998) indeed showed that these enzymes attack this high molecular weight substrate in an endo-like manner. The seemingly exo-type action reported for these enzymes must therefore result from the preferred use of low molecular weight and/or highly branched substrates in laboratory activity assays, which are easier to work with than high molecular weight amylose. Interestingly, CGTases and also G4α and G2α all possess the E-domain, which is absent in most α-amylases (see Fig. 7). The strong interaction of this domain with starch and various oligosaccharides (Knegtel et al. 1995; Strokopytov et al. 1996; Penninga et al. 1996) may cause some physical constraints in the degradation of these compounds, which also could lead to the seemingly exo-type of attack. Furthermore, binding of amylpectin to the E-domain, or the mere presence of this domain, may limit the accessibility of the regions between branching points leading to incomplete degradation of starch, which can also be interpreted as an exo-type of attack. No experiments supporting or rejecting this hypothesis have been described so far.

The formation of large cyclic α(1-4) glucans from amylose by CGTase was also interpreted as opposing generally held views on cycloextrin product specificity of CGTases (see below). The same study, however, also revealed significantly higher peaks for α- and β-cycloextrin in the reaction with α- and β-CGTase, respectively, present even in the early stages of the reaction (Terada et al. 1997). Furthermore, the action of potato disproportionating enzyme (DE) (Takahara et al. 1996) and B. stearothermophilus branching enzyme (BE) (Takata et al. 1996) on amylose indicate that the CGTase specific cyclization reaction is not required for the formation of the large cyclic products. DE catalyzes intermolecular transglycosylation reactions similar to the disproportionation reaction of CGTase, while BE cleaves an α(1-4) bond in one oligosaccharide molecule or starch chain and links the cleaved off part via an α(1-6) linkage to another molecule or chain. Although for neither enzyme intramolecular transglycosylation reactions had been reported before, both enzymes were found to produce large cyclic α(1-4) glucans similar to those formed by CGTase (with one α(1-6) bond in the ring for BE) (Takahara et al. 1996; Takata et al. 1996). In these experiments low concentrations of high molecular weight amylose (0.4, 0.2, and 0.3 % for CGTase, DE, and BE, respectively) were used, amounting to concentrations in the µM range. Therefore the “cyclization of amylose molecules” is not necessarily a novel reaction catalyzed by the enzymes, but is a direct effect of the limited availability of acceptor molecules. For DE and BE the smallest cyclic glucans formed consisted of 17 and 18 glucose residues, respectively, indicating that the specific CGTase cyclization reaction is only required for production of smaller cyclic oligosaccharides (cycloextrins: 6-8 glucose residues mainly). Although the preferred use of low molecular weight and/or highly branched substrates for the determination of cycloextrin formation and the rather simple HPLC methods generally used for the detection of the produced cycloextrins have probably limited observations of large cycloextrins, production of δ-, ε-, ζ-, and η-cycloextrins (consisting of 9, 10, 11, and 12 glucose residues, respectively) from starch has been reported (Pulley and French, 1961; Penninga et al. 1995).
III.2. CGTase versus α-amylase action

Since the first description of a *B. macerans* strain capable of producing cyclodextrins from starch (Schardinger, 1911), numerous CGTase enzymes, mostly from gram positive bacteria, have been purified and characterized. The question what precisely determines the difference in reaction specificity between α-amylases and CGTases (see Fig. 10) has received much attention. When retrieving CGTase amino acid sequences from the SWISS-PROT/EMBL protein data base the following description of CGTases is kindly provided:

“CGTase may consist of two protein domains: the one in the amino-terminal side cleaves the alpha-1,4-gluconic bond in starch, and the other in the C-terminal side catalyzes other activities, including the reconstitution of an alpha-1,4-gluconic linkage for cyclizing the maltooligosaccharide produced.”

Indeed, the major difference between CGTases and α-amylases is the presence of additional C-terminal domains in the former enzyme. It has been hypothesized that these additional domains are involved in catalyzing the formation of cyclodextrins (Kimura et al. 1987). Experiments with CGTase from alkalophilic *Bacillus* sp. 1011 of which 10 or 13 amino acid residues were deleted from the C-terminus were considered to support this hypothesis (Kimura et al. 1989). Later experiments, however, failed to confirm these findings (Hellman et al. 1990; Bender, 1990) and indicated that the observed effects may have been caused by interference of the deletions with the structural integrity of the enzyme (Hellman et al. 1990).

As explained above, the C-terminal domain (E-domain) is responsible for binding to (raw) starch. The differences in reaction specificity between CGTases and α-amylases, therefore, appear to be based on specific differences in the active centers.

