Structure of the *Aspergillus oryzae* α-Amylase Complexed with the Inhibitor Acarbose at 2.0 Å Resolution[†,‡]

A. Marek Brzozowski and Gideon J. Davies*

Department of Chemistry, University of York, Heslington, York YO1 5DD, England

Received March 10, 1997; Revised Manuscript Received June 9, 1997[⊗]

ABSTRACT: The three-dimensional structure of the *Aspergillus oryzae* α-amylase (TAKA-amylase), in complex with the inhibitor acarbose, has been determined by X-ray crystallography at a resolution of 1.98 Å. The tetrasaccharide inhibitor is present as a hexasaccharide presumably resulting from a transglycosylation event. The hexasaccharide occupies the −3 to +3 subsites of the enzyme, consistent with the known number of subsites determined by kinetic studies, with the acarbose unit itself in the −1 to +3 subsites of the enzyme. The transition state mimicking unsaturated pseudo-saccharide occupies the −1 subsite as expected and is present in a distorted $^2H_3$ half-chair conformation. Careful refinement plus extremely well-resolved unbiased electron density suggest that the hexasaccharide represents a genuine transglycosylation product, but the possibility that this apparent species results from an overlapping network of tetrasaccharides is also discussed. Catalysis by α-amylase involves the hydrolysis of the α-1,4 linkages in amylose with a net retention of the anomeric configuration, *via* a double-displacement mechanism, as originally outlined by Koshland [Koshland, D. E. (1953) *Biol. Rev. 28*, 416−336]. The enzymatic acid/base and nucleophile, residues Glu230 and Asp206, respectively, are appropriately positioned for catalysis in this complex, and the hexasaccharide species allows mapping of all the noncovalent interactions between protein and ligand through the enzyme's six subsites.

α-Amylases (α-1,4-glucan 4-glucanohydrolase, EC 3.2.1.1) catalyze the hydrolysis of the α-1,4 glycosidic linkages in starches, such as amylose and amylopectin. They are found in family 13 of the classification of glycosylhydrolases according to Henrissat (Henrissat, 1991; Henrissat & Bairoch, 1993, 1996). Family 13 is the largest of the glycosylhydrolase families (the sequence families are now available at the web site http://expasy.hcuge.ch/cgi-bin/lists?glycosid.txt; Henrissat & Bairoch, 1996) and, in addition to α-1,4-hydrolases, also contains enzymes with distinct substrate and product specificities, notably cyclodextrin glycosyltransferases (CGTases)[1] and enzymes specific for α-1,6 and α-1,1 glycosidic linkages [for review, see Svensson (1994)]. The structures of family 13 enzymes consist of a characteristic $(\beta/\alpha)_8$-barrel catalytic core domain with a varying number of extra domains. Fungal α-amylases are typically three domain structures consisting of the catalytic core domain linked to a C-terminal eight-stranded β-sheet domain. A third domain, termed domain B, is found inserted between the third β-sheet and the helix of the $(\beta/\alpha)_8$-barrel (Swift et al., 1991) and may play a role in both enzyme stability and substrate binding. Domain B is also known to confer a number of isoenzyme specific properties in the barley α-amylases (Juge et al., 1995; Rodenburg et al., 1994). A single structural calcium ion is found bridging the A and B domains in all family 13 amylase structures. A weakly binding, inhibitory, calcium site has also been described for the enzymes from *Aspergillus oryzae* and *Aspergillus niger* (Boel et al., 1990).

The substrate for amylases, amylose, is a polymer of 1,4-linked α-D-glucopyranoside units. Starch itself has a radial structure in which the terminal nonreducing sugars are displayed on the outside of the starch globule, providing an easily accessible food storage system. The hydrolysis of starch is performed by a consortium of free enzymes which display the full range of macroscopic properties possessed by glycosylhydrolases. Exoenzymes such as glucoamylase remove the terminal nonreducing sugar, and β-amylase liberates disaccharides from the nonreducing end (with inversion of the anomeric configuration, hence the term "β"-amylase); endo-acting amylases cleave at internal glycosidic linkages, while processive enzymes perform a processive degradation (initially termed "multiple attack") of the substrate after an initial endo attack (Robyt & French, 1967). The *A. oryzae* α-amylase (hereafter TAKA-amylase) falls into the endo-acting category of amylases. It was one of the first glycosylhydrolases to receive extensive biochemical characterization. The increase in k_{cat}/K_M for maltooligosaccharides with an increasing degree of polymerization (Nitta et al., 1971) suggested approximately six or seven subsites contributing to catalysis (Figure 1). At the three-dimensional level, the structure of the native TAKA-amylase has been described (Matsuura et al., 1984), together with an analysis of both the structural and inhibitory calcium sites (Boel et al., 1990; Swift et al., 1991). Despite the wealth of biochemical data on this system, oligosaccharide complexes of TAKA-amylase have so far escaped structural characterization. A number of oligosaccharide inhibitors have, however, been successfully used to elucidate details of

[†] This work was funded, in part, by the Biotechnology and Biological Sciences Research Council of Great Britain. G.J.D. is a Royal Society University Research fellow.

[‡] Coordinates for the structures described in this paper have been deposited with the Brookhaven Protein Data Bank (accession references 6TAA and 7TAA).

* Corresponding author. Telephone: -44-1904-432596. Fax: -44-1904-410519. E-mail: davies@yorvic.york.ac.uk.

[⊗] Abstract published in *Advance ACS Abstracts,* August 1, 1997.

[1] Abbreviations: TAKA-amylase, *Aspergillus oryzae* α-amylase; DP, degree of polymerization; CGTase, cyclodextrin glycosyltransferase.

