CHAPTER 2

CRYSTAL STRUCTURE OF *BACILLUS STEAROTHERMOPHILUS* α-AMYLASE: POSSIBLE FACTORS DETERMINING THE TERMOSTABILITY

2.1. Introduction

α-Amylase (α-1,4-glucan-4-glucanohydrolase, EC 3.2.1.1) catalyzes the hydrolysis of α-D-(1,4)-glucosidic linkages in starch and related carbohydrates, releasing α-anomeric products. To date α-amylases from various sources have been isolated and characterized. An α-amylase from *Bacillus stearothermophilus* is a thermostable enzyme, which we previously used to study the effect of mutation of an amino acid residue near the catalytic site on the activity (Takase, 1993). Since the crystal structure of BSTA was not known, we interpreted the results by referring to that of Taka-amylase A (Matsuura *et al.*, 1984). This left us to determine the structures of BSTA and its mutants in a later study. Another important motivation for solving the structure of BSTA resides in the fact that its amino acid sequence is homologous to those of a less thermostable α-amylase from *B. amyloliquefaciens* (66% sequence identity) and a more thermostable α-amylase from *B. licheniformis* (65% sequence identity). The half-lives of irreversible thermoinactivation at 90℃ and pH 6.5 (pH optimum) are 2 min (BAA), 50 min (BSTA) and 270 min (BLA) (Tomazic & Klibanov, 1988). The temperature optima of these amylases are 60, 75 and 90℃, respectively (Vihinen & Mäntsäla, 1989). Since the crystal structure of BLA (but not of BAA) is now known (Machius *et al.*, 1998), solution of the structure of BSTA will aid in further understanding the mechanism of thermostability. A two-residue insertion seen in BAA and also in another homologous α-amylase from *Bacillus* sp. KSM-1378 (LAMY), relative to BAA in sequence alignment, was shown by site-directed mutagenesis to be a cause of their thermal instability (Suzuki *et al.*, 1989; Igarashi *et al.*, 1998). Because BSTA also has an insertion at the same position, the comparison of the structures of BSTA and BLA would

provide a structural basis for the effect of the insertion on thermostability. Recently we have successfully obtained crystals of recombinant wild-type BSTA suitable for X-ray diffraction analysis (Suvd *et al.*, 2000). In this paper we describe the three-dimensional structure of BSTA, compare it with that of BLA and discuss the structural characteristics contributing to the difference in thermostability between the two α-amylases.

2.2. Methods

Purification, crystallization and data collection

Purification and crystallization of BSTA was carried out as described previously (chapter 1, Suvd *et al.*, 2000). The crystal used in this study belonged to the monoclinic space group $P2_1$ with unit cell dimensions of a= 53.5, b=92.8, c=53.1Å, and β=109.4°. X-ray intensity data were collected at room temperature from a crystal piece mounted in a glass capillary at the Photon Factory. Details of the data collection statistics are summarized in Table 1-2.

Structure determination

The structure of BSTA was determined by the molecular replacement method, using the program AMoRe (Navaza, 1994) in the CCP4 suite (Collaborative Computational Project, 1994). Structure of BLA (1BPL; Protein Data Bank entry; Diamond, 1997) was used for the initial reference model. Both the rotation- and translation- function searches gave a single solution. The translation function had a correlation coefficient of 0.41 and an R-factor of 0.450 in the 6-3.0Å resolution range. After the rigid-body refinement which resulted in an R-factor of 0.424 in the 6-2.5Å resolution range, iterative process of positional and simulated annealing refinement was followed using the program X-PLOR (Brünger *et al.*, 1987), using the parameters of Engh & Huber (1991). Ten percent of the observed reflections were randomly removed for the cross-validation (Brünger, 1992). Coordinates were adjusted repeatedly against the renewed $2F_{obs} - F_{calc}$ and $F_{obs} - F_{calc}$ electron density maps with the program QUANTA (Molecular Simulations Inc.). During the course of refinement, three

calcium ions and one sodium ion were identified for the large peaks in a difference Fourier map. Solvent molecules were identified as positive peaks in electron density maps within hydrogen-bonding distance of appropriate protein atoms or another solvent atoms. The results of analysis are summarized in Table 2-1.