III.3. Sequence similarities in CGTases

In general, CGTases show a clear similarity in amino acid sequence, ranging from 47 to 99%, which should be sufficient to allow identification of residues responsible for the differences in α-amylase and CGTase action. Fig. 11 shows the amino acid sequence of the *Bacillus circulans* strain 251 CGTase and the conserved residues deduced from a sequence alignment of the 21 CGTases listed in Table 3. The structural features (α-helices and β-sheets) of the enzyme are indicated to allow a thorough comparison with α-amylases, which are the members of the family 13 of glycosylases most closely related to CGTases. For this comparison an alignment including 30 amylases from various sources (fungi, plants, bacteria) performed by Finn Drablos (personal communication) was used. The first specific CGTase residues (unique and completely conserved) are found between β-strand 1 and α-helix 1 and consist of the stretches 27DG and 32NNPXG and the single residues 46L and 53D. Of these residues only Pro34 is not completely conserved; it is absent in the CGTase of *K. pneumoniae*, the most dissimilar of the CGTases included in the alignment. Asp27, Asn32, Asn33, and Asp53 are ligands of a calcium binding site observed in CGTases (Lawson et al. 1994; Klein and Schulz, 1991), but not in α-amylases. Gly28, Pro34 and Gly36 probably serve as structural support for this calcium binding site. Leu46 is not involved in calcium binding, but
Table 3. CGTases used in the amino acid sequence alignments. Indicated are the sources, the main product formed by the enzyme and the reference were the enzyme characteristics are described. Sequences are obtained from the SWISS-PROT/EMBL protein data base.

| Abbreviation | Bacterial source                     | Main product | Reference                                |
|--------------|--------------------------------------|--------------|------------------------------------------|
| KLEPN        | K. pneumoniae strain M5a1             | α            | (Binder et al. 1986)                     |
| BMAC2        | B. macerans                          | α            | (Sakai et al. 1987)                      |
| BMACE        | B. macerans strain NRRL B388          | α            | (Fujihara et al. 1992b)                  |
| TBNVO        | Thermoanaerobacter sp. ATCC53,627     | α/β          | (Dijkhuizen et al. 1996)                 |
| TABIUM       | T. thermosulfurigenes EM1             | β/α          | (Wind et al. 1994)                       |
| BSTEAM       | B. stearothermophilus strain NO2      | α/β          | (Fujihara et al. 1992b)                  |
| BLICH        | B. licheniformis                      | α/β          | (Hill et al. 1990)                       |
| BCIR8        | B. circulans strain 8                | β            | (Nitschke et al. 1990)                   |
| BC192        | B. circulans strain E192             | β            | (Bovetto et al. 1992)#                   |
| B663         | Bacillus sp. strain 6,6,3             | β            | *                                       |
| BF2          | B. circulans strain F2               | β            | (Kim et al. 1992)                        |
| BC251        | B. circulans strain 251              | β            | (Lawson et al. 1994)                     |
| B1018        | Bacillus sp. strain B1018            | β            | (Itkore et al. 1990)                     |
| B1011        | alkalophilic B. sp. strain 1011       | β            | (Kimura et al. 1987)                     |
| B382         | alkalophilic B. sp. strain 38,2       | β            | (Hamamoto et al. 1987; Kaneko et al. 1988) |
| B171         | alkalophilic B. sp. strain 17,1       | β            | (Kaneko et al. 1989)                     |
| BKC201       | Bacillus sp. strain KC201            | β            | (Kitamoto et al. 1992)                   |
| BSP11        | alkalophilic B. sp. strain 1,1        | β (no α)     | (Schmid et al. 1988)                     |
| BOHB         | B. ohbensis (strain C-1400)           | β (no α)     | (Sin et al. 1991)                        |
| BREV         | Brevibacillus brevis strain CD162     | γ/β          | (Kim et al. 1998)                        |
| BF290        | B. firmus/tentus strain 290-3        | γ/β          | (Englbrecht et al. 1988)#                 |

# Sequence obtained from Roquette comp.

* Akhmetzjanov, A.A., ENTREZ-NCBI seq ID: 39839 (1992)

the neighbouring residue 47 is involved in binding (semi)cyclic oligosaccharides (see below) and is typically an Arg, Lys, or His in CGTases. More unique CGTase residues are found in the B-domain: Phe136 (Tyr in K. pneumoniae), Phe/Tyr151, Glu153, Gly165 (Thr in K. pneumoniae, but typically an aromatic amino acid in α-amylases), Tyr167, Phe175 (Tyr in K. pneumoniae), His177, Gly180, and the stretch 192-K/R,N,L,F/Y,D-196, of which only Leu194 is observed also in α-amylases. In K. pneumoniae this stretch starts with His and ends with Asn, but these are conservative modifications, which are quite different from the corresponding residues in α-amylases. Also residues Ile190 and Tyr191 are conserved in all CGTases, except in the enzyme from K. pneumoniae. Noticably, Tyr191 forms a contact between domains B and D (Harata et al. 1996) and domain D is partially deleted in the CGTase from K. pneumoniae. Apart from amino acid residue 195, usually a smaller amino acid (Gly, Ser or Val) in α-amylases, the functions of none of the conserved residues in the B-domain has been studied thus far. The phenyl group of residue 195 is located at the center of the CGTase active site and thus might be involved, by hydrophobic interaction with the carbohydrate residues, in bending the non-reducing end towards the reducing end of the bound oligosaccharide, resulting in cyclodextrin formation. Mutations at this position (Fujihara et al. 1992a; Nakamura et al. 1994a; Penninga et al. 1995; Wind et al. 1998) indeed showed that
CGTase, a biocatalyst with unique properties