S0006-2960(97)00539-4 CCC: $14.00

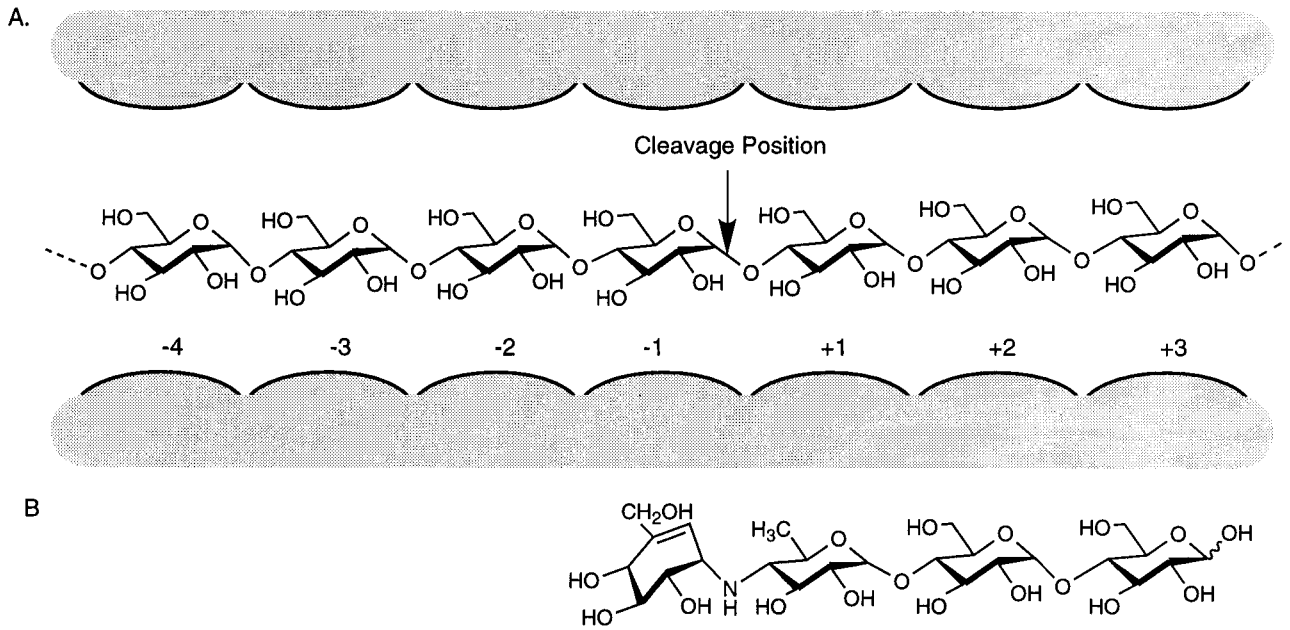

FIGURE 1: (A) Subsite structure of TAKA-amylase. Kinetic studies (Nitta et al., 1971) suggest approximately seven subsites, labeled −4 to +3 [nomenclature according to Davies et al. (1997)]. (B) Chemical structure of acarbose.

substrate binding on other amylolytic enzymes. The trestatin family of compounds, originally isolated from *Streptomyces* species and developed for clinical use in the treatment of diabetes and other sugar metabolism disorders (Schmidt et al., 1981), have proven to be extremely popular both in therapeutic terms and in the effective crystallographic analysis of protein−oligosaccharide interactions (Aleshin et al., 1994; Gilles et al., 1996; Qian et al., 1994; Stoffer et al., 1995; Strokopytov et al., 1995, 1996). One particular trestatin-derived tetrasaccharide, acarbose, warrants special attention since it is an extremely potent amylase inhibitor. Acarbose has a nonreducing end pseudo-disaccharide, consisting of a valienamine unit linked to 4-amino-4,6-dideoxy-α-D-glucose, which is then linked to maltose (Figure 1b). The strong inhibition is widely attributed to the enhanced binding of the nonreducing valienamine, whose half-chair conformation mimics the substrate distortion expected in the transition state. The adjacent glycosidic bond is *N*-linked, preventing enzymatic hydrolysis. In all studies, except that with the CGTase from *Bacillus circulans* (Strokopytov et al., 1995; where the lack of hydrolysis can be ascribed to the absence of the exocyclic hydroxyl group of the 6-deoxy sugar), acarbose is found bound with the transition state-mimicking ring in the −1 subsite.

α-Amylases catalyze the cleavage of the glycosidic bond with a net retention of the anomeric configuration. This occurs by a double-displacement reaction with the formation and hydrolysis of a covalent glycosyl−enzyme intermediate *via* oxocarbenium ion transition states (Figure 2). The covalent glycosyl−enzyme intermediate has recently been unambiguously trapped by the use of 2-deoxy-2,2-difluoroglucosides (Braun et al., 1995) and 5-fluoroglycosyl fluorides (McCarter & Withers, 1996). In the TAKA-amylase, the catalytic acid/base and the nucleophile are residues Glu230 and Asp206, respectively. Although for hydrolytic enzymes in aqueous solution the glycosyl−enzyme intermediate is usually deglycosylated by an incoming water nucleophile, it is possible that an alternate nucleophile such as an oligosaccharide species attacks, resulting in the formation of an extended sugar chain by transglycosylation. Indeed, such a transglycosylation event is the normal mode of action for a number of family 13 enzymes. Other possible reaction mechanisms for these enzymes, such as ring opening or the stabilization of a long-lived oxocarbonium ion intermediate, have been discussed but lack serious experimental support [for review, see Sinnott (1990) and McCarter and Withers (1994)].

In this paper, we describe the structure of the TAKA-α-amylase in complex with the inhibitor acarbose, at 1.98 Å resolution. The observed electron density suggests the presence of an apparently transglycosylated species with at least six well-ordered saccharide units, as has been demonstrated in other related studies (Gilles et al., 1996; Qian et al., 1994; Strokopytov et al., 1996). The hexasaccharide binding is entirely consistent with the subsite description derived from kinetic studies in the early 1970s (Nitta et al., 1971) prior to the crystallographic structure determination. Crystallographic analysis, in particular the use of maximum likelihood weighted electron density maps calculated prior to the incorporation of any saccharide species in the refinement protocol, has been performed to identify the bound species with the minimum bias.

MATERIALS AND METHODS

Crystallization, Data Collection, and Processing. Pure TAKA-α-amylase was kindly supplied by Novo-Nordisk A/S. Prior to crystallization, the enzyme was washed extensively on Centricon 10K membranes with a buffer consisting of 10 mM Tris-HCl (pH 7.5) and 5 mM $CaCl_2$ and concentrated to 30 mg mL^{-1}. Crystallization of native TAKA-α-amylase was performed with substantial modification of the original conditions. The crystals have been grown with the use of the sitting drop method (plates available from Charles Supper) at 18 °C. Mother liquor was composed of 0.1 M Hepes/HCl buffer, 5 mM $CaCl_2$ (pH 7.5), and 16−18% (w/v) 5000 MW monomethyl polyethylene glycol [for details of these precipitants, see Brzozowski (1993)]. In solution, the enzyme is active at this pH, although we cannot say whether it has activity in crystal. The sitting drop consisted of 2 μL of protein solution together with 2 μL of mother liquor. Large (1.0 × 0.3 × 0.2 mm) crystals of the

