The 2.0-Å Resolution Structure of Soybean β-Amylase Complexed with α-Cyclodextrin†,‡

Bunzo Mikami,§,‖ Edward J. Hehre,⊥ Mamoru Sato,# Yukiteru Katsube,# Masaaki Hirose,§ Yuhei Morita,§,○ and James C. Sacchettini*

Departments of Biochemistry and of Microbiology and Immunology, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461, Research Institute for Food Science, Kyoto University, Uji, Kyoto 611, Japan, and the Institute for Protein Research, Osaka University, Suita, Osaka 565, Japan

Received November 20, 1992; Revised Manuscript Received February 22, 1993

ABSTRACT: New crystallographic findings are presented which offer a deeper understanding of the structure and functioning of β-amylase, the first known exo-type starch-hydrolyzing enzyme. A refined three-dimensional structure of soybean β-amylase, complexed with the inhibitor α-cyclodextrin, has been determined at 2.0-Å resolution with a conventional *R*-value of 17.5%. The model contains 491 amino acid residues, 319 water molecules, 1 sulfate ion, and 1 α-cyclodextrin molecule. The protein consists of a core with an $(\alpha/\beta)_8$ supersecondary structure, plus a smaller globular region formed by long loops (L3, L4, and L5) extending from β-strands β3, β4, and β5. Between the two regions is a cleft that opens into a pocket whose floor contains the postulated catalytic center near the carboxyl group of Glu 186. The annular α-cyclodextrin binds in (and partly projects from) the cleft with its glucosyl O-2/O-3 face abutting the $(\alpha/\beta)_8$ side and with its α-D(1→4) glucosidic linkage progression running clockwise as viewed from that side. The ligand does not bind deeply enough to interact with the carboxyl group of Glu 186. Rather, it occupies most of the cleft entrance, strongly suggesting that α-cyclodextrin inhibits catalysis by blocking substrate access to the more deeply located reaction center. Of the various α-cyclodextrin interactions with protein residues in loops L4, L5, L6, and L7, most notable is the shallow inclusion complex formed with Leu 383 (in L7, on the core side of the cleft) through contacts of its methyl groups with the C-3 atoms of four of the ligand's D-glucopyranosyl residues. All six residues of the bound α-cyclodextrin are of $^4C_1$ conformation and are joined by α-1,4 linkages with similar torsional angles to form a nearly symmetrical torus as reported for crystalline inclusion complexes with α-cyclodextrin. We envision a significant role for the methyl groups of Leu 383 at the cleft entrance with respect to the productive binding of the outer chains of starch.

This paper reports the detailed three-dimensional structure of β-amylase (EC 3.2.1.2) (specifically, isozyme 2 from soybeans, *Glycine max*), bringing this enzyme into the select group of proteins whose structures have been resolved at or below 2-Å resolution. β-Amylase has long been of special interest to carbohydrate enzymologists because of its unique action in catalyzing the hydrolytic cleavage of successive maltosyl residues (to form β-maltose) from the nonreducing ends of outer chains of starch and maltooligosaccharides. The enzyme is produced by higher plants and some microorganisms, and the amino acid sequences from several sources show distinct conserved regions but little or no relation to those of other starch-degrading enzymes (Svensson, 1988; Mikami *et al.*, 1988). Early studies conducted with the crystalline sweet potato enzyme (Balls *et al.*, 1948) brought new fundamental information to light on the physicochemical properties and catalytic function of β-amylase. In a probe of the enzyme's apparent need for substrates having a terminal D-glucosyl residue with a free 4-OH group, Thoma and Koshland (1960) found α-cyclodextrin (α-CD)[1] to be a competitive inhibitor and presented this as supporting evidence for the induced fit theory of catalysis. Thoma *et al.* (1963) reported evidence for the binding of one α-cyclodextrin molecule per catalytic center and for the existence of one center per (ca. 50 000 Da) protein subunit. Sweet potato β-amylase crystals were subjected to an X-ray diffraction study by Colman and Matthews (1971); however, due to the high molecular weight of this tetrameric enzyme and the large dimensions of the unit cell, only weak diffraction patterns were observed.

In 1975, Morita *et al.* reported the preparation of good-quality trigonal crystals of soybean β-amylase (SBA),[1] a monomer of 57 000 Da. Initial diffraction data at low resolution (Aibara *et al.*, 1984) indicated that the protein consists of a large and a small domain with a cleft between them. Elucidation of the full amino acid sequence of SBA (Mikami *et al.*, 1988) permitted more refined studies. Thus, Nitta *et al.* (1989) identified the carboxyl group of Glu 186 as the catalytically active residue in a unique peptide sequence labeled by an affinity reagent. Moreover, from a low- (6-Å) resolution model of the three-dimensional structure of a SBA/maltose complex, Mikami *et al.* (1991) identified the location of two maltose binding sites at the base of the cleft. Better diffraction data were obtained for the β-amylase complexed

† This study was supported by Research Grant GM-45859 from the National Institutes of Health (to J.C.S.) and Research Grant DMB 89-04332 from the National Science Foundation (to E.J.H.).

‡ Crystallographic coordinates have been deposited in the Brookhaven Protein Data Bank under the file name 1BTC.

* Address correspondence to this author at the Department of Biochemistry, Albert Einstein College of Medicine.

§ Research Institute for Food Science, Kyoto University.

‖ Research Associate in Biochemistry and in Microbiology and Immunology, Albert Einstein College of Medicine, on leave from the Research Institute for Food Science, Kyoto University.

⊥ Department of Microbiology and Immunology, Albert Einstein College of Medicine.

Institute for Protein Research, Osaka University.

○ Present affiliation: Research and Development Center, Fuji Oil Co., Ltd., Yawara, Tsukubagun, Ibaraki 300-24, Japan.

[1] Abbreviations: SBA, soybean β-amylase; α-CD, α-cyclodextrin (cyclohexaamylose).

0006-2960/93/0432-6836$04.00/0 © 1993 American Chemical Society

with α-cyclodextrin, and its initial model was built at 3 Å and an *R*-factor of 48% (Mikami *et al.*, 1992). The model indicated that the enzyme has an $(\alpha/\beta)_8$ core structure similar to that reported for α-amylases and cyclodextrin glycosyltransferase but differs from them in its loop structures. The model showed the α-CD bound in the cleft between the enzyme's core and loops L3–L5 but could not define the orientation of the α-D-glucosyl residues of the ligand to the enzyme.

Present efforts have succeeded in refining to high resolution the crystal structure of the β-amylase/α-CD complex. We describe the structure and the stereochemical orientation and binding interactions of α-CD to residues of the protein. Comparison is made with other enzymes with respect to overall structure and to the locus of substrate and analogue binding. The spatial disposition of the bound ligand relative to the presumed catalytic reaction center is discussed as well as its bearing on the mechanism whereby α-CD inhibits β-amylase. Studies on the relation of structure to other aspects of catalysis by β-amylase, e.g., to the stereochemistry of reactions promoted with β-maltose (Hehre *et al.*, 1969) and maltal (Hehre *et al.*, 1986; Kitahata *et al.*, 1991), are in progress with comparably refined structures of SBA complexed with β-maltose and with maltal (Mikami, Hehre, and Sacchettini, unpublished results).

EXPERIMENTAL PROCEDURES

Trigonal crystals of soybean β-amylase, isozyme 2, prepared in 1983 as described by Morita *et al.* (1975), were used; they had been stored at 4 °C for 8 years in 0.1 M acetate buffer, pH 5.4, containing 50% saturated ammonium sulfate, 0.1 M NaCl, 18 mM 2-mercaptoethanol, and 1 mM EDTA. Assays of catalytic activity, carried out by incubating a buffered solution of these crystals with amylopectin at pH 5.4 according to Morita *et al.* (1975), gave hydrolysis rates that were less than 5% of that shown by fresh, native SBA. Activity approaching that of the latter could, however, be recovered by treating the crystals with 2-mercaptoethanol at pH 7 as described for the purified solution enzyme (Mikami *et al.*, 1982).