it is important for the cyclization process. However, even Tyr195Gly mutations retained 10-25% of cyclization activity, indicating that the aromatic character of this residue is not crucial for this reaction. An alternative role for Tyr195 might be keeping water from the active site, thus preventing hydrolysis (Penninga et al. 1995). Mutating Tyr(Phe)195 into residues found at the corresponding position in α-amylases, however, did not result in drastically increased hydrolyzing activity (Penninga et al. 1995; Wind et al. 1998). Relatively few unique residues are found in the second part of the A-domain: Tyr210, Trp218, Ile226, and Trp258. The last residue is positioned directly next to the catalytic Glu257 and is usually a small hydrophobic residue (A, V, I, or L) in α-amylases. Mutation of this residue in the CGTase from B. stearothermophilus (W254V) resulted in a fourfold decrease in cyclization activity, while the hydrolyzing activity was hardly affected (Fujiwara et al. 1992a).

Figure 11. Amino acid sequence of the catalytic domain of B. circulans strain 251
CGTase. Residues indicated below the sequence are conserved in CGTases as determined by an alignment of the CGTases listed in Table 3; @ = Tyr or Phe; . = conserved replacements in other CGTases; α1-α8 indicate the position of the α-helices; β1-β8 indicate the position of the β-sheets. Bold residues are completely conserved, underlined residues are unique for CGTases, italic residues are (completely) conserved in CGTases and at least one group of α-amylases, italic underlined residues are (completely) conserved in all CGTases and α-amylases.
Chapter 1

III.4. CGTase three-dimensional structures

The increasing availability of X-ray crystallographic structures of CGTase proteins, especially with inhibitors, substrates, or products bound at the active site, can provide more insights into the factors determining the unique cyclodextrin producing activity of these enzymes. The first requirement for the formation of cyclodextrins is binding of a substrate of sufficient length in the active site. As described above, this requirement is well met in CGTases and the structure of the maltononaose inhibitor bound in the active site has allowed determination of the subsite architecture in the substrate binding groove (see Fig. 9) (Strokopytov et al. 1996). Especially subsites +1,-1 and -2, where the bond-cleavage process takes place, have an architecture identical to A. oryzae α-amylase (Matsuura et al. 1984; MacGregor, 1993), except for the presence of Tyr195, located at the center of the active site. Of the amino acid residues forming subsites -3 to -7 Asp196, hydrogen bonding to the glucose residue at subsite -3, and Asn193, hydrogen bonding to the glucose residue at subsite -6, have been identified as specific CGTase residues. These amino acids thus may specifically contribute to the cyclization reaction, but no mutants clarifying their roles have been reported, and their involvement therefore is unclear. The amino acid residues interacting with the sugar residue at subsite -7 (Ser145, Ser146, Asp147) are not specifically conserved in CGTases. The loop of the B-domain in which they are located, however, contains proline 143, which is present in most CGTases, but not in α-amylases (see Figs. 11 and 12). At subsite +2 a special binding mode of the glucose residue is observed. Besides hydrogen bonding to Lys232 (almost completely conserved in CGTase, but also at least functionally conserved in α-amylases) hydrophobic interactions with both Phe183 and Phe259 are observed. Although these residues have been reported to be typical for CGTases (Nakamura et al. 1994a), in most α-amylases at least one and many times both of the corresponding residues is highly hydrophobic (Phe, Tyr, or Trp). For instance in PPA the residue corresponding to Phe183 is Tyr151, and this residue interacts with the 2-cyclodextrin bound near the catalytic residues (Larson et al. 1994). In A. oryzae α-amylase, complexed with an acarbose-derived maltotetraose inhibitor bound in the active site, the hydrophobic moiety of Leu232 (equivalent to Phe259 in CGTases) has stacking interactions with the glucose residue at subsite +2 (Brzozowski and Davies, 1997). This indicates that also at subsite +2 substrate binding of CGTase is similar to that of α-amylase. Mutation Phe183Leu in the CGTase from alkalophilic Bacillus sp. 1011 resulted in fourfold and sixfold decreases in cyclization and starch degrading activities, respectively (Nakamura et al. 1994a). Mutation Phe259Leu resulted in a similar decrease in cyclization activity and a threefold decrease in starch degrading activity, while a similar mutation in the CGTase from Bacillus stearothermophilus (Phe255Ile) resulted in complete removal of the cyclization activity, with a fourfold reduction in the starch degrading activity and a doubling of the saccharifying activity (hydrolysis) (Fujiiwara et al. 1992a). A double mutant Phe183Leu/Phe259Leu in the CGTase from alkalophilic Bacillus sp. 1011 displayed a cyclization activity which was 0.5 % of that of the wild type, which is a much larger decrease in activity than would be predicted from the combination of the two single mutations (Nakamura et al. 1994a). Probably these residues play a cooperative role in binding the non-reducing end of the linear chain when it assumes the circular conformation required for
CGTase, a biocatalyst with unique properties

