X-ray Structure of Novamyl, the Five-Domain "Maltogenic" α-Amylase from *Bacillus stearothermophilus*: Maltose and Acarbose Complexes at 1.7 Å Resolution[†,‡]

Zbigniew Dauter,[§] Miroslawa Dauter,[§] A. Marek Brzozowski,[§] Søren Christensen,[||] Torben V. Borchert,[||] Lars Beier,[||] Keith S. Wilson,[§] and Gideon J. Davies*[,§]

Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, York Y010 5DD, U.K., and Novo Nordisk A/S, Novo Allé, DK-2880 Bagsværd, Denmark

Received February 3, 1999; Revised Manuscript Received April 30, 1999

ABSTRACT: The three-dimensional structure of the *Bacillus stearothermophilus* "maltogenic" α-amylase, Novamyl, has been determined by X-ray crystallography at a resolution of 1.7 Å. Unlike conventional α-amylases from glycoside hydrolase family 13, Novamyl exhibits the five-domain structure more usually associated with cyclodextrin glycosyltransferase. Complexes of the enzyme with both maltose and the inhibitor acarbose have been characterized. In the maltose complex, two molecules of maltose are found in the −1 to −2 and +2 to +3 subsites of the active site, with two more on the C and E domains. The C-domain maltose occupies a position identical to one previously observed in the *Bacillus circulans* CGTase structure [Lawson, C. L., et al. (1994) *J. Mol. Biol. 236*, 590−600], suggesting that the C-domain plays a genuine biological role in saccharide binding. In the acarbose−maltose complex, the tetrasaccharide inhibitor acarbose is found as an extended hexasaccharide species, bound in the −3 to +3 subsites. The transition state mimicking pseudosaccharide is bound in the −1 subsite of the enzyme in a $^2H_3$ half-chair conformation, as expected. The active site of Novamyl lies in an open gully, fully consistent with its ability to perform internal cleavage via an endo as opposed to an exo activity.

The enzyme-catalyzed hydrolysis of starch plays a key role in the utilization of storage polysaccharides in many organisms. It is fundamental to many important industrial reactions, with enzymatic hydrolysis being the basis of both the liquefaction and saccharification processes during the production of high-glucose syrups from corn starch. Starch-active enzymes are also frequently utilized in the textile and detergent industries and are increasingly playing a role in the baking industry as "antistaling" agents. The enzymatic hydrolysis of the α-1,4 linkages in the starch polymer may be catalyzed by many different glycoside hydrolases. *O*-Glycoside hydrolases have been classified into more than 60 families on the basis of their amino acid sequences (*1*−*4*) with amylolytic enzymes found in many of these families. Starch-hydrolyzing enzymes display the full range of macroscopic properties for polysaccharide hydrolases, with true exo enzymes (such as glucoamylase, family 15, and *β*-amylase, family 14), true endo enzymes (such as α-amylase, family 13), and "multiple-attack" or "processive" enzymes (*5*). Prolonged Novamyl-catalyzed hydrolysis of starch results principally in the production of the disaccharide maltose, hence its description as a "maltogenic α-amylase" (*6*). This led to speculation that it was an exo enzyme with a requirement for nonreducing chain ends, but this is now believed not to be the case. The initial catalytic event may be an endo attack, anywhere in the polymer chain, since Novamyl[1] exhibits no specificity or requirement for polymer chain ends (*7*); it can catalyze the hydrolysis of terminally modified maltodextrins and cyclodextrins (Figure 1A).

The sequence of the five-domain α-amylase from *Bacillus stearothermophilus* (Novamyl) places it in glycoside hydrolase family 13 (*8*). Family 13 is the largest of all the families and contains many starch-active enzymes, most importantly α-amylase and cyclodextrin glycosyltransferase (CGTase), but also enzymes specific for other α-linked glucose-based oligosaccharides such as α-1,6 and α-1,1 hydrolases (for a review, see ref *9*). These enzymes catalyze glycoside hydrolysis with net retention of anomeric configuration by a double-displacement mechanism essentially as outlined by Koshland in 1953 (*10*, *11*). The basic framework for enzymes from family 13 consists of three domains: a catalytic core $(\beta/\alpha)_8$ structure termed domain A, a C-terminal eight-stranded *β*-sheet domain C, and a small loop excursion between the third sheet and helix of the catalytic core forming domain B (*9*). This is the structure exhibited by the majority of the fungal and bacterial α-amylases (*12*). CGTase is more complex; it has two additional domains, D and E, with the

[†] This work was funded, in part, by the Biotechnology and Biological Sciences Research Council, the University of York, and Novo-Nordisk A/S. 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 (file names 1qho and 1qhp).

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

[§] University of York.

[||] Novo Nordisk A/S.

[1] Abbreviations: CGTase, cyclodextrin glycosyltransferase; DP, degree of polymerization; Novamyl, five-domain α-amylase from *B. stearothermophilus.*