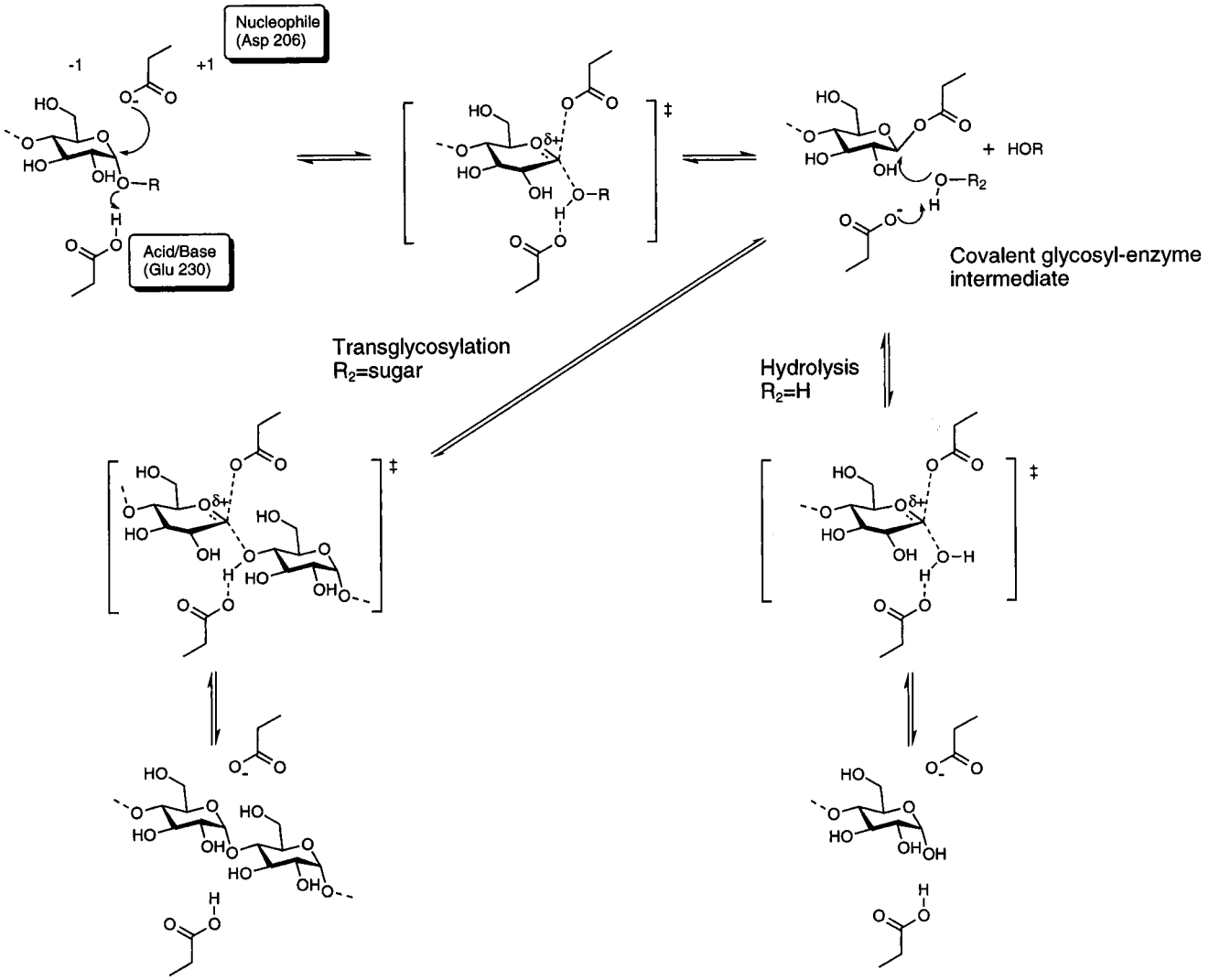

FIGURE 2: Catalytic mechanism of TAKA-amylase. Two potential reaction schemes, those leading to hydrolysis and transglycosylation, are shown.

orthorhombic form of TAKA-α-amylase crystals grow within 1 week. Acarbose for soaking experiments was kindly supplied by H. Driguez (CERMAV, Grenoble, France) and was judged to be a pure tetrasaccharide species by reverse-phase HPLC analysis (unpublished results). The soaking solution was prepared by dissolving acarbose in the mother liquor at a concentration of 15 mM. The first acarbose soaking experiment was done for 12 h but all TAKA-amylase crystals dissolved after this period of time. The next soaking was therefore repeated at the same acarbose concentration but with frequent manual inspection of the crystals, every 10−15 min. Crystal surfaces deteriorated visibly after 5 min. Perfectly formed crystal faces started to show regularly distributed small holes, and the crystals themselves were slowly dissolving. Crystals were allowed to shrink to up to half of their original size, which took approximately 6 h, and the best sample was used for X-ray data collection.

Data were collected from a single crystal mounted in a glass capillary using a MAR research imaging plate and Cu rotating anode. Data (100°) were collected in 1° frames from a sample mounted with the crystallographic *b** axis roughly parallel to the spindle axis of the camera, but mis-set slightly to reduce the loss of data in the "blind region". Data were processed with the DENZO program (Otwinowski, 1993) and further reduced with programs from the CCP4 suite (Collaborative Computational Project Number 4, 1994). The crystals belong to space group $P2_12_12_1$, with the following cell dimensions: $a = 50.76$ Å, $b = 67.06$ Å, $c = 131.62$ Å, and $\alpha = \beta = \gamma = 90°$; the crystals are isomorphous with respect to those used for the native structure determination (Swift et al., 1991). There is a single molecule of TAKA-amylase in the asymmetric unit.

Refinement. Since this cell is similar to the previously determined native unit cell, the native coordinates were used as the starting model for refinement. The starting crystallographic *R* factor was 0.49. At this point, 5% of the reflections were set aside for cross-validation analysis (Brünger, 1992) and used not only to monitor various refinement strategies but also for the basis of the maximum likelihood weighting used in the refinement procedure. The initial *R* factor of 0.49 for all data was sufficiently poor for the cross-validation subset to be considered truly "free". Initial refinement consisted of rigid body refinement using the program AMoRe (Navaza, 1994), treating the whole amylase structure as a single rigid body. Following this, manual rebuilding with the X-FIT routines of QUANTA (Molecular Simulations, Inc.) were interspersed with standard least-squares refinement using the maximum likelihood program REFMAC (Murshudov et al., 1997) with target stereochemistry from the corrected set of Engh and Huber (1991). As all observed data from 20 to 1.98 Å were

Table 1: Data Quality and Completeness, Given in Bins of Resolution, for the TAKA-Amylase−Acarbose Complex

resolution (Å)	R_{merge}[a]	mean $I/\sigma I$	completeness (%)	multiplicity
20.0−6.00	0.042	23.5	97.0	3.5
4.34	0.048	24.5	99.6	3.5
3.57	0.051	23.5	99.2	3.6
3.11	0.062	18.9	98.5	3.6
2.79	0.085	14.5	98.1	3.6
2.55	0.113	11.9	97.5	3.7
2.36	0.142	9.8	96.9	3.7
2.21	0.165	8.4	96.5	3.7
2.09	0.223	6.4	95.8	3.7
1.98	0.291	5.1	94.2	3.7
totals	0.088	12.6	97.0	3.7

[a] $R_{merge} = \sum_{hkl}\sum_i|I_{hkli}-\langle I_{hkl}\rangle|/\sum_{hkl}\sum_i\langle I_{hkl}\rangle$.

employed in the refinement, a low-resolution 2 G bulk solvent correction was applied. Since the nature of the oligosaccharide ligand in amylase complexes has been the matter of some contention, care was taken not to include the ligand until the density was clear and unambiguous, although, even with the starting phases, when the crystallographic R value was 0.49, clear continuous density for the whole hexasaccharide could be seen. Upon incorporation of the hexasaccharide ligand, refinement continued, as above, but all the O6 atoms of the ligand were given zero occupancy to provide further substantiation of the exact nature of the saccharide species using difference maps. Coordinates and observed structure factor amplitudes have been deposited with the Brookhaven Protein Data Bank (Bernstein et al., 1977).

RESULTS AND DISCUSSION

Quality of the Final Model Structure. A summary of the data quality and completeness is given in Table 1. The data consist of 113 411 observations of 31 062 unique reflections and merged with an R_{merge} of 0.088 for all reflections and have a mean multiplicity of observation of 3.7 observations/reflection. The overall completeness is 97%. The final model structure, consisting of 3686 protein atoms, 64 ligand atoms, a single Ca^{2+} ion, and 460 solvent water molecules, has a crystallographic R factor of 0.159 with an R_{free} of 0.236. All the non-glycine residues have conformational angles (φ and ψ) in permitted regions of the Ramachandran plot (Ramachandran et al., 1963) with 91% of these in the "most favored regions" as defined by PROCHECK (Laskowski et al., 1993). A plot of the R factor against resolution or the σ_A method gives an upper estimate for the mean coordinate error (assuming model errors to be the sole determinant of F_{obs} vs F_{calc} discrepancies) of approximately 0.15 Å (Read, 1986). The protein structure is very similar with the native enzyme structure (6TAA), with the exception of a few small corrections in side chain orientations. Final refinement statistics for the TAKA-amylase−acarbose structure are given in Table 2.