More recently two structures of CGTases with cyclodextrin products bound in the active site have been elucidated (Schmidt et al. 1998; Uitdehaag et al. 1999a). A γ-cyclodextrin bound in the active site of the CGTase from B. circulans strain 251 revealed a binding mode similar to that of a linear maltoooligosaccharide at subsites +1 to -2, resulting in distortion of the cyclodextrin ring. At subsite +2 the hydrophobic interactions of the glucose residue with Phe183 and Phe259 are modified; whereas the linear substrate has better stacking interactions with Phe183, the cyclodextrin product stacks better with Phe259. Furthermore, the hydrogen bonding interaction with Lys232 is absent in the cyclodextrin structure. These findings support a specific role of the residues at subsite +2. At subsite -3 the glucose residue of the γ-cyclodextrin hydrogen bonds with Arg47, an interaction not observed with linear oligosaccharides (Uitdehaag et al. 1999a). As mentioned above, residue 47 is functionally conserved in CGTases (Arg, Lys, or His) and may therefore play a role in the specific CGTase catalyzed reactions. A similar interaction between Lys47 and a glucose at subsite -3 was found in the structures of mutant Glu257Ala of the CGTase from B. circulans strain 8 with a β-cyclodextrin bound in the active site (Schmidt et al. 1998) and of the Thermus thermosulfurigenes strain EM1 CGTase with a maltohexaose inhibitor bound in a semicyclic conformation in the active site (Wind et al. 1998). Site-directed mutagenesis experiments have revealed the involvement of Arg47 in the conformational change of the oligosaccharide chain during the cyclization and coupling reactions catalyzed by CGTase (van der Veen et al. 2000a).

III.5. Cyclodextrin product specificity of CGTases

It has been suggested that the size of the aromatic amino acid (Phe or Tyr), present in a dominant position in the center of the active site cleft of CGTases (see above), influences the preferred cyclodextrin size. Sin et al. (Sin et al. 1993) proposed a mechanism in which the starch chain folds around this residue. Substitution of this central amino acid by a tryptophan, Tyr188Trp in the B. ohbensis CGTase (Sin et al. 1994) and Tyr195Trp in the B. circulans strain 8 CGTase (Parsiegla et al. 1998) indeed doubled the relative production of γ-cyclodextrin. However, several other Tyr188 mutations (Sin et al. 1994), as well as the substitution of Tyr195 of the B. circulans 251 CGTase by other amino acids (Penninga et al. 1995), and the mutation F191Y at the similar position in the CGTase of B. stearothermophilus NO2 (Fujiiwara et al. 1992a) do not support this proposed mechanism. Furthermore, natural α-, β-, and γ-CGTases all have Tyr or Phe at this position, indicating that this residue is not involved in the differences in product specificity observed for these naturally occurring enzymes.

As explained above product specificity of α-amylases and CGTase may depend largely on the number of subsites available for binding glucose units in the active site. The structure of the CGTase from B. circulans 251 complexed with a maltononaose inhibitor (Strokopytov et al. 1996) (Fig. 9) has revealed several amino acid residues involved in binding of the maltononaose. Most of these amino acids have been shown to be conserved in all CGTases (Phe183, Phe259, Asn193, Asp196), all CGTases and several α-amylases (Lys232,
His233, Asp371, Arg375), or in the whole α-amylase family of glycosylases (Asp229, Glu257, His327, Asp328) (see above). These residues are therefore not considered to be involved in product specificity of natural CGTases, although mutations can result in altered product specificity as has been shown for the histidins (Nakamura et al. 1993) and the phenylalanines (Nakamura et al. 1994a). Less conserved are the residues involved in the strong hydrogen bonding network between the enzyme and the glucose bound at subsite -7 (Fig. 12). Mutation Ser146Pro, aimed at disturbing this hydrogen bonding network, resulted in a decreased preference for β-cyclization (van der Veen et al. 2000b), confirming that the inhibitor binding mode resembles the mode of substrate binding required for the formation of β-cyclodextrin. The residues involved in this hydrogen bonding network are located in a loop at the start of the B-domain, between the completely conserved residues His140 (conserved in the whole α-amylase family) and Glu153 (unique for CGTases) (Fig. 11).