10.1021/bi990256l CCC: $18.00
Published on Web 06/06/1999

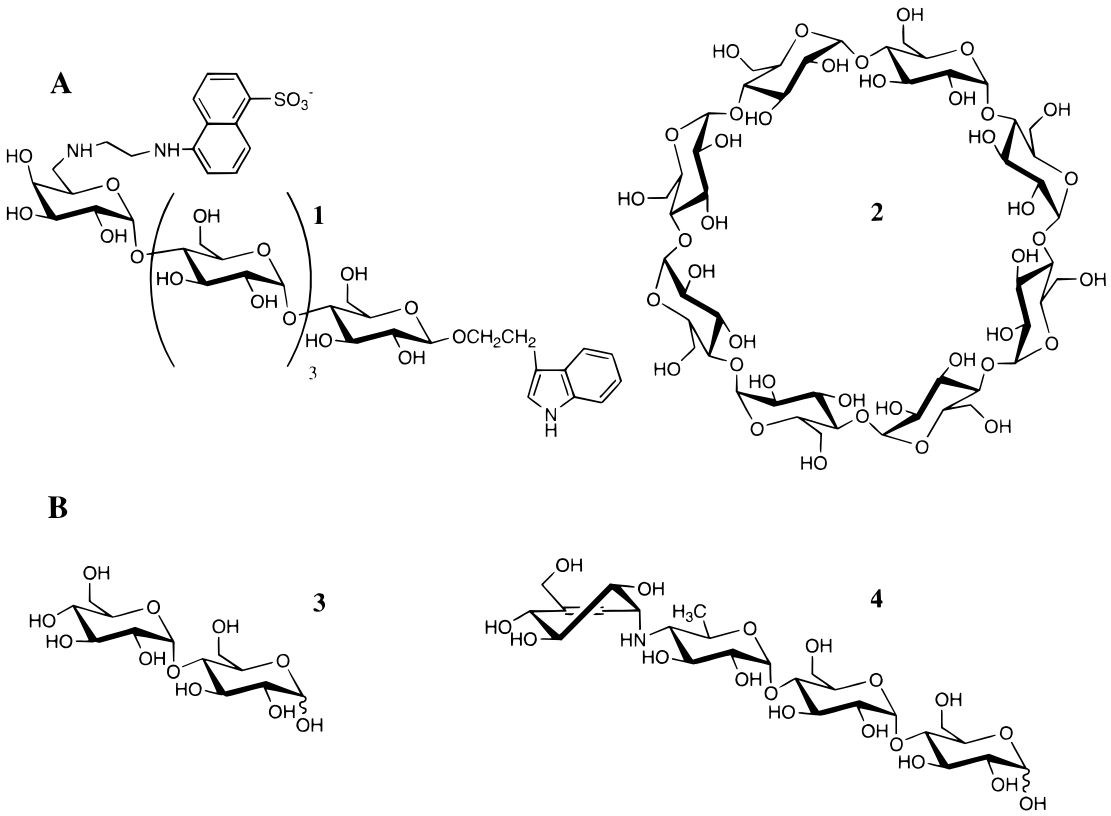

FIGURE 1: (A) Oligosaccharide substrates used for the characterization of Novamyl. (**1**) Indolylethyl 5-[(2-aminoethyl)amino]-1-naphthylenesulfonate-substituted maltopentaside (*35*) and (**2**) γ-cyclodextrin. (B) The disaccharide maltose (**3**) and tetrasaccharide inhibitor acarbose (**4**) used in this study.

E-domain implicated in binding to granular starch. Numerous α-amylase and CGTase structures of both the free enzyme and the enzyme in complex with oligosaccharide inhibitors have been described (for examples, see refs *13* and *14*).

In this paper, we describe the structure of the *B. stearothermophilus* "maltogenic" α-amylase. The structure has been determined both in a complex with maltose and in a complex with maltose and acarbose (Figure 1B) at 1.7 Å resolution. Unusually, this enzyme exhibits the five-domain organization extremely similar to that exhibited by the family 13 CGTases. The substrate-binding surface, however, is an open groove which may accommodate both cyclodextrins as well as linear substrates and which presents no obvious structural barrier to endo catalytic activity.

MATERIALS AND METHODS

Expression and Purification. Four grams of a commercial "Novamyl BG 10.000 granulate" (from Novo Nordisk A/S) was dissolved in 25 mM HEPES buffer (pH 7.5) with 200 mg/mL ammonium sulfate. The sample was applied to an 80 mL Toyopearl butyl column equilibrated in the same buffer and washed. Novamyl was eluted in one step by changing the buffer to 25 mM HEPES (pH 7.5). The pooled material was diluted until the conductivity was below 2 mS cm^{-1} and applied to a Q-Sepharose column equilibrated in 25 mM HEPES (pH 7.5). Novamyl was collected from the run-through. The pH of the pool was adjusted to 9.5, and the sample was reapplied to Q-Sepharose equilibrated in 25 mM piperazine (pH 9.5). For fractionation, a gradient to 1 M NaCl in 10 column volumes was used. The fractions were analyzed for Novamyl activity and pooled according to purity and concentration. The pool was adjusted to pH 7.5, concentrated, and subsequently dialyzed against 10 mM Tris (pH 7.5).

α-Cyclodextrin (α-CD) was immobilized on divinyl sulfone-activated agarose by the method described by the manufacturer (Kem-En-Tec). A ligand concentration of 6 g of α-CD to 25 mL of gel was used during coupling to a medium-activated gel type (14 mmol/L reactive sites). A 1 mL sample of the gel was packed in a XK16 column (Pharmacia) and equilibrated with 25 mM sodium acetate (pH 5.0). One milliliter of purified Novamyl (A_{280} = 3.17) was applied to the column at a flow rate of 2 mL/min. After washing with the equilibration buffer followed by washing with the same buffer with the addition of 0.5 M NaCl, bound material was eluted by including 1% α-CD. Purified Novamyl was dialyzed against 10 mM Tris (pH 7.5).

Crystallization and Data Collection. Crystals of the enzyme were grown by the hanging drop method. The protein solution contained 19 mg/mL enzyme in 10 mM Tris buffer (pH 6.0), 0.2 M NaCl, and 5 mM $CaCl_2$. The well solution contained 0.9 M Li_2SO_4, 2.5% (w/v) PEG 1450, 50 mM TEA (triethanoloamine) buffer (pH 6.5), and 100 mM maltose. In the crystallization drops, these solutions were mixed 1:1. Before the crystals were frozen for data collection, they were immersed in a cryoprotectant solution consisting of the well solution with the addition of 30% (v/v) glycerol. Acarbose complex crystals were obtained by soaking such maltose-containing crystals for 4 days in the well solution containing 20 mM acarbose.

