

Refined Crystal Structures of Glucoamylase from *Aspergillus awamori* var. X100

Alexander E. Aleshin¹, Chad Hoffman²,
Leonid M. Firsov¹ and Richard B. Honzatko^{2†}

¹Department of Molecular
St Petersburg Nuclear Physics Institute,
St Petersburg, Russia

²Department of Biochemistry and Biophysics
Iowa State University
Ames IA 50011, U.S.A.

The refined crystal structures of a proteolytic fragment of glucoamylase from *Aspergillus awamori* var. X100 have been determined at pH 6 and 4 to a resolution of 2.2 Å and 2.4 Å, respectively. The models include the equivalent of residues 1 to 471 of glucoamylase from *Aspergillus niger* and a complete interpretation of the solvent structure. The *R*-factors of the pH 6 and 4 structures are 0.14 and 0.12, respectively, with root-mean-square deviations of 0.014 Å and 0.012 Å from expected bondlengths. The enzyme has the general shape of a doughnut. The “hole” of the doughnut consists of a barrier of hydrophobic residues at the center, which separates two water-filled voids, one of which serves as the active site. Three clusters of water molecules extend laterally from the active site. One of the lateral clusters connects the deepest recess of the active site to the surface of the enzyme.

The most significant difference in the pH 4 and 6 structures is the thermal parameter of water 500, the putative nucleophile in the hydrolysis of maltooligosaccharides. Water 500 is associated more tightly with the enzyme at pH 4 (the pH of optimum catalysis) than at pH 6. In contrast to water 500, Glu179, the putative catalytic acid of glucoamylase, retains the same conformation in both structures and is in an environment that would favor the ionized, rather than the acid form of the side-chain.

Glycosyl chains of 5 and 8 sugar residues are linked to Asparagines 171 and 395, respectively. The conformations of the two glycosyl chains are similar, being superimposable on each other with a root-mean-square discrepancy of 1.9 Å. The *N*-glycosyl chains hydrogen bond to the surface of the protein through their terminal sugars, but otherwise do not interact strongly with the enzyme. The structures have ten serine/threonine residues, to each of which is linked a single mannose sugar. The structure of the ten *O*-glycosylated residues taken together suggests a well-defined conformation for proteins that have extensive *O*-glycosylation of their polypeptide chain.

Keywords: glycoprotein structure; enzyme catalysis; X-ray crystallography; carbohydrase mechanisms; amylolytic enzymes

1. Introduction

Glucoamylase (α -1 $\rightarrow$ 4 glucan glucohydrolase, EC 3.2.1.3) catalyzes the release of β -glucose from the non-reducing end of maltodextrin chains (Weill *et al.*, 1954; Pazur & Ando, 1960; Hiromi *et al.*, 1983). Glucoamylases from various species of *Aspergillus* are used extensively in industry in the production of glucose (Saha & Zeikus, 1989). The

glucose in turn serves as feedstock to biological fermentations or is used in the manufacturing of high fructose sweeteners. Unfortunately, glucoamylase is not suited ideally to its industrial roles, having a pH of optimum catalysis and thermostability that are too low. In order to engineer an enzyme with superior properties, however, one requires at the very least a thorough understanding of the relationship of the enzyme's structure to its function.

Various species of *Aspergillus* elaborate a parent protein which, after cleavage of a signal peptide and

† Author to whom all correspondence should be addressed.

glycosylation, appears in the extracellular growth medium. The secreted protein (usually called glucoamylase I) has three functional domains (Svensson *et al.*, 1983, 1989): (1) an N-terminal catalytic domain of approximately 440 residues, (2) an *O*-glycosylated linker domain of approximately 70 residues and (3) a C-terminal starch binding domain of approximately 100 residues. Typically the parent protein, which in species of *Aspergillus* is 615 or 616 amino acids (Boel *et al.*, 1984; Nunberg *et al.*, 1984; Hayashida *et al.*, 1989a; Shibuya *et al.*, 1990; Hata *et al.*, 1991), undergoes limited proteolysis, resulting in the appearance of glucoamylases (often called glucoamylase II) of smaller molecular weight. The shortened forms of glucoamylase I lack the entire starch binding domain and, in some instances, all or part of the *O*-glycosylated domain (Svensson *et al.*, 1986; Hayashida *et al.*, 1989a,b). Glucoamylase II has properties of catalysis identical to those of glucoamylase I, using soluble maltodextrins as substrates. The k_{cat} increases and the K_m decreases for maltooligosaccharides of increasing length up to six residues (Abdullah *et al.*, 1963; Hiromi *et al.*, 1983). However, only glucoamylase I can act directly on raw starch granules (Hayashida *et al.*, Svensson *et al.*, 1982).