Pro143 is highly conserved in this loop and only present in CGTases, indicating that it may have an important role in preserving a suitable loop conformation. The absence of this proline and the much shorter loop in the CGTase of K. pneumoniae again indicate structural differences of this enzyme and do not allow a functional comparison with the other CGTases. The differences in product specificity of the CGTases are reflected in subtle differences in the amino acids following Pro143. In CGTases producing predominantly β-cyclodextrin these are either ASSD or AMET, while in CGTases producing more equal amounts of α- and β-cyclodextrin the stretch ASET is found, and in CGTases producing little or no α-cyclodextrin ALET. The primarily γ-cyclodextrin forming CGTase from B. firmus/lentus strain 290-3 completely lacks the residues in this region involved in substrate binding. This situation has recently been copied into the β-CGTase from B. circulans strain 8 by replacing residues 145-

Figure 12. Amino acid sequence alignment of the region 140-153 in CGTases. Included are the CGTases listed in Table 3. The CGTases are ordered according to their cyclodextrin product specificity; α-CGTases at the top, γ-CGTase at the bottom (Table 3).
CGTase, a biocatalyst with unique properties

151 by a single aspartate (Δ(145-151)D mutant) (Parsiegla et al. 1998), resulting in the stretch HTSPADAE (HTSPVDIE in the B. firmus/lentus strain 290-3 CGTase, see Fig. 12), which indeed resulted in increased γ-cyclodextrin production. The higher γ-cyclization activity of this mutant has been explained by inducing a further opening of the active site cleft to produce more space for the bound glucosyl chain (Parsiegla et al. 1998). However, from the structure of the maltonaose inhibitor and the comparison with α-amylases (see above), it is evident that the loop region 145-151 flanks the end of the substrate binding cleft, which opens up to the medium, indicating that there is no space needed to accommodate additional sugar residues. An alternative explanation for the increased production of γ-cyclodextrin by this mutant can be deduced from the action of CGTases on amylose (see above). Although large cyclic α(1-4)glucans were formed, the preference for formation of α- or β- cyclodextrin was clearly visible, even in the early stages of the reaction (Terada et al. 1997). For the B. firmus/lentus strain 290-3 wild type and the B. circulans strain 8 Δ(145-151)D mutant CGTases, missing the residues involved in product specificity in loop region 145-151, cyclodextrin production will therefore automatically shift to formation of the larger cyclodextrins; not by the creation of more space for the bound glucosyl chain, but due to the lack of specific interactions. These findings support the involvement of this loop region in cyclodextrin product specificity. However, the small differences in this loop between enzymes with different specificities and the fact that the Ser146Pro mutation in the B. circulans strain 251 CGTase did not result in a shift in product specificity comparable to the naturally occurring variation, although the hydrogen bonding network at subsite -7 was effectively disturbed (van der Veen et al. 2000b), suggest that there is more to cyclodextrin product specificity than modified subsite specificities.

A second region which may be involved in product specificity is found at subsite -3. Tyr89 in the maltonaose structure (Fig. 9) has hydrophobic interactions with the glucose bound at this subsite, an interaction also observed between Tyr75 and the glucose bound at subsite -3 in the structure of A. oryzae α-amylase complexed with an acarbose derived maltohexaose inhibitor (Brzozowski and Davies, 1997). In the more thermostable CGTases this residue is typically an aspartate, which has been shown to create a novel salt bridge with Lys47 in the structure of the T. thermosulfurigenes EM1 CGTase (Knegtel et al. 1996). Actually the whole loop containing residue 89 is remarkably different in these enzymes (Fig. 13) and has been proposed to contribute with novel hydrogen bonds and apolar contacts to the stabilization of the T. thermosulfurigenes EM1 CGTase (Knegtel et al. 1996). That these differences may also inflict changes in product specificity has been shown in the CGTase from B. circulans strain 251, where mutation Tyr89Asp resulted in a slight shift towards α-cyclodextrin production (van der Veen et al. 2000b). A Tyr89Ser mutation in the CGTase from alkalophilic Bacillus strain I-5, however, did not affect production profiles. Mutant Tyr89Phe of the same enzyme showed enhanced β-cyclodextrin specificity, but also resulted in a decreased conversion of starch into cyclodextrins, while an Asn94Ser mutation in the same loop (resembling the situation in the thermophilic enzymes) enhanced α-cyclodextrin specificity, and resulted in an increased conversion of starch into cyclodextrins (Kim et al. 1997). The α-CGTases from the B. macerans strains, however, are very similar to the β-CGTases in this loop region. The only remarkable difference is the substitution of the aromatic
residue 84, which is also replaced in the α-CGTase of *K. pneumoniae* and the γ-CGTase from *B. firmus/lentus* strain 290-3 (see Fig. 13), but otherwise a residue unique for CGTases. This residue may therefore be specifically involved in β-cyclization. Another characteristic of this loop is the stretch 78-QPV-80 immediately following β-strand 2, of which Gln78 is a residue unique for CGTases. Again the *K. pneumoniae* α-CGTase and the *B. firmus/lentus* strain 290-3 γ-CGTase are exceptions, specifically containing the sequence PPI for this stretch in both enzymes. Another similarity between these two enzymes is the fact that the loop is shorter than in most other CGTases. The α-CGTase of *K. pneumoniae*, however, shows clear homology to the α/β-CGTase of *B. stearothermophilus*, whereas the γ-CGTase from *B. firmus/lentus* strain 290-3 is homologous to the CGTases producing virtually no α-cyclodextrin, which are the only other CGTase enzymes showing the same reduction in loop size (Fig. 13).

