J. Biochem. **125**, 1120-1130 (1999)

Crystal Structure of β-Amylase from *Bacillus cereus* var. *mycoides* at 2.2 Å Resolution

Takuji Oyama,*,[1] Masami Kusunoki,[†] Yoji Kishimoto,*,[2] Yoshiyuki Takasaki,[‡] and Yasunori Nitta*

**Laboratory of Enzyme Chemistry, College of Agriculture, University of Osaka Prefecture, Sakai, Osaka 599-8531;* [†]*Institute for Protein Research, Osaka University, Suita, Osaka 565-0871; and* [‡]*Department of Materials Science, Faculty of Engineering, Miyazaki University, Miyazaki 889-2155*

Received January 8, 1999; accepted March 17, 1999

The crystal structure of β-amylase from *Bacillus cereus* var. *mycoides* was determined by the multiple isomorphous replacement method. The structure was refined to a final *R*-factor of 0.186 for 102,807 independent reflections with $F/\sigma(F) \geq 2.0$ at 2.2 Å resolution with root-mean-square deviations from ideality in bond lengths, and bond angles of 0.014 Å and 3.00°, respectively. The asymmetric unit comprises four molecules exhibiting a dimer-of-dimers structure. The enzyme, however, acts as a monomer in solution. The β-amylase molecule folds into three domains; the first one is the N-terminal catalytic domain with a $(\beta/\alpha)_8$ barrel, the second one is the excursion part from the first one, and the third one is the C-terminal domain with two almost anti-parallel β-sheets. The active site cleft, including two putative catalytic residues (Glu172 and Glu367), is located on the carboxyl side of the central β-sheet in the $(\beta/\alpha)_8$ barrel, as in most amylases. The active site structure of the enzyme resembles that of soybean β-amylase with slight differences. One calcium ion is bound per molecule far from the active site. The C-terminal domain has a fold similar to the raw starch binding domains of cyclodextrin glycosyltransferase and glucoamylase.

Key words: active site cleft, *Bacillus cereus*, β-amylase, crystal structure, raw starch binding domain.

β-Amylase [EC 3.2.1.2] is an exo-enzyme which catalyzes the hydrolysis of the α-1,4-glucosidic linkages of substrates such as starch, and liberates β-maltose from the non-reducing end of a substrate. The distribution of β-amylase in nature is limited to higher plants and bacteria. The amino acid sequences of more than ten β-amylases of both origins have been determined by sequence analysis of peptides or by deduction from the nucleotide sequences of genes. Sequence alignments revealed that several conserved regions exist in all β-amylases in spite of the low sequence identity between bacterial and plant β-amylases (about 30%) (*1-4*). Glu186 of soybean β-amylase and Glu187 of the sweet potato enzyme, which was proposed to be catalytic residues based on the results of an affinity labeling experiment involving 2,3-epoxypropyl α-D-glucopyranoside (α-EPG), are included in the conserved regions (*5, 6*). Recently, the crystal structures of soybean β-amylase in complexes with substrate analogues were reported (*7*). The authors proposed the catalytic residues were Glu186 and Glu380, and the hypothetical binding of a substrate, maltotetraose, in the active site. The amino acid residues involved in the oligosaccharide binding in the crystal structures of soybean β-amylase are commonly found in the sequences of both bacterial and plant β-amylases. Hence, the structures of the active sites of β-amylases from both kingdoms are expected to be very similar to each other. On the other hand, Nitta *et al.* reported that the kinetic behavior of bacterial and plant β-amylases differs (*8*). For example, the intrinsic molecular activity of β-amylase from *Bacillus cereus* var. *mycoides* (abbreviated as BCM β-amylase) is five times greater than that of soybean β-amylase, whereas the binding affinity of BCM β-amylase for a substrate is about 0.8 kcal/mol lower than that of soybean β-amylase. The binding affinity for α-EPG, an affinity labeling reagent for BCM β-amylase, is 1-1.6 kcal/mol greater than those of the soybean and sweet potato β-amylases, although the concomitant inactivation rate constants for the three enzymes are similar. The report attributed the differences in kinetic parameters to a subtle structural difference in the active site between bacterial and plant β-amylases. The sequence alignments also indicated that bacterial β-amylases have additional residues in a C-terminal region as compared to plant β-amylases. These additional residues exhibit sequence homology with those in the raw starch binding domains of other carbohydrate-degrading enzymes from bacteria (*3, 9*). This corresponds to the fact that bacterial β-amylases have the ability to bind raw starch whereas plant β-amy-

[1] To whom correspondence should be addressed. E-mail: chicago@biochem.osakafu-u.ac.jp

[2] Present address: Minase Research Institute, Ono Pharmaceutical Co., Ltd., Mishima, Osaka 618-0014.

Abbreviations: α-EPG, 2,3-epoxypropyl α-D-glucopyranoside; BCM, *Bacillus cereus* var. *mycoides*; CGTase, cyclodextrin glycosyltransferase [EC 2.4.1.19]; MIR, multiple isomorphous replacement; SR, synchrotron radiation.

lases do not.

BCM β-amylase was found by Takasaki in 1976 (*10*), and its amino acid sequence has been deduced from the nucleotide sequence of the gene (*11*). Structural comparison at the atomic level between bacterial and plant β-amylases would explain the differences in kinetic behavior between them and reveal how bacterial β-amylases bind raw starch. No structure determination, however, has been reported for bacterial β-amylases. We therefore determined the crystal structure of the enzyme and here report its three-dimensional structure at 2.2 Å resolution. The coordinate data have been deposited with the Protein Data Bank.

EXPERIMENTAL PROCEDURE

BCM β-amylase, which consists of a single polypeptide chain of 516 amino acids and has a molecular weight of 58,000, was purified and crystallized as described (*12*). Crystals of the enzyme grown at 293 K from PEG6000 belong to the monoclinic space group, $C2$, with unit cell dimensions of $a=177.9$ Å, $b=112.9$ Å, $c=146.2$ Å, and $\beta=105.8°$. There are four molecules per asymmetric unit, which corresponds to a Matthews parameter (*13*), V_m, of 3.03 Å^3·Da^{-1} and a solvent content of 59.4%. Heavy-atom derivatives were prepared by the soaking method. Native crystals were first transferred to a solution containing 15% (w/v) PEG6000 buffered at pH 6.5 with 50 mM Bistris or with 50 mM MES to avoid precipitation of heavy-atom compounds, which may occur at a weak alkaline pH (9.0). Then, crystals were soaked into various solutions of heavy-atom compounds. Four heavy-atom compounds, HCOOTl, mersaryl acid, $(CH_3)_3CCOOPb$ and $K_2Pt(NO_2)_4$ were found to be useful for phasing.

Diffraction data sets for native crystals and for a Tl derivative were collected by synchrotron radiation (SR) at the beamline BL-6A of the Photon Factory (PF) operated at 2.5 GeV at the High Energy Accelerator Research Organization, Tsukuba. The X-ray beam was monochromatized to 1.00 Å with a Si(111) monochromator and was collimated with a 0.2 mm square aperture. Oscillation and Weissenberg photographs were recorded on 200 mm × 400 mm imaging plates mounted on a screenless Weissenberg camera for macromolecular crystals (*14*) with a cylindrical cassette of 430 mm radius. Data for Hg, Pb, and Pt derivatives were collected with a RIGAKU R-AXIS IIc imaging plate diffractometer mounted on a RIGAKU RU-200 rotating-anode generator with graphite-monochromatized Cu$K\alpha$ radiation operated at 40 kV and 100 mA. The statistics for intensity data sets are given in Table I.