X-ray diffraction data for both the maltose and acarbose complexes were collected using a Rigaku RU200 rotating

Table 1: Data Collection Statistics[a]

	maltose	acarbose
space group	$P3_121$	$P3_121$
cell dimensions (Å)	$a = 89.76$	$a = 89.82$
	$c = 185.71$	$c = 185.75$
resolution (Å)	20−1.7 (1.73−1.7)	20−1.7 (1.73−1.7)
total no. of reflections measured	460513	630805
no. of unique reflections	87912	94506
completeness (%)	95.8 (92.5)	99.1 (97.2)
$R(I)$ merge (%)	8.0 (45.4)	8.2 (49.4)
$I/\sigma(I)$	17.8 (2.5)	19.7 (3.1)
$I > 2\sigma(I)$ (%)	82 (54)	84 (58)

[a] Statistics for the highest-resolution shell are given in parentheses.

anode source ($\lambda = 1.5418$ Å) together with a MAR Research 300 mm imaging plate scanner and long focusing-mirror optics. Crystals were mounted in a rayon-fiber loop and frozen in a stream of nitrogen gas at 100 K. For both complexes, two sets of images were collected, at "high" resolution (1.7 Å) with 30 min of exposure per 0.75° of rotation and "low" resolution with 5 min of exposure per 1° of rotation. The data were processed and merged using the HKL suite (*15*) (Table 1). The low-resolution data, up to 2.8 Å, are 100% complete.

Structure Solution and Refinement. The structure of the maltose complex was determined by molecular replacement using the program AMoRe (*16*, *17*) and the coordinates of the *Bacillus circulans* strain 251 cyclodextrin glucosyltransferase (PDB entry 1CDG) (*18*) with data in the resolution range of 10−4.0 Å. This and all other programs used were from the CCP4 suite (*19*) unless explicitly stated otherwise. The solution in space group $P3_121$ was clear after rigid body fitting, with a correlation coefficient of 36.3% and an R factor of 46.6%, whereas the next (incorrect) solution had a correlation coefficient of 13.5% and an R factor of 52.9%. There was no clear solution in space group $P3_221$.

The model obtained from molecular replacement and containing the structure of CGTase, with all atoms changed to "dummy" water molecules, was subjected to 10 cycles of ARP (*20*) in tandem with unrestrained maximum likelihood refinement using REFMAC (*21*). The full resolution range of the data, 20−1.7 Å, was used with 5% set aside for cross validation (*22*). In each cycle, ARP was allowed to reject and add up to 50 atoms. The R factor fell from 52.1 to 36.6% and the free R from 54.2 to 48.3%. ARP was not driven to completion but terminated when a substantially improved and easily interpretable electron density map was obtained. Most of the Novamyl protein chain with its correct sequence (572 out of 686 residues) was built into this map using QUANTA (Molecular Simulations Inc., 1997).

The initial model of Novamyl was refined by REFMAC with stereochemical restraints applied to the protein chain (*23*). ARP was employed to handle the solvent atoms on the basis of Fourier syntheses with maximum likelihood-weighted $2F_{obs} - F_{calc}$ and $F_{obs} - F_{calc}$ coefficients, estimated from σ_A values by REFMAC on the basis of the cross validation "free" subset of reflections. In the initial stages, it was realized that the active site of the enzyme contains two molecules of maltose and that additional maltose molecules were present on the C and E domains. The refinement converged with an R factor of 14.7% and an R_{free} of 19.4%. The refinement of the acarbose complex started from the final maltose complex coordinates since the two were highly isomorphous. All nonprotein atoms (maltose and waters) were first removed, and the new solvent was built using ARP. This procedure resulted in a high-quality map with unambiguous electron density for the oligosaccharide ligands.

Table 2: Refinement and Structure Quality Statistics for the Maltose and Maltose−Acarbose Complexes

	maltose	maltose−acarbose
resolution (Å)	20−1.7	20−1.7
no. of protein atoms/molecule (residues 1−686)	5352[a]	5356[a]
no. of solvent waters	1138	1106
R_{cryst} (outer resolution shell)	0.153 (0.252)	0.151 (0.225)
R_{free} (outer resolution shell)	0.191 (0.272)	0.175 (0.278)
rms deviation for 1−2 bond (Å)	0.008	0.007
rms deviation for 1−3 angle distance (Å)	0.023	0.022

[a] This apparent discrepancy results from different modeling of alternative conformations in the two structures.

RESULTS AND DISCUSSION

Quality of the Final Model Structure of the Novamyl Maltose Complex. The gene encoding Novamyl exhibits a 709-residue open-reading frame, including a 33-residue Gram-positive translocation signal which is post-translationally removed to leave a mature Novamyl species of 686 residues (*7*, *8*). A few discrepancies with respect to the previously published sequence of Novamyl have been observed by DNA sequencing (L. Beier and T. V. Borchert, unpublished) and are confirmed by the X-ray structure described here. There is a three-residue insertion, ADG, after residue Gly 221 (published numbering of the mature enzyme); the region from residue 334 (old numbering) now is NSNKANLHQALAFILTSRGTPSI (as opposed to NSKNKANLHQRLLSFSLRGVRPPI in the published sequence), and Asn 76 should have been Asp 76. Sequence alignments revealed that the closest degree of similarity, over the whole length of the protein, was observed with the CGTases from family 13 with a maximal identity of approximately 60%. The structure was therefore determined by molecular replacement using the known family 13 CGTase from *B. circulans* strain 251 (PDB entry 1CDG) (*18*) as the search model.

The final data set for the maltose complex consists of 460 513 observations of 87 912 unique data points, 95.6% complete to 1.7 Å with a merging R value of 0.08 (Table 1). The final model consists of residues 1−686 (5352 protein atoms), four molecules of maltose, a single sulfate ion, three Ca^{2+} ions, and 1083 water molecules. The crystallographic R value is 0.153 with a free R value of 0.191, and 90% of the non-glycine residues lie in the most favored regions of the Ramachandran plot (*24*) with no residues in forbidden regions, as defined by PROCHECK (*25*). A summary of the model statistics for the maltose and maltose−acarbose complexes is given in Table 2.