The TAKA-amylase structure consists of three domains. The catalytic core A domain is comprised of a $(\beta/\alpha)_8$-barrel; the catalytic center of the enzyme being located on the C-terminal side of the barrel is the case in all $(\beta/\alpha)_8$-barrel enzymes. A short decoration occurs between the third barrel strand and helix. This forms domain B (residues 122−176), whose main structural feature is a short three-stranded antiparallel β-sheet. The interface between the A and B domains is maintained by a structural, conserved, "high-affinity" calcium site, whose ligands are donated by both the A and B domains (see below for details). The final domain in the TAKA-amylase structure [other family 13 members include a number of extra domains; see Svensson (1994) for details] is an eight-stranded β-sandwich domain (residues 384−478). The substrate binding groove is formed by the fissure along the A and B subunit interface (Figures 3 and 4).

Oligosaccharide Binding. In order to identify the exact nature of the bound species, the differences between the pseudo-disaccharide found in acarbose and a genuine maltose must be analyzed. In addition to the distorted unsaturated ring species, the salient difference between the pseudo-disaccharide unit of acarbose and maltose is the 4-amino-4,6-dideoxyglucose unit of the former, which lacks the exocyclic hydroxyl of glucose (see Figure 1B). Identification of the genuine species found in the crystallographic analysis frequently rests on the ability to positively identify the

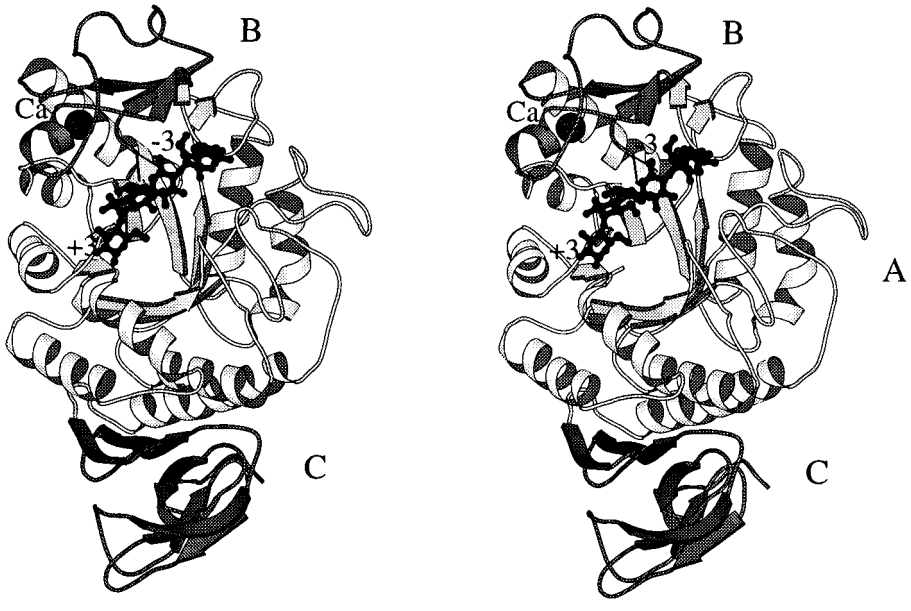

FIGURE 3: Schematic diagram of the TAKA-amylase structure drawn with the MOLSCRIPT program (Kraulis, 1991). The hexasaccharide ligand is shown in ball-and-stick mode and the single calcium ion as a shaded sphere, and the A−C domains are labeled.

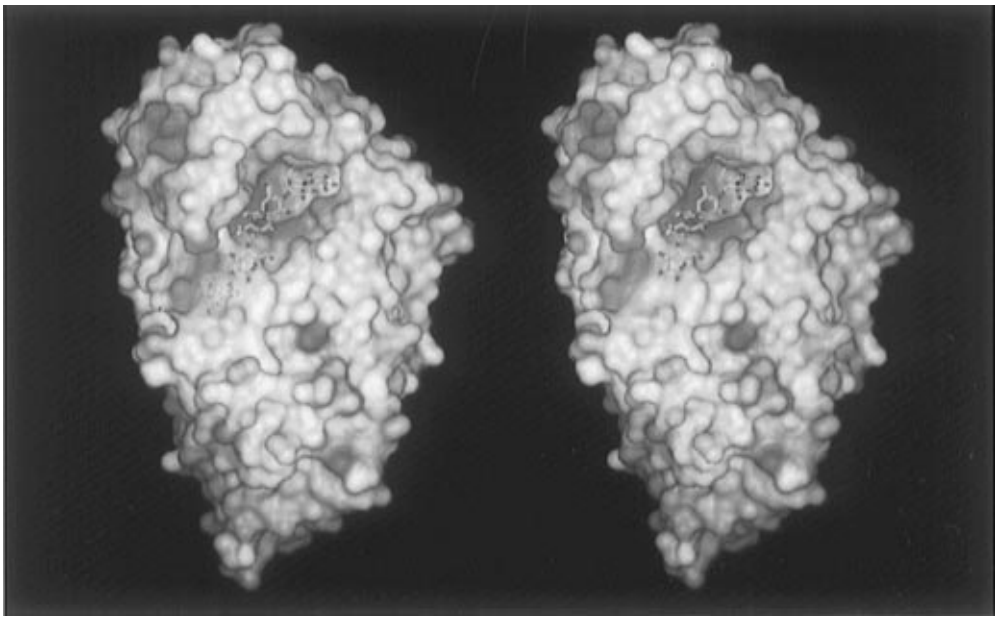

FIGURE 4: Electrostatic surface figure calculated with the MOLVIEWER program (M. Hartshorn, unpublished). The hexasaccharide ligand is shown in "licorice" representation.

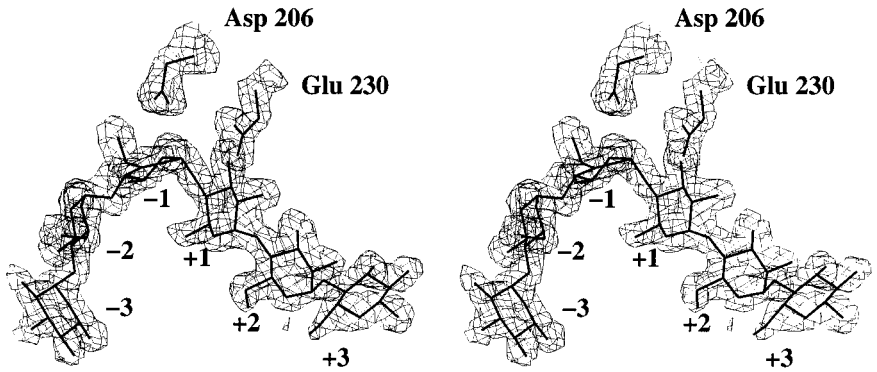

FIGURE 5: Electron density for the hexasaccharide ligand, calculated with σ_A maximum likelihood weights (Murshudov et al., 1997), at a contour level of 0.35 eÅ^{-3}. The refined coordinates are shown, but the map was generated with phases calculated prior to the incorporation of the oligosaccharide. The catalytic acid/base and nucleophile, Glu230 and Asp206, are shown.

presence or absence of this O6 hydroxyl. This is often difficult, since the O6 hydroxyl groups frequently exhibit high temperature factors and one can rarely be absolutely sure of their presence. This is especially true for sugars on the periphery of the subsite-binding cleft where the C6 hydroxyl interacts with bulk solvent. Crystallographically, phase bias and memory effects caused by the incorporation of an incorrect sugar model at an early stage of refinement could easily lead to the persistence of misleading density throughout the refinement procedure and may well account for some of the ambiguity and confusion concerning these species that exist in the published literature. In order to determine the exact nature of this transglycosylated species, careful examination of the unbiased electron density was undertaken. Electron density maps, with σ_A/likelihood weighting [see Murshudov et al. (1997) and Read (1986) for details] calculated *prior* to the incorporation of any sugar species in any refinement or phase calculation, were used exclusively since they represent the only nonbiased view of the oligosaccharide.