The crystals (0.8 × 0.5 × 0.5 mm) belong to space group $P3_121$ with unit cell dimensions $a = b = 86.1$ Å and $c = 144.3$ Å (Morita *et al.*, 1975; Mikami *et al.*, 1992). They comprised ca. 52% protein and showed one molecule per asymmetric unit (Morita *et al.*, 1975). V_M (Mathews, 1968) = 2.76 Å^3/Da. Prior to X-ray diffraction analysis, the crystal was soaked at 20 °C for 20 h in 0.1 M acetate buffer, pH 5.4, containing 50% saturated ammonium sulfate and 9 mM α-cyclodextrin.

Data Collection and Phase Calculations. The data on a $K_3UO_2F_5$ derivative and two CH_3HgOH [Hg(1) and Hg(3)] derivatives of the crystal were used to solve the protein structure to 3-Å resolution as described by Mikami *et al.* (1992). Data on the crystal complexed with α-CD were collected to a resolution of 2 Å using a Rigaku R-AXIS IIc imaging plate system (Sato *et al.*, 1992). The image data frames were converted to indexed intensity data sets by Higashi's auto-indexing procedure (Higashi, 1990) based on Wonacott's strategy (Wonacott, 1980). The full data set was obtained by merging and scaling the indexed data sets according to the method of Fox and Holmes (1966). Table I summarizes the data collection and statistics. The heavy atom parameters were refined by using PHASIT from the PHASES package (W. Furey, VA Medical Center and University of Pittsburgh, Pittsburgh, PA). The mean figure of merit was 0.57 for 11 396 reflections.

Solvent flattening and phase combination (Wang, 1985) were applied to the isomorphous replacement phase using the PHASES package. Sixteen cycles of solvent flattening were carried out. The solvent fraction of the unit cell was set to 45% for the first four cycles, 50% for the next four cycles, and 55% for the final eight cycles. The final *R*-factor and figure-of-merit were 0.263 and 0.811 for 11 396 reflections, respectively.

Table I: Summary of Data Collection and Reduction

X-ray source	Cu Kα
X-ray optics	Franks double mirror optics
monochromatization	Ni foil and mirror
IP size	200 mm × 200 mm
pixcel size	105 μm
no. of crystals used	1
resolution limit	2.0 Å
oscillation range per frame	1.5°
total oscillation range	31.5°
exposure time	45 min/frame
no. of observed reflections	61927
no. of independent reflections	32958
completeness	77%
merge *R*	
full reflections	4.28%
partial reflections	6.27%
total reflections	4.66%
no. of rejected reflections	
full reflections	0
partial reflections	2

Model Building and Refinement. The reported (Mikami *et al.*, 1992) model was rebuilt using the solvent-flattened electron density map which revealed several improperly positioned main-chain residues in loop regions. The positions of 177 amino acids were corrected by using the TOM molecular graphics package, a derivative of FRODO (Jones, 1988), on an IRIS 4D computer (Silicon Graphics). Extra electron density was observed about four of the protein's six cysteine residues. The additional density, apparently due to covalent modification, was fitted as a 2-mercaptoethanol ($-SCH_2CH_2-OH$) adduct in each case. Refinement of this new model was then carried out by simulated annealing, using XPLOR (Brünger *et al.*, 1987) as described in Table II. The results obtained in the course of refinement are given in Table III. The temperature factors for all atoms of the starting model were set to 18 Å^2. After each macrocycle the structure was reevaluated and improved manually by using the program TOM. The multiple isomorphous replacement (MIR) electron density map was used to fit the atoms of β-amylase together with current model phased $|F_o| - |F_c|$ and $2|F_o| - |F_c|$ electron maps through macrocycle 3. The structure of the α-CD molecule was clear in the MIR map, and it was incorporated (as separate D-glucopyranosyl residues) after macrocycle 1. Water molecules were incorporated from macrocycle 3 through macrocycle 9, wherever the density in $|F_o| - |F_c|$ map was higher than 4σ and the distances between the waters and hydrogen bond donors or acceptors were between 2.5 and 3.2 Å. The most difficult positions to build were the four N-terminal residues and residues 96–104 of loop 3 (L3), owing to incomplete electron density. On completion of the fifth round of simulated annealing, positional refinement was carried out without including loop 3 in order to obtain an omit map. This electron density map showed sufficiently distinct density to allow the positioning of the model's coordinates for loop 3 (except for Gly 97). The four N-terminal residues were removed after macrocycle 6. For the last stage of refinement, overall anisotropic *B*-factor refinement was used.

RESULTS

Description of the Refined Model. The 2.0-Å resolution structure contains 4012 non-hydrogen protein atoms including 4 modified cysteine residues, 319 ordered water molecules,

Table II: Protocol of XPLOR Refinement of the α-CD/SBA Complex

cycle	WA[a] (Mcal/mol)	POSI[b] cycles	PREP[c] cycles	3000 K[d] ps	POSI[e] cycles	BREF[f] cycles	POSI cycles	BOVER[g] cycles
1	250	20	40	0.5	20			
2	280	20	40	0.5	30	10	20	
3	400				40	10	40	
4	350				100	10	60	
5	350				100	10	60	
6	350				100	5	60	
7	300				100	5	80	10
8	300				100	5	80	10
9	300				120	5	120	10

[a] WA, weight for the effective energy term accounting for the diffraction. [b] POSI, conjugated gradient minimization, tolerance ΔF = 0.05 Å. If any atom movement exceeds ΔF, the first derivatives of the effective energy term are recalculated. [c] PREP, conjugated gradient minimization with soft repulsive potential, ΔF = 0.05 Å. [d] 3000 K, molecular dynamics at a "temperature" of 3000 K, time step = 0.5 fs, ΔF = 0.2 Å. [e] POSI, conjugated minimization, ΔF = 0. [f] BREF, individual *B*-factor refinement. [g] BOVER, overall anisotropic *B*-factor refinement.

Table III: Course of Refinement Using XPLOR

	macrocycle no.									
	0[a]	1	2	3	4	5	6	7	8	9
resolution range (Å)	10–2.2	10–2.2	10–2.2	9–2.1	9–2.0	9–2.1	9–2.1	9–2.0	9–2.0	9–2.0
no. of reflections	25221	25221	25221	28372	30555	28105	28105	30555	30080	30045
R-factor (%)	53.0	27.6	23.9	22.6	21.4	18.3	17.4	17.8	17.7	17.5
no. of amino acids	495	495	495	495	495	495	495	495	491	491
α-CD included	–	–	+	+	+	+	+	+	+	+
no. of solvent molecules	0	0	0	109	190	301	311	321	321	320[b]
rms deviation of										
bond length (Å)	0.025	0.021	0.018	0.020	0.019	0.016	0.017	0.015	0.015	0.014
bond angle (deg)	3.25	4.25	3.91	3.88	3.70	3.31	3.35	3.20	3.11	3.08
dihedral (deg)	28.3	26.0	25.1	25.5	25.6	25.2	25.2	25.1	24.8	24.9
improper angle (deg)	1.93	1.84	1.52	1.67	1.52	1.42	1.36	1.26	1.23	1.16
no. of corrected amino acids		37	31	18	6	3	15	5	1	0

[a] Data before simulated annealing refinement. [b] 320 refers to 319 water molecules and 1 sulfate ion.