Recently we reported the preliminary crystal structure of a proteolytic fragment of glucoamylase I from *Aspergillus awamori* var. *X100* (Aleshin *et al.*, 1992). The proteolytic fragment, named glucoamylase-II(471), corresponds to residues 1 to 471 of glucoamylase from *Aspergillus niger*. Glucoamylase-II(471) is catalytically active with properties of catalysis nearly identical to that of glucoamylase II from *A. niger*. Furthermore, the crystalline fragment contains the entire catalytic domain and the N-terminal half of the *O*-glycosylated domain. Aleshin *et al.* (1992) have described the fold of the polypeptide chain as a non-trivial variation of the α/β -barrel, naming it accordingly an α/α -barrel. Juy *et al.* (1992) have observed a similar fold of the polypeptide chain for the catalytic domain of cellulase-D from *Clostridium thermocellum*, calling it a twisted α -barrel.

The initial crystallographic study of glucoamylase had been done at pH 6, which is approximately 0.2 pH units above the apparent pK_a (5.8) of an essential catalytic acid of the enzyme (Hiromi *et al.*, 1966; Savel'ev & Firsov, 1982). In order to explore the possible role of pH on the conformation of the enzyme, we have determined the structure of glucoamylase-II(471) at pH 4, near the pH of optimum catalysis. We present here complete models of the protein, glycosyl chains and solvent structure of crystalline glucoamylase-II(471) at pH 4 and 6.

2. Material and Methods

(a) Enzyme purification and crystallization

Glucoamylase-II(471) was prepared from *A. awamori* var. *X100* as described by Neustroev & Firsov (1990). Conditions for the growth of the fungus were adjusted so

that the glucoamylase I ($M_r=100,000$ by SDS/polyacrylamide gel electrophoresis), produced initially, was reduced in size by the action of fungal acid proteases to a mixture of glucoamylases of average $M_r=72,000$. This mixture of proteins migrated as 2 partially resolved bands on SDS/polyacrylamide gel electrophoresis.

The apparent heterogeneity of the protein did not prevent the reproducible growth of crystals of glucoamylase-II(471) by vapor diffusion. A solution of 15 mg/ml protein, 50 mM potassium phosphate (pH 5.95), and 13% (w/v) polyethylene glycol 6000 (Fluka) was equilibrated against a solution of 50 mM phosphate (pH 5.95), 30% (w/v) polyethylene glycol 6000. Crystals of space group $P2_12_12_1$ appeared in 1 to 2 weeks and grew slowly to a size of 0.2 to 0.3 mm. These crystals of glucoamylase-II(471) are stable in 50 mM potassium phosphate (pH 5.95), 28% (w/v) polyethylene glycol 6000. Crystalline specimens used in the structure determination at pH 6 grew from enzyme treated by an α -mannosidase as described by Golubev *et al.* (1992). However, the enzyme used in the growth of crystals for the pH 4 study had not been treated by α -mannosidase. As described below, no significant difference was observed in the structure of the glycosyl chains, as a consequence of treatment of the enzyme with α -mannosidase. We observe, however, that protein exposed to α -mannosidase crystallizes in approximately 1 week, whereas enzyme not treated by α -mannosidase crystallizes only after 3 weeks. Crystals at pH 4 were produced by soaking crystals, grown at pH 6, for 2–3 days in a solution containing 28% (w/v) polyethylene glycol 6000 and 100 mM sodium acetate (pH 4).

(b) Data collection

Data collection for the pH 6 structure was carried out in the laboratory of Dr Arthur Arnone, Department of Biochemistry, University of Iowa, Iowa City, IA on a Hamlin area detector and at Siemens Analytical X-Ray Instruments, Inc., Madison, WI on a Siemens detector. As described by Aleshin *et al.* (1992), the data from 2 crystals at pH 6 were merged and averaged together, using weights of $(1/\sigma(F))^2$. Data collection for the structure at pH 4 was done on a Siemens area detector at Iowa State University, Ames, IA. The refined unit cell parameters for the crystals at pH 6 and 4 were nearly identical. For the pH 6 crystals $a=116.7$, $b=104.3$, $c=48.49$ and for the pH 4 crystal $a=116.8$, $b=104.4$, $c=48.55$. Statistics relevant to the data collection are in Table 1.

(c) Refinement

The starting model for both the pH 6 and the pH 4 structures was that of Aleshin *et al.* (1992). That initial model did not include an interpretation of the solvent structure, a description of side-chains in alternative conformations or a complete interpretation of electron density associated with the *N*-linked glycosyl chains. A Silicon Graphics 4D25 and the program TOM (Cambillau & Horjales, 1987) were used for additional model building against electron density maps based on the Fourier coefficients $(|F_{obs}| - |F_{calc}|) \exp\{i\alpha_{calc}\}$ and $(2|F_{obs}| - |F_{calc}|) \exp\{i\alpha_{calc}\}$.