Clear connected density is found for six saccharide units in the active site of the TAKA-amylase (Figure 5). The active center of TAKA-amylase is located in a deep pocket on the C-terminal side of the $(\alpha/\beta)_8$-barrel domain. The substrate binding surface is not a long groove across the enzyme as is observed with endo-acting enzymes hydrolyzing linear β-linked oligosaccharides [for review, see Davies and Henrissat (1995)] but is instead found at the center of a V-shaped surface depression formed by the interface between the A and B domains. The substrate binding region extends sharply up to the surface from the catalytic center (Figure 4). The sugars at the nonreducing end of the chain enter *via* a steep couloir, until reaching the −1/+1 subsites at the catalytic center where the sugar chain makes a severe kink and changes direction. The sharp change in direction results from a change in the glycosidic bond torsion angle (O5−C1−O4′−C4′ torsion angle = 33°) coupled with a distortion of the sugar ring from a normal $^4C_1$ chair toward a half-chair conformation. This results in the O2−O3′ hydrogen bond, a feature of α-1,4-linked oligosaccharides, becoming broken. These features are essentially as observed by others (Gilles et al., 1996; Qian et al., 1994; Strokopytov et al., 1995), although there is no substantial loop movement upon binding as has been observed for the pancreatic enzymes (Qian et al., 1994). A summary of the torsion angles and O2−O3′ hydrogen bond distances is given in Table 3. A diagram of the main protein-ligand interactions is given in Figure 6 and described below.

Since the inhibitor acarbose is a pseudo-tetrasaccharide, we, like others (Gilles et al., 1996; Qian et al., 1994;

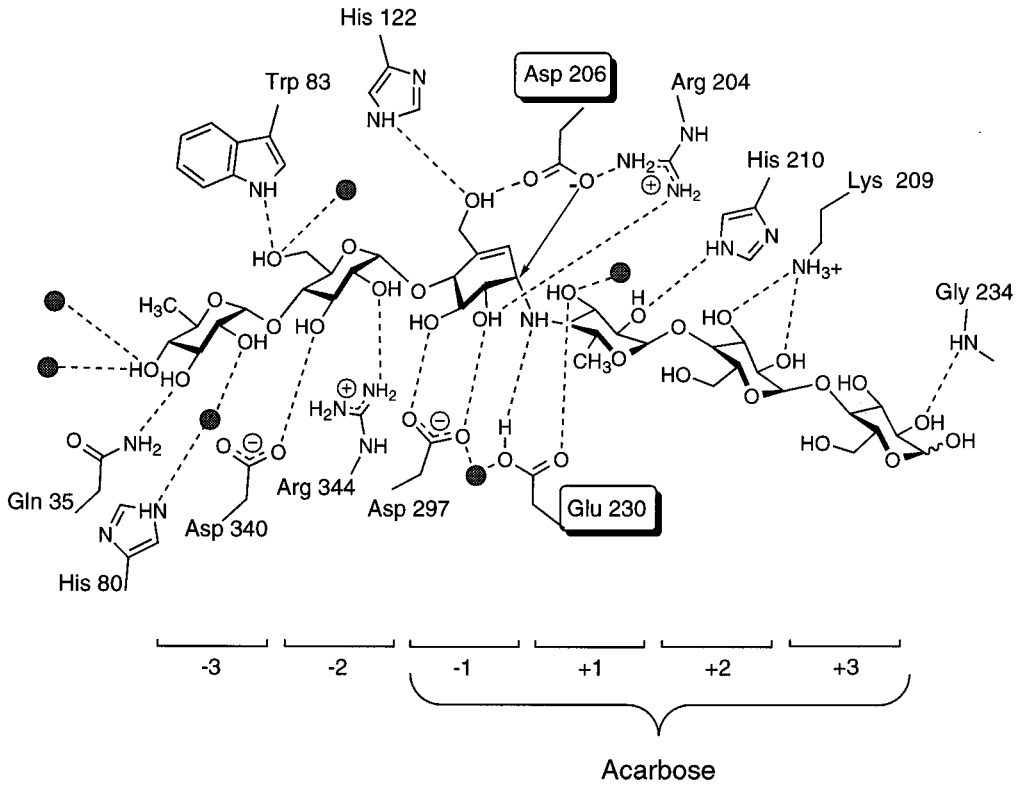

FIGURE 6: Schematic diagram of the interactions of acarbose with TAKA-amylase. Hydrogen bonds are shown as dotted lines and water molecules as shaded spheres.

Table 2: Refinement and Structure Quality Statistics for the TAKA-Amylase−Acarbose Complex

no. of protein atoms			3686			
no. of ligands/ion atoms			64 ligand/1 Ca^{2+}			
no. of solvent waters			460			
resolution used in refinement (Å)			20−1.98			
R_{cryst}			0.159			
R_{free}			0.236			
rms deviation of 1−2 bonds (Å)			0.014			
rms deviation of 1−3 angles (Å)			0.033			
rms deviation of chiral volumes ($Å^3$)			0.138			
average main chain B ($Å^2$)			16.8			
average side chain B ($Å^2$)			19.7			
average solvent B ($Å^2$)			38			
			subsite			
mean ligand B ($Å^2$)	−3	−2	−1	+1	+2	+3
	26.8	13.4	8.2	11.2	18.8	27.3
main chain ΔB, bonded atoms ($Å^2$)				3.3		

Table 3: Glycosidic Bond Torsion Angles and O2−O3′ Hydrogen Bond Distances for the TAKA−Acarbose Complex

glycosidic linkage	torsion angle (φ) (O5−C1−O4′−C4′)	torsion angle (ψ) (C1−O4′−C4′−C5′)	O2−O3′ hydrogen bond distance (Å)
−3/−2	143.6	−105.5	2.88
−2/−1	120.7	−101.9	2.93
−1/+1	33.24	−151.2	no H bond
+1/+2	109.6	−119.8	2.71
+2/+3	121.3	−107.4	2.65

Strokopytov et al., 1996), interpret the extended density in terms of a transglycosylation event to form a longer, more potent inhibitor. The electron density certainly substantiates this proposal. Density for all the glycosidic linkages is unambiguous; neither do spurious density features exist in either the initial or final maps, nor do aberrant B values result from refinement, both of which would have been more indicative of an overlapping network of tetrasaccharide species. Certainly, for the purposes of stable crystallographic refinement and dissection of the protein−carbohydrate interactions, refinement of a single hexasaccharide provides the most appropriate sugar model. The possibility that a transglycosylation has taken place is, however, an area of some contention which is discussed further, below. It is undoubtedly true that the best analysis would involve mass spectrometry of redissolved crystals, but the sensitivity of current analyses, together with the very small volume of a protein crystal, has so far made all attempts at these analyses, by ourselves and others, unfruitful. Density extends for the six subsites of the enzyme, from −3 to +3, consistent with the number of subsites and the preferred pattern of binding for a hexasaccharide proposed from kinetic studies, prior to even the native structure solution (Nitta et al., 1971).