The major heavy-atom sites of Tl, Hg, and Pt derivatives were located on difference Patterson maps at 5.0 Å, and then multiple isomorphous replacement (MIR) phases were calculated with the MLPHARE program (*15*) in the CCP4 program suite (*16*) using these heavy atom sites. Difference Fourier maps at 5.0 Å with the MIR phases disclosed further sites for the Pt and Pb derivatives, which were incorporated into the phasing. After heavy-atom parameters had been further refined, final best phase angles based on the four derivatives were calculated (Table I). The 2.2 Å MIR phases were improved by solvent flattening and histogram matching with the DM program (*17*) in the CCP4 suite. The improved electron density map clearly revealed the non-crystallographic symmetry (NCS) among the four molecules in the asymmetric unit together with the knowledge of heavy-atom binding sites. Recalculation of DM further improved the 2.2 Å MIR phases through a combination of solvent flattening, histogram matching and NCS averaging, giving an interpretable electron density map (Fig. 1). A molecular model of BCM β-amylase was built on a Silicon Graphics Indigo2 workstation with the O program (*18*). Crystallographic refinement of the model with the X-PLOR program (*19*) and manual model correction with the O program were iterated until the convergence of refinement. NCS-restraints were applied for the four molecules in the asymmetric unit tightly in the early stages of refinement, and then loosely in the middle stages. Finally, the four molecules were released from NCS-restraints. In the course of refinement, an $(|F_{obs}|-|F_{calc}|)^{\alpha}$ electron density map showed one strong peak of higher than $10\sigma(\rho)$ per enzyme molecule. X-ray fluorescence analysis

TABLE I. **Statistics for intensity data and heavy atom refinement.**

Heavy atom compound	Native	HCOOTl	Mersaryl-acid	$K_2Pt(NO_2)_4$	$(CH_3)_3CCOOPb$
Soaking conc. (mM)	—	0.5	0.4	0.5	5.4
Soaking time (d)	—	2	8	2.5	3.5
X-ray source	SR[a]	SR[a]	CuK_α	CuK_α	CuK_α
No. of crystals	3	1	1	1	1
Resolution range (Å)	95.3-2.2	95.3-3.0	95.3-2.8	95.3-3.5	95.3-3.2
No.of observed reflections	649,891	114,433	104,640	72,226	72,100
No. of unique reflections	113,913	48,930	54,699	34,237	40,193
Completeness[b]	81.5	87.3	80.1	96.2	85.9
R_{merge} (I) (%)[c]	6.90	4.8	8.1	4.5	7.2
$R_{iso}(F)$ (%)[d]	—	13.2	14.8	18.9	10.5
No. of sites	—	5	5	9	3
R_{Kraut}[e]	—	0.73	0.83	0.79	0.96
R_{Cullis}[f]	—	0.81	0.86	0.88	0.99
Phasing power$_{Centric}$[g]	—	0.95	0.78	0.68	0.22
Phasing power$_{Acentric}$[g]	—	1.18	1.00	0.86	0.29
Mean figure of merit (overall)	0.437 for 33,444 reflections in the range 15.0-2.8 Å.				

[a]Synchrotron radiation. [b]The completeness of the data to the indicated resolution is given. [c]$R_{merge}=\Sigma_{hkl}\Sigma_i|I(hkl)_i-\langle I(hkl)\rangle|/\Sigma_{hkl}\Sigma_i I(hkl)_i$, where $I(hkl)_i$ is the ith measurement of reflection hkl, $\langle I(hkl)\rangle$ the mean value of equivalent reflections, and i runs through the symmetry-related reflections. [d]$R_{iso}=\Sigma||F_{PH}|-|F_P||/\Sigma|F_{PH}|$ where $|F_{PH}|$ and $|F_P|$ are the derivative and native structure factor amplitudes. [e]$R_{Kraut}=\Sigma||F_{PH}|-|F_P+F_H||/\Sigma|F_{PH}|$ where F_H is the calculated heavy-atom structure factor amplitude. [f]$R_{Cullis}=\Sigma||F_{PH}\pm F_P|-|F_H|/\Sigma|F_{PH}\pm F_P|$ where F_H is the calculated heavy-atom sturcture factor amplitude. [g]The phasing power is defined as the mean heavy-atom structure amplitude divided by the mean lack-of-closure error.

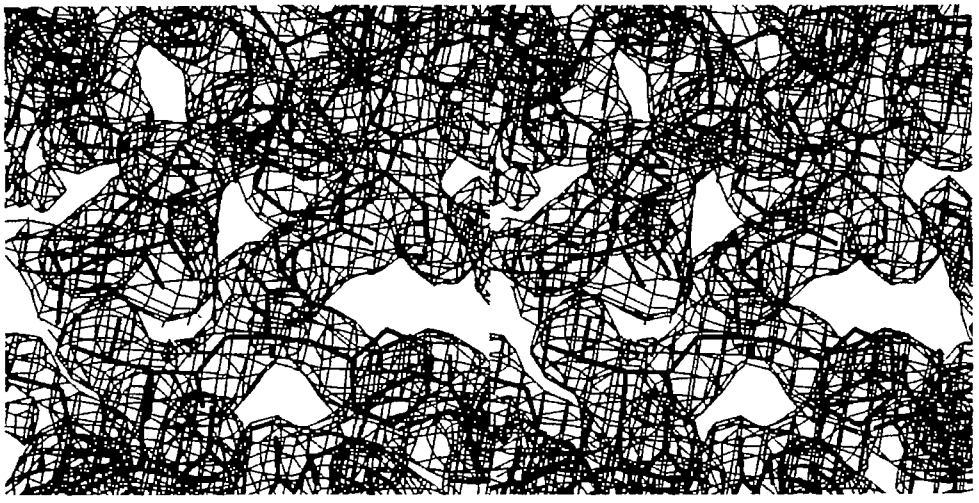

Fig. 1. **Stereo view of the MIR phased electron density map.** The map calculated at 2.8 Å is shown contoured at $1\sigma(\rho)$ with the final model.

of an aqueous solution of the enzyme with a spectrometer, RIX2100 (RIGAKU), significantly showed the presence of calcium ions. We therefore assigned the strong peaks in the difference map as calcium ions. After solvent molecules had been incorporated into the model, refinement was performed with the X-PLOR and REFMAC programs (*20*) alternatively.

RESULTS AND DISCUSSION

Quality of the Final Model—The refined model contains four BCM β-amylase molecules, four calcium ions and 748 water molecules per asymmetric unit. The whole model was built into the electron density map, and its sequence is completely consistent with that deduced from the nucleotide sequence (*11*). The final crystallographic R-factor is 0.186 for 102,871 independent reflections with $F \geq 2\sigma(F)$ in the resolution range of 8.0 to 2.2 Å, and the free R-factor is 0.240 for 5% of the independent reflections. The refinement statistics are given in Table II. The mean positional error of the atoms estimated from a Luzzati plot (*21*) is 0.25 Å. A Ramachandran plot (*22*) showed that 87.2% of the non-glycine or non-proline residues are in the most favored regions defined with the PROCHECK program (*23*), 12.4% in additionally allowed regions, and 0.4% in generously allowed regions.

Non-Crystallographic Symmetry (NCS)—Figure 2 shows the arrangement of four molecules in the asymmetric unit. These four molecules exhibit a unique "dimer-of-dimers" structure. We refer to these molecules as Mol1, Mol2, Mol3, and Mol4. Mol1 and Mol2 are related by a non-crystallographic diad (M_1) and form a dimer. Similarly, Mol3 and Mol4 are related by another local diad (M_2) and form another dimer. The interaction between Mol1 and Mol2 is mainly hydrophobic; Asn243 and Thr299 of Mol1 interact with Trp449 and Trp495 of Mol2, and *vice versa*. The hydrogen bonding interaction between Mol1 and Mol2 is through only two water molecules located between them, there being no direct hydrogen bonding. The interactions between Mol3 and Mol4 are the same as the Mol1-Mol2 interactions. The Mol1-Mol2 dimer and the Mol3-Mol4 dimer are further related by a non-crystallographic diad (D). The D diad relates Mol1 to Mol3, and Mol2 to Mol4. In contrast to monomer-monomer interactions, dimer-dimer interactions consist mainly of hydrogen bonding (10 directly between protein molecules and 14 through water molecules located between the two dimers). The D-axis makes an angle of 68.1° with a crystallographic diad in parallel with the b-axis, whereas the M_1 and M_2 axes are almost at right angles (90.6°) with each other. These three axes do not intersect with each other.