The models for crystals of glucoamylase-II(471) at pH 4 and 6 were refined using restrained least squares (Hendrickson & Konnert, 1980). Target values for the standard deviations of specific categories of stereochemical restraints appear in Table 2. The high-mannose

Table 1
Data collection statistics

Data set	Number of measurements	Unique reflections	Unique reflections with $F > 2\sigma(F)$	Total possible reflections	R_{sym}^c	Resolution (Å)
Native, pH 6 ^a	233,315	34,496	33,871	35,431	0.044	2.1
Native, pH 4 ^b	57,478	20,905	20,213	24,005	0.035	2.4

^a Data collected on a Hamlin area detector in the laboratory of Dr A. Arnone, Dept of Biochemistry, University of Iowa, Iowa City, IA and averaged, using weights of $(1/\sigma(F))^2$, with data collected on a Siemens area detector at Siemens Analytical Instruments, Inc. Madison, WI.

^b Data collected on a Siemens area detector at Iowa State University, Ames IA.

^c $R_{\text{sym}} = \sum_j \sum_i |I_{i,j} - \langle I_j \rangle| / \sum_j \sum_i I_{i,j}$, where the index i runs over multiple observations of the same intensity and the index j runs over all crystallographically unique intensities.

glycosyl chain (Einsphar, 1989) was used as the basis for modeling the *N*-glycosyl chains of glucoamylase, whereas the density of *O*-glycosyl residues was represented by α -D-mannose, in accordance with the findings of Gunnarsson *et al.* (1984). Standard groups for α -D-mannose and *N*-acetyl-D-glucosamine were based on the coordinates of monosaccharides obtained by X-ray diffraction studies (Chu & Jeffrey, 1968; Longchambon *et al.*, 1976). The glycosidic bond angles were set initially to 115°. Standard groups were developed separately for serine-mannose, threonine-mannose and asparagine-glycosyl chains. Glycosyl residues were restrained to the ⁴C₁-chair conformation by specification of all 2, 3 and 4-center distance restraints, except for those 4-center distance restraints involving the O6 atom of each glycosyl residue or atoms of the *N*-acetyl group of *N*-acetyl-D-glucosamine. All

4-center distance restraints which influence the dihedral angles associated with the oxygen atoms of the glycosidic bonds were omitted also. In addition to the distance restraints, the chiral volumes were restrained for atoms C1, C2, C3, C4 and C5 of each residue in all the glycosyl chains. The *N*-acetyl group of *N*-acetyl-D-glucosamine residues was subject to planar restraints. Restraints on all possible non-bonded contacts between atoms of different residues of the *N*-glycosyl chains were in force during refinement as well as restraints involving non-bonded contacts between the O6 atom and atoms O5 and C4 of the same residue.

Refinement proceeded over the course of 65 cycles for the pH 6 structure and 46 cycles for the pH 4 structure. Adjustments were made in the coordinates (x , y and z) and isotropic thermal parameters (B) for each atom of the

Table 2
Agreement between target and observed stereochemistry for native pH 4 and pH 6 structures of glucoamylase^a

Structure	pH 4	pH 6
Total number of atoms	4404	4517
Total number of solvent sites	561	673
Number of structure factors used in refinement	20,407 ^b	30,205 ^c
<i>R</i> -factor	0.122 ^b	0.141 ^c
Mean <i>B</i> (Å ²)		
Protein and glycans	10.3	9.75
Protein, main-chain	8.9	8.3
Protein, side-chain (including glycans)	11.7	11.2
<i>O</i> -glycans	25.9	22.6
<i>N</i> -glycans	29.1	28.4
R.m.s. ^d deviation (in Å) for distances		
Two-center	0.012 (0.02)	0.014 (0.02)
Three-center	0.028 (0.03)	0.029 (0.03)
Four-center	0.038 (0.05)	0.039 (0.05)
R.m.s. deviation (in Å) for planar groups	0.011 (0.02)	0.012 (0.02)
R.m.s. deviation (in Å ³) for chiral volumes	0.138 (0.15)	0.138 (0.15)
R.m.s. deviation (in Å) for non-bonded contacts		
Contacts determined by single torsion angles	0.222 (0.25)	0.221 (0.25)
Contacts determined by multiple torsion angles	0.152 (0.25)	0.155 (0.25)
Contacts determined by hydrogen bond	0.182 (0.25)	0.192 (0.25)
R.m.s. deviation (in degrees) for torsion angles		
Planar peptides ($\omega = 180$)	2.1 (3)	2.3 (3)
Staggered side-chains	14.0 (15)	13.6 (15)
Orthonormal	28.1 (20)	27.6 (20)
R.m.s. ^d ΔB (Å ²)		
Main-chain bonded	0.590 (1.0)	0.628 (1.0)
Main-chain angle	1.004 (1.5)	1.018 (1.5)
Side-chain bonded	1.842 (2.0)	1.898 (2.0)
Side-chain angle	2.781 (3.0)	2.739 (3.0)

^a Target root-mean-square deviations given in parentheses.