The domain structure of Novamyl is different from that seen in other α-amylases. It displays the five-domain organization typical of the CGTases (Figure 2). The catalytic core $(\beta/\alpha)_8$ domain consists of residues 1−131 and 201−

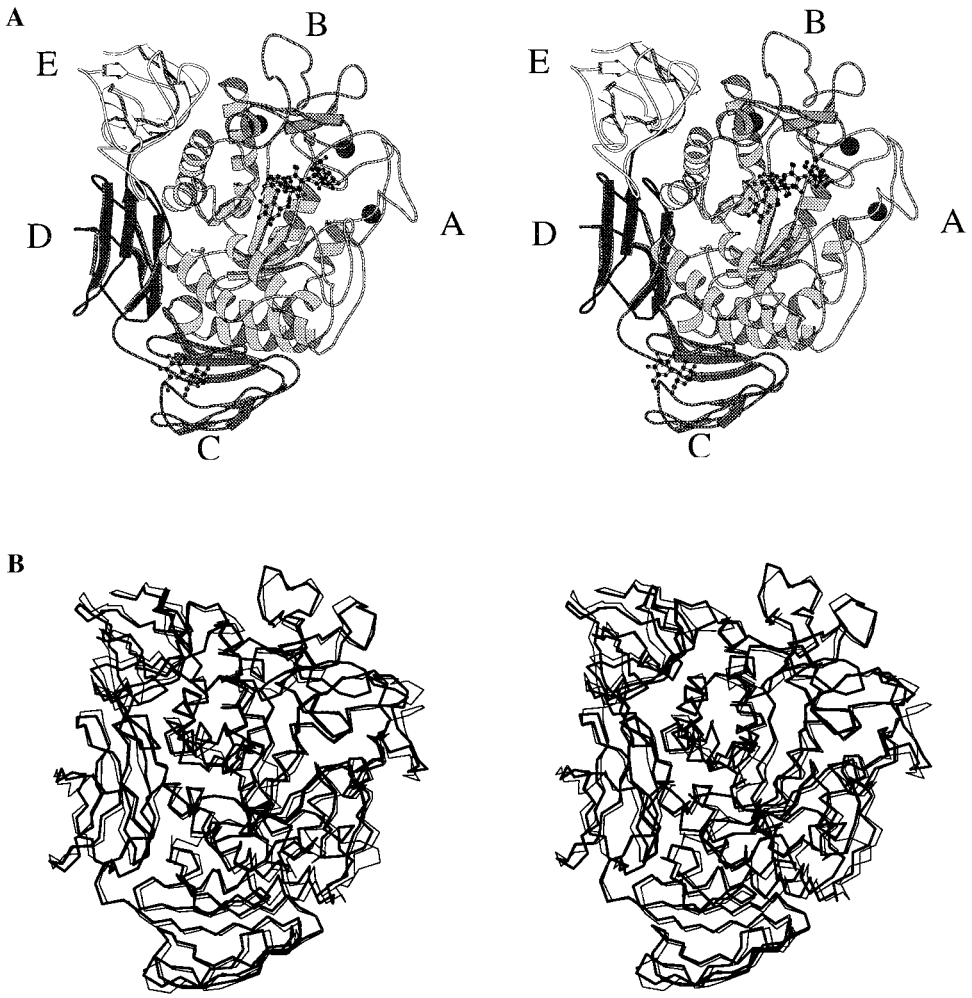

FIGURE 2: (A) Divergent stereo schematic diagram of the maltogenic α-amylase, Novamyl, from *B. stearothermophilus* drawn with MOLSCRIPT (*36*). The hexasaccharide ligand in the active site and the maltose residue on domain C are shown in ball-and-stick representation and the calcium ions as shaded spheres. (B) Overlay of Novamyl with the CGTase from *B. circulans* strain 251 (*18*).

408. The large excursion, residues 131−201, between barrel strand β-3 and helix α-3, forms domain B. The conserved structural calcium site is located at the interface between the A- and B-domains where it stabilizes the substrate-binding cleft. Domain C is an eight-stranded β-barrel formed by residues 408−496. Over the "catalytic domains" (A−C), the sequence of Novamyl is approximately 50% identical with known CGTase sequences compared to approximately 20−28% with α-amylase A−C-domains. In addition to these three domains which are typical of the α-amylase family, Novamyl has two further domains, more traditionally associated with the CGTases: domain D consisting of residues 496−577 and the starch-binding domain E consisting of residues 577−686.

The topology of Novamyl is similar to that of the CGTases. A total of 647 (out of the 686) residues overlap with the *B. circulans* CGTase with an rms of 1.1 Å on the equivalent Cα positions (Figure 2B). There are a number of significant differences between the two structures. A loop insertion (residues 615−622) in the E-domain of Novamyl, compared to CGTase, provides extensive additional interactions between the E- and D-domains and, most importantly, between the putative starch-binding E-domain and the catalytic core A-domain. The loop makes extensive H-bonding interactions with helix α-5 from the A-domain. In common with other family 13 enzymes, such as porcine pancreatic amylase (1PPI) (*26*), helix α-5 also contains a short insertion of two amino acids in the region consisting of residues 262−267 that is not present in CGTase. This, together with slight changes in the adjacent loop orientation, provides a different environment at the reducing end of the active site cleft, more amenable to extended linear polymers, which is closed in CGTases by a loop from residue 260 to 264 (*B. circulans* numbering). Another difference between CGTase and Novamyl occurs at the "nonreducing" end of the active site cleft. Novamyl contains a five-residue insertion in the B-domain (residues 190−194) which partially encloses the active site at the −3 subsite. Additionally, this insertion provides residues involved in both −2 and −3 subsite binding (see below). The absence of this insertion in CGTases allows for additional subsites, with up to −7 characterized crystallographically (*27*). Novamyl contains an additional Ca^{2+} site that is absent in CGTase. Novamyl displays a single maltose molecule bound to the "starch-binding" E domain, in contrast to the related *B. circulans* CGTase which displayed two molecules of maltose on this domain. Novamyl also exhibits a C-domain maltose positioned exactly as previously observed in CGTase, pointing to a genuine biological role, in maltooligosaccharide binding, for the C-domain as proposed by Dijkstra and colleagues (*18*). The implications for this are discussed below.