It was found from the profile obtained in a gel-filtration column experiment that the enzyme is mono-disperse (*10*, our unpublished data). We also confirmed that the molecules exist as monomers at a concentration of 20 mg/ml (working solution for crystallization) by means of dynamic light-scattering measurements (*24*) with a DynaPro-801 Molecular Sizing Instrument (Protein Solutions) (our unpublished data). It is therefore concluded that the tetrameric structure of the enzyme in the crystal line state is not related to its enzyme function in solution.

As described under "EXPERIMENTAL PROCEDURE," NCS restraints were not applied in the later stages of model refinement. Nevertheless, the final structures of the four molecules in the asymmetric unit resemble each other with r.m.s. discrepancies of 0.194 Å for main-chain atoms and 0.344 Å for side-chain atoms. Figure 3 shows the average r.m.s. discrepancies of main-chain atoms from the mean structure averaged over the four molecules in the asymmetric unit as a function of the residue number. In this figure, there are six regions with high r.m.s. discrepancies, labeled a to f. Regions a, d, e, and f (residues 94-96, 432-435, 462-465, and 488-489, respectively) are located on the molecular surface and exposed to the solvent with high B-factors (on average 45.7 Å^2 for main-chain atoms). The residues in these regions are more flexible and hence may exhibit conformational variations. On the other hand, the high r.m.s. discrepancies in regions b and c (residues 109-112 and 210-226) reflect differences in packing environments among the four molecules. All other regions are well superimposable among the four molecules in the asymmetric unit with low r.m.s. discrepancies, indicating that the four molecules have rigid and essentially the same structures. Hence, we only describe the structure of Mol1, as a representative of the four molecules.

Overall Structure—Figure 4 shows the main chain fold-

ing of BCM β-amylase. The enzyme folds into three domains, A (residues 1-87, 135-165, 248-287, and 313-417), B (residues 88-134, 166-247, and 288-312), and C (residues 418-516). The N-terminal 417 residues fold into domains A and B, and the remaining C-terminal 99 residues into domain C. Domains A and B form one globular structure with dimensions of 70 A × 45 A × 40 A, while domain C forms another with dimensions of 35 A × 25 A × 25 A. The overall enzyme molecule has the shape of a kidneybean with a dent at the boundary between domains A and C. Domain A has a $(\beta/\alpha)_8$ barrel or TIM barrel structure (*25*), while domain B consists of three loops extending from domain A. Domain C is made up of two β-sheets of almost anti-parallel pairs of β-strands. Figure 5 shows the topology of the secondary structures and the domain arrangement of the enzyme.

Domain A, the central and catalytic domain of BCM β-amylase, folds into a $(\beta/\alpha)_8$ barrel. A $(\beta/\alpha)_8$ barrel is found invariably in plant β-amylases (*7, 26*), α-amylases of various origins (*27 34*), cyclodextrin glycosyltransferase (CGTase) (*35*), and maltotetraose-forming exo-amylase (*36*). The polypeptide segments between α-helices and β-strands of the $(\beta/\alpha)_8$ barrel are defined as loops (Fig. 5). Loops located on the carboxyl side of the β-strands are longer than those on the amino side. This feature is commonly found in $(\beta/\alpha)_8$ barrel proteins. The deep active site cleft is formed on the carboxyl side of the β-sheet of the $(\beta/\alpha)_8$ barrel. Domain A has three α-helices (α1, α2, and α7), other than those of the $(\beta/\alpha)_8$ barrel, and two 3_{10}-helices ($A3_{10}$-1 and $A3_{10}$-2). Domain B consists of three loops, L3, L4, and L5, which extend from the β-strands, Aβ3, Aβ4, and Aβ5, of the $(\beta/\alpha)_8$ barrel, respectively (Fig. 5). Domain B has some secondary structures, consisting of one 3_{10}-helix ($B3_{10}$-1) and two short anti-parallel β-sheets (Bβ1-Bβ2 and Bβ3-Bβ4) in L3, three α-helices (α3, α4, and α5) and four 3_{10}-helices ($B3_{10}$-2, $B3_{10}$-3, $B3_{10}$-4, and $B3_{10}$-5) in L4, and an α-helix (α6) in L5. The fold of domain B does not belong to any typical supersecondary structure. The sole disulfide bond in the enzyme is formed between Cys91 and Cys99 in loop L3. This S-S bond is conserved in the β-amylases from *B. polymyxa* (*37*) and *B. cereus* BQ-10 S1 SpoII (*38*). Domain B constitutes part of the active site cleft on the carboxyl side of the β-sheet of the $(\beta/\alpha)_8$ barrel. Domain C is in contact with domain A roughly on the opposite side of domain B (Fig. 4). Domain C has no contact with domain B. Domain C consists of two β-sheets packed against each other with residues 418-516. One β-sheet is composed of four β-strands, Cβ1, Cβ3, Cβ4, and Cβ8, where only Cβ1 and Cβ8 form a parallel β-strand pair. The other sheet also consists of four β-strands, Cβ2, Cβ5, Cβ6, and Cβ7, with anti-parallel hydrogen bonding. The fold of domain C is similar to those of the raw starch binding domains of CGTase (*35*) and glucoamylase (*39*).

Calcium Binding The highest electron density in the Fourier map was assigned to one calcium ion per molecule (Fig. 6a). Because no binding of calcium in β-amylases has been reported, we confirmed the existence of calcium in the enzyme by means of X-ray fluorescence experiments. This is the first finding of the binding of a metal ion in β-amylases. The calcium ion is located between loop L2 and helix

TABLE II. **Refinement statistics.**

No. of protein atoms	4 × 4,119
No. of amino acid residues	4 × 516
No. of calcium ions	4 × 1
No. of water molecules	748
Average B-factor (A²) for	
Main-chain	23.4
Side-chain	25.1
Resolution range (A)	8.0-2.2
No. of independent reflections	107,457
Completeness (%)	79.7
R-factor[a]	0.186
R_{free}[b]	0.240
Amplitude cutoff	$\|F\| > 2.0\sigma(F)$
R.m.s. deviations from ideality for	
Bonds (A)	0.014
Angles ( )	2.998
Dihedrals ( )	25.361

[a] R-factor $= \Sigma(|F_{obs}| - |F_{calc}|)/\Sigma|F_{obs}|$ where $|F_{obs}|$ and $|F_{calc}|$ are the observed and calculated structure factor amplitudes. [b] R_{free} factor is calculated using 5% of the data.

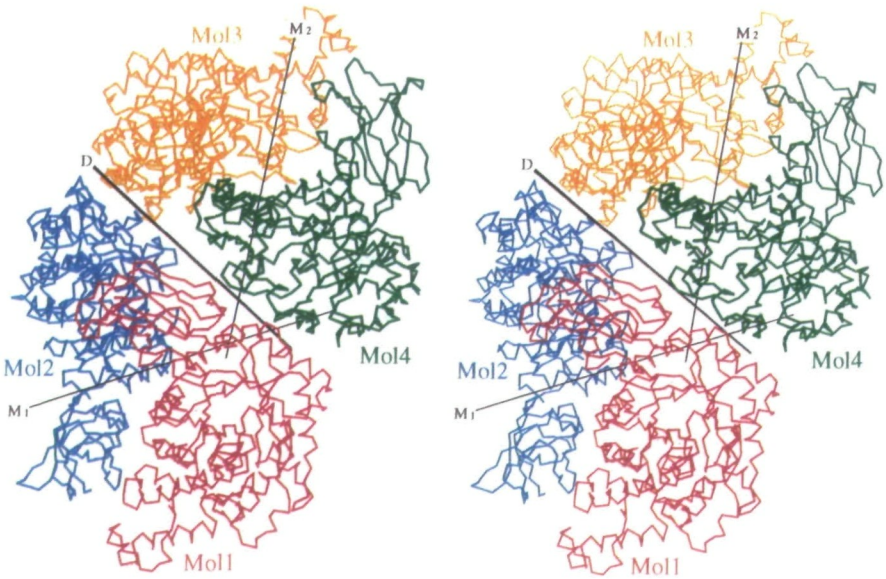

Fig. 2. **Stereo view of the "dimer of dimers" structure in the asymmetric unit.** Mol1 is shown in magenta, Mol2 in cyan, Mol3 in orange, and Mol4 in green, respectively. Three non-crystallographic symmetry axes, D, M_1, and M_2, are denoted by black lines.