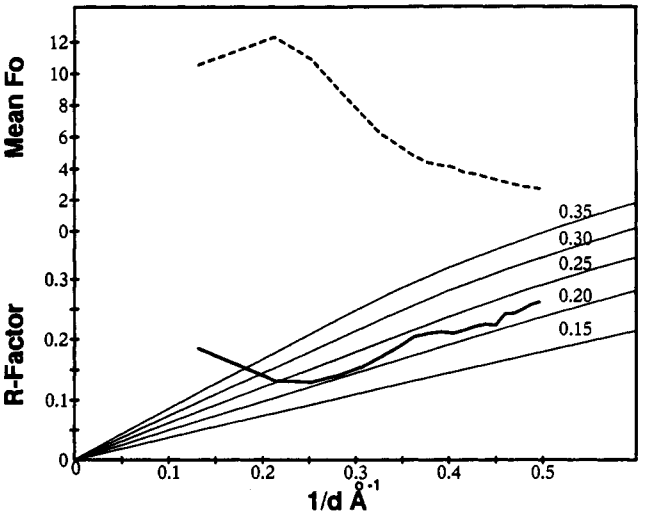

FIGURE 1: Luzzati plot for the refined SBA structure with bound α-CD. Curves are calculated with all data in the resolution range 9–2 Å for rms coordinate errors of 0.2–0.5 Å.

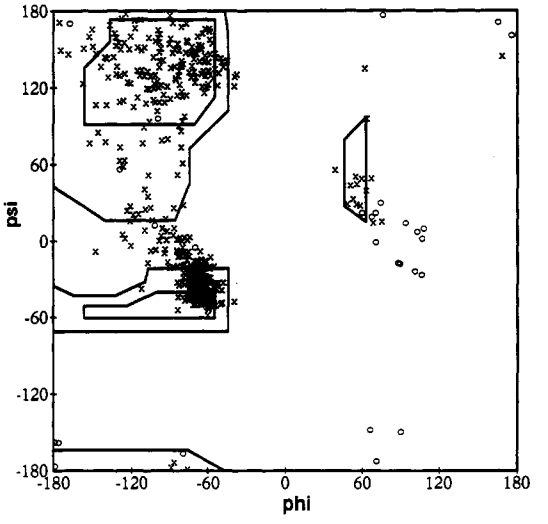

FIGURE 2: Ramachandran plot for the refined SBA structure with bound α-CD. Glycines are plotted as circles; all other residues, as crosses. The ϕ and ψ angles of all residues are in the "allowed" regions of low energy with the exception of Arg 420 (ϕ, 65.58°; ψ, 135.93°).

and 1 sulfate ion. The final *R*-factor for the refined model was 17.5 using data from 9.0- to 2.0-Å resolution (30 045 reflections). Root mean square deviations from standard bond geometry were 0.014 Å for bond lengths, 3.1° for bond angles, 24.9° for torsion angles, and 1.16° for planarity (see Table III).

Figure 1 provides a Luzzati (1952) plot of the *R*-factor and the mean F_o versus $2 \sin\theta/\lambda$ (resolution). The plot also reveals a direct correlation between *R*-factor and structure factor amplitudes. For 9–6-Å resolution data the high *R*-factors are likely due to the fact that we did not employ a solvent continuum factor during the course of the refinement. The sharp falloff in the structure factor amplitude at high resolution probably contributes to the increased *R*-factor at this resolution range. The curve allowed us to estimate an average positional standard deviation of between 0.20 and 0.25 Å from the theoretical curves calculated according to Luzzati (1952).

Figure 2 is a Ramachandran plot of the main-chain torsion angles of the final SBA model. Most of the residues fall in or near the energetically preferred regions. One non-glycine residue, Arg 420 (ϕ, 66°; ψ, 136°), is the exception. It has well-defined density, is located at the end of a β-strand (residues 415–419), and is in the $i + 1$ position of a distorted type 1 turn.

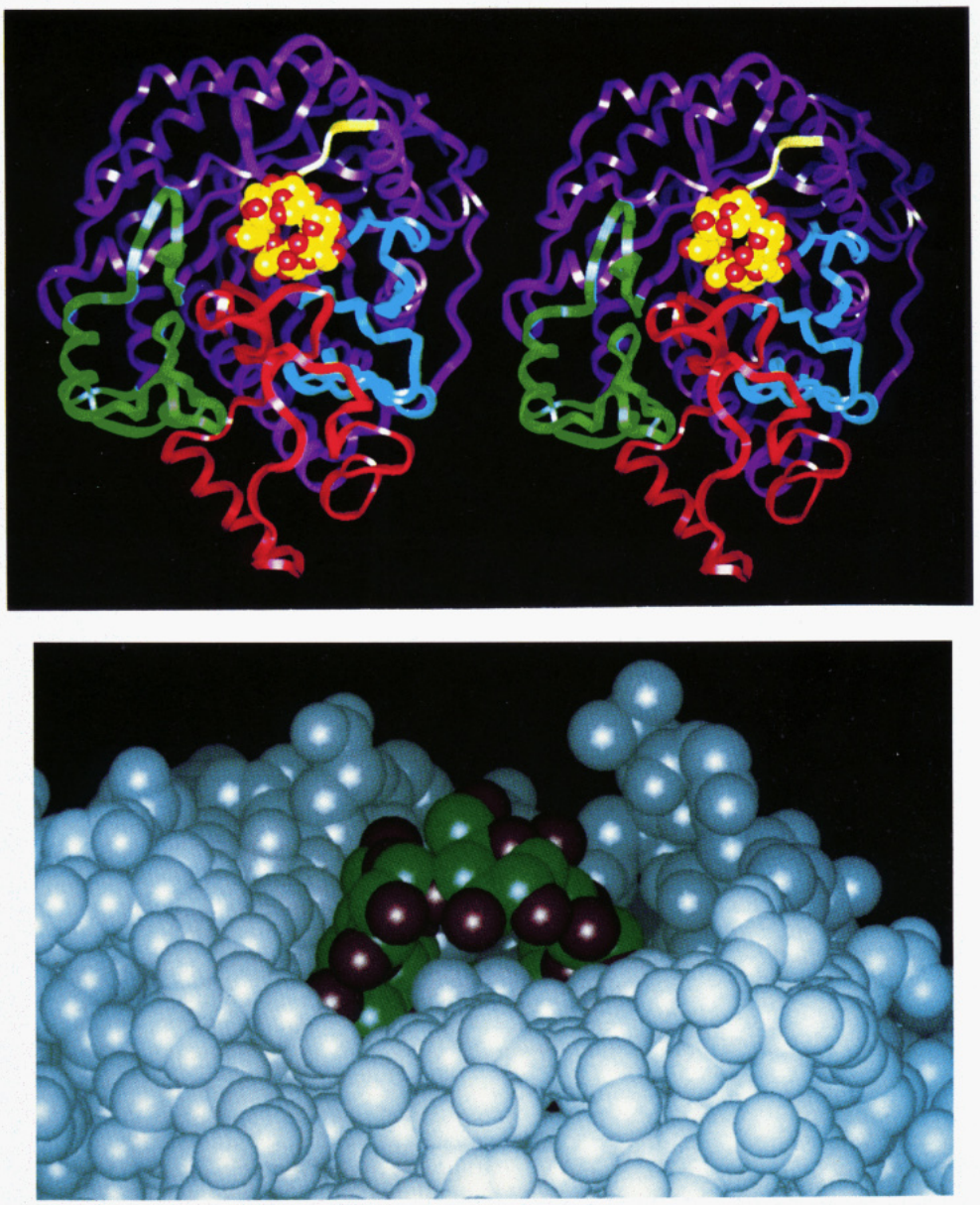

FIGURE 3: (Top) Ribbon representation of the SBA structure with α-CD bound in the pocket between the $(\alpha/\beta)_8$ core (purple) and the lobe formed by loops L3 (green), L4 (red), L5 (blue-green), L6 (blue), and L7 (white). The view looks to the carboxyl end of the β-barrel and the O-6 face of the ligand. (Bottom) Representation of the surface of the protein with α-CD bound in and partly blocking the cleft. Penetration of Leu 383 (in L7) into the large face of α-CD is evident.

One cis-peptide bond has been identified in SBA, between Phe 200 and Pro 201. This bond is situated in loop 4 and forms a sharp kink that is stabilized by a hydrogen bond between the carboxyl oxygen of Glu 199 and the main-chain nitrogen of Arg 202. Another hydrogen bond between the carboxyl oxygen of Pro 201 and the indole nitrogen of Trp 302 also stabilizes the cis-peptide bond.