^b Comprises all observed moduli from 10 to 2.2 Å resolution, with $|F| \geq \sigma(|F|)$ and $|F| > 0.0$.

^c Comprises all observed moduli from 10 to 2.4 Å resolution, with $|F| \geq \sigma(|F|)$ and $|F| > 0.0$.

^d R.m.s., root-mean-squared.

model. Sites for water molecules were added to the models provided that (1) electron density was present at a level of at least 3σ in maps based on Fourier coefficients $(|F_{\text{obs}}| - |F_{\text{calc}}|) \exp \{i\alpha_{\text{calc}}\}$ and $(2|F_{\text{obs}}| - |F_{\text{calc}}|) \exp \{i\alpha_{\text{calc}}\}$; (2) at least one hydrogen bond with an associated donor-acceptor distance of 2.4 to 3.2 Å could be made between the new site and an existing atom of the model; (3) no close non-bonded contacts (less than 3.0 Å) could be made between the new site and existing atoms of the model and (4) the appearance of the electron density was approximately spherical. As a final criterion, sites for waters were omitted if their associated B -parameters exceeded 75 Å^2 after refinement. Occupancy parameters for all waters sites were not refined, because of the high correlation between occupancy and thermal parameters for data below 2.0 Å resolution. Thus, solvent sites with B values between 50 Å^2 and 75 Å^2 probably represent water molecules with occupancy parameters below 1.0 and true thermal parameters substantially lower than those reported from the refinement.

Side-chains were screened for multiple conformations on the basis of the following criteria: (1) electron density must be present at a level of at least 3σ in maps based on Fourier coefficients $(|F_{\text{obs}}| - |F_{\text{calc}}|) \exp \{i\alpha_{\text{calc}}\}$ and $(2|F_{\text{obs}}| - |F_{\text{calc}}|) \exp \{i\alpha_{\text{calc}}\}$; (2) alternative conformers must have reasonable contacts with nearby atoms; (3) each alternative must adopt or be close to a conformation with staggered χ angles. The presence of an alternative conformer for a side-chain was confirmed by inspection of "omit" electron density maps.

3. Results and Discussion

(a) Secondary structure

The models for glucoamylase-II(471) at pH 6 and 4 have been deposited with the Protein Data Bank, Brookhaven National Laboratory and are available on request. The accession number is not available at the time of publication. In keeping with prior work (Aleshin *et al.*, 1992), we have adopted the amino acid numbering of glucoamylase from *A. niger* (Svensson *et al.*, 1983; Boel *et al.*, 1984). Thus, the sequence numbers skip from 101 to 103 in the deposited coordinate sets. The sequence used here is identical to that used by Aleshin *et al.* (1992). Statistics relevant to the refined models appear in Table 2. An analysis due to Luzzati (1952) infers an

overall root-mean-square uncertainty in atomic coordinates of between 0.1 and 0.2 Å.

The secondary structures of the two models are identical to within experimental uncertainty and are represented in Figure 1. Helical regions are residues 1 to 21, 53 to 68, 72 to 90, 126 to 145, 148 to 169, 186 to 206, 211 to 227, 246 to 255, 272 to 287, 318 to 338, 345 to 354, 368 to 391 and 416 to 430. The helices are defined, for the most part, by a series of at least four residues, the ϕ and ψ angles of which are within $\pm 30^\circ$ of values typical for an α -helix ($\phi = -64^\circ$, $\psi = -40^\circ$). However, we also arbitrarily include residues that deviate from the criteria above, provided they are embedded within a helical span or belong to the C terminus of a helix. The final turn of helices in glucoamylase is often distorted toward a 3_{10} -conformation. Residues belonging to the distorted C-terminal ends of helices, if not included as part of the helix, would be classified as reverse turns. As a consequence, the assignment of some helical regions are longer here, than reported by Aleshin *et al.* (1992).