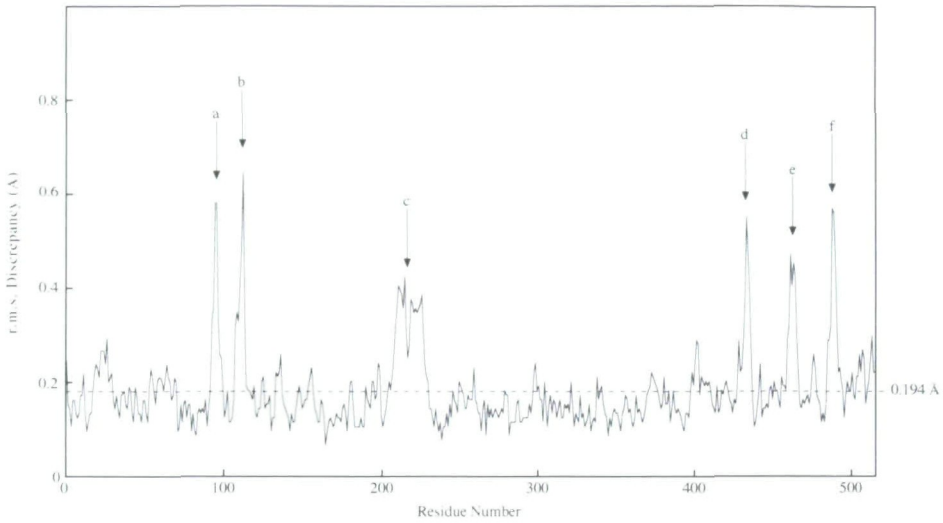

Fig. 3. **Average r.m.s. discrepancies of main-chain atoms from the mean structure averaged over the four molecules in the asymmetric unit as a function of the residue number.** The dashed line shows the average value for all the r.m.s. discrepancies of main-chain atoms. Regions with high average r.m.s. discrepancies are labeled as a to f. See the text for details.

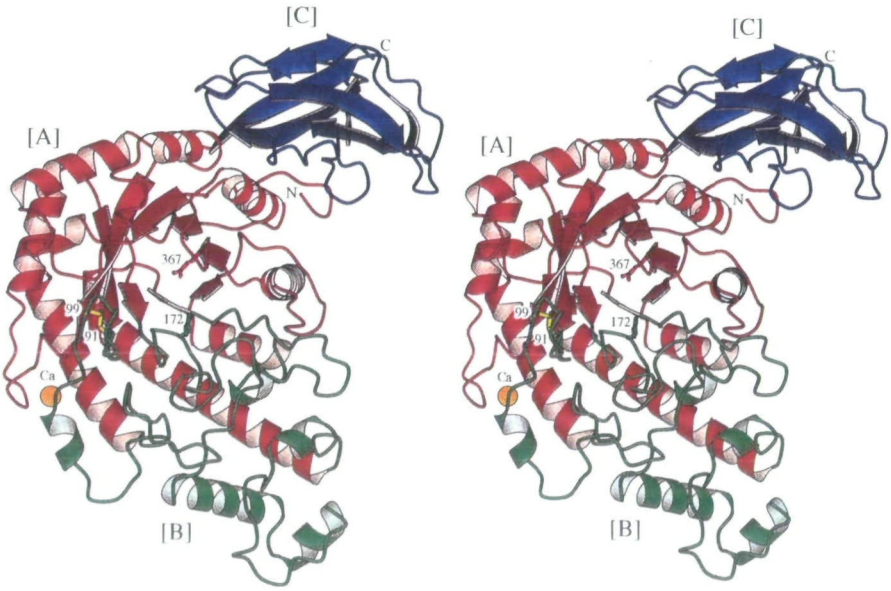

Fig. 4. **Ribbon representation of a BCM β-amylase.** Domain A is shown in red, domain B in green, and domain C in blue. The calcium ion is represented by a orange sphere. The S-S bond and the catalytic residues, Glu172 and Glu367, are shown as a ball-and-stick model. The figure was prepared with MOLSCRIPT (*44*).

Aα3 of domain A, and liganded by the five side chain oxygens of Glu56, Asp60, Asn61, Glu141, and Glu144, and a water molecule in an octahedral geometry (Fig. 6b). This calcium ion seems not to be involved in catalysis, because the ion is 20.8 and 32.2 A apart from the putative catalytic residues, Glu172 and Glu367, respectively. The function of the calcium ion is unknown at present, but it probably stabilizes the structure of the enzyme. No calcium ion is found in the crystal structure of soybean β-amylase. In the soybean β-amylase, the position of the calcium ion in BCM β-amylase is occupied by the guanidinium ion of Arg148 (*7*). Sequence alignment (our unpublished data) indicated that the calcium binding residues of the enzyme are also conserved in the β-amylases from *B. circulans* (*40*), *B. polymyxa* (*41*), and *B. cereus* BQ-10 S1 SpoII (*3*), but not plant enzymes, suggesting the former three enzymes may have a calcium ion.

Structural Comparison of Raw Starch Binding Domains We inferred that domain C of the enzyme has raw starch binding ability on comparison of the structure of domain C of the enzyme with those of raw starch binding domains, the E domain of CGTase (cyclodextrin glycosyltransferase) (PDB ID code: 1CDG, *35*), and SBD (starch binding domain) of glucoamylase (PDB ID code: 1AC0, *39*). Figure 7a shows ribbon models for domain C of the enzyme, and the raw starch binding domains of CGTase and glucoamylase. The amino acid sequence of domain C of the enzyme exhibits 31.3% identity with the raw starch binding domain of CGTase, and 28.2% identity with that of glucoamylase. The sequence identity between the raw starch binding domains of CGTase and glucoamylase is 34.0% (Fig. 7b). Despite the low sequence identities between the three homologous domains, their three dimensional structures exhibit very similar folding. The topologies of the β-strands for the three domains are almost the same, consisting of almost anti-parallel β-sheet structures. The major difference is that CGTase lacks the third β-strand of BCM β-amylase and glucoamylase. Domain C of the enzyme can be superimposed on the CGTase E domain and glucoamylase SBD with r.m.s. discrepancies of 0.59 and 1.74 A, respectively, for the C_α atoms in the β-strands and a successive loop immediately after Cβ2. Domain C of BCM