Protein−Ligand Interactions at −3 and −2 Subsites. The primary feature of binding of the nonreducing end sugar in the −3 subsite is the aromatic stacking, as has been observed with many protein−carbohydrate interactions (Vyas, 1991), with Tyr75. The sugar makes only a single direct H bond to the protein from O3 to the amide nitrogen of Gln35. A further water-mediated H bond exists, bridging His80 with the sugar O2 hydroxyl. The O4 hydroxyl makes two strong H bonds to what appear to be well-defined solvent molecules, although at very low (<0.1 $eÅ^{-3}$) contour levels one could interpret this density as the beginning of an additional disordered sugar ring. No electron density is observed for the O6 atom, leading one to believe either that this sugar unit is a 6-deoxyglucose, present because of a transglycosylation event, or that it is a glucose unit and the O6 group is highly disordered. The −2 sugar makes a greater number of direct H bonds to the protein: the O3 atom with the carboxylate of Asp340, the O2 with the guanadino head group of Arg344, and the O6 with the NE1 atom of Trp83.

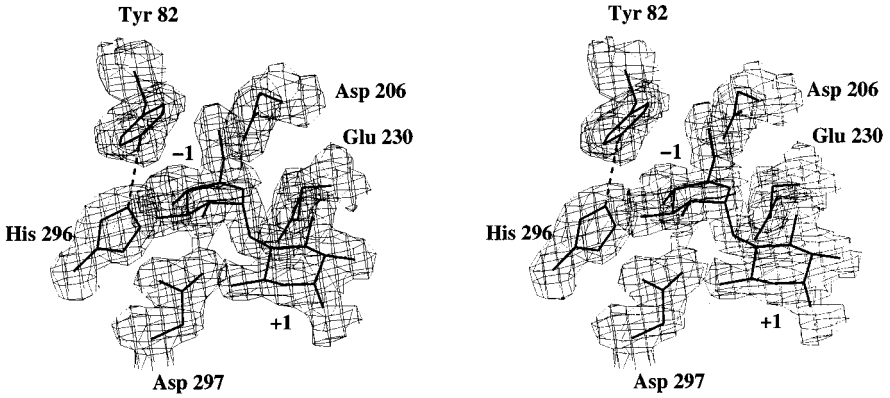

FIGURE 7: Electron density for the acarbose pseudo-disaccharide unit present in the −1 and +1 subsites at the catalytic center. Key residues interacting with the substrate are shown, and the hydrogen bond between His296 and Tyr82 is shown as a dotted line. The electron density shown is calculated with σ_A maximum likelihood weights (Murshudov et al., 1997), at a contour level of 0.35 eÅ^{-3}. The refined coordinates are shown, but the map was generated with phases calculated prior to the incorporation of the oligosaccharide.

The O6 hydroxyl makes an additional interaction with a well-defined solvent water molecule.

Protein−Ligand Interactions at +2 and +3 Subsites. These leaving group subsites are notable for their relative paucity of direct hydrogen bonding interactions with the protein, a feature we have observed with a number of other glycosylhydrolase complexes (Davies et al., 1995; Sulzenbacher et al., 1996) and which presumably exists to improve the k_{off} rate for product release. Direct interactions exist only with the +2 sugar O2 and O3 groups with the terminal amino of Lys209 and between the +3 sugar O3 and the main chain amide of Gly234. The role of Lys209 in binding the +2 subsite sugar is consistent with the observation that this residue is conserved in family 13 enzymes acting on α-1,4-linked sugars, but not those enzymes acting on α-1,6-linked sugars whose aglycons would be orientated differently (Svensson, 1994). Other sugar hydroxyls make interactions only with solvent water. The pyranoside ring in the +2 subsite stacks above the hydrophobic moiety of Leu232 as previously suggested (Svensson, 1994). No O6 density is observed for the +3 subsite sugar, again suggesting that it is presumably highly mobile in the crystal (see above).

Catalytic Center with −1 and +1 Subsites. The −1 subsite contains the unsaturated pseudo-saccharide ring, whose distortion is believed to mimic that found in the transition state (Figure 7). As such, its presence in the −1 subsite is to be expected and has indeed been observed in all analogous studies, except for that of the *B. circulans* CGTase−acarbose complex (Strokopytov et al., 1995). The presence of the C5═"O5" double bond ("O5" since this oxygen has in fact been replaced by a carbon) results in planarity at C4−C5−"O5"−C1, closely resembling the planarity that must exist in the oxocarbenium ion transition states that occurs during hydrolysis [see Sinnott (1990) and references therein]. Planarity at C4−C5−"O5"−C1 results in distortion toward a $^2H_3$ half-chair, with C4, C5, "O5", and C1 defining a four-atom plane (rms deviation of 0.03 Å), with C2 and C3 above and below this plane by 0.43 and 0.44 Å, respectively. In the unusual case of the CGTase−acarbose complex, the electron density is instead interpreted in terms of the 4,6-dideoxyglucose ring occupying the −1 subsite rather than the unsaturated cyclitol, observed in other studies. The conformation of the −1 subsite ring in the CGTase structure is described as "a full $^4C_1$ chair conformation", but overlap of the CGTase-acarbose structure (Brookhaven Protein Data Bank entry 1CXG) with that described here reveals an almost identical conformation for this −1 subsite ring, which we would prefer to describe as a half-chair conformation. In the case of the CGTase structure, the C4, C5, O5, and C1 again define a four-atom plane (rms deviation of 0.02 Å), with C2 and C3 above and below this plane by 0.51 and 0.37 Å, respectively. When the two oligosaccharide models are overlapped (considering just the −2, −1, and +1 subsites), the rms deviation for the ring atoms and glycosidic linkages is 0.24 Å. Since the −1 and +1 subsite pyranoside rings overlap with an rms deviation of just 0.09 Å, the main differences between the CGTase and TAKA sugar conformations resides in the −2 subsite, whose rings overlap with an rms deviation of 0.24 Å. In the TAKA complex, we interpret this as a standard ring in its preferred $^4C_1$ chair conformation, whereas in the CGTase complex, the unsaturated ring is believed to occupy this site in a distorted half-chair conformation.