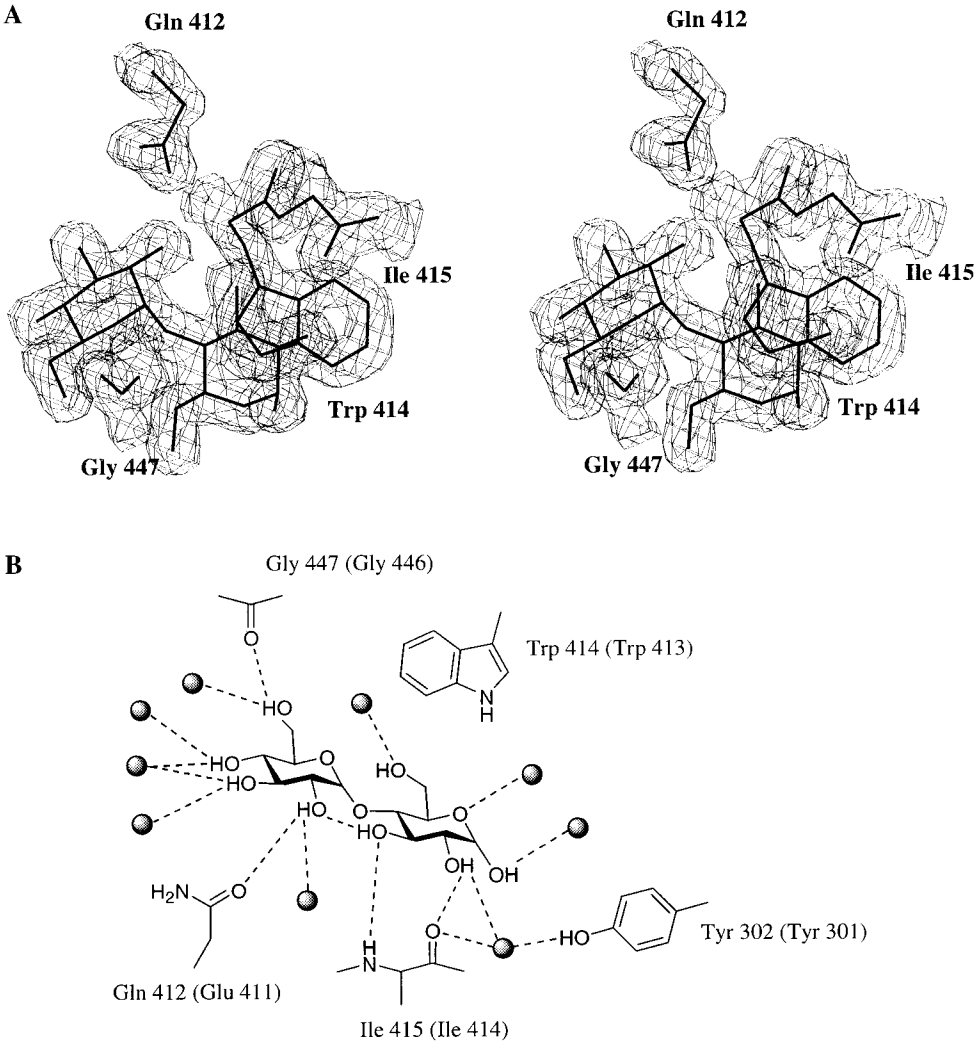

FIGURE 3: Maltose binding to the C-domain of Novamyl. (A) Maximum likelihood-weighted $2F_o - F_c$ electron density, contoured at 0.4 eÅ^{-3}. (B) Schematic diagram of the interactions with the C-domain. Equivalent residues involved in maltose binding to the C-domain of the *B. circulans* CGTase (*18*) are shown in parentheses.

Novamyl has three calcium sites. One is the conserved "high-affinity" structural calcium found at the interface between the A- and B-domains where residues involved in its coordination also contribute to the substrate-binding surface around the +1 subsite, as reviewed in refs *14* and *28*. The coordination is provided by a bidentate interaction with the carboxylate of Asp 198, main chain carbonyls from Gln 184 and His 232, the side chain carbonyl of Asn 131, and three water molecules, resulting in a total of eight ligands. The two remaining calciums are in domain A. One is coordinated by the main chain carbonyl of residue Asn 77, the carboxylates of residues Glu 101 and Glu 102 (both in a bidentate manner), one of the carboxylate oxygens of Asp 76 and Asp 79, and a single water molecule, for a total of eight ligands. These residues lie in an extended loop between strand β-2 and helix α-2. The inner surface of this loop is involved in the +2 and +3 subsite formation, but the calcium lies on the external face away from the active site and is not directly involved in substrate binding. This calcium ion is absent in CGTase, which possesses a similar main chain conformation but lacks the appropriate side chain ligands. The third calcium helps to stabilize a loop between the first β-strand and helix. It is heptacoordinate with ligands provided by the main chain carbonyls of Asp 23 and Gly 48, a single carboxylate oxygen from Asp 21 and Asp 50, side chain carbonyls from Asn 26 and Asn 27, and a single water molecule. There is a sulfate ion, presumably as a result of the presence of lithium sulfate in the crystallization medium, at the interface between domains C and D where it is bound to the main chain amide groups of Ala 496 and Gly 521 and the side chain amide of Asn 544.

Maltose Binding. The structure of Novamyl grown in the presence of 100 mM maltose exhibits four molecules of the disaccharide. One is bound on the C-domain (Figure 3**),** where a maltose residue was also found in *B. circulans* CGTase (*18*). The equivalent maltose in Novamyl exhibits binding interactions almost identical to that in CGTase. Four of the five H-bonds are totally conserved, and the fifth, between Glu 411 and O-3A of maltose in CGTase, is conservatively changed to an interaction between the carbonyl of Gln 412 in Novamyl (Figure 3B). The other predominant interaction, the stacking with Trp 413 in CGTase, is also completely conserved in Novamyl. Since the crystal packing environments of CGTase and Novamyl are quite different, the presence of an identically positioned maltose, coupled to its identical interactions with the C-domain, points to a genuine biological role in maltooligosaccharide binding for this domain, as suggested by