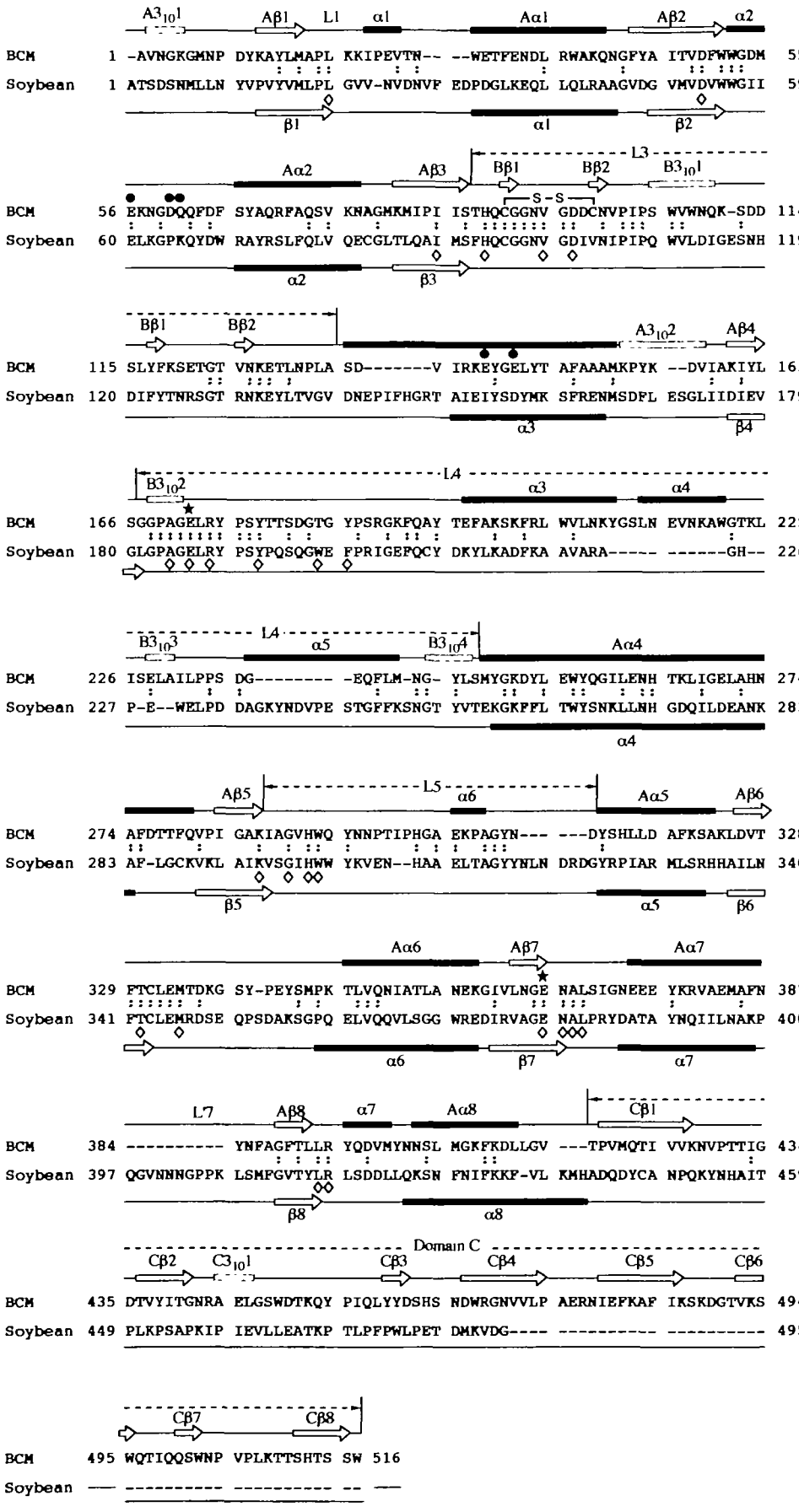

Fig. 5. **Comparison of the primary structures of BCM and soybean β-amylases.** Black bars represent α-helices, gray bars 3_{10}-helices, and arrows β-strands. For soybean β-amylase, only the secondary structure elements which comprise the $(\beta/\alpha)_8$ barrel are shown. Two catalytic residues are denoted by stars, residues involved in calcium binding of the BCM β-amylase by circles, and residues involved in substrate binding in the crystal structure of soybean β-amylase/maltose complex (PDB code: 1BYC) by diamonds.

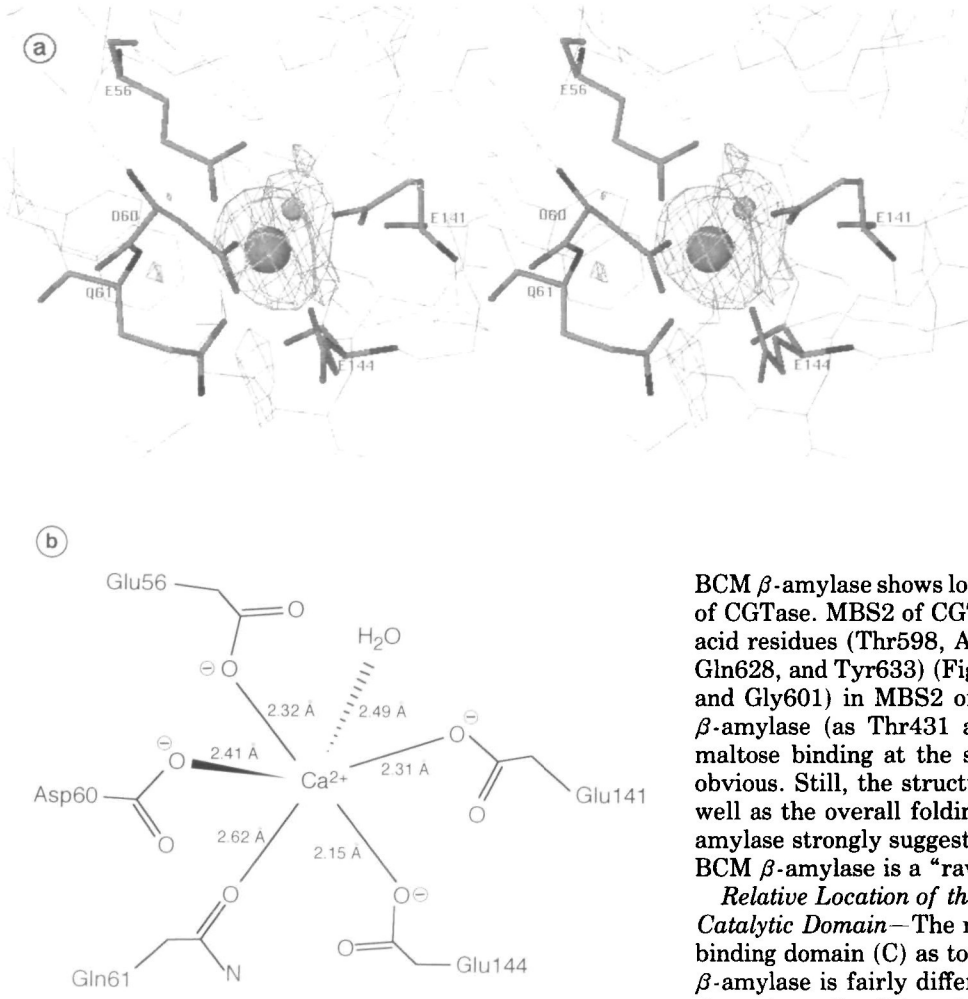

Fig. 6. **Calcium binding site of BCM β-amylase.** (a) An omit map without the contribution from calcium ions and solvent molecules. (b) Schematic representation of residues and a water molecule around the calcium ion.

β-amylase resembles the raw starch binding domain of CGTase more closely than that of glucoamylase.

In CGTase, two maltose binding sites, MBS1 (Maltose binding site 1) and MBS2 (Maltose binding site 2), in the raw starch binding domain were identified in the crystal structure of the CGTase/maltose complex, as shown in Fig. 7a (*35*). Site-directed mutation analysis of CGTase revealed that MBS1 is the most important site for raw starch binding (*42*). We compared the structure of domain C of BCM β-amylase with that of the raw starch binding domain of CGTase, paying special attention to the maltose binding sites. The first site, MBS1, of CGTase is formed from five amino acid residues, Trp616, Lys651, Trp662, Glu663 and Asn667. The counterparts in BCM β-amylase are Trp449, Lys482, Trp495, Gln496, and Gln499 (Fig. 7b). These five residues of CGTase and BCM β-amylase have similar locations and orientations in the framework of the raw starch binding domain. This indicates that maltose could bind to BCM β-amylase in a similar manner to CGTase. The crystal structure of the enzyme shows that Trp449 and Trp495 are involved in a major hydrophobic interaction between two monomers in a dimer. Trp449 and Trp495 may bind maltose when the enzyme is adsorbed on starch. In the case of glucoamylase, NMR spectroscopy demonstrated that the site corresponding to MBS1 interacts with carbohydrate (*39*). Compared to MBS1, the second site of BCM β-amylase shows lower structural similarity to MBS2 of CGTase. MBS2 of CGTase is formed from seven amino acid residues (Thr598, Ala599, Gly601, Asn603, Asn627, Gln628, and Tyr633) (Fig. 7b). Only two residues (Thr598 and Gly601) in MBS2 of CGTase are conserved in BCM β-amylase (as Thr431 and Gly434). The probability of maltose binding at the second site of the enzyme is not obvious. Still, the structural similarity of the first site as well as the overall folding to those of CGTase and glucoamylase strongly suggest that possibility that domain C of BCM β-amylase is a "raw starch binding domain."