The quality of the structure was analyzed using the programs WHATIF (Vriend, 1990) and PROCHECK (Laskowski *et al.*, 1993). Accessible surface areas were calculated using the program GRASP (Anthony *et al.*, 1991). Cavities were searched for using the program VOIDOO (Kleywegt & Jones, 1994) on both proteins using a number of different parameters and protein orientations. We also used the program VOIDOO to calculate the packing densities (the ratio of the van der Waals volume of an atom to the actual space it occupies (Richards, 1974)). The real volumes were calculated by the Voronoi procedure using the programs ACCESS and VOLUME (Richards, 1977).

Protein Data Bank accession number

The coordinates for BSTA were deposited in the RCSB Protein Data Bank (accession number 1hvx).

2.3. Results and discussion

Quality of model

The quality of the structure was assessed using the program PROCHECK (Laskowski *et al.*, 1993) and WHATIF (Vriend, 1990). In the Ramanchandran plot (Ramanchandran & Sasisekharan, 1968), 87.4% of non-glycine residues have their backbone torsion angles in the most favoured region and 11.8% residues in the additional allowed region. Only one residue Tyr151 (ϕ=83°, ψ=–40.5°) lay in the disallowed area, though it has well-defined electron density map (Figure 2-1). In the omit map, where Tyr151 and neighboring residues were omitted from the model, the difference electron density confirmed that they are correctly placed. The mean positional error of the atoms as estimated from the Luzatti plot (Luzatti, 1952) is 0.21Å. Crystal parameters and refinement statistics are listed in Tables 1-2, 2-1. Among the residues from -2 to 515 deduced from the DNA sequence (Sohma *et al.*, 1987) and the N-terminal amino acid sequence analysis (Suvd *et al.*, 2000), residues 1-483 were visible in electron density maps, but the rest of residues were not, probably because of their disordered structure. Acarbose, a substrate analogue inhibitor is not visible at the active site in electron density maps, although it is added during crystallization (Suvd *et al.*, 2000). The final model resulted in an R-factor of 0.156 and in an R_{free}-factor of 0.197 in the resolution range of 8.0-2.0Å. The average of B-factors is 9.6Å^2 for all atoms and 9.1Å^2 for protein atoms. These unusually low values are consistent with the fact that the crystals diffracted beyond 2.0Å resolution considering very small crystals with the maximum dimensions of 0.1x0.05x0.02mm (Suvd *et al.*, 2000).

Overall structure of BSTA

BSTA consists of a single polypeptide chain, folded into three distinct domains A, B and C as commonly observed in α-amylases (Figures 2-2, 2-3). Domain A represents a compact core structure comprising a $(\beta/\alpha)_8$-barrel. Domain B shows a variable structure in α-amylases. It is composed of a sheet of four antiparallel β-strands and a pair of antiparallel β-strands, and fairly long loops between β-strands. The Ca^{2+}-Na^{+}-Ca^{2+} binding site is located in the interior of domain B, and spanning to the interface with domain A, as described later. Domain C forms a distinct globular unit which consists of eight strands folded into a Greek key motif and forms the third Ca^{2+} binding site in cooperation with domain A (Figure 2-2).

Although an acarbose was added to the crystallization condition of BSTA (Suvd *et al.*, 2000), in the refined structure of the protein no distinct able sign of bound acarbose was detected in its active-site cleft.

Figure 2-4 shows a superposition with BLA. The root mean square (RMS) difference for Cα atom pairs is 0.83Å (Figure 2-5). The residues involved in the secondary structure elements (α-helices and β-strands) are shown in the topological alignment of these enzymes (Figure 2-3). The main secondary structure elements are essentially the same in both enzymes. The two amylases display 65% identity in their amino acid sequence. From the 35% of total difference (169 residues) between two amylases 37% (63 residues) have taken place in α-helices structures.

Active site

The active site is located at the C terminal side of the β-strands of the $(\beta/\alpha)_8$-barrel in domain A. From the structural comparison with other α-amylases, Asp234, Glu 264 and Asp331 are assigned as the catalytic residues in BSTA (Figure 2-3). Figure 2-6 shows the active site superimposed with that of BLA. The residues in the active site are strictly conserved, but some positional differences are observed, particularly in the catalytic residues. This may reflect the flexible nature of catalytic residues, which may be important for catalytic reactions.