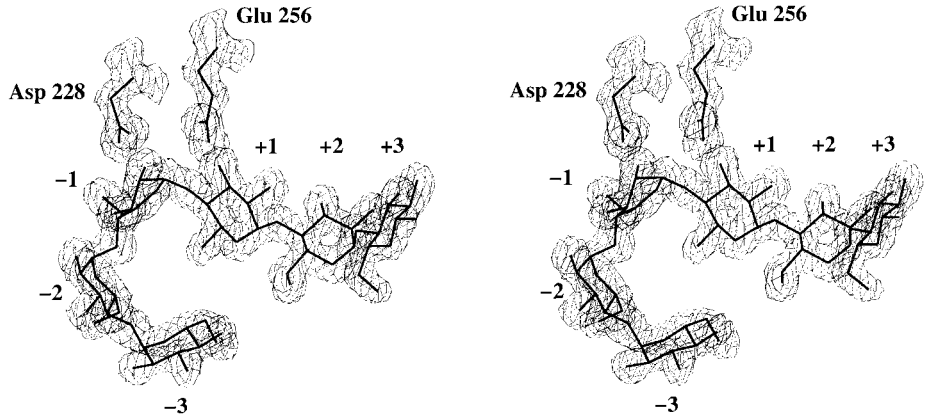

FIGURE 4: Electron density for the acarbose-derived hexasaccharide bound to the −3 to +3 sites of Novamyl. The map shown is a maximum likelihood-weighted $2F_o - F_c$ electron density, contoured at 0.4 eÅ^{-3} in divergent stereo.

Dijkstra and co-workers (*18*). We believe that the role of the C-domain is likely to be solely in localizing the enzyme on the substrate. A role in processivity cannot be ruled out, but it seems unlikely that the same polysaccharide chain that enters the catalytic site could also interact simultaneously with the C-domain.

There are two other molecules of maltose in the active site. They are bound in the −2 to −1 and +2 to +3 sites (nomenclature described in ref *29*). The interactions of the well-defined maltose molecule in the −1 to −2 subsites are essentially identical to that of the equivalent subsites in the acarbose complex, described below. The only notable difference is that the unbiased density reveals that the O-1 hydroxyl on the anomeric carbon of the maltose is present as a mixture of both anomers as a result of mutarotation with the α-anomer present with an occupancy of about 0.66 and the β-anomer 0.33. The +2 to +3 subsite maltose is less well-defined in the electron density (*B* values of approximately 30 Å^2 when modeled with 60% occupancy) and occupies a slightly different position compared to that seen for the equivalent sugars in the acarbose complex, described below.

Novamyl Acarbose−Maltose Complex. A complex of Novamyl with the pseudotetrasaccharide inhibitor acarbose was obtained by soaking the maltose-grown crystals overnight in 20 mM acarbose. Acarbose is a potent amylase inhibitor, the tight binding resulting from the nonreducing end valienamine pseudosaccharide whose half-chair conformation mimics the oxocarbenium ion transition state (*11*). Acarbose has proved to be particularly useful for dissecting protein−carbohydrate interactions in a variety of amylolytic enzymes (*13*, *14*, *26*, *27*, *30*, *31*). Acarbose is a tetrasaccharide, yet in the published structures of family 13 enzymes, it is frequently observed as a species with a higher degree of polymerization, leading to speculation that the enzyme is able to transglycosylate or condense molecules to make a longer and hence tighter-binding inhibitor (*13*, *14*, *26*, *27*).

After one session of refinement of the Novamyl acarbose−maltose complex, it was clear that the active site contained a hexasaccharide species. In addition, maltose is observed on the C and E domains, as before. The electron density for two modified sugars in the active site allowed them to be identified as the pseudodisaccharide moiety of acarbose. The acarbose was extended at its nonreducing end by two α-1,4-linked glucopyranose rings. The hexasaccharide was incorporated into the model, and the refinement converged with an *R* factor of 14.82% and an R_{free} of 17.79% (Table 2). The electron density at the glycosidic bonds between all six sugars was clear, and the density at the two ends finished abruptly, indicating that the active site contains a single hexameric chemical species and not several smaller overlapped oligomers. Discrimination between these two interpretations may be realized by refining the individual crystallographic occupancies with a program such as SHELXL97 (*32*). Although at 1.7 Å resolution the results should be interpreted with caution, the occupancies refined to the following values (from the nonreducing end): (−3) 0.89, 1.03, 1.05, 1.00, 0.93, and 0.87 (+3). This suggests that the hexasaccharide is a genuine single chemical species and not an overlap of smaller sugars.

The acarbose-derived hexasaccharide binds in subsites −3 to +3. The two additional glucose units at the nonreducing end of the acarbose moiety are unambiguously glucopyranoside rings and display good density for the hydroxymethyl group (Figure 4). In this sense, they are genuinely different from previously observed complexes where the absence of density for the exocyclic group of the additional glucose units confused their interpretation. This is important, since it would point to hexasaccharide formation via condensation of maltose and acarbose and not via a transglycosylation mechanism, since cleavage of acarbose followed by transglycosylation would result in a second 6-deoxy species in the product (*14*) (Figure 5). The −2 to −1 subsite maltose present under the original conditions is perfectly positioned to undergo condensation with the added acarbose as an acceptor to form the present complex. We have tried to wash maltose off the protein both in crystal and in solution to prevent this reaction, without success. Extended species have been seen in many previous family 13 complexes (*13*, *26*, *27*, *31*) and the arguments for and against such mechanisms discussed at length (*11*, *14*). They will not be considered further here. Whether condensation reactions play any role in the natural environment is clearly open to speculation. We think it is unlikely since uptake and utilization systems will exist for removal of maltose by the organism and the predominant sugar species will be longer polysaccharides, where hydrolysis and perhaps transglycoyslation mechanisms are more likely.