Relative Location of the Starch Binding Domain as to the Catalytic Domain—The relative location of the raw starch binding domain (C) as to the catalytic domain (A) in BCM β-amylase is fairly different from that in CGTase. In the β-amylase, the C-terminal domain, i.e. the putative raw starch binding domain, is connected directly to the N-terminal catalytic domain in contact with helices Aα7 and Aα8 of the $(\beta/\alpha)_8$ barrel. On the other hand, CGTase has two additional domains, C and D, between the N-terminal catalytic domain and the C-terminal raw starch binding domain along the polypeptide chain (*35*). The C-terminal raw starch binding domain of CGTase is in contact with the fourth helix of the barrel. It has been noted that raw starch granules are hydrolyzed in the active site cleft of carbohydrate-degrading enzymes, not in the raw starch binding domain.

Structural Comparison between BCM and Soybean β-Amylases—To determine structural similarities and differences between bacterial and plant β-amylases, we compared the structure of BCM β-amylase with that of soybean β-amylase. In this study, the comparison was made with the structure of free soybean β-amylase (PDB ID code: 1BYA, *7*). The Cα atoms of the β-sheets in the $(\beta/\alpha)_8$ barrels were least-squared fitted, leading to a r.m.s. discrepancy of 0.66 Å for 46 equivalent C_α atoms. The superimposed Cα structures of the whole molecules of BCM and soybean β-amylases are shown in Fig. 8a. Although BCM β-amylase exhibits low sequence identity of 31.3% with soybean β-amylase (Fig. 5), its fold is very similar to that of the soybean enzyme, except for the C-terminal region. The structures of the C-terminal regions of the two enzymes are totally different. BCM β-amylase

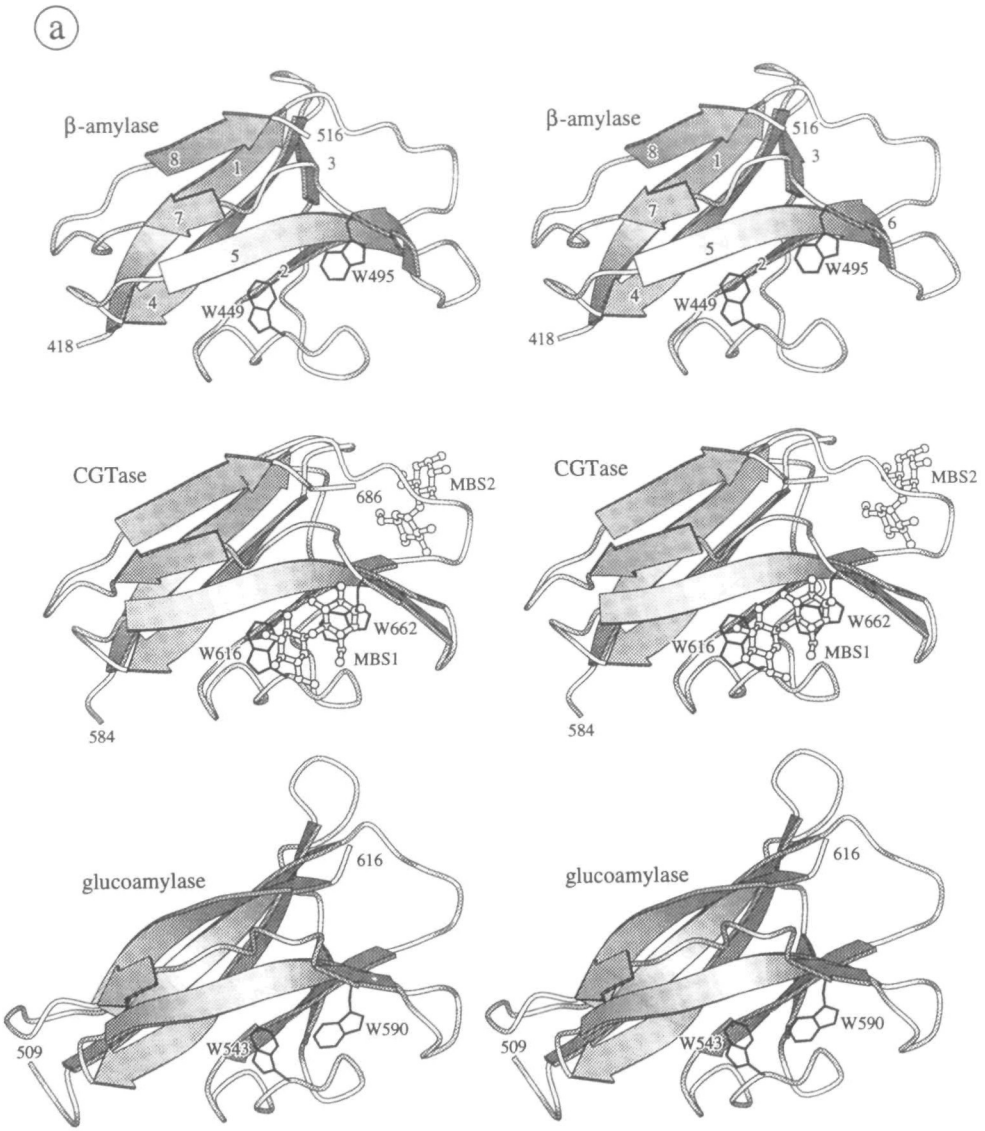

Fig. 7. **Raw starch binding domains.** (a) Ribbon drawings of BCM β-amylase (top), CGTase (center), and glucoamylase (bottom) prepared with MOLSCRIPT (*44*). β-Strands are drawn as arrows, and the two maltose molecules in MBS1 and MBS2 of CGTase as a ball-and-stick model (center). The two tryptophan residues in MBS1 of CGTase, and the corresponding residues in BCM β-amylase and glucoamylase are also shown as a wire model. (b) Sequence alignment between the raw starch binding domains. The residues in MBS1 (●) and MBS2 ( ) of CGTase are marked. The residues conserved in the three enzymes are denoted by asterisks.

b

```
                    Cβ1                 Cβ2                 Cβ3
β-amylase     418 -TP---VMQT IVVKNVPTTI GDTVYITGNR AELGSWDTKQ YPIQ-LYYDS HSN------D 466
CGTase        584 ---GDQVSVR FVVNNATTAL GQNVYLTGSV SELGNWDP-A KAIGPMYNQV VYQY----PN 635
glucoamylase  509 CTTPTAVAVT FDLTA-TTTY GENIYLVGSI SQLGDWET-S DGIA-LSAD- --KYTSSDPL 562

                    Cβ4       Cβ5       Cβ6       Cβ7             Cβ8
β-amylase     467 WRGNVVLPAE RNIEFKAFIK SKDGTVKSWQ TI-QQSWNPV PL----KTTS HTSSW----- 516
CGTase        636 WYYDVSVPAG KTIEFKFLKK Q-GSTV-TWE GGSNHTF-TA PSS---GTAT INVNWQP--- 686
glucoamylase  563 WYVTVTLPAG ESFEYKFIRI ESDDSV-EWE SDPNREY-TV PQACGTSTAT VTDTWR---- 616
```

has a putative raw starch binding domain (C domain) in its C-terminal region whereas soybean β-amylase does not have any domain structure. In domain A, the eight α-helices and eight β-strands of the $(\beta/\alpha)_8$ barrels are similar in length and geometry in the two β-amylases. The structures of the loops connecting these secondary structure elements in the two enzymes are roughly similar except for two loops, L1 and L'7. These structural features shared by the two β-amylases could be the structural foundation of β-amylases. The residues of domain B comprising the active site have similar positions in the two β-amylases, and consists of three segments of residues 88-

