SUMMARY AND CONCLUDING REMARKS

Starch is the major storage carbohydrate in many plants like maize, potato, wheat, barley, rice, cassava and sorghum. Up till now, more than 20 different types of starch-converting enzymes have been identified. In the past decades, the number of applications for these enzymes as well as their production levels have increased enormously. As a result, these enzymes now comprise more than 30% of the world enzyme production. The starch industry is, however, still searching for suitable, more efficient, new starch-converting enzymes. Especially thermostable enzymes are of industrial interest since liquefaction of starch is performed at high temperatures (up to 110°C), by which the solubility of starch is increased and its viscosity in water is reduced. New enzymes may be found by the screening of enzymes from novel microorganisms or by rational design via site-directed mutagenesis. Both approaches were described in this thesis.

Most of the starch-converting enzymes belong to one protein family, family 13 or the \( \alpha \)-amylase family, a group of \((B/\alpha)_{13}\)-barrel enzymes. Cyclodextrin glycosyltransferase (CGTase) and \( \alpha \)-amylase belong to this \( \alpha \)-amylase family. These enzymes are functionally and structurally closely related. Both enzymes catalyze the degradation of starch by the cleavage of \( \alpha-1,4 \)-glycosidic bonds, resulting in the formation of linear saccharides such as maltose and glucose by \( \alpha \)-amylases and the formation of cyclodextrins by CGTases. Cyclodextrins (CD's) are cyclic compounds composed of 6, 7 or 8 glucose units linked via \( \alpha-1,4 \)-bonds, which have the ability to form inclusion complexes with a wide range of small hydrophobic molecules used in the food, cosmetic, pharmaceutical and agrochemical industries. The primary structures of \( \alpha \)-amylases and CGTases show about 30% similarity, whereas the three-dimensional structures of both enzymes are quite similar. To improve knowledge of structure-function relationships within the \( \alpha \)-amylase family, the factors determining the product specificity of CGTases and \( \alpha \)-amylases were studied in this thesis by means of protein engineering.

In the general introduction (Chapter 1), the functions and applications of the four groups of enzymes involved in the degradation of starch are described, as well as their structure-function relationships. The overall similarity in the primary structure of enzymes belonging to the \( \alpha \)-amylase family is low; however, in all of these enzymes four conserved regions have been identified, containing the active site residues and several substrate binding residues. The catalytic mechanism for starch hydrolysis is most probably that of a general acid-base catalysis mechanism similar to that proposed for hen egg-white lysozyme. Differences in carbohydrate-protein interactions between the different enzymes are believed to account for the differences in product specificity.

Factors determining the thermostability of proteins are also reviewed in the general introduction. Many different factors have been proposed to contribute to the thermostability of proteins including hydrophobic interactions, ionic and hydrogen bonding, disulfide
SUMMARY

bonds, metal binding, packing efficiency, α-helix stabilization and loop stabilization; in general, thermostability cannot be explained by a single factor. Thermophilic bacteria are important sources of thermostable enzymes; in the initial stages of this study two novel starch-converting enzymes isolated from thermophilic bacteria were characterized biochemically: the thermostable α-amylase from Bacillus stearothermophilus MFF4 and the thermostable CGTase from Thermanaerobacterium thermosulfurigenes EM1.

A novel strain of B. stearothermophilus was isolated from samples of a potato-processing industry. This isolated strain was found to produce a highly thermostable α-amylase, compared to known α-amylases from other B. stearothermophilus strains (Chapter 2). Furthermore, the α-amylase was active at acidic pH values. The α-amylase was partially purified and displayed a molecular weight of 54 kDa on SDS-PAGE. The temperature optimum for α-amylase activity was 95°C and the pH optimum for activity was in the range of pH 3.5-5.0. This relatively low pH optimum of the α-amylase compared to pH optima of commercially used α-amylases, is an advantage when considering possible use in the starch liquefaction process. The second step in this process is a saccharification of the maltodextrins formed, with glucoamylase at pH 4.0-4.5. Starch liquefaction enzymes active at the saccharifying pH of 4.5 allow savings to be made on the amount of acid and base needed for pH adjustment, thereby also reducing the costs of ion exchange media and chemicals for purification of the syrup.