In Table 3 we provide a complete list of interaxial angles of all helices of glucoamylase-II(471). The inner set of helices (labeled H2, H4, H6, H8, H10 and H13 in Figure 1) are approximately parallel to each other, having an average interaxial angle between contact neighbors of 49.5° . The outer helices taken as a group (labeled H1, H3, H5, H7, H9 and H12 in Figure 1) are approximately antiparallel to the set of inner helices. The average interaxial angle between neighboring helices in sequence (excluding helix H11, which is not part of the α/α -barrel) is 144° . The axes of the peripheral set of helices are essentially parallel to each other, making an average angle of 28° to their packing neighbors. The average packing angles listed above, are near one of the preferred values described by Chothia *et al.* (1978).

The assignment of β -sheet structure by Aleshin *et al.* (1992) remains unchanged as a consequence of the additional refinement. Most of the sheet structures in glucoamylase-II(471) are two-stranded antiparallel loops: (1) residues 21 to 23 and 35 to 38, (2) residues 91 to 93 and 96–97, (3) residues 173 to

Table 3
Interaxial angles in degrees of helices of glucoamylase-II(471)

	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H12	H13	H11 ^a
H1	—	116	46	95	96	68	136	89	94	116	54	137	118
H2		—	150	21	144	50	107	32	119	42	119	37	74
H3			—	132	51	102	97	118	88	117	73	134	133
H4				—	157	30	128	17	127	48	109	52	85
H5					—	142	48	144	72	109	93	108	106
H6						—	155	20	142	54	105	69	112
H7							—	135	55	102	97	85	64
H8								—	144	39	118	50	103
H9									—	149	43	128	45
H10										—	157	21	109
H12											—	153	67
H13												—	90

^a Helix 11 is not part of the α/α -barrel.

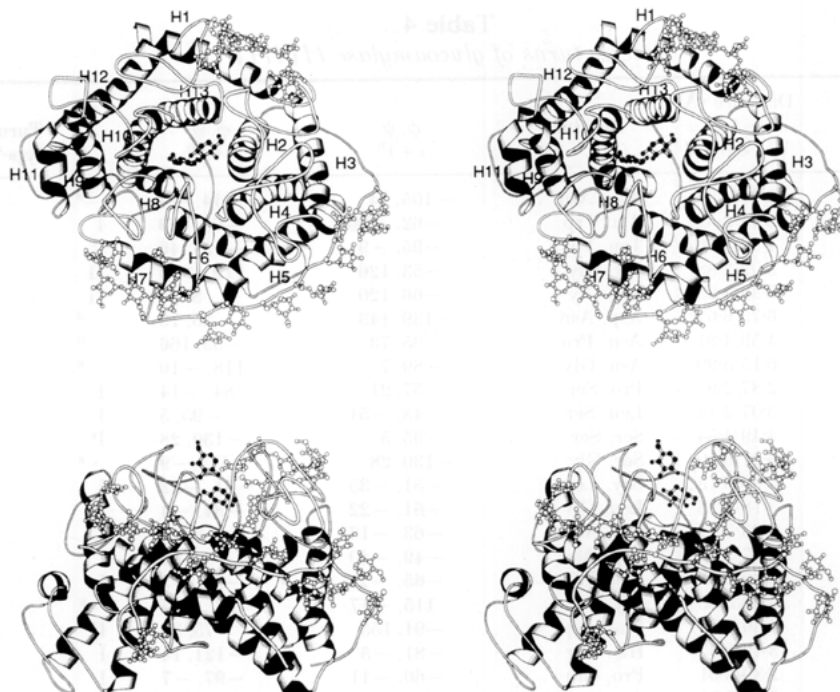


Figure 1. Schematic of glucoamylase-II(471) as viewed (top) into the packing void of the α/α -barrel and (bottom) perpendicular to the barrel. Atoms of the glycosyl chains are represented explicitly by open circles, whereas atoms of 2 molecules of bound 1-deoxynojirimycin, as observed by Harris *et al.* (1993), are represented by filled circles. The illustrations were produced by the software MOLSCRIPT (Kraulis, 1991).

176 and 181 to 183, (4) residues 184–185 and 235 to 237, (5) residues 262 to 264 and 267–268, (6) residues 339 to 344 and 360 to 365 (7) residues 392–393 and 396–397 and (8) residues 398 to 403 and 407 to 415. The only exception to the above is a three-stranded antiparallel β -sheet comprised of residues 49 to 52, 107 to 110 and 113 to 116.