The TAKA-amylase −1 subsite ring, whose conformation mimics that found in the transition state, makes a number of noncovalent contacts with the protein. The O3 and O2 hydroxyls form H bonds with Asp297; the interaction between O2 and the OD2 atom of Asp297 would appear to be a particularly strong, "low-barrier" possibly charged, H bond with a distance of only 2.50 Å. Strong H bonds have been proposed to stabilize key transition state interactions for a number of glycosylhydrolases [see Namchuk and Withers (1995) and references therein] and other systems (Frey et al., 1994), and these observations are in keeping with those proposals. In addition to this key role in binding the O2 and O3 hydroxyls, Asp297 makes a water-mediated H bond with the catalytic acid/base Glu230, consistent with the belief that it may also play a role in maintaining an elevated pK_a for the acid/base for the glycosylation step of the reaction. The O6 group is present in the *gauche−gauche* conformation (C4−C5−C6−O6 torsion angle = 83.8°) which permits two interactions, one with the totally conserved His122, confirming the proposals made by Strokopytov et al. (1995) and with the carboxylate group of the enzymatic nucleophile Asp206. Mutation of His122 is known not only to decrease k_{cat}/K_M for the real substrate but also to increase the K_i for acarbose, implying that it plays a role in transition state stabilization, presumably by enhanced binding of the distorted sugar (Svensson, 1994). The catalytic nucleophile, Asp206, is perfectly poised for nucleophilic attack at C1; it is 2.9 Å distant with its *syn* lone pairs directed toward the anomeric carbon as expected by

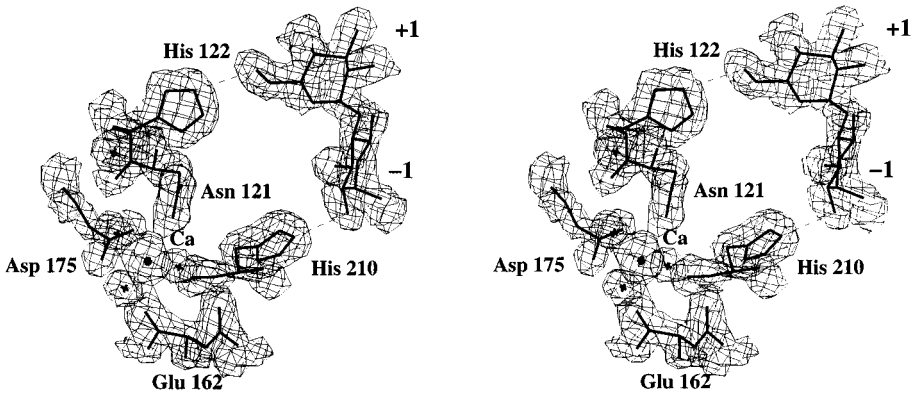

FIGURE 8: Electron density for the single calcium ion and ligands calculated at a contour level of 0.35 eÅ^{-3}. The proximity of the sugars at the catalytic center is shown.

theory (Gandour, 1981). The −1 subsite makes an important aromatic stacking contact with Tyr82. The orientation of this tyrosine is maintained by a H bond between the Tyr OH with NE2 of His296. In this complex, we propose that this H bond, stabilizing the aromatic stacking of Tyr82 over the −1 subsite, is the major role of His296, rather than hydrogen bonding to substrate that has previously been proposed for other family 13−acarbose complexes (Qian et al., 1994; Strokopytov et al., 1995). In our structure, the His296 NE2−O3 and −O2 distances are 3.3 and 3.5 Å, respectively, which coupled with the extremely poor H bonding geometry for these interactions, and the fact that the sugar hydroxyls already have a full complement of hydrogen bonds, would tend to rule out a major role for this histidine in direct H bonding to the −1 subsite sugar. The conformational flexibility of this residue in the various amylase structures has been noted (Machius et al., 1995), however, and the different orientations and hydrogen bonds observed in various studies may reflect pH effects. The enzymatic acid/base, Glu230, makes a good "side-on" hydrogen bond with the interglycosidic nitrogen at the point of cleavage with a distance of 2.8 Å. The other oxygen of this residue interacts with the +1 subsite sugar O3 group. Other than an H bond between His210 and the O2 hydroxyl, all other interactions in the +1 subsite are *via* solvent water molecules. The +1 subsite is occupied by the 6-deoxyglucose ring; one can postulate that if a hydroxyl group were present here it would interact either with solvent water or with Asp297.

Calcium Binding. The amylase family contains a conserved calcium binding site [described as the "high-affinity" site in Boel et al. (1990)] which is located at the interface between the catalytic A domain and the B domain. The ligand stabilizes the loop conformation which forms a large lip over the substrate binding groove (Figure 3). The Ca^{2+} ion is bound by eight ligands (Figure 8): the carboxylate oxygens of Asp175 in a bidentate fashion (distances of 2.54 and 2.85 Å, respectively), the side chain carbonyl oxygen of Asn121 (2.51 Å), three water molecules (2.41, 2.46, and 2.76 Å), and the main chain carbonyls of Glu162 (2.51 Å) and His210 (2.54 Å). Interestingly, this calcium site is indirectly linked to the important −1 and +1 sugar binding subsites of the enzyme. The main chain carbonyl of His210 is involved in Ca^{2+} binding, and the side chain of this residue provides a direct hydrogen bond to the O2 hydroxyl of the +1 subsite sugar. Less directly, Ca^{2+} binding involves the side chain of Asn121, this time with the adjacent residue His122 providing a crucial hydrogen bond to the O6 hydroxyl of the −1 subsite sugar. These interactions thus provide a means by which the effects of Ca^{2+} binding may be directly transmitted through to substrate binding and catalysis. The "inhibitory" Ca^{2+} site, described for the native structure (Boel et al., 1990), maps to the location of the anomeric C1 atom at the catalytic center (−1 subsite) in the TAKA-amylase−acarbose complex, providing a structural rationale for its predominantly competitive inhibition.

Catalytic Mechanism. Catalysis by α-amylase is widely believed to take place *via* a double-displacement reaction, in which a covalent glycosyl−enzyme intermediate is formed and subsequently hydrolyzed *via* oxocarbenium ion transition states (Figure 2) (Koshland, 1953). The specific roles for each of the catalytic residues were for some time a matter of debate. At a structural level, Dijkstra and colleagues proposed roles for the three catalytic acid groups in the active site of the homologous CGTase which, together with the elegant chemistry by the Withers group (Braun et al., 1995; McCarter & Withers, 1996), clarified the functions of the catalytic residues. Glu230 functions as the Brønsted acid/base, initially protonating the glycosidic bond to promote leaving group departure and later serving as general base to activate the incoming nucleophile for attack at C1. Asp206 is the catalytic nucleophile. Nucleophilic attack at C1 assists leaving group departure and results in the formation of a stable covalent−enzyme intermediate (which has been trapped and identified by mass spectrometry; McCarter & Withers, 1996). The third conserved aspartate, Asp297, may play a role in maintaining an appropriate pK_a for the catalytic acid/base, but it certainly plays a key role in substrate binding, especially *via* its strong H bond to the O2 group in the distorted complex.