The thermostable α-amylase from T. thermosulfurigenes EM1, formerly Clostridium thermosulfurogenes EM1, was purified by affinity chromatography. The enzyme displayed a molecular weight of 68 kDa on SDS-PAGE, which is unusually high for an α-amylase. α-Amylases generally have a molecular weight of 45-55 kDa, whereas CGTases generally have a molecular weight of 70-75 kDa. α-Amylases are known to consist of three domains (A, B and C) whereas CGTases are known to possess two additional domains (D and E) absent from α-amylases. Characterization of the purified enzyme from T. thermosulfurigenes EM1 demonstrated that the enzyme in fact is a CGTase (Chapter 3). Several other examples of incorrect classification of CGTases as α-amylases are reported in the literature, mainly due to superficial biochemical characterization. Product analysis after incubation of the Thermanaerobacterium enzyme with starch revealed the formation of α-, β- and γ-CD (28:58:14), as well as linear sugars. Maximum CGTase cyclization activity was observed at 80-85°C and at pH 4.5-7.0. Highly thermostable CGTases may find industrial applications for starch solubilization in the cyclodextrin industry, eliminating the need for α-amylase pretreatment.

The crystal structure of the thermostable CGTase from T. thermosulfurigenes EM1 was determined at 2.3 Å resolution (Chapter 4). The final model consisted of 683 amino acids, two calcium ions and 343 water molecules. The structure of this thermostable CGTase was quite similar to those of mesophilic CGTases; differences were only observed at surface loop regions, which contributed to the stabilization of the enzyme with novel hydrogen bonds, metal binding, packing efficiency, α-helix stabilization and loop stabilization; in general, thermostability cannot be explained by a single factor.
and loop stabilization; in
thermophilic bacteria are
characterized by high thermostable α-
thraerothermophilus strains of this study two novel
bacteria were characterized
Novaram thermopolus MFF4 and
osulfurigenes EM1.
thermophilic strains
thermophillus MFF4 and
thermophillus EM1. The cr-amylase
values. The α-amylase samples of a potato-
thermophilus strains Concentration of the enzyme was
activity was considered possible
was determined at pH 4.0-4.5. Starch
also reducing the costs
M1, formerly Clostridium
the enzyme displayed an
CGTases generally
CGTase (Chapter 3).
-amyrase family. Other features possibly contributing to the
novel hydrogen bonds and apolar contacts. Other features possibly contributing to the
thermostability of the CGTase were five new salt-bridges and three Gly to Ala/Pro
substitutions.

CGTases are known to catalyze four different reactions: cyclization, coupling,
disproportionation and hydrolysis. In the coupling reaction, the cyclodextrin ring structure is
opened and transferred to a linear acceptor oligosaccharide chain. In the
disproportionation reaction, part of a linear oligosaccharide is transferred to another linear
acceptor oligosaccharide chain. The specific enzyme activities for cyclization, coupling and
disproportionation of the T. thermosulfurigenes EM1 CGTase were found to be rather
similar to those of other CGTases, whereas the specific activity for hydrolysis was
relatively high. The relatively high hydrolytic activity of this CGTase may be due to a
reduced number of carbohydrate-protein interactions. Compared to the mesophilic
CGTase from Bacillus circulans 251, the CGTase from T. thermosulfurigenes lacks a
stacking interaction between the aromatic side chain of Tyr89 (B. circulans CGTase
numbering) and glucose molecules at the sugar binding subsites 3 and 4. Bound
oligosaccharides hence miss an important stabilizing interaction, by which cyclization will
be reduced. A decrease in cyclization implies an increase in hydrolytic activity. Indeed,
the CGTase from T. thermosulfurigenes has a twofold decreased cyclization activity and
a sixfold increased hydrolysis activity when compared to the B. circulans 251 CGTase.