We adopted the criteria of Lewis *et al.* (1973) in assigning reverse turns of glucoamylase-II(471). The reverse turns listed in Table 4 consist of a sequence of four residues, for which the α -carbon of the first residue is within 7.0 Å of the α -carbon of the fourth residue, and for which the second and third residues do not belong to an α -helix. Reverse turns comprise 32% of all residues of the enzyme. By far, the type I turn is most prevalent, occurring in 28 out of 41 instances. Approximately one-half (22 out of 41) of the turns listed in Table 4 have hydrogen bonds between carbonyl i and amide $i+3$, which is consistent with trends observed for other proteins (Crawford *et al.*, 1973; Rose *et al.*, 1985; Wilmot & Thornton, 1989). Sites of *N*-glycosylation in glucoamylase are at the third position of type I turns, with a hydrogen bond between the carbonyl of residue i and the amide of residue $i+3$.

Disulfide linkages exist between the following pairs of cysteines: (1) 222 and 469, (2) 210 and 213 and (3) 262 and 270. The dihedral angle of the disulfide linkage between residue 210 and 213 is right-handed, having an average value of 69° for the pH 4 and 6 structures. The dihedral angles for the linkages between residues 262 and 270 and residues 222 and 469 are left-handed, averaging to -96° and -83° , respectively. The dihedral angles

observed for disulfide linkages in glucoamylase-II(471) are near the typical values of -90° and $+90^\circ$ (Richardson, 1981). Distances between the α -carbons of the cysteines involved in disulfide linkages range from 4.83 Å to 5.45 Å. These distances are again typical for pairs of disulfide-linked cysteines (Richardson, 1981).

In addition to the above elements of secondary structure, other noteworthy features are in the refined models. We observe *cis*-peptide linkages between Gly23-Ala24, Asn45-Pro46 and Arg122-Pro123. In addition, Pro157 and Pro218 reside in helices H5 and H7, respectively, imparting a kink or bend to the axis of each helix. The plot of ϕ and ψ angles (Ramachandran *et al.*, 1963) for the pH 4 and 6 structures (not shown) reveals two residues (Asn313 and Ser411) that lie significantly beyond the allowed regions of the plot. Asn313 and Ser411 are maintained by their local hydrogen bonding environments in conformations which bring non-bonded atoms into close contact. Aleshin *et al.* (1992) have provided the details of the hydrogen bond interactions responsible for the conformations of Asn313 and Ser 411.

(b) Domain organization and solvent structure

Glucoamylase-II(471) consists of two domains (Figure 1). An α/α -barrel, described by Aleshin *et al.* (1992), comprises the catalytic domain (residues 1 to 440). As noted previously (Aleshin *et al.*, 1992; Harris *et al.*, 1993), the active site of the enzyme lies in the packing void of the inner ring of helices (Figure 1). The active site of glucoamylase lies at

Table 4
Reverse turns of glucoamylase-II(471)^a

<i>i</i>	Distance (Å) O _i -N _{i+3} pH 4/pH 6	Residues <i>i</i> + 1, <i>i</i> + 2	ϕ, ψ <i>i</i> + 1 ^b	ϕ, ψ <i>i</i> + 2 ^b	Turn type ^c
23	5.27/5.29	Asn, Asn	-105, 14	-114, -4	— ^d
26	3.41/3.52	Ala, Trp	-62, -30	-95, -9	I
27	4.80/4.86	Trp, Val	-95, -9	-127, 46	I ^e
29	3.02/2.88	Ser, Gly	-53, 126	119, -25	II
33	3.33/3.29	Ser, Gly	-66, 120	87, 5	II
43	6.75/6.67	Asp, Asn	-139, 143	35, 73	— ^d
44	4.59/4.81	Asn, Pro	35, 73	-93, 166	— ^d
68	6.15/5.99	Asn, Gly	-89, 7	118, -10	— ^g
93	2.87/2.91	Pro, Ser	-57, 21	-84, -14	I
97	3.07/3.14	Leu, Ser	-48, -51	-95, 5	I
98	4.49/4.54	Ser, Ser	-95, 5	-130, 28	I ^e
99	6.37/6.50	Ser, Gly	-130, 28	113, -9	— ^g
101	3.00/3.13	Gly, Leu	-51, -35	-61, -22	I
103	3.18/3.23	Leu, Gly	-61, -22	-84, -4	I
110	3.03/3.01	Val, Asp	-63, -16	-91, 6	I
169	3.07/3.18	Trp, Asn ^f	-49, -40	-84, 25	I
176	3.00/2.90	Leu, Trp	-65, -23	-86, -3	I
229	5.82/5.80	Gly, Ser	115, -17	-127, -35	— ^g
236	4.30/4.39	Phe, Asp	-91, 155	75, 31	II ^e
253	3.49/3.37	His, Thr	-81, -5	-121, 14	I
257	2.89/3.01	Pro, Glu	-60, -11	-97, -7	I
263	3.58/3.71	Asp, Ser	-73, -33	-64, -32	I
264	3.22/3.15	Ser, Thr	-64, -32	-106, 5	I
268	2.74/2.93	Pro, Cys	-50, -39	-91, -4	I
286	4.62/4.55	Ser, Ile	-102, 7	-107, -10	I ^e
289	2.80/2.73	Thr, Leu	-55, -32	-62, -23	I
290	3.14/3.13	Leu, Asn	-62, -23	-95, -3	I
292	3.33/3.20	Asp, Gly	-55, 130	89, -16	II
296	2.95/2.93	Asp, Ser	-64, -8	-103, 3	I
306	3.13/3.16	Pro, Glu	-60, -22	-89, -5	I
310	2.66/2.78	Tyr, Tyr	-57, 115	59, 24	II
311	2.76/2.74	Tyr, Asn	59, 24	96, -6	I ^e
353	5.77/5.49	Leu, Tyr	-102, -30	-139, 95	VIII
355	2.89/2.86	Ser, Gly	-57, -20	-90, 17	I
364	3.11/3.16	Ser, Ser	-70, -11	-98, -4	I
393	2.73/2.84	Ser, Asn ^f	-59, -31	-91, 5	I
403	3.79/3.76	Lys, Ser	-62, -20	-111, -48	I ^e
404	5.67/5.74	Ser, Asp	-111, -48	-110, -6	I ^e
412	4.36/4.41	Arg, Asp	-82, 159	39, 58	II ^e
438	3.03/3.12	Glu, Thr	-56, -32	-66, -16	I
439	2.94/2.98	Thr, Ser	-66, -16	-116, 17	I
440	5.14/5.11	Ser, Ala	-116, 17	-144, 43	I ^d or VIII ^d
446	4.95/5.07	Gly, Thr	-63, -33	-136, 129	VIII