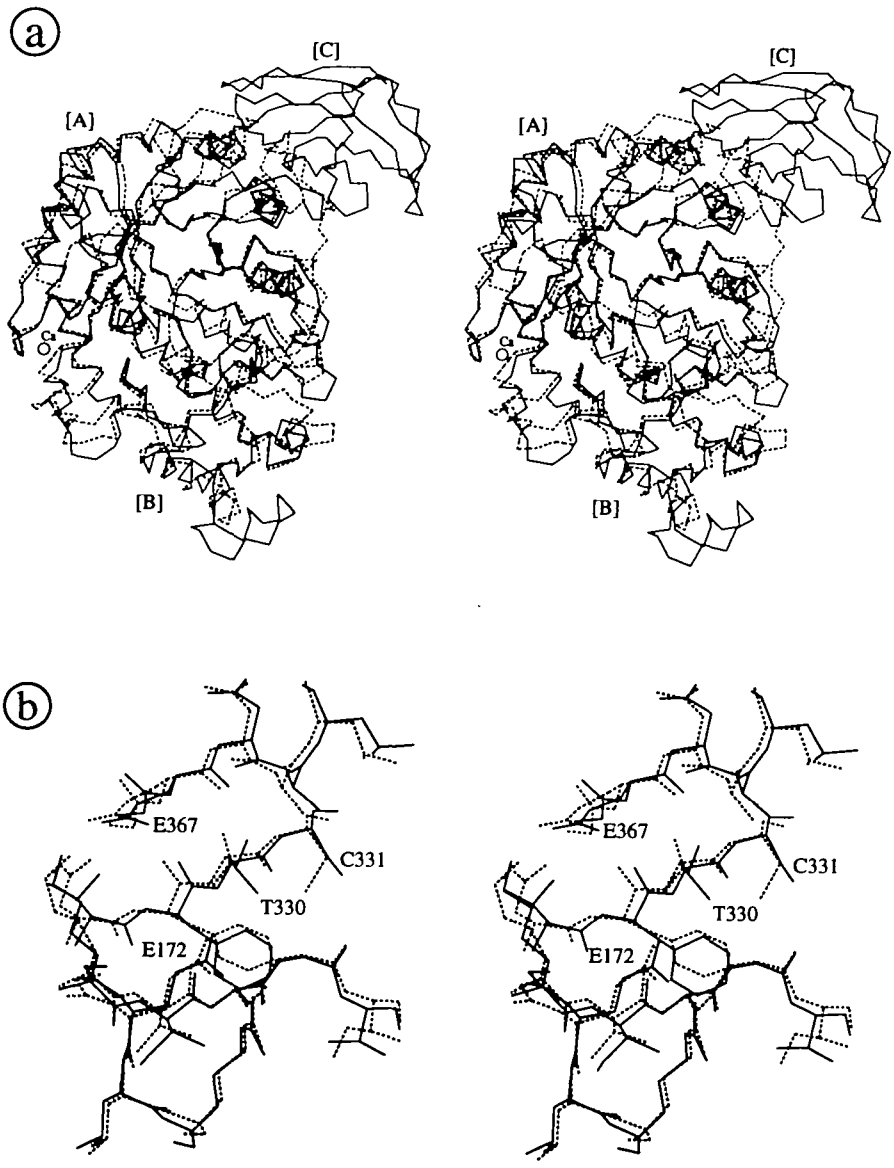

Fig. 8. **Structural comparison of BCM and soybean β-amylase.** (a) C_α representation of BCM β-amylase superimposed on soybean β-amylase. BCM β-amylase is drawn as solid lines and soybean β-amylase as dashed lines. The figure is a view from the carboxyl side of the $(\beta/\alpha)_8$ barrel along its axis. (b) A non-hydrogen atoms drawing of residues in the active site cleft for BCM β-amylase superimposed on soybean β-amylase. BCM β-amylase is drawn as solid lines and soybean β-amylase as dashed lines. Glu-172, Thr330, Cys331, and Glu367 of BCM β-amylase are labeled.

94, 166-177, and 288-293, each following β-strands Aβ3, Aβ4, and Aβ5 of the $(\beta/\alpha)_8$ barrel. Other residues of domain B have different positions in the two enzymes, corresponding to deletions or insertions in the two sequences. In spite of these structural differences, the domain Bs of the two enzymes have similar dimensions of 50 Å × 40 Å × 20 Å.

In soybean β-amylase, two glutamic acid residues (Glu186 and Glu380) are proposed to be catalytic residues (*5, 7*). Glu172 and Glu367 of BCM β-amylase, the putative catalytic residues, as judged on sequence alignment, are well superimposable on the corresponding residues of soybean β-amylase, as shown in Fig. 8b. Site-directed mutagenesis also revealed the roles of Glu172 and Glu367 as catalytic residues (our unpublished data). The crystal structure of soybean β-amylase in a complex with maltose had two bound saccharides in the active site cleft, indicating that 24 amino acid residues including the two catalytic ones are in contact with the bound saccharides through hydrogen bonds or hydrophobic interactions at subsites 1 through 4 (Fig. 5) (*7*). These 24 residues of soybean β-amylase are almost completely conserved in the sequence of and also structurally in BCM β-amylase. It is hence inferred that the binding manner of BCM β-amylase as to saccharides is similar to that of soybean β-amylase.

Conspicuous differences between the two β-amylases, however, were observed in the side-chain conformation of two residues in loop L6 located between subsites 2 and 3. The χ_1 angles of Thr330 and Cys331 of BCM β-amylase differ by about 180° and 90° from those of the corresponding residues of Thr342 and Cys343 of soybean β-amylase, respectively. Thr342 of soybean β-amylase interacts with a bound saccharide through hydrogen bonding in a crystal of a complex with maltose (*7*). Chemical modification of Cys343 of soybean β-amylase caused a loss of activity (*43*), although the residue does not interact directly with a bound saccharide in the soybean β-amylase/maltose complex (*7*). Such observed differences in the side-chain conformation of the two residues between BCM and soybean β-amylases may cause the differences in kinetic behavior between them

(5). The tertiary structures of BCM β-amylase in complexes with saccharides are needed to fully understand the interactions with saccharides in the active site. Our preliminary X-ray analysis of BCM β-amylase crystals soaked in several saccharide solutions showed bound saccharides in the active site cleft in electron density maps. We are now conducting structure analyses of these crystals and will published the results elsewhere.

This work was performed with the approval of the Photon Factory Program Advisory Committee (Proposal No. 95G064). We wish to thank Dr. N. Sakabe and his colleagues for their cooperation and help with the data collection at the Photon Factory.