| 78          | 89        | 97              |
|-------------|-----------|-----------------|
| KLEPN       | PPIENVNNT--DAAG--NTGY  |
| BMAC2       | QPVENITAVINY-SGVN-NTAY  |
| BMACE       | QPVENITSVIKYG-SGVN-NTSY  |
| TBNovo      | QPVENYAVLDP-STGGSTSY  |
| TABIUM      | QPVENYAVLDP-STGGSTSY  |
| BSTEa       | QPVENFVMNDASG--ASY  |
| BLICH       | QPVENIFATINY-SGVT-NTAY  |
| BCIR8       | QPVENIFATINY-SGVT-NTAY  |
| BC192       | QPVENIFATINY-GGVT-NTAY  |
| B663        | QPVENIFATINY-SGVT-NTAY  |
| BF2         | QPVENYSVINYSGVN-NTAY  |
| BC251       | QPVENYSVINYSGVN-NTAY  |
| B1018       | QPVENYSVINYSGVN-NTAY  |
| B1011       | QPVENYSVINYSGVN-NTAY  |
| B382        | QPVENYSVINYSGVN-NTAY  |
| B171        | QPVENYSVINYSGVN-NTAY  |
| BKC201      | QPVENYALHP--SGY--TSY  |
| BSP11       | QPVENYALHP--SGY--TSY  |
| BOHB        | QPVENYALHP--SGY--TSY  |
| BREV        | QPVENYALHP--SGY--TSY  |
| BF290       | PPIENVMELHP--GSP--ASY  |

**Figure 13. Amino acid sequence alignment of the region around residue 89 in CGTases.** Included are the CGTases listed in Table 3. The CGTases are ordered according to their cyclodextrin product specificity; α-CGTases at the top, γ-CGTase at the bottom (Table 3).

New insights in factors determining CGTase cyclodextrin product specificity came from crystal soaking experiments with the α-CGTase from *T. thermosulfurigenes* EM1 which resulted in the structure of the enzyme complexed with a maltohexaose inhibitor bound in the active site. The conformation of this maltohexaose was more bent compared to the maltononaose conformation and it was suggested to represent a specific intermediate in cyclization for the formation of α-cyclodextrin (Wind et al. 1998). Also in the *B. circulans* strain 251 CGTase double mutant Y89D/S146P a maltohexaose bound in a similar conformation in the active site cleft was observed. Although this double mutant produced significantly more α-cyclodextrin compared to the wild-type enzyme, with the cyclodextrin product ratio changing from 14:66:20 (α:β:γ) for the wild-type enzyme to 30:51:19 for the
CGTase, a biocatalyst with unique properties

mutant enzyme, this mutant still favours β-cyclodextrin production (van der Veen et al. 2000b). The bent conformation of the ligands in these enzymes thus probably represent an intermediate in (α-, β-, and γ-) cyclization. The correlation of (increased) α-cyclodextrin specificity and the preference of binding a linear oligosaccharide in a bent conformation can be related to the rate limiting step in the cyclization reaction, which is probably the conformational change from linear substrate to circular product, especially for the smallest (α-)cyclodextrin (van der Veen et al. 2000c). Stabilization of an intermediate conformation may therefore result in a generally increased cyclization activity, which is most significant for α-cyclization (van der Veen et al. 2000b). Mutations Asp371Arg and Asp197His of T. thermosulfurigenes EM1, aiming to hinder and to stabilize, respectively, the maltohexaose bent conformation (Wind et al. 1998), support this hypothesis. Mutation Asp371Arg resulted in enhanced levels of β- and γ-cyclodextrins produced: from 25:58:14 (α:β:γ) for the wild type to 6:68:26 for the mutant enzyme, however, with drastically decreased specific activities, supporting the role of the bent intermediate in formation of all cyclodextrins. Mutant Asp197His produced enhanced levels of α-cyclodextrin: from 28:58:14 (α:β:γ) for the wild type to 35:49:16 for the mutant. This also explains the effects of mutating Tyr195, which leads to increased production of larger cyclodextrins when changed into Trp (Sin et al. 1994; Parsiegla et al. 1998; Penninga et al. 1995), but even more so when changed to Leu (Penninga et al. 1995). All these mutants resulted in decreased production of cyclodextrins, indicating that the efficiency of the cyclization reaction was negatively affected. Combination with the studies described above indicates that the increased ratio of larger cyclodextrins directly results from interference with the role of Tyr195 in cyclization. This role can be partially taken over by Trp, resulting in a rather slight shift towards larger cyclodextrins. In the Tyr195Leu mutant, however, the function of residue 195 is lost, resulting in complete absence of α-cyclodextrin production.