Crystals of the T. thermosulfurigenes CGTase were soaked with the inhibitor
acarbose and with maltohexaose. The structure of the resulting protein-carbohydrate
complex was determined at 2.6 Å resolution, revealing a maltohexaose inhibitor bound
in the active site of the CGTase protein (Chapter 5). An identical experiment with the
CGTase from Bacillus circulans strain 251 revealed a maltononaose inhibitor bound in the
active site. At sugar-binding site +3 the binding mode of the maltohexaose inhibitor was
radically different from the maltononaose binding mode, the former being more bent. The
new maltohexaose bent conformation was found to be related to the enhanced α-
cyclodextrin production by the enzyme. The cyclodextrin product specificity of the CGTase
from T. thermosulfurigenes was subsequently engineered by site-directed mutagenesis.
Mutation D371R was aimed at hindering the maltohexaose conformation and resulted in
enhanced production of the larger size cyclodextrins. The α:β:γ cyclodextrin product ratio
was changed from 28:58:14 for the wild-type CGTase to 6:68:26 for mutant D371R.
Mutation D197H was aimed at stabilization of the new maltohexaose conformation and
resulted in increased production of α-CD. The cyclodextrin product ratio was 35:49:16 for
mutant D197H.

The pH optimum of the T. thermosulfurigenes CGTase was engineered by
changing amino acids close to Glu258. Residue Glu258 is highly conserved within the α-
amylase family. The residue is involved in catalysis in CGTases as well as α-amylases
and is the proton donor in the first step of the reaction. Phe284 was replaced by Lys
Asn327 by Asp. The pH optima for cyclization and hydrolysis were decreased to more acidic pH values in both mutants when compared to the wild-type enzyme. The mutants showed changes in both the high and low pH slopes of the optimum curve, which suggests that the pH optimum curve of CGTase is determined only by residue Glu258.

Furthermore, the factors determining α-amylase and cyclodextrin glycosyltransferase specificity were engineered in the CGTase from *T. thermosulfurigenes*. Amino acid residue 196 occupies a dominant position in the active site cleft and is a small residue in α-amylases but a bulky aromatic residue (Tyr or Phe) in CGTases. Characterization of the constructed mutant F196G revealed that, for unknown reasons, apart from the F196G mutation domain E as well as a part of domain D had become deleted (Chapter 6). This nevertheless did not prevent the purification of a stable and active mutant CGTase protein, similar to an α-amylase protein in domain structure and identity of residue 196. The truncated mutant F196G showed drastically reduced cyclization, coupling and disproportionation activities compared to the wild-type enzyme, but doubled hydrolysis activities, reaching the highest level ever reported for a CGTase. Due to the loss of domain E, the mutant CGTase displayed a strongly reduced raw starch binding capacity, similar to most α-amylases.

Natural enzymes may not be ideally suited for industrial applications under strongly unnatural conditions. Thus protein engineering can be a powerful tool to adapt enzymes to their specific applications, but only when the three-dimensional structure of the protein of interest and knowledge of structure-function relationships is available. The elucidation of the three-dimensional structure of the thermostable CGTase from *T. thermosulfurigenes* provided a structural rationale for the differences in thermostability and product specificity compared to the CGTase from the mesophilic *B. circulans* strain 251. The structure of a protein-carbohydrate complex subsequently allowed engineering of the *T. thermosulfurigenes* CGTase. The results presented in this thesis show that the cyclodextrin product specificity, the pH optimum curve as well as the ratio between cyclization and hydrolysis activity of the CGTase from *T. thermosulfurigenes EM1* can be significantly changed by site-directed mutagenesis, illustrating the feasibility of CGTase protein engineering. The three-dimensional structure and site-directed mutations described in this thesis are used for the improvement of a commercial thermostable CGTase protein.