^a If the distance CA_{*i*} to CA_{*i*+3} is less than 7.0 Å and if residues *i* + 1 and *i* + 2 do not belong to an α -helix, then residues *i* to *i* + 3 are a reverse turn.

^b ϕ and ψ angles in degrees are averages of angles from the pH 4 and 6 structures.

^c Listed angles for residues *i* + 1 and *i* + 2 are usually within $\pm 30^\circ$ of ideal angles for the corresponding type of turn with perhaps at most one of the ϕ, ψ angles deviating by as much as $\pm 45^\circ$. Ideal values are from Wilmot & Thornton (1989).

^d Turn involves a *cis*-peptide linkage. No classification was made.

^e Two or more of the ϕ, ψ angles deviate by more than $\pm 45^\circ$ from the ideal values used by Wilmot & Thornton (1989).

^f Site of glycosylation.

^g At least one angle deviates by more than 90° from any set of ideal values.

the opposite end of its barrel in comparison to other amylolytic enzymes which have an α/β -barrel for their catalytic domains (Aleshin *et al.*, 1992). The catalytic domain also has two sites of *N*-glycosylation, described more completely below. The second domain of glucoamylase-II(471) consists of the last 30 residues of the protein and represents the N-terminal half of the *O*-glycosylated domain. The second domain of glucoamylase-II(471) bears a total of ten sites of *O*-glycosylation, each glycosyl chain being a single residue of mannose.

One of the striking features of the α/α -barrel is the packing void created by the inner set of helices. The centre of the "void" is filled by the side-chains of leucines 58, 130, 177 and 319, tryptophans 178 and 417, and Phe187. However, on both sides of this "hydrophobic plug" are large pockets of water (Figure 2). One pocket in the vicinity of Leu58 contains at least a dozen water molecules. The other pocket is the deepest recess of the active site, which contains at least seven ordered water molecules. Interestingly, the glucose transport protein, which