REFERENCES

1. Mikami, B., Morita, Y., and Fukazawa, C. (1988) Primary structure and function of β-amylase (in Japanese). *Seikagaku* **60**, 211-216
2. Toda, H. (1993) Sequence analysis of sweet potato β-amylase (in Japanese). *Denpun Kagaku* **36**, 87-101
3. Nanmori, T., Nagai, M., Shimizu, Y., Shinke, R., and Mikami, B. (1993) Cloning of the β-amylase gene from *Bacillus cereus* and characteristics of the primary structure of the enzyme. *Appl. Environ. Microbiol.* **59**, 623-627
4. Pujadas, G., Ramírez, F.M., Valero, R., and Palau, J. (1996) Evolution of β-amylase: patterns variation and conservation in subfamily sequences in relation to parsimony mechanisms. *Proteins* **25**, 456-472
5. Nitta, Y., Isoda, Y., Toda, H., and Sakiyama, F. (1989) Identification of glutamic acid 186 affinity-labeled by 2,3-epoxypropyl α-D-glucopyranoside in soybean β-amylase. *J. Biochem.* **105**, 573-576
6. Toda, H., Nitta, Y., Asanami, S., Kim, J.P., and Sakiyama, F. (1993) Sweet potato β-amylase. Primary structure and identification of the active-site glutamyl residue. *Eur. J. Biochem.* **216**, 25-38
7. Mikami, B., Degano, M., Hehre, E.J., and Sacchettini, J.C. (1994) Crystal structures of soybean β-amylase reacted with β-maltose and maltal: active site components and their apparent roles in catalysis. *Biochemistry* **33**, 7779-7787
8. Nitta, Y., Shirakawa, M., and Takasaki, Y. (1996) Kinetic study of active site structure of β-amylase from *Bacillus cereus* var. *mycoides*. *Biosci. Biotech. Biochem.* **60**, 823-827
9. Svensson, B., Jespersen, H., Sierks, M.R., and MacGregor, E.A. (1989) Sequence homology between putative raw-starch binding domains from different starch-degrading enzymes. *Biochem. J.* **264**, 309-311
10. Takasaki, Y. (1976) Purifications and enzymatic properties of β-amylase and pullulanase from *Bacillus cereus* var. *mycoides*. *Agric. Biol. Chem.* **40**, 1523-1530
11. Yamaguchi, T., Matsumoto, Y., Shirakawa, M., Kibe, M., Hibino, T., Shunji, K., Takasaki, Y., and Nitta, Y. (1996) Cloning, sequencing, and expression of a β-amylase gene from *Bacillus cereus* var. *mycoides* and characterization of its products. *Biosci. Biotech. Biochem.* **60**, 1255-1259
12. Oyama, T., Kusunoki, M., Kishimoto, Y., Takasaki, Y., and Nitta, Y. (1998) Crystallization and preliminary X-ray analysis of β-amylase from *Bacillus cereus* var. *mycoides*. *Protein Pept. Lett.* **5**, 349-354
13. Matthews, B.W. (1968) Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491-497
14. Sakabe, N. (1991) X-ray diffraction data collection system for modern protein crystallography with a Weissenberg camera and an imaging plate using synchrotron radiation. *Nuclear Instr. Methods Physics Res.* **A303**, 448-463
15. Otwinowski, Z. (1991) Isomorphous replacement and anomalous scattering in *Proceedings of the CCP4 Study Weekend* (Wolf, W., Evans, P.R., and Leslie, A.G.W., eds.) pp. 80-88, SERC Daresbury Laboratory, Warrington, UK
16. Collaborative Computational Project Number 4 (1994) The CCP4 Suite: Programs for protein crystallography. *Acta Crystallogr.* **D50**, 760-763
17. Cowton, K. (1994) CCP4 density modification package in *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* Vol. 31, pp. 34-38
18. Jones, T.A. (1978) A graphics model building and refinement system for macromolecules. *J. Appl. Cryst.* **11**, 268-272
19. Brünger, A.T., Kuriyan, J., and Karplus, M. (1987) Crystallographic R factor refinement by molecular dynamics. *Science* **235**, 458-460
20. Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr.* **D53**, 240-255
21. Luzzati, P.V. (1952) Traitement statistique des erreurs dans la determination des structures cristallines. *Acta Crystallogr.* **5**, 802-810
22. Ramachandran, G.N. and Sasisekharan, V. (1968) Conformation of polypeptides and proteins. *Adv. Protein Chem.* **23**, 283-437
23. Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993) *J. Appl. Crystallogr.* **26**, 283-291
24. Harding, S.E. (1994) Determination of diffusion coefficients of biological macromolecules by dynamic light scattering. *Methods Mol. Biol.* **22**, 97-108
25. Banner, D.W., Bloomer, A.C., Petsko, G.A., Phillips, D.C., Pogson, C.I., Wilson, I.A., Corran, P.H., Furth, A.J., Milman, J.D., Offord, R.E., Priddle, J.D., and Waley, S.G. (1975) Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 angstrom resolution using amino acid sequence data. *Nature* **255**, 609-614
26. Cheong, C.G., Eom, S.H., Chang, C., Shin, D.H., Song, H.K., Min, K., Moon, J.H., Kim, K.K., Hwang, K.Y., and Suh, S.W. (1995) Crystallization, molecular replacement solution, and refinement of tetrameric β-amylase from sweet potato. *Proteins* **21**, 105-117
27. Boel, E., Brady, L., Brzozowski, A.M., Derewenda, Z., Dodson, G.G., Jensen, V.J., Petersen, S.B., Swift, H., Thim, L., and Wöldike, H.F. (1990) Calcium binding in α-amylases: an X-ray diffraction study at 2.1 Å resolution of two enzymes from *Aspergillus*. *Biochemistry* **29**, 6244-6249
28. Brady, L., Brzozowski, A.M., Derewenda, Z., Dodson, E.J., and Dodson, G.G. (1991) Solution of the structure of *Aspergillus niger* acid α-amylase by combined molecular replacement and multiple isomorphous replacement methods. *Acta Crystallogr.* **B47**, 527-535
29. Qian, M., Haser, R., and Payan, F. (1993) Structure and molecular model refinement of pig pancreatic α-amylase at 2.1 Å resolution. *J. Mol. Biol.* **231**, 785-799
30. Kadziola, A., Abe, J., and Svensson, H.R. (1994) Crystal and molecular structure of barley α-amylase. *J. Mol. Biol.* **239**, 104-121
31. Machius, M., Wiegand, G., and Huber, R. (1995) Crystal structure of calcium-depleted *Bacillus licheniformis* α-amylase at 2.2 Å resolution. *J. Mol. Biol.* **246**, 545-559
32. Ramasubbu, N., Paloth, V., Luo, Y., Brayer, G.D., and Levine, M.J. (1996) Structure of human salivary α-amylase at 1.6 Å resolution: Implications for its role in the oral cavity. *Acta Crystallogr.* **D52**, 425-446
33. Fujimoto, Z., Takase, K., Doui, N., Momma, M., Matsumoto, T., and Mizuno, H. (1998) Crystal structure of a catalytic-site mutant α-amylase from *Bacillus subtilis* complexed with maltopentaose. *J. Mol. Biol.* **277**, 393-407
34. Strobl, S., Maskos, K., Betz, M., Wiegand, G., Huber, R., Gomis-Rüth, F.X., and Glockshuber, R. (1998) Crystal structure of yellow meal worm α-amylase at 1.64 Å resolution. *J. Mol. Biol.* **278**, 617-628
35. Lawson, C.L., van Montfort, R., Strokopytov, B., Rozeboom, H.J., Kalk, K.H., de Vries, G.E., Penninga, D., Dijkhuizen, L., and Dijkstra, B.W. (1994) Nucleotide sequence and X-ray structure of cyclodextrin glycosyltransferase from *Bacillus circulans* strain 251 in a maltose-dependent crystal form. *J. Mol. Biol.* **236**, 590-600
36. Morishita, Y., Hasegawa, K., Matsuura, Y., Katsube, Y.,

Kubota, M., and Sakai, S. (1997) Crystal structure of a maltotetraose-forming exo-amylase from *Pseudomonas stutzeri*. *J. Mol. Biol.* **267**, 661-672
37. Uozumi, N., Matsuda, T., Tsukagoshi, N., and Udaka, S. (1991) Structural and functional roles of cysteine residues of *Bacillus polymyxa* β-amylase. *Biochemistry* **30**, 4594-4599
38. Nomura, K., Yoneda, I., Nanmori, T., Shinke, R., Morita, Y., and Mikami, B. (1995) The role of SH and S-S groups in *Bacillus cereus* β-amylase. *J. Biochem.* **118**, 1124-1130
39. Sorimachi, K., Jacks, A.J., Le Gal-Coëffet, M.F., Williamson, G., Archer, D.B., and Williamson, M.P. (1996) Solution structure of the granular starch binding domain of glucoamylase from *Aspergillus niger* by nuclear magnetic resonance spectroscopy. *J. Mol. Biol.* **259**, 970-987
40. Siggens, K.W. (1987) Molecular cloning and characterization of the β-amylase gene from *Bacillus circulans*. *Mol. Microbiol.* **1**, 86-91
41. Kawazu, T., Nakanishi, Y., Uozumi, N., Sasaki, T., Yamagata, H., Tsukagoshi, N., and Udaka, S. (1987) Cloning and nucleotide sequence of the gene coding for enzymatically active fragments of the *Bacillus polymyxa* β-amylase. *J. Bacteriol.* **169**, 1564-1570
42. Penninga, D., van der Veen, B.A., Knegtel, R.M., van Hijum, S.A., Rozeboom, H.J., Kalk, K.H., Dijkstra, B.W., and Dijkhuizen, L. (1996) The raw starch binding domain of cyclodextrin glycosyltransferase from *Bacillus circulans* strain 251. *J. Biol. Chem.* **271**, 32777-32784
43. Mikami, B., Nomura, K., and Morita, Y. (1994) Two sulfhydryl groups near the active site of soybean β-amylase. *Biosci. Biotech. Biochem.* **58**, 126-132
44. Kraulis, P.J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946-950