Blocked Cysteine Residues. Four of the six cysteine residues in SBA are blocked with 2-mercaptoethanol to form S–S–$(CH_2)_2OH$ as described previously (Mikami *et al.*, 1992). These blocked cysteine residues—Cys 95, Cys 288, Cys 343, and Cys 448—are well defined in the $2|F_o| - |F_c|$ map. Cys 288 and Cys 448 are located on the N-terminal side of the β-barrel with their modified side chains extended into solvent. In contrast, Cys 95 and Cys 343 are located on the C-terminal side of the β-barrel.

Solvent Molecules. The final model contains 319 water molecules and 1 sulfate ion. The average electron density, from omit maps, and the average refined *B*-factor of the waters were 1.88σ and 35.4 Å^2, respectively. There are 227 water molecules forming hydrogen bonds to the protein; 150 are hydrogen bonded to the main-chain O or N and 136 to the side-chain O, N, or S. There are about 40 waters in the cavity between the core and lobe region, some 20 of them near the bound α-CD.

Conformational Features. The fold of the polypeptide backbone of SBA in the presence of bound α-CD is shown in the ribbon diagram of Figure 3, and a schematic summary of the protein's secondary structure elements, assigned by hydrogen bonds and torsion angles as calculated by XPLOR, is presented in Figure 4. The hydrogen bonds are defined by a donor–acceptor distance between 2.5 and 3.2 Å.

The enzyme is composed of a large $(\alpha/\beta)_8$ core (residues 13–91, 150–181, 258–296, and 325–442), a smaller lobe made up of three long loops [L3 from β-strand β3 (residues 92–149), L4 from β4 (residues 182–257), and L5 from β5 (residues 297–324) of the $(\alpha/\beta)_8$ barrel], and a long C-terminal loop of the protein formed by residues 443–495. Two short 3_{10} helices are found in the N-terminal region (residues 7–10) and after the $(\alpha/\beta)_8$ barrel (residues 451–453). The core of

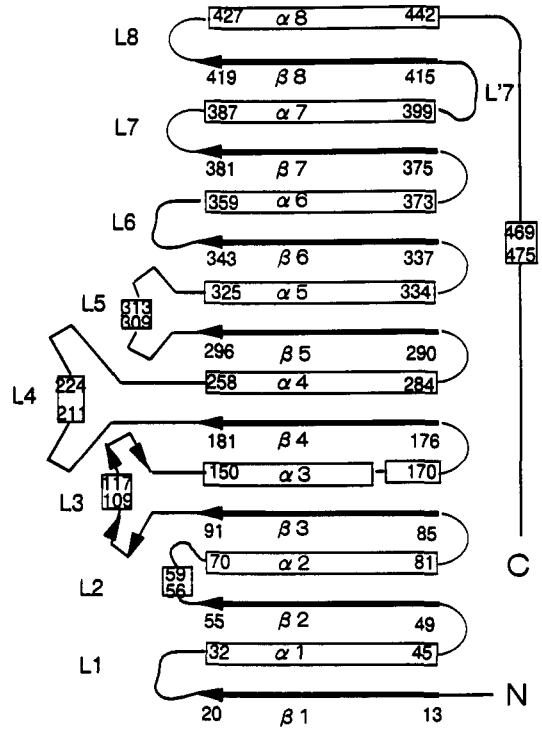

FIGURE 4: Schematic diagram showing the topology and secondary structural elements of SBA. Solid arrows represent β-strands; open rectangles represent α-helices. Elements of the $(\alpha/\beta)_8$ structural core are numbered consecutively; loops are designated by loop numbers.

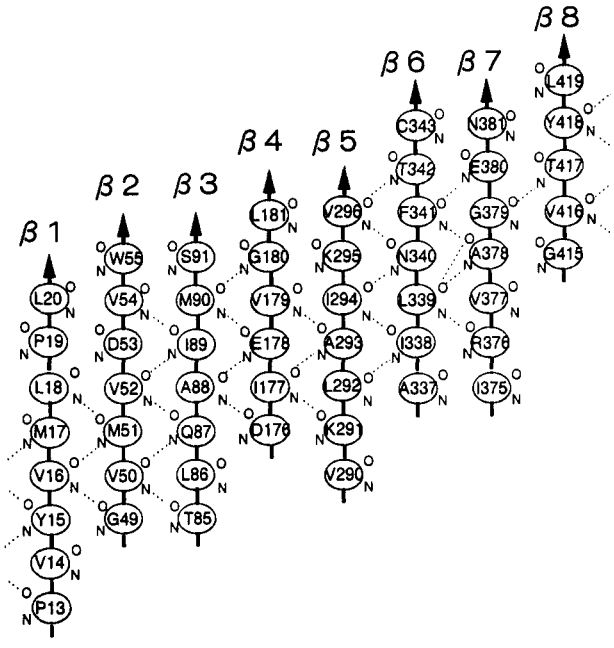

FIGURE 5: Hydrogen-bonding diagram on the β-barrel of the $(\alpha/\beta)_8$ domain. Hydrogen bonds less than 3.2 Å are shown as dotted lines.

β-amylase was earlier referred to as the large domain of the protein, with loops L3 and L4 defined as a second small domain (Mikami *et al.*, 1991, 1992).

The $(\alpha/\beta)_8$ Structure. The core of β-amylase is constructed as a prototypical $(\alpha/\beta)_8$ fold, similar to triosephosphate isomerase (Banner *et al.*, 1975). The eight parallel β-strands contain 54 amino acid residues and the eight helices contain 128 residues. Figure 5 describes the secondary structural elements of the protein. Five of the eight β-strands are seven amino acids in length; β1, β4, and β8 are eight, six, and five residues long, respectively. There are 31 hydrogen bonds formed between component residues of the parallel strands, using as criterion a maximum of 3.2 Å between hydrogen bond donors and acceptors. There are two additional weak hydrogen bonds between the nitrogen of Gly 379 in β7 and the oxygen of Gly 415 in β8 (distance, 3.33 Å) and also between the nitrogen of Ser 91 in β3 and the oxygen of Val 54 of β2 (3.32 Å). The parallel β-barrel is an ellipsoid whose major axis direction is that of β2. The interior of the β-barrel contains seven charged residues (Asp 53, Asp 176, Glu 178, Lys 291, Lys 295, Arg 376, and Glu 380) and eight polar residues (Tyr 15, Thr 85, Gln 87, Ser 91, Asn 340, Thr 342, Asn 381, and Thr 417). Five of these 15 residues are located on the N-terminal side of the barrel, five on the C-terminal side, and five in the middle of the barrel. In contrast to these charged or polar residues in the interior of the barrel, most of the side chains interacting with the helices on the outside of the $(\alpha/\beta)_8$ barrel are hydrophobic.

The eight α-helices of the structure vary in length between 10 and 27 residues (Figure 4). Six of the eight (α1, α2, α5, α6, α7, α8) are three or four turns in length; α-helix α3 contains 20 residues or about seven turns of helix, and α4 contains 27 residues or eight turns of helix. Of the eight α-helices only α3 is kinked. It shows a sharp bend of about 100° between Met 165 and Ser 166 that separates the helix into α3A (Ala 150–Asn 164) and α3B (Ser 166–Glu 170).

Loops. There are 15 loops connecting the alternating β-strands and α-helices of the core structure of β-amylase. These loops are defined as being on the N-terminal side if they connect one of the component helices to the N-terminal of a component β-strand; conversely, loops which connect the C-terminal of a β-strand to the N-terminal of an α-helix are termed C-terminal loops. This convention is further described in Figure 4. The loops on the N-terminal side (L'1–L'7) are generally shorter than those on the C-terminal side (L1–L8). With one exception the N-terminal loops are at most 5 residues long; the last, L'7, contains 15 residues. On the other hand, six loops on the C-terminal side contain 14 or more amino acid residues. The three longest loops, L3 (59 residues), L4 (77 residues), and L5 (28 residues), form a small lobe extending from the C-terminal end of the β-barrel core. Each of these loops includes an α-helix. There is a similarity of folding of L3, L4, and L5; the common motif is loop_1–helix–loop_2.