The structure of the active site is also similar to that of Taka-amylase A, which rationalizes the mutant designs previously made by us (Takase, 1993) to investigate the effect of mutation near the catalytic site. In the previous design we referred to the structure of Taka-amylase A (Kusunoki *et al.*, 1984) and found that Asn295 (Asn329 in BSTA) is located in close proximity (4-5Å distance between side-chain atoms) to Asp297 (Asp331 in BSTA), one of the catalytic residues. The current structure of BSTA shows that the corresponding residues in BSTA are arranged in similar positions (Figure 2-7). In the previous study we mutated Asn329 to Lys, Asp and Val, resulting in altered specific activity, pH and temperature dependence of activity (Takase, 1993). Worthy of note is the Asp329→Val mutant that showed much higher activities than the wild-type enzyme at low temperatures (5–15°C). We modeled the structures of these mutants and found that they are not deviated much from the wild-type structure. Thus, the mutated residues are close to Asp331 so that they could affect the catalytic activity. This validates our previous interpretation that the

altered catalytic properties of the mutant enzymes are the results of the charge or hydrophobic effects of the mutated residues on Asp331.

Metal ion binding sites

In the course of refinement, three peaks with a high level of electron density appeared in the interior of domain B and another at the interface between domain A and domain C. These peak positions correspond to those of three Ca^{2+} ions and one Na^{+} ion in BLA (Machius *et al.*, 1998), and therefore shown as CaI, CaII, CaIII and Na in Figures 2-2 and 2-3.

CaI is strictly conserved in all α-amylases. CaI and CaII form a linear triad with a sodium ion as described in BLA, but there are some differences between the two α-amylases, as follows. 1) The side-chain OD1 of Asp204 in BLA is involved in binding of CaII, but the side chain of the corresponding residue Asp207 in BSTA is positioned away from binding to CaII. Instead, a water molecule participates in coordination to CaII (Figure 2-8a and Table 2-2). 2) Asn104 in BLA is substituted by Asp105 in BSTA, but both residues are involved in the coordination to CaI.

CaIII contributes to bridging between two loops, one of which comes from Aα6 of domain A and the other protruded between Cβ1 and Cβ2 of domain C (Figure 2-8b and Table 2-2). There are some differences in coordinating residues to CaIII: Tyr302 in BLA is substituted by the Phe305 in BSTA and His406 in BLA by Ser406 in BSTA.

Interpretation of the thermostability difference based on structural comparison with BLA

The nature of thermostability of various enzymes have been extensively investigated and the factors affecting greatly to the thermostability of the enzymes are roughly defined as follows:

- *Tighter packing* (reduced surface area and an increased number of buried residues (Knapp *et al.*, 1997.)).

- *Higher order of secondary structure* as more number of helices, strands, and 3_{10}-helices and shorter loops and fewer cavities (Vogt *et al.*, 1997).

- *Improved stability of helices* as having higher helices propensity or via ion pair, hydrogen-bonding interactions and hydrophobic interactions between the side chains within α-helices (Facchiano *et al.*, 1998).

- *More prolines* that decrease the entropy of unfolded state (Matthews *et al.*, 1987; Delboni *et al.*, 1995)

- *Fewer asparagines and glutamines,* thus less deamidation at slight elevated temperatures (Klibanov *et al.*, 1985).

- *Large network of ion pairs and salt bridges increase in net number of hydrogen bonds,* that stabilize overall tertiary structure. Vogt *et al.*, (1997) detected an increase in an average of 11.7 internal hydrogen bonds for 13 of the 16 families of enzymes compared for the thermostability, for every 10°C of thermal stability.

- Some authors noted also that *Arg/(Arg+Lys) ratio* is higher in the thermostable enzymes, than in their counterpart mesophile homologous (Mrabet *et al.*, 1992).

- *Stabilization by metal ions for an metal-dependent enzymes* (Salminen *et al.*, 1996).

Although certain criteria stand for the thermostability of proteins in general, it appears that the thermostability of a particular enzyme is rather reflected by its unique structure. Here we compare the structures of BSTA and BLA and discuss their thermostability difference.

The superimposition of Cα-models (Figure 2-4) and topological alignment (Figure 2-3) of BSTA and BLA showed that the length and position of individual secondary structural elements are almost identical in these two enzymes. Moreover, as shown in Table 2-4, total solvent accessible surface area, volume and molecular packing densities of both molecules are very similar to each other. Total void's volumes of BSTA and BLA are very small compared with other protein molecules and may not be considered as a large overall difference (Tables 2-3, 2-4). Thus, it is difficult to account for the thermostability difference on the basis of the overall structure. It is likely that the thermostability difference is ascribed to the local structural environment.