The hexasaccharide allows description of all the noncovalent interactions for the −3 to +3 subsites (Figure 6). The

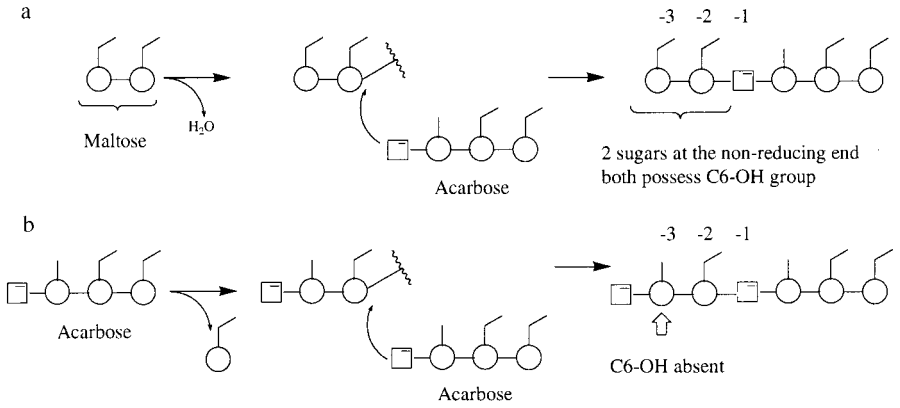

FIGURE 5: Schematic figure showing the formation of an extended acarbose-derived oligosaccharide via (a) condensation of maltose and (b) transglycosylation of acarbose. Only formation of the oligosaccharide by condensation with maltose yields a product in which the −3 and −2 subsite sugars possess the characteristic C-6-OH hydroxymethyl substituent.

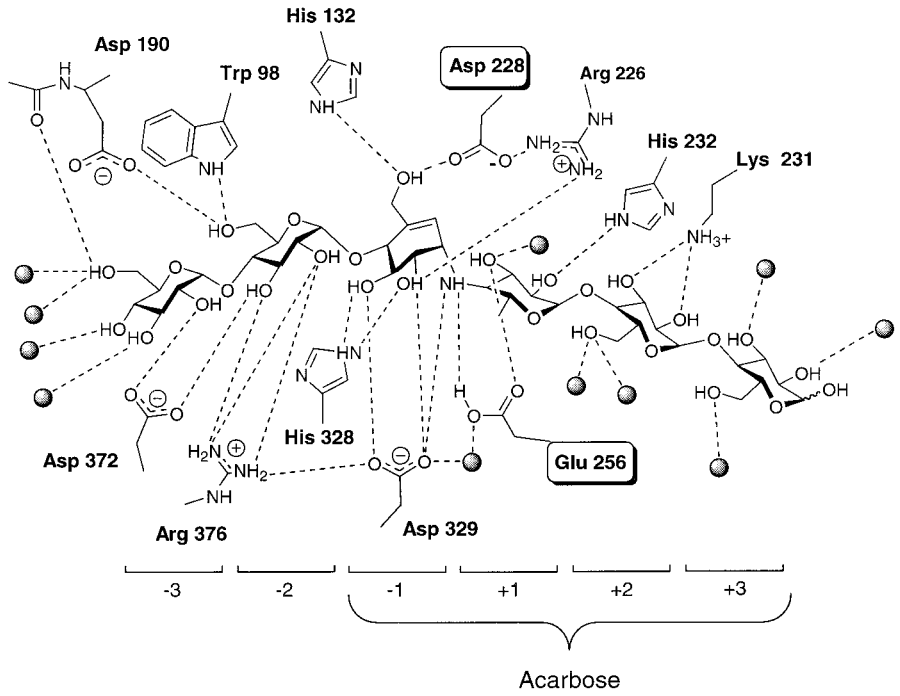

FIGURE 6: Schematic diagram of the interactions between Novamyl and the modified acarbose hexasaccharide inhibitor.

unsaturated pseudosaccharide ring lies in the −1 subsite as expected and as observed in most related studies. The ring is in the $^2H_3$ half-chair conformation, imposed by the C6−C5 double bond, believed to mimic the half-chair transition state. The majority of the interactions we have described previously for the hexasaccharide complex of the *Aspergillus oryzae* (TAKA) α-amylase (*14*) are retained in the equivalent hexasaccharide complex of Novamyl. The hydrogen-bonding interactions in the −1 to +3 subsites are essentially identical, with the +2 hydrophobic stacking interaction being provided by Tyr 258 in Novamyl as opposed to Leu 232 in TAKA. At the critical −1 subsite, the only difference is in the flexible His 328. In the family 13 structures, this histidine is found in a range of conformations. It may interact with the tyrosine residue which forms the hydrophobic stacking in the −1 subsite, as observed in TAKA amylase (*14*) and is often seen directly H-bonding to the O-2 and O-3 hydroxyls of the −1 subsite sugar (*33*). In Novamyl, His 328 lies in an intermediate position between these two extremes and appears to be hydrogen bonded to both the −1 sugar and Tyr 92. The surface of the active site through the −2 and −3 subsites in Novamyl results in a much tighter curvature of the hexasaccharide than was observed in the previous α-amylase structures. A result of this is that the linear hexasaccharide species in Novamyl exhibits a conformation reminiscent of part of a cyclodextrin, consistent with its high hydrolytic activity on these substrates compared to α-amylases.

Conclusions. Early work with this *B. stearothermophilus* α-amylase revealed a number of unusual, apparently contradictory, properties that distinguished it from other amylases (*6*). Notably, the apparent hydrolysis product was the disaccharide maltose. This was interpreted as evidence that, like glucoamylase and *β*-amylase, Novamyl was an exo enzyme with a requirement for nonreducing chain ends. Unlike *β*-amylase, the maltose product had the α-configuration, indicative of an enzyme acting with net retention of anomeric configuration. Many properties of the enzyme indicated that, although maltose was the final reaction product, the enzyme exhibited no requirement for chain ends. It was, for example, active on cyclodextrins. More recently,

the mechanism of Novamyl has been revisited (*7*). The enzyme shows high catalytic activity on cyclodextrins (especially γ-cyclodextrin (**2**), Figure 1) and on terminally modified substrates such as compound **1** (Figure 1). In addition, hydrolysis of amylose leads to a dramatic decrease in the degree of polymerization, indicative of endo attack within the polymer chain. Thus, the enzyme displays superficially contradictory properties: the product is maltose, but the enzyme is not specific for chain ends. It is likely that maltose represents the final reaction product from prolonged hydrolysis, but a "multiple-attack" (*5*) mechanism in which the enzyme catalyzes the hydrolysis of numerous bonds before it dissociates from the substrate cannot fully be ruled out.