Deformation in the Protein Surface. As indicated in the ribbon diagram of the protein (Figure 3), a cavity-type deformation is present in the surface of the lobe region of the protein adjoining the C-terminal end of the β-barrel. This cavity involves residues of loops L3, L4, L5, L6, and L7 as bounding structures. Specifically, one wall of the cavity (against the β-barrel) is formed by residues 382–384 of L7 and residues 343–358 of L6. Going clockwise, the wall next comprises residues 298–308 of L5, residues 190–200 of L4, residues 96–102 of L3, and residue 20 at the carboxyl end of β-strand β1. The cavity floor consists of residues 295–296 of β5, residue 297 of L5, and residues 186–189 of L4. The length of the rhomboidal cavity, from the Cα of Leu 20 (β1) to that of Ala 355 (L6), is about 30 Å; the width, from Gln 194 (L4) to Leu 383 (L7), is approximately 15 Å; the depth, measured from Ser 195 (L4) to the Cα of Glu 186 (L4), is about 18 Å. The cavity is, of course, of far smaller volume than these distances would suggest, as it is mostly filled with the side chains of the boundary residues. This cavity is the place for binding substrates and their analogues and for the functioning of the protein as a specific catalyst. The α-CD inhibitor, for example, binds partly into the cavity entrance, whereas the carboxyl group of Glu 186 which marks the presumed catalytic center is found at the base of the cavity. This carboxyl group is 5.7 Å below the nearest atom of the bound α-CD (O-3 of

Table IV: α-Cyclodextrin/Protein Interactions

	L4 and L5 side of the cleft			L6, L7, and core side of the cleft		
type of interaction	sugar residue	protein residue	Å	sugar residue	protein residue	Å
hydrogen bonds	O-2 Glc 4	NE2 His 300	3.3	O-2 Glc 2	O Ala 382	2.7
	O-3 Glc 4	NE2 His 300	3.0	O-3 Glc 3	O Ala 382	3.1
van der Waals C–C contacts	C-1 Glc 3	CH2 Trp 301	4.1	C-3 Glc 4	CE Met 346	3.9
	C-1 Glc 4	CH2 Trp 301	4.1	C-3 Glc 3	CB Ala 382	3.9
	C-1 Glc 4	CE1 Phe 200	4.0	C-3 Glc 4	CD1 Leu 383	3.9
	C-1 Glc 4	CG Phe 200	3.6	C-3 Glc 5	CD1 Leu 383	4.2
	C-1 Glc 4	CD1 Phe 200	3.7	C-3 Glc 1	CD2 Leu 383	4.2
	C-1 Glc 4	CD2 Phe 200	3.8	C-3 Glc 2	CD2 Leu 383	3.9
	C-2 Glc 4	CE1 Phe 200	3.9	C-2 Glc 1	CG Pro 384	4.1
	C-2 Glc 4	CD1 Phe 200	3.8	C-3 Glc 1	CG Pro 384	3.7
	C-2 Glc 4	CZ Phe 200	3.9			
	C-6 Glc 4	CZ3 Trp 198	3.9			
	C-6 Glc 4	CH2 Trp 198	4.1			

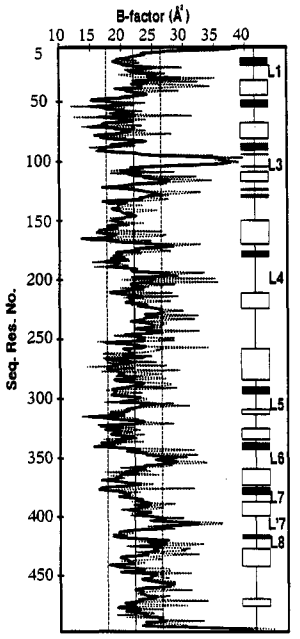

FIGURE 6: Plot of *B*-factor (in Å²) versus residue number. The thick line refers to the average *B*-factor of the main-chain atoms; the thin dotted line refers to that of the side-chain atoms. Dark bands represent β-strands; open rectangles, α-helices.

Glc 3) and 7.5 Å from the α-CD's nearest glycosidic bridge oxygen atom (O-4 of Glc 3).

Chain Flexibility. The average isotropic temperature factors of the main chain and side chains are 22.0 and 23.7 Å², respectively. Figure 6 presents a residue by residue plot of the *B*-factors. A striking observation is that segments of loops L3, L4, and L6 forming the walls of the pocket into which α-CD binds have higher than average *B*-factors. In particular, residues 96–103 in L3 ($loop_1$) showed an average *B*-factor value of 35.3 Å², with the position of the Cα atom of Gly 97 showing weak density in the final $2|F_o| - |F_c|$ map. Another high *B*-factor segment is in L'7 (residues 403–408) between α7 and β8 on the opposite side of the protein. In general, all higher *B*-value regions are located at the surface of the protein. The most rigid parts are in the deep interior of the protein, including the parallel β-barrel and most of the helices of the $(\alpha/\beta)_8$ barrel (Figure 6).

The amino acid sequences of residues in β-amylases from higher plants and microorganisms are compared in Figure 7. It is evident that most of the conserved residues of SBA are located at the C-terminal of the β-barrel (β1–β7) and in the successive loop regions, especially L3 and L4. In the segment of L3 noted above for its flexibility, the sequence of residues 95–101 is identical in all reported β-amylase sequences (Figure 7).

C-Terminal Loop. After the last helix (α8) of the $(\alpha/\beta)_8$ structure a long loop, comprising residues 443–495, leads to the C-terminus of the peptide chain (Figure 4). Most of this loop (residues 443–485) surrounds the N-terminal side of the $(\alpha/\beta)_8$ barrel; its final residues (486–495) are slanted toward but do not reach the C-terminal side of the $(\alpha/\beta)_8$ barrel. The C-terminal loop interacts with β1, β8, α4, α5, α6, α7, and α8. It contains one 3_{10}-helix and one α-helix, at residues 451–453 and 469–475, respectively; the φ and ψ angles of the main-chain configuration, except for the helices, correspond to β-strand or nearly β-strand angles. The amino acid sequence of this C-terminal loop shows no conserved residues between plant and bacterial β-amylases (Figure 7), suggesting the structural variability of this region.

Conformation of α-Cyclodextrin. The glucose units of the α-CD bound to SBA are in the $^4C_1$ pyranoid form and α-anomeric configuration (Figure 8), as reported for free α-CD by X-ray crystallography (Saenger *et al.*, 1983). The following parameters (averages and standard errors) are found for the six D-glucose residues of the bound α-CD: glycosidic bridge atom angle, +120 ± 0.4°; φ torsion angle (O4···C1–O4′–C4′), +162 ± 2.5°; ψ torsion angle (C1–O4′–C4′···O4″), −168 ± 4.5°. The distances between oxygen atoms O-2 and O-3 of adjacent glucose residues are between 2.7 and 3.2 Å, allowing them to be linked by hydrogen bonds; and the O-4 atoms of all six glucose units lie in a plane with a least squares deviation of 0.2 Å. These features demonstrate the nearly symmetric toroidal structure of the bound α-CD, as reported for the inclusion complexes of α-CD with various guest molecules (Saenger *et al.*, 1976). Crystalline α-CD in the "water-adduct" state shows diminished symmetry, as indicated by the rotation of one glucose residue, a maximum deviation from the O-4 atom plane of 0.98 Å, and O-2 to O-3 distances involving the distorted glucose residue of 4.66 and 3.35 Å (Manor & Saenger, 1972, 1974).