BSTA has insertions and deletions of four very short one to two residue in amino acid sequence compare to BLA (Figure 2-3). As noted by Shortle *et al.*, (1995), small insertion of two or one amino acid residues could generate propagated structural changes as forming local bulges, in which the inserted residues are pushed outward producing small confined structural distortion or result in lateral displacements, which leads to the expansion of loop. The two-residue insertion in BSTA, Ile181-Gly182,

pushes away the spatially contacting region including Asp207 (Figure 2-9). As the result, the side-chain of Asp207 cannot coordinate the Ca^{2+}, although a water molecule contributes to the restoration of Ca^{2+} ion coordination geometry (Figure 2-8a , Table 2-2).

BAA and a semi-alkaline α-amylase LAMY also possess the two-residue insertion at exactly the same position as BSTA: Glu178-Gly179 in BAA (Figure 2-3) and Thr183-Gly184 in LAMY (Figure 2-10). Previously at the time when the three-dimensional structures of any of these α-amylases were not known, Suzuki *et al.* (1989) and Igarashi *et al.* (1998) erroneously regarded the insertion as Arg179-Gly180 (BSTA numbering), which is two-residue upstream and conserved among the three α-amylases. Curiously, however, the deletion of Arg-Gly in BAA and LAMY by site-directed mutagenesis resulted in an increase of thermostability. Apparently, the deletion of either Arg-Gly or the next two residues downstream, the actual insertion site, would remove the bulge created by the extra two residues (Figure 2-9) and thus seem to have similar consequences on the thermostability by affecting CaII binding. However, in BSTA Arg179 forms a salt bridge with Glu129 (Figure 2-9), which may contribute to thermostability. Therefore, deletion of Ile181-Gly182 rather than Arg179-Gly180 is expected to result in a higher thermostability. We predict a similar enhancement of thermostability by deleting the actual insertion instead of Arg-Gly for BAA and LAMY.

As for other insertion and deletion sites BSTA possesses one residue insertion site at its N-terminus, which does not border much the structural order of the enzyme and a three residue deletion site in the loop joining helices Aα7 and Aα8 (Figure 2-3). In BLA this extended loop is stabilized with the ten hydrogen bonds, two salt bridges and eight hydrogen bonds and does not introduce structural de stabilization. As for

BSTA this loop forms four hydrogen bonds with surrounding secondary structure elements.

Table 2-7 shows a total number of hydrogen bond interactions in each domain of BSTA and BLA, where the number of their hydrogen bond interactions between main-chain side-chain atoms and between side-chain side-chain atoms are simply compared between BSTA and BLA. Table 2-7 was derived from the Tables 2-5 and 2-6. Totally, BSTA has nine hydrogen bonds less than BLA, which costs about 12kcal/mol, assuming 1.3(±0.6) kcal/mol contribution to the stability for every intra molecular hydrogen bond (Shirley *et al.*, 1992). This tendency is prominent in domain A comprising of the $(\beta/\alpha)_8$-barrel, where 10 more hydrogen bonds were observed in BLA, forming a denser hydrogen bond network in the inter-helical region of the $(\beta/\alpha)_8$-barrel. This may contribute partly to higher thermostability in BLA compared with BSTA. On the other hand, in domain B, BSTA has six more hydrogen bonds (including four more salt bridges) than BLA. Therefore, domain B of BSTA appears more stable than that of BLA, but it may not much affect the thermostability of the entire molecule.

By careful inspection of the structures we found a region which differs in local hydrophobic stacking and hydrogen bonding interactions between BSTA and BLA as shown in Figure 2-11 (a, b, c, d). BLA forms sixteen hydrogen bonds around Aα1, Aα2 and Aα8. Substitution of Gln20, Lys23, Arg24 of Aα1 and Glu82 of Aα2 of BLA with Thr21, Thr24, Lys25 and Gln83 in BSTA respectively resulted in the loss of seven hydrogen bonds Figure 2-11 (a, b). Furthermore a hydrophobic core is observed in BLA between four tyrosine rings (Tyr10, Tyr31, Tyr363 and Tyr367) and hydrophobic residues from Aα8, Aα1, Aα2, Aβ2 and Aβ3 (Figure 2-11 c). The corresponding region in BSTA generates two small contiguous voids

(Figure 2-11 d, arrows) with total volume of 26Å^3, which are obviously derived from the change of residues Asp28, Tyr31, Leu90 and Met366 in BLA to Lys25, Asn32, Ala91 and Tyr369 in BSTA respectively. The local differences in this region may contribute to the thermostability difference between BSTA and BLA.