Another interesting amino acid found near subsite -3 is residue 47, which interacts with (semi)cyclic compounds, but not with linear oligosaccharides. It is found next to the unique CGTase residue Leu46 in the loop between β-strand 1 and α-helix 1 which has ligands of the CGTase specific Ca²⁺ binding site on both ends (see above, Fig. 11). Except in the α-CGTases and the thermostable CGTases from Thermoanaerobacter sp. and T. thermosulfurigenes EM1, its position is even more defined by the sulfide bridge formed by Cys43 and Cys50. The nature of residue 47 shows a clear discrimination between the different groups of CGTases as defined in Fig. 14. In α- and αβ-CGTases it is a lysine. In β-CGTases it is either a lysine or an arginine. In the CGTases producing virtually no α-cyclodextrin it is a histidine. Finally, in the γ-CGTase from B. firmus/lentus strain 290-3 a threonine is found at this position. Interestingly, in the maltohexaose structure of the αβ-CGTase from T. thermosulfurigenes EM1 Lys47 hydrogen bonds to the glucose residue at subsite -3, an interaction which is not observed for Arg47 in the maltohexaose structure of the double mutant Y89D/S146P of the β-CGTase from B. circulans strain 251. These differences may be induced by the specific residues (Lys or Arg, respectively) or by the absence or presence of the sulfide bridge, respectively. In either case the observations again support the hypothesis that stabilization of (the) intermediate conformation(s) of the oligosaccharide chain during the cyclization reactions have the most stimulating effect on α-cyclization. This is also shown by mutations Arg47Leu and Arg47Gln in the CGTase from B. circulans strain 251, supporting
the involvement of this residue in the change of the oligosaccharide conformation. Both mutants showed a generally decreased cyclization activity, but also a shift towards the production of larger cyclodextrins (van der Veen et al. 2000a).

|        | 43 | 50 |
|--------|----|----|
| KLEPN  | DPNLKKYT |
| BMAC2  | HS-NLKLKYF |
| BMACE  | RS-NLKLKYF |
| TBN0VO | HT-SLKKYF |
| TABIUM | HT-SLKKYF |
| BSTEA  | CT-NLRRKYC |
| BLICH  | CS-NLKLKYC |
| BCIR8  | CS-NLKLKYC |
| BC192  | CS-NLKLKYC |
| B663   | CSTNLKLYC |
| BF2    | CT-NLRLYC |
| BC251  | CT-NLRLYC |
| B1018  | CT-NLRLYC |
| B1011  | CT-NLRLYC |
| B382   | CT-NLRLYC |
| B171   | CT-NLRLYC |
| BKC201 | CI-DLHKYC |
| BSP11  | CI-DLHKYC |
| BOHB   | CS-DLHKYC |
| BREV   | CS-DLHKYC |
| BF290  | CL-DLTKYC |

**Figure 14.** Amino acid sequence alignment of the region around residue 47 in CGTases. Included are the CGTases listed in Table 3. The CGTases are ordered according to their cyclodextrin product specificity; α-CGTases at the top, γ-CGTase at the bottom (Table 3).