Soybean β-Amylase/α-Cyclodextrin Interaction. The bound α-CD interacts with SBA at L4, L5, L6, and L7, by way of hydrogen bonds and van der Waals contacts. Table IV summarizes these interactions. As listed there and illustrated in Figure 9, only four distinct hydrogen bonds were assigned between hydroxyl groups of the ligand and residues of the main-chain protein (imidazole NE2 of His 300 in L5 and O of Ala 382 in L7). Glucosyl hydroxyl groups are also linked to water by hydrogen bonds at five sites. One water molecule, hydrogen bonded to O-6 of Glc 4, is located near the α-CD cavity. In contrast to the paucity of hydrogen bonds between the α-CD and protein, many van der Waals contacts

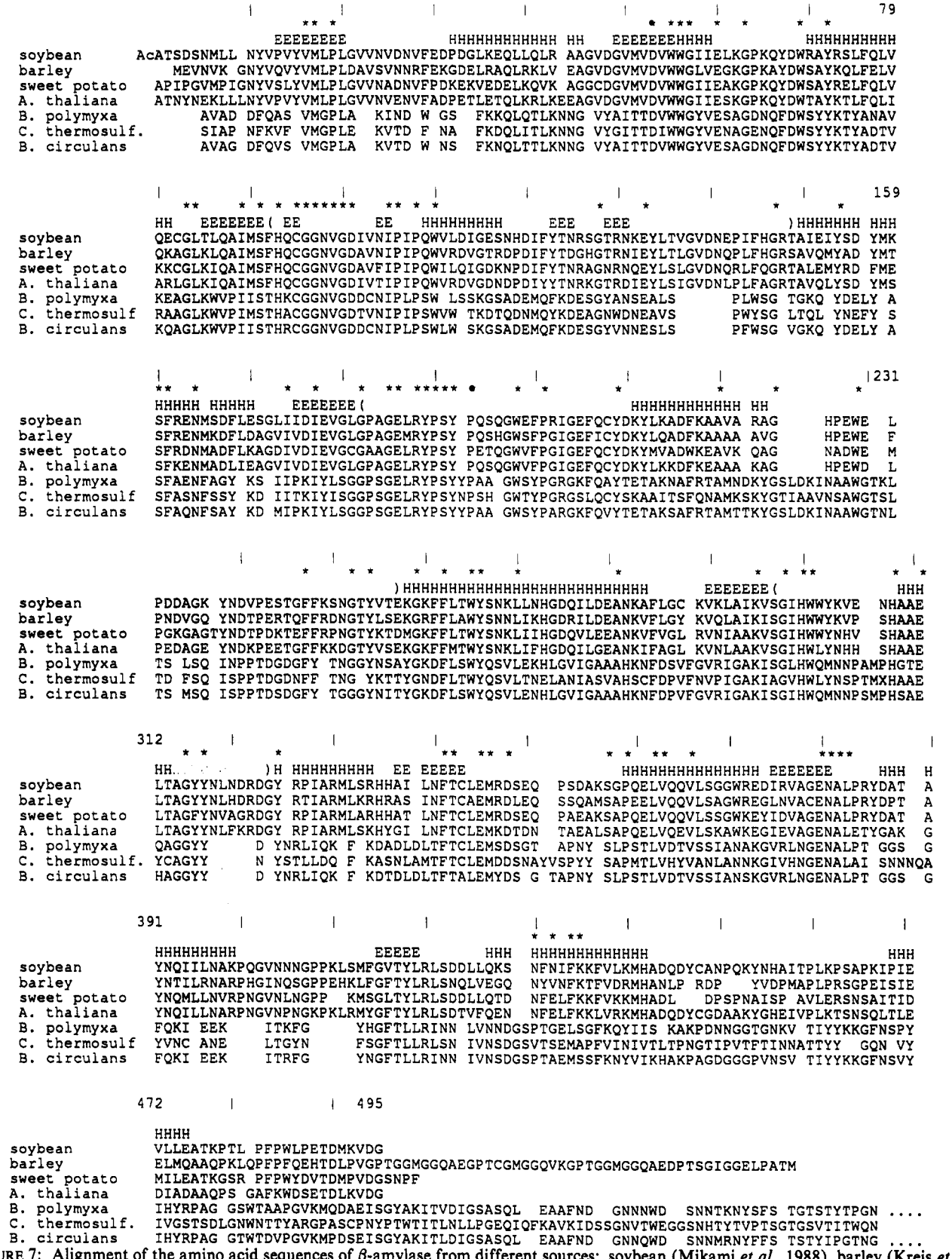

FIGURE 7: Alignment of the amino acid sequences of β-amylase from different sources: soybean (Mikami *et al.*, 1988), barley (Kreis *et al.*, 1987), sweet potato (Toda, 1989), *Avabidopsis thaliana* (Monroe *et al.*, 1991), *Bacillus circulans* (Siggins, 1987), *Bacillus polymyxa* (Kawazu *et al.*, 1987), and *Clostridium thermosulfurogenes* (Kitamoto *et al.*, 1988). H and E represent α-helix and β-sheet residues, assigned manually, in SBA. Segments in the small lobe structure are designated in parentheses. Noteworthy among the conserved amino acids (marked with an asterisk) are most of the residues that interact with the α-CD ligand. Trp 198, His 300, Trp 301, Met 346, Ala 382, and Leu 383 are conserved in the enzyme from all seven biological sources. Phe 200 is conserved in the four plant enzymes (Trp is present in the three bacterial enzymes); Pro 384 is found in three of the four plant enzymes and in two of the bacterial β-amylases examined.

are found between them. Five of the six α-CD glucosyl moieties have such contact with residues of L7, including Ala 382, Leu 383, and Pro 384. Gly 4 shows, in addition, multiple contacts with residues of L4 (Trp 198, Phe 200), L5 (His 300, Trp 301), and L6 (Met 346), making Glc 4 the most tightly bound of the six units.

The most interesting interaction between SBA and α-CD involves Leu 383, both of whose δ-methyl groups are found

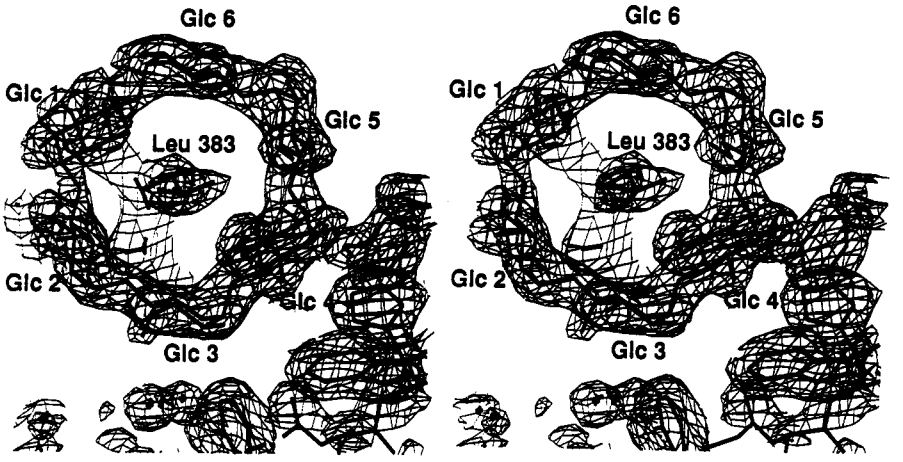

FIGURE 8: Stereoscopic view, from the small lobe side of the cleft (O-6 face of bound α-CD looking toward the core structure), of the final $(2F_o - F_c)\exp(i\phi_c)$ map around the bound α-CD, together with the refined model. The depicted contour level is 1.1σ. The Leu 383 side chain, originating from the opposite $(\alpha/\beta)_8$ side of the cleft, is seen through the α-CD cavity, as is its interaction with glucosyl C-3 atoms of α-CD. (The appearance of extensive intrusion into the hydrophobic cavity is deceptive; cf. Figure 9.)