As seen from the Table 2-4, percentage of total accessible surface area of charged residues in BLA is about two times greater than in BSTA. Histidine residues largely contribute to this difference, because BLA has 13 more histidine residues than BSTA, of which seven histidines substitute hydrophobic residues partially exposed in BSTA, assuming Tyr as a hydrophobic residue. The rest six histidine residues are substituted with polar residues. Since hydration of nonpolar groups apparently destabilizes proteins (Spassov *et al.*, 1995), the above substitutions may contribute partly to stabilization of BLA (Figures 2-12 a, b, c, d).

Other criteria determining the thermostability of the proteins are investigated for BSTA, BLA and BAA in Table 2-8. These criteria do not show significant correlation with the thermostability of the three amylases, therefore, do not contribute to the thermostability differences.

Recently, Declerck *et al.*, (2000) investigated the relationship between thermostability and structure using the mutation analysis of BLA, and found that some of the amino acid changes contribute to the differences of thermostability between BLA, BAA and BSTA. A reduction in thermostability caused by replacement of Ala269 by lysine in BLA may be explained using superposition of BLA over BSTA (Figure 2-13). In BSTA the corresponding residue, Lys272 forms a stable salt bridge with Asp190. However, Lys269 in the mutated model of BLA has no such a salt bridge and moreover has short contact with Gln264 which is bulky compared with the corresponding residue, Ser267 in BSTA, and may cause

structural rearrangement around its local region. In these regards, it is understandable that BAA with less thermostability may have similar structure with the above mutant, Ala269Lys of BLA.

As for salt bridge interaction Asp121-Arg127 in BLA disturbance of which causes decrease in thermostabilty of BLA in BSTA it is substituted with the hydrogen bond, formed between Asn122 and Gln128. Both residues are subjective to the deamidation phenomena and therefore this bond might be unfavorable for BSTA. These two residues are conserved in BAA.

In the present comparison between crystal structures of BSTA and BLA, we found significant differences which may account for their thermostability difference, as follows. 1) The two-residue insertion in BSTA, Ile181-Gly182, pushes away the spatially contacting region including Asp207, which corresponds to Ca^{2+} coordinating Asp204 in BLA. As the result, the side-chain of Asp207 cannot coordinate the Ca^{2+}, although a water molecule contributes to restoration of Ca^{2+} coordination geometry. 2) BSTA contains nine hydrogen bonds fewer than BLA, which costs about 12kcal/mol. This tendency is prominent in the $(\beta/\alpha)_8$-barrel, where 10 fewer hydrogen bonds were observed in BSTA. BLA forms denser hydrogen bond network in the inter helical region which may stabilize α-helices in the barrel. 3) A few small voids observed in the α-helical region of the $(\beta/\alpha)_8$-barrel domain in BSTA decrease inter helical compactness and hydrophobic interactions. 4) Percentage of solvent-accessible surface area of charged residues in BLA is about two times larger than in BSTA.

2.4. Summary

The X-ray crystal structure of a thermostable α-amylase from *Bacillus stearothermophilus* (BSTA) has been determined at 2.0 Å resolution. The main-chain fold is almost identical with that of the known crystal structure of *Bacillus licheniformis* α-amylase (BLA), as might be expected by their high degree of sequence homology (65% sequence identity). BLA is known to be more stable than BSTA. In structural comparison between crystal structures of BSTA and BLA, we found significant differences which may account for the thermostability differences, as follows. 1) The two-residue insertion in BSTA, Ile181-Gly182, pushes away the spatially contacting region including Asp207. As the result, the side-chain of Asp207 cannot coordinate the Ca^{2+} , although a water molecule contributes to restoration of Ca^{2+} coordination geometry. 2) BSTA contains nine hydrogen bonds fewer than BLA, which costs about 12kcal/mol. This tendency is prominent in the $(\beta/\alpha)_8$-barrel, where 10 fewer hydrogen bonds were observed in BSTA. BLA forms denser hydrogen bond network in the inter helical region which may stabilize α-helices in the barrel. 3) A few small voids observed in the α-helical region of the $(\beta/\alpha)_8$-barrel domain in BSTA decrease inter helical compactness and hydrophobic interactions. 4) A total of fraction of polar solvent-accessible surface area in BLA is about two times larger than that in BSTA.