**III.6. Cyclodextrin product inhibition**

During incubation of starch with CGTase a maximum of only 40% of the starch is converted into cyclodextrins. The cyclization activity of CGTase was found to be inhibited by its cyclodextrin products, removal of which from the reaction mixture via ultrafiltration enhanced the conversion of starch into cyclodextrins (up to 60%) (Bergsma et al. 1988). Although this product inhibition has been recognized, its mechanism has not been clarified until recently. Different studies have revealed a mixed type of inhibition (Lee and Kim, 1992; Penninga et al. 1996) or uncompetitive inhibition (Lee and Tao, 1995). In the latter study, however, a poor substrate was used, resulting in substrate inhibition at concentrations higher than 0.055%. Therefore these product inhibition studies may have been severely affected by the limited amount of substrate used. New insights in the product inhibition of CGTase have been obtained by analysis of the inhibitory effect of cyclodextrin on glucoamylase activity. Degradation of raw starch by glucoamylase was found to be severely inhibited by cyclodextrins (Fukuda et al. 1992; Fagerstrom, 1994). This inhibitory effect was less significant in the action on soluble starch and completely absent when short oligosaccharides were used as a substrate (Fukuda et al. 1992). Apparently inhibition of glucoamylase starch degradation by cyclodextrins is linked to the raw starch binding domain of glucoamylase, which is supported by the fact that this domain has a high affinity for β-cyclodextrin (Svensson
and Sierks, 1992; Kusnadi et al. 1994). β-Cyclodextrin can bind to the raw starch binding domain at two sites similar to the MBS’s on the E-domain of CGTase (Sorimachi et al. 1996). These MBS’s of CGTase from *B. circulans* strain 251 were found to interact with maltose, longer oligosaccharides and cyclodextrins (Knegtel et al. 1995). Mutational analysis of these MBS’s revealed that the non-competitive component of the mixed inhibition by cyclodextrins function of MBS 2 (with Tyr633) in the E-domain, interfering with amylose binding and blocking the groove leading to the active site (Penninga et al. 1996). The remaining competitive product inhibition, interfering with catalysis in the active site can be affected by changing residues which specifically interact with cyclodextrins (van der Veen et al. 2000a). However, no mutants have been described thus far that show enhanced starch conversion due to decreased product inhibition.
IV. Aims and outline of this thesis

Cyclodextrin glycosyltransferase (CGTase) is a unique enzyme, catalyzing the formation of cyclodextrins from starch. Aim of our work is to gain insight in the factors determining the specific CGTase properties. As stated in the introduction, cyclodextrin product specificity and cyclodextrin product inhibition are major drawbacks for the use of CGTase in the industrial production of cyclodextrins. Here the features of the enzyme determining reaction specificity, product specificity, and product inhibition are investigated.

In Chapter 2 the role of the E-domain in CGTase is analyzed. For glucoamylase an identical domain is responsible for binding to raw starch and inhibition of glucoamylase activity by β-cyclodextrin. CGTase contains three maltose binding sites, two of which are situated in the E-domain. Trp616 and Trp662 of maltose binding site 1 and Tyr633 of maltose binding site 2 were replaced by alanines. The results show that maltose binding site 1 is most important for raw starch binding whereas maltose binding site 2 is involved in guiding linear starch chains into the active site. β-Cyclodextrin causes product inhibition by interfering with catalysis in the active site and with the function of maltose binding site 2 in the E-domain.

In Chapter 3 the three transglycosylation reactions catalyzed by CGTase are kinetically analyzed. Cyclization (cleavage of an α-glycosidic bond in amylose or starch and subsequent formation of a cyclodextrin) is a single-substrate reaction with an affinity for the high molecular weight substrate used which was too high to allow elucidation of the kinetic mechanism. Previous studies, however, have revealed Michaelis-Menten kinetics when using shorter amylose chains. Coupling (cleavage of an α-glycosidic bond in a cyclodextrin ring and transfer of the resulting linear maltooligosaccharide to an acceptor substrate) is a two-substrate reaction yielding one product and proceeds according to a random ternary complex mechanism. Disproportionation (cleavage of an α-glycosidic bond of a linear maltooligosaccharide and transfer of one part to an acceptor substrate) is a two-substrate reaction yielding two products and proceeds according to a ping-pong mechanism. The different kinetic mechanisms observed for the various reactions suggest that it is possible to manipulate CGTase in such a manner that a single reaction is affected most strongly.

In Chapter 4 cyclodextrin product specificity was investigated based on the X-ray crystal structure of CGTase complexed with a maltononaose. The maltononaose revealed specific glucose binding subsites (+2 to -7) in the CGTase active site. To probe the importance of these substrate binding sites for the α-, β-, and γ-cyclodextrin product ratios, three single and one double mutant were constructed. Residues constituting subsites involved in initial substrate binding or in the subsequent circularization resulting in formation of the cyclodextrins contribute strongly to the size of cyclodextrin products formed and thus to CGTase product specificity.
In Chapter 5 the role of Arg47, a residue involved in binding cyclic compounds only, is investigated. Characterization of two site-directed mutants of Arg47 show that this residue is involved in the cyclization and coupling reactions catalyzed by CGTase. By interacting with cyclodextrins it is involved in the competitive product inhibition exerted by these cyclodextrins. The mutants allowed identification of several enzyme features which are important factors in the production of cyclodextrins.

In Chapter 6 the acceptor binding site of CGTase, which plays an important role in all CGTase catalyzed reactions is investigated. Several mutants were constructed, again based on the X-ray crystal structure of CGTase complexed with a maltononaose. The results allowed identification of important features of CGTase determining reaction specificity.