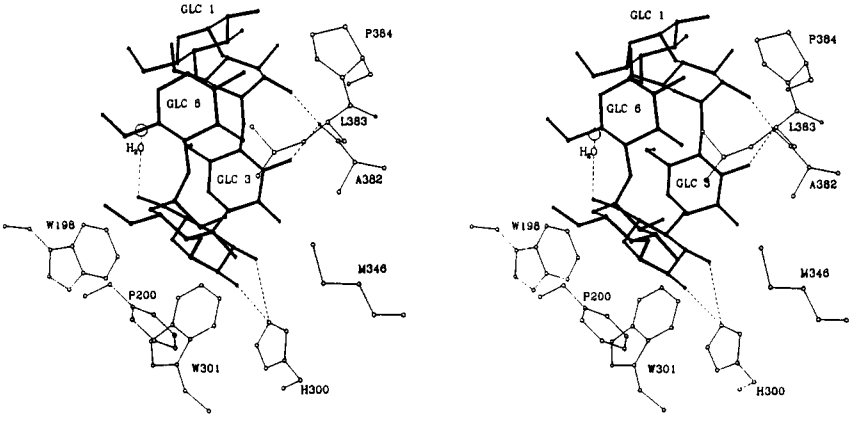

FIGURE 9: Binding site of α-CD in the cleft of SBA, showing the broad (O-2/O-3) base of the ligand facing the core side of the cleft. The dotted lines indicate potential hydrogen bond interactions within 3.2 Å between the protein and O-2 and O-3 atoms of glucosyl residues and between a water molecule (centered at the narrow end of the α-CD cavity) and the O-6 atom of Glc 4. Protein residues having apparent van der Waals (3.6–4.2 Å) contact with the ligand are also shown. The view is similar to that of Figure 3 (bottom) and nearly orthogonal to that of Figure 8.

at the wider base of the α-CD cavity in van der Waals contact with the C-3 atoms of Glc 1, 2, 4, and 5. This interaction resembles at least marginally that found in diffraction studies of crystallized α-CD inclusion complexes with various compounds by Saenger *et al.* (1976), so perhaps the Leu 383 interaction may account for the observed symmetry of the bound α-CD. However, it is not certain if complexation with the Leu 383 disrupts the water/α-CD interactions reportedly responsible (Saenger *et al.*, 1976) for the tensed (water-adduct) conformation. The possibility that α-CD in the soak solution might preexist in relaxed conformation is also not excluded (Saenger *et al.*, 1977).

α-CD Binding Orientation Relative to the Protein. It is clear from the present model of the SBA/α-CD complex that the direction of α-(1→4) glucosidic linkage progression in the ligand bound to enzyme runs clockwise when viewed from the wider (glucosyl O-2/O-3) face of the α-CD. This finding provides a clue to the possible binding orientation of the outer nonreducing chains of starch to β-amylase, a problem currently under investigation. In addition, it is clear for the first time from the specific binding interactions of α-CD to residues of L4, L5, L6, and L7 (Table IV and Figure 9) that both the narrower glucosyl O-6 face and the wider O-2/O-3 face of the annular ligand are closely associated with protein residues that define most of the cleft entrance. This finding is relevant to the question of how α-CD functions to inhibit starch hydrolysis by the enzyme.

DISCUSSION

Soybean β-amylase belongs to the broad class of $(\alpha/\beta)_8$ barrel enzymes. Its $(\alpha/\beta)_8$ core structure plus smaller adjacent region formed by loops extending from the carboxyl end of several β-strands basically resembles the single-domain structure of triosephosphate isomerase (Banner *et al.*, 1975) or adenosine deaminase (Wilson *et al.*, 1991) rather than the multidomain structures of α-amylases (Matsuura *et al.*, 1984; Buisson *et al.*, 1987) and cyclodextrin glycosyltransferase (Kubota *et al.*, 1991). The one-domain structure of SBA, however, is not the only feature that distinguishes it from amylases of other types. The $(\alpha/\beta)_8$ core itself differs in detail from the $(\alpha/\beta)_8$ barrels of α-amylases (Matsuura *et al.*, 1984; Buisson *et al.*, 1987; Boel *et al.*, 1990) or cyclodextrin glycosyltransferase (Klein & Schulz, 1991; Kubota *et al.*, 1991), and it is substantially different from the (α/α) barrel of glucoamylase (Aleshin *et al.*, 1992). It will be of interest to learn if the apparently flexible loop 3 sequence that forms one boundary of the substrate binding pocket in β-amylase may have a functional role in catalysis, as has been reported for an apparently similar loop in triosephosphate isomerase (Alber *et al.*, 1981; Pompliano *et al.*, 1991).

The 2-Å-resolution three-dimensional structure of β-amylase complexed with α-cyclodextrin opens the way to a deeper understanding of catalysis by this classic exoglucanase. For example, analysis of the structure provides new insight into

the basis for the finding that α-cyclodextrin is a competitive inhibitor of β-amylase, K_i of 0.3–0.5 mM, whether in hydrolyzing starch (Thoma & Koshland, 1960; Marshall, 1973; Mikami *et al.*, 1983) or α-maltosyl fluoride (Mizokami and Hehre, unpublished results). Basic information on inhibition of catalysis might perhaps not be expected from a structure involving an SH-blocked β-amylase of suppressed catalytic activity. However, measurements of ultraviolet difference spectra show that the α-cyclodextrin binding constants (K_d of 0.34–0.35 mM) with SH-blocked soybean β-amylase preparations of slight catalytic activity are indistinguishable from that (K_d of 0.35 mM) with fully active enzyme (Mikami *et al.*, 1983). Furthermore, a low-resolution structure of unblocked SBA/α-CD shows a nearly identical structure. As evident from Figures 3 and 9, α-cyclodextrin binds in a specific stereochemical orientation in a cleft in the protein surface and occupies much of the entrance to the cavity. The α-CD does not extend deeply enough into the cavity to make contact with the enzyme's presumed catalytic center. Its nearest atom, O-3 of one of the deeper glucose units, is 5.7 Å from the carboxyl group of Glu 186 which Nitta *et al.* (1989) identified as a possible reactive group through studies of SBA labeled with the active site directed inhibitor 2,3-epoxypropyl α-D-[^{14}C]glucopyranoside. The nearest α-CD glycosidic bridge atom is 7.5 Å from Glu 186. These findings suggest that the β-amylase catalytic center is located in a pocket below the bound α-CD. Preliminary observations of complexes of SBA with maltose and with α-CD at low (6-Å) resolution showed one electron density peak of size to accommodate maltose to be located deeper in the cleft than α-CD and a second maltose-sized density peak to share a locus with the α-CD (Mikami *et al.*, 1991, 1992). More definitive information on the enzyme's active site structure and substrate binding interactions has recently been obtained in studies of 2-Å-resolution structures of SBA/β-maltose and of SBA/maltal complexes (Mikami, Hehre and Sacchettini, unpublished results).

The location of bound α-CD remote from Glu 186 is in accord with the failure of α-CD to be hydrolyzed. The binding of α-CD above the presumed catalytic center so as to occupy much of the cleft entrance (Figure 3) indicates that α-CD inhibits catalysis by physically blocking the access of α-(1→4)-linked glucose residues to the reaction center. This simple explanation was not considered by Thoma and Koshland (1960), who interpreted the ability of α-CD to competitively inhibit starch hydrolysis by β-amylase as evidence for the induced fit hypothesis.

A newly observed feature is the shallow α-CD inclusion complex formed by the δ-methyl group carbon atoms of Leu 383 (a strictly conserved residue in β-amylases; Figure 7) making van der Waals contacts with the C-3 atoms of four of the α-CD glucose residues (Table IV). This unusual interaction of a "hydrophobic" residue with a substrate analog at the surface of the protein may reflect the existence of a substrate recognition function for Leu 383, involving a complex with a loose helical turn at the nonreducing end of a substrate molecule. The structural analogy between α-CD and the helices of starch chains in aqueous solution has long been recognized through their I_2 complexes (cf. French, 1975) and is clearly evident in V-amylose crystals (Zugenmeier & Sarko, 1976). There is an ordered water molecule centered in the α-CD ring cavity opposite Leu 383 and located at the narrow (O-6) end of the cavity. It is hydrogen bonded to O-6 of Glc 4.