Nature of the Hexasaccharide Species. The observation that sugar species bound to family 13 enzymes often appear longer than one would expect is frequently interpreted in terms of an enzymatic transglycosylation event, giving rise to an extended oligosaccharide inhibitor. A simple transglycosylation scheme leading to a hexasacchaide species would involve an initial hydrolysis of acarbose (bound in the −3, −2, −1, and +1 sites) liberating glucose as leaving group and generating a trisaccharide glycosyl−enzyme intermediate. Hydrolysis of acarbose by α-amylases has been described, although the TAKA-amylase itself finds even normal maltotetraose to be a very poor substrate (Nitta et al., 1971). The enzymatic deglycosylation, with a second molecule of acarbose acting as an acceptor, would result in the formation of a heptasaccharide species. Presumably, following a migration in the active site cleft, six units remain

bound and well-defined in the six subsites of the enzyme. Evidence for this comes from the observation that transglycosylation by retaining enzymes is certainly a well-defined phenomenon and longer oligosaccharides are known to bind more tightly and would dominate binding in the crystal, even if present at only very low concentrations (Nitta et al., 1971). Under normal circumstances, however, transglycosylation is considered to be thermodynamically unfavorable and often requires the use of extremely large substrate concentrations and a nonaqueous environment. In crystal, one is therefore forced to propose substantially different k_{on} and k_{off} rates for saccharide binding, overcoming bulk thermodynamics, or that contaminant species exist which are more easily transglycosylated. While all published crystallographic analyses of these apparently extended species in family 13 enzymes interpret the density as a genuine transglycosylation, no group has yet published any experimental evidence showing that these enzymes are capable of such a reaction; indeed, many have tried and have not yet been successful. We have not been able to detect any enzymatic transglycosylation of acarbose, in solution (unpublished results). A very persuasive argument against interpretation of these extended species as transglycosylated products comes from analysis of oligosaccharide complexes of inverting cellulases. Inverting enzymes are unable to perform transglycosylation reactions, yet oligosaccharide complexes of the endoglucanase CelA from *Clostridium thermocellum* reveals density for oligosaccharides longer than those used in soaking experiments (P. Alzari, personal communication; Alzari et al., 1996). In this case, since this enzyme is catalytically incapable of performing transglycosyation reactions, the extended density *must* result from overlapping networks of oligosaccharides. The authors have been able to verify this and refine the individual components only by collecting data at genuinely atomic, 0.9 Å, resolution (P. Alzari, personal communication). In the absence of similar atomic resolution data for the α-amylase complexes, the extended ligand is certainly most appropriately refined as a single extended species. One must consider, however, that for the α-amylase family structures the case for transglycosylation in crystal remains, as yet, unproven.

ACKNOWLEDGMENT

The authors thank Novo-Nordisk A/S for providing pure TAKA-amylase for studies in our laboratory and Dr. Hugues Driguez (Centre de Recherches sur les Macromolécules Végétales, Grenoble, France) for kindly providing of acarbose.

REFERENCES

Aleshin, A. E., Firsov, L. M., & Honzatko, R. B. (1994) *J. Biol. Chem. 269*, 15631−15639.
Alzari, P. M., Souchon, H., & Dominguez, R. (1996) *Structure 4*, 265−275.
Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. T., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535−542.
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.
Braun, C., Brayer, G. D., & Withers, S. G. (1995) *J. Biol. Chem. 270*, 26778−26781.
Brünger, A. T. (1992) *Nature 355*, 472−475.
Brzozowski, A. M. (1993) *Acta Crystallogr. D 49*, 352−354.
Collaborative Computational Project Number 4 (1994) *Acta Crystallogr. D50*, 760−763.
Davies, G., & Henrissat, B. (1995) *Structure 3*, 853−859.
Davies, G. J., Tolley, S. P., Henrissat, B., Hjort, C., & Schülein, M. (1995) *Biochemistry 34*, 16210−16220.
Davies, G. J., Wilson, K. S., & Henrissat, B. (1997) *Biochem. J. 321*, 557−559.
Engh, R. A., & Huber, R. (1991) *Acta Crystallogr. A 47*, 392−400.
Frey, P. A., Whitt, S. A., & Tobin, J. B. (1994) *Science 264*, 1927−1930.
Gandour, R. D. (1981) *Bioinorg. Chem. 10*, 169−176.
Gilles, C., Astier, J.-P., Marchis-Mouren, G., Cambillau, C., & Payan, F. (1996) *Eur. J. Biochem. 238*, 561−569.
Henrissat, B. (1991) *Biochem. J. 280*, 309−316.
Henrissat, B., & Bairoch, A. (1993) *Biochem. J. 293*, 781−788.
Henrissat, B., & Bairoch, A. (1996) *Biochem. J. 316*, 695−696.
Juge, N., Rodenburg, K. W., Guo, X.-J., Chaix, J.-C., & Svensson, B. (1995) *FEBS Lett. 363*, 299−303.
Koshland, D. E. (1953) *Biol. Rev. 28*, 416−436.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Laskowski, R. A., McArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 282−291.
Machius, M., Wiegand, G., & Huber, R. (1995) *J. Mol. Biol. 246*, 545−559.
Matsuura, Y., Kusunoki, M., Harada, W., & Kakudo, M. (1984) *J. Biochem. (Tokyo) 95*, 695−702.
McCarter, J. D., & Withers, S. G. (1994) *Curr. Opin. Struct. Biol. 4*, 885−892.
McCarter, J. D., & Withers, S. G. (1996) *J. Biol. Chem. 271*, 6889−6894.
Murshudov, G. N., Vagin, A. A., & Dodson, E. J. (1997) *Acta Crystallogr. D 53*, 240−255.
Namchuk, M. N., & Withers, S. G. (1995) *Biochemistry 34*, 16194−16202.
Navaza, J. (1994) *Acta Crystallogr. A 50*, 157−163.
Nitta, Y., Mizushima, M., Hiromi, K., & Ono, S. (1971) *J. Biochem. 69*, 567−576.
Otwinowski, Z. (1993) in *Data Collection and Processing: proceedings of the CCP4 study weekend* (Sawyer, L., Issacs, N., & Bailey, S., Eds.) Science and Engineering Research Council, Daresbury, U.K.
Qian, M., Haser, R., Buisson, G., Duée, E., & Payan, F. (1994) *Biochemistry 33*, 6284−6294.
Ramachandran, G. N., Ramakrishnan, C., & Sasisekharan, V. (1963) *J. Mol. Biol. 7*, 95−99.
Read, R. J. (1986) *Acta Crystallogr. A 42*, 140−149.
Robyt, J. F., & French, D. (1967) *Arch. Biochem. Biophys. 122*, 8−16.
Rodenburg, K. W., Juge, N., Guo, X.-J., Søgaard, M., Chaix, J.-C., & Svensson, B. (1994) *Eur. J. Biochem. 221*, 277−284.
Schmidt, D. D., Frommer, W., Junge, B., Müller, L., Wingender, W., & Truscheit, E. (1981) in *First international symposium on acarbose* (Creutzfeldt, W., Ed.) pp 5−15, Excerpta Medica, Amsterdam.
Sinnott, M. L. (1990) *Chem. Rev. 90*, 1171−1202.
Stoffer, B., Aleshin, A. E., Firsov, L. M., Svensson, B., & Honzatko, R. B. (1995) *FEBS Lett. 358*, 57−61.
Strokopytov, B., Penninga, D., Roseboom, H. J., Kalk, K. H., Dijkhuizen, L., & Dijkstra, B. W. (1995) *Biochemistry 34*, 2234−2240.
Strokopytov, B., Knegtel, R. M. A., Penninga, D., Rozeboom, H. J., Kalk, K. H., Dijkhuizen, L., & Dijkstra, B. W. (1996) *Biochemistry 35*, 4241−4249.
Sulzenbacher, G., Driguez, H., Henrissat, B., Schülein, M., & Davies, G. J. (1996) *Biochemistry 35*, 15280−15287.
Svensson, B. (1994) *Plant Mol. Biol. 25*, 141−157.
Swift, H. J., Brady, L., Derewenda, Z. S., Dodson, E. J., Dodson, G. G., Turkenburg, J. P., & Wilkinson, A. J. (1991) *Acta Crystallogr. B 47*, 535−544.
Vyas, N. K. (1991) *Curr. Opin. Struct. Biol. 1*, 732−740.

BI970539I