The three-dimensional structure of Novamyl is fully consistent with its ability to carry out an endo attack within the polysaccharide chain. The active site is found in a steep-sided depression on the enzyme surface, very similar to that seen in endo α-amylases, and imposes no structural barrier to an initial endo attack. In this sense, it is wholly different from the exo amylolytic enzymes β-amylase (family 14) and glucoamylase (family 15) whose active sites are in pockets on the enzyme surface, ensuring entry of just the polysaccharide nonreducing chain ends (*34*). In addition, the acarbose-derived hexasaccharide inhibitor binds in the -3 to $+3$ subsites, wholly inconsistent with a maltogenic (-2) exo enzyme.

It is feasible that many of the macroscopic properties of Novamyl reside in its granular starch-binding domain, which could retain the amylose chain and prevent it from dissociating fully from the enzyme following a catalytic event. This cannot explain the rather confusing observation that, on small soluble substrates, Novamyl catalyzes the hydrolysis of α-cylodextrin in a processive manner but linear maltohexaose in a nonprocessive fashion (*7*). These results cannot be explained by the presence of a starch-binding domain but must instead reside in the relative binding energies of the respective subsites. The three-dimensional structure of Novamyl is entirely consistent with its description as an endo-acting amylase; there is no structural barrier to an initial endo attack. This, coupled to its observed activities on cyclodextrins, terminally modified linear maltodextrins, and amylose, certainly points to its reclassification as an α-amylase with unusual properties.

ACKNOWLEDGMENT

We thank Hugues Driguez for kindly providing oligosaccharide analogues.

REFERENCES

1. Henrissat, B. (1991) *Biochem. J. 280*, 309−316.
2. Henrissat, B., and Bairoch, A. (1993) *Biochem. J. 293*, 781−788.
3. Henrissat, B., and Bairoch, A. (1996) *Biochem. J. 316*, 695−696.
4. Henrissat, B., and Davies, G. J. (1997) *Curr. Opin. Struct. Biol. 7*, 637−644.
5. Robyt, J. F., and French, D. (1967) *Arch. Biochem. Biophys. 122*, 8−16.
6. Outtrup, H., and Norman, B. E. (1984) *Starch/Staerke 36*, 405−411.
7. Christophersen, C., Otzen, D. E., Norman, B. E., Christensen, S., and Schäfer, T. (1997) *Starch/Staerke 50*, 39−45.
8. Diderichsen, B., and Christensen, L. (1988) *FEMS Microbiol. Lett. 56*, 53−60.
9. Svensson, B. (1994) *Plant Mol. Biol. 25*, 141−157.
10. Koshland, D. E. (1953) *Biol. Rev. 28*, 416−436.
11. Davies, G., Sinnott, M. L., and Withers, S. G. (1997) in *Comprehensive Biological Catalysis* (Sinnott, M. L., Ed.) pp 119−209, Academic Press, London.
12. Jespersen, H. M., MacGregor, E. A., Sierks, M. R., and Svensson, B. (1991) *Biochem. J. 280*, 51−55.
13. Strokopytov, B., Penninga, D., Roseboom, H. J., Kalk, K. H., Dijkhuizen, L., and Dijkstra, B. W. (1995) *Biochemistry 34*, 2234−2240.
14. Brzozowski, A. M., and Davies, G. J. (1997) *Biochemistry 36*, 10837−10845.
15. Otwinowski, Z., and Minor, W. (1997) in *Methods in Enzymology: Macromolecular Crystallography, Part A* (Carter, C. W., Jr., and Sweet, R. M., Eds.) pp 307−326, Academic Press, London and New York.
16. Navaza, J. (1994) *Acta Crystallogr. A50*, 157−163.
17. Navaza, J., and Saludijan, P. (1997) *Methods Enzymol. 276*, 581−594.
18. Lawson, C. L., Montfort, v. R., Strokopytov, B., Rozeboom, H. J., Kalk, K. H., de Vries, G. E., Penninga, D., Dijkhuizen, L., and Dijkstra, B. W. (1994) *J. Mol. Biol. 236*, 590−600.
19. Collaborative Computational Project Number 4 (1994) *Acta Crystallogr. D50*, 760−763.
20. Lamzin, V. S., and Wilson, K. S. (1993) *Acta Crystallogr. D49*, 129−147.
21. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. D53*, 240−255.
22. Brünger, A. T. (1992) *Nature 355*, 472−475.
23. Engh, R. A., and Huber, R. (1991) *Acta Crystallogr. A47*, 392−400.
24. Ramachandran, G. N., Ramakrishnan, C., and Sasisekharan, V. (1963) *J. Mol. Biol 7*, 95−99.
25. Laskowski, R. A., McArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 282−291.
26. Qian, M., Haser, R., Buisson, G., Duée, E., and Payan, F. (1994) *Biochemistry 33*, 6284−6294.
27. Strokopytov, B., Knegtel, R. M. A., Penninga, D., Rozeboom, H. J., Kalk, K. H., Dijkhuizen, L., and Dijkstra, B. W. (1996) *Biochemistry 35*, 4241−4249.
28. Boel, E., Brady, L., Brzozowski, A. M., Derewenda, Z., Dodson, G. G., Jensen, V. J., Petersen, S. B., Swift, H., Thim, L., and Woldike, H. F. (1990) *Biochemistry 29*, 6244−6249.
29. Davies, G. J., Wilson, K. S., and Henrissat, B. (1997) *Biochem. J. 321*, 557−559.
30. Aleshin, A. E., Firsov, L. M., and Honzatko, R. B. (1994) *J. Biol. Chem 269*, 15631−15639.
31. Gilles, C., Astier, J.-P., Marchis-Mouren, G., Cambillau, C., and Payan, F. (1996) *Eur. J. Biochem. 238*, 561−569.
32. Sheldrick, G. M., and Schneider, T. R. (1997) *Methods Enzymol. 277*, 319−343.
33. Machius, M., Wiegand, G., and Huber, R. (1995) *J. Mol. Biol. 246*, 545−559.
34. Davies, G., and Henrissat, B. (1995) *Structure 3*, 853−859.
35. Payre, N., Cottaz, S., and Driguez, H. (1995) *Angew. Chem., Int. Ed. 34*, 1239−1241.
36. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

BI990256L