The catalytic site of β-amylase is situated in a cul-de-sac, consistent with the early hypothesis of Myrbäck (1950). This structure allows endwise hydrolysis of unbranched maltosaccharide or amylose-type molecules as well as the endwise cropping of the outer chains of branched dextrin, amylopectin, or glycogen molecules. The small cavity catalytic site of this exoglycanase differs from the long active site clefts reported for endoglycanases such as hen's egg lysozyme (Blake *et al.*, 1967) and α-amylases (Matsuura *et al.*, 1984; Buisson *et al.*, 1987; Boel *et al.*, 1990) and from the tunnel-shaped catalytic site reported for the core of cellobiohydrolase II exocellulase by Rouvinen *et al.* (1990). Our findings do not support the proposal of the latter authors that a tunnel-shaped active site structure is typical for exoglucanases.

ACKNOWLEDGMENT

The authors are grateful to Drs. Vern L. Schramm, Fred Brewer, and John Blanchard, Albert Einstein College of Medicine, for valuable discussions.

REFERENCES

Aibara, S., Yamashita, H., & Morita, Y. (1984) *Agric. Biol. Chem. 48*, 1575–1579.
Aleshin, A., Golubev, A., Firson, L. M., & Honzatko, R. B. (1992) *J. Biol. Chem. 267*, 19291–19298.
Balls, A. K., Walden, M. K., & Thompson, R. R. (1948) *J. Biol. Chem. 173*, 9–19.
Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D. C., Pogson, C. I., & Wilson, I. A. (1975) *Nature 255*, 609–614.
Blake, C. C. F., Johnson, L. N., Mair, G. A., North, A. C. T., Phillips, D. C., & Sarma, V. R. (1967) *Proc. R. Soc. London, Ser. B 167*, 378–388.
Boel, E., Brady, L., Brzozowski, A. M., Derewenda, Z., Dodson, G., G., Jensen, V. J., Petersen, S. B., Swift, H., Thim, L., & Woldike, H. F. (1990) *Biochemistry 29*, 6244–6249.
Brünger, A. T., Kuriyan, J., & Karplus, M. (1987) *Science 35*, 458–460.
Buisson, G., Duee, E., Haser, R., & Payen, F. (1987) *EMBO J. 6*, 3909–3916.
Colman, P. M., & Matthews, B. W. (1971) *J. Mol. Biol. 60*, 162–168.
Fox, G. C., & Homes, K. C. (1966) *Acta Crystallogr. 20*, 886–891.
French, D. (1975) in *MTP International Review of Science, Biochemistry Series One* (Whelan, W. J., Ed.) Vol. 5, pp 267–336, Butterworths, London.
Hehre, E. J., Okada, G., & Genghof, D. S. (1969) *Arch. Biochem. Biophys. 135*, 75–89.
Hehre, E. J., Kitahata, S., & Brewer, C. F. (1986) *J. Biol. Chem. 261*, 2147–2153.
Higashi, T. (1990) *J. Appl. Crystallogr. 23*, 253–257.
Jones, T. A. (1985) *Methods Enzymol. 115*, 157–171.
Kawazu, T., Nakanishi, Y., Uozumi, N., Sasaki., T., Yamagata, H., Tsukagoshi, N., & Udaka, S. (1987) *J. Bacteriol. 169*, 1564–1570.
Kitahata, S., Chiba, S., Brewer, C. F., & Hehre, E. J. (1991) *Biochemistry 30*, 6769–6775.
Kitamoto, N., Yamagata, H., Kato, T., Tsukagoshi, N., & Udaka, S. (1988) *J. Bacteriol. 170*, 5848–5854.
Klein, C., & Schulz, G. E. (1991) *J. Mol. Biol. 217*, 737–750.
Kreis, M., Williamson, Y., Buxton, B., Pywell, J., & Svendsen, I. (1987) *Eur. J. Biochem. 169*, 517–525.
Kubota, M., Matsuura, Y., Sakai, S., & Katsube, Y. (1991) *Denpun Kagaku 38*, 141–146.
Luzzati, V. (1952) *Acta Crystallogr. 5*, 802–810.
Manor, P. C., & Saenger, W. (1974) *J. Am. Chem. Soc. 96*, 3630–3639.
Marshall, J. J. (1973) *Eur. J. Biochem. 33*, 494–499.
Mathews, B. W. (1968) *J. Mol. Biol. 33*, 491–492.
Matsuura, Y., Kusunoki, H., Harada, W., & Kakudo, M. (1984) *J. Biochem. 95*, 697–702.

Mikami, B., Aibara, S., & Morita, Y. (1982) *Agric. Biol. Chem. 46*, 943–953.
Mikami, B., Nomura, K., & Morita, Y. (1983) *J. Biochem. 94*, 107–113.
Mikami, B., Morita, Y., & Fukazawa, C. (1988) *Seikagaku 60*, 211–216.
Mikami, B., Shibata, T., Hirose, M., Aibara, S., Sato, M., Katsube, Y., & Morita, Y. (1991) *Denpun Kagaku 38*, 147–151.
Mikami, B., Shibata, T., Hirose, M., Aibara, S., Sato, M., Katsube, Y., & Morita, Y. (1992) *J. Biochem. 112*, 541–546.
Monroe, J. D., Salminen, M. D., & Preiss, J. (1991) *Plant Physiol. 97*, 1599–1601.
Morita, Y., Aibara, S., Yamashita, H., Yagi, F., Suganuma, T., & Hiromi K. (1975) *J. Biochem. 77*, 343–351.
Myrbäck, K. (1950) *Ark. Kemi 2*, 417–422.
Nitta, Y., Kunikata, T., & Watanabe, T. (1983) *J. Biochem. 93*, 1195–1201.
Nitta, Y., Isoda, Y., Toda, H., & Sakiyama, F. (1989) *J. Biochem. 105*, 573–576.
Nitta, Y., Tomita, K., Kohno, M., Nakashima, T., & Matsumoto, Y. (1991) *Denpun Kagaku 38*, 159–164.
Rouvinen, J., Bergfors, T., Teeri, T., Knowles, J. K. C., & Jones, T. A. (1990) *Science 249*, 380–386.
Saenger, W. (1983) *Biochem. Soc. Trans. 2*, 136–139.
Sato, M., Yamamoto, M., Imada, K., Katsube, Y., Tanaka, N., & Higashi, T. (1992) *J. Appl. Crystallogr. 25*, 348–357.
Siggens, K. W. (1987) *Mol. Microbiol. 1*, 86–91.
Svensson, B. (1988) *FEBS Lett. 230*, 72–76.
Svensson, B. (1991) *Denpun Kagaku 38*, 125–135.
Thoma, J. A., & Koshland, D. E. (1960) *J. Am. Chem. Soc. 82*, 3329–3333.
Thoma, J. A., Koshland, D. E., Ruscica, J., & Baldwin, R. (1963) *Biochem. Biophys. Res. Commun. 12*, 184–188.
Toda, H. (1989) *Denpun Kagaku 36*, 87–101.
Wang, B. C. (1985) *Methods Enzymol. 115*, 90–112.
Wilson, D. K., Rudolph, F. B., & Quiocho, F. A. (1991) *Science 252*, 1278–1284.
Wonacott, A. J. (1980) *IDXREF. A Suite of Programs for On-Line Evaluation and Analysis of Integrated Intensities on Small-Angle Rotation/Oscillation Photographs*, Blackett Laboratory, Imperial College, London, England.
Zugenmaier, P., & Sarko, A. (1976) *Biopolymers 15*, 2121–2136.