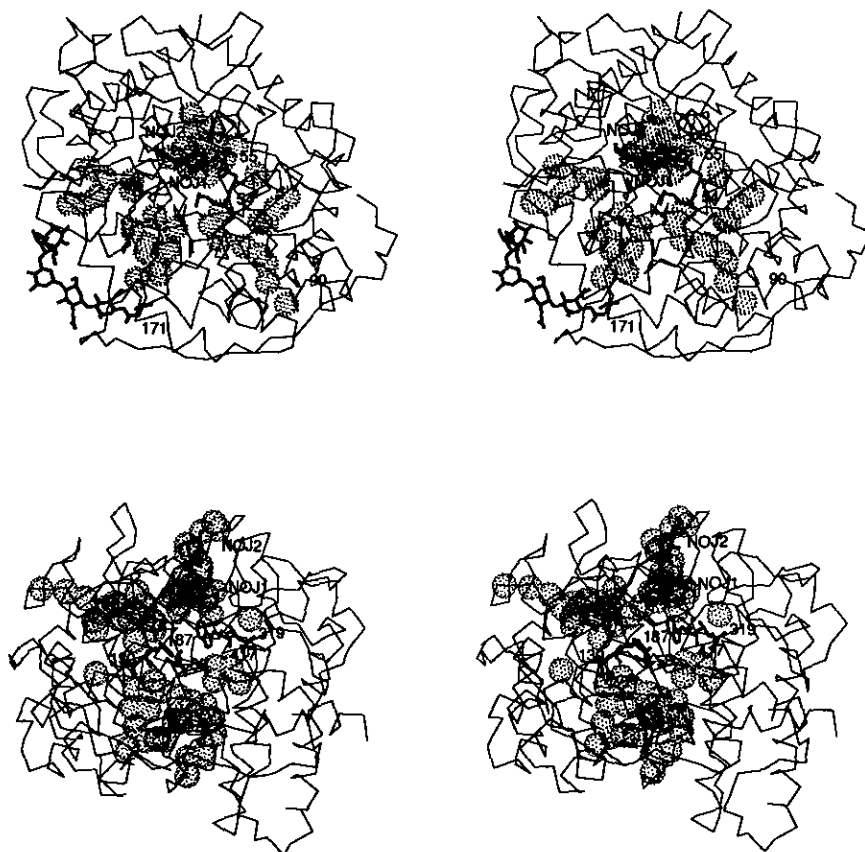


Figure 2. Stereoview revealing the water molecules that lie within the radius of gyration of glucoamylase. The dots represent a surface of radius 1.5 Å centered about each water molecule. The top view is into the active site, showing the 3 clusters of water molecules that extend laterally from the active site. The bottom view is perpendicular to the inner set of helices that define the packing void. Explicitly shown in bold lines are some of the residues of the hydrophobic barrier that separates the 2 water-filled ends of the packing void. Also shown in bold lines are the 2 molecules of 1-deoxynojirimycin (NOJ1 and NOJ2) bound to the active site as revealed by the study of Harris *et al.* (1993).

mediates the transport of glucose across cell membranes, is a membrane spanning protein, comprised putatively of 12 α -helices (Mueckler *et al.*, 1985; Gould & Bell, 1990; Nikaido & Saier, 1992). Although we observe little sequence similarity between glucoamylase and the glucose transport proteins, the organization of helices in glucoamylase may be relevant to the structure of glucose transport proteins. Many proteins which have diverse function and no discernible sequence similarity adopt the same polypeptide fold, as best illustrated by the class of proteins which have the α/β -barrel motif.

In addition to the packing voids located within the inner set of helices, three clusters of water molecules extend laterally from the active site (Figure 2). One of the clusters, in fact, is a bifurcated chain of water molecules beginning at two locations within the active site in the vicinity of Asp55 and Arg54, merging near Thr53 into a single chain, and ending in the vicinity of Gly90 at the surface of the enzyme. The remaining two clusters of water molecules are separated from the bulk solvent by the glycosyl chain of Asn171. Harris *et al.* (1993) have shown that the binding of 1-deoxynojirimycin displaces six of seven water molecules

from the deepest recess of the active site. The displaced water molecules may diffuse through one of the lateral passages, rather than slide by the substrate as it enters the "front-end" of the active site. The disorder observed for Asp55 (described below) suggests that the water molecules nearby can move to alternative sites.

(c) Thermal parameters and alternative conformations

A plot of average thermal parameters as a function of residue number reveals high values for residues at the N and C terminus (Figure 3). The high thermal parameters persist over the first three to four residues of the N terminus. As noted by Aleshin *et al.* (1992), the partial proteolysis of the first three residues from the N terminus may be largely responsible for the high thermal parameters. Svensson *et al.* (1982) discovered that 50% of the molecules of glucoamylase from *A. niger* lack the first three residues. The high *B* parameters of the last eight to ten residues of glucoamylase-II(471) may result from partial proteolysis as well. We observe a large jump in *B* parameters of the main-chain following 462. In fact, the thermal parameters

climb from 5 to 10 Å² to 25 to 30 Å² in the pH 6 structure and from 5 to 10 Å² to 35 to 40 Å² in the pH 4 structure over a span of two residues. The large increase in the *B* values of the last eight to ten residues for the pH 4 and 6 structures may be due to the existence in the crystal of a mixture of polypeptide chains that differ in length by eight to ten residues. Furthermore, the average *B* parameters for the N and C-terminal residues are approximately 10 Å² higher in the pH 4 structure than in the pH 6 structure, suggesting that the enzyme used in the pH 4 study suffered more proteolytic degradation than the enzyme used in the pH 6 study. The difference in *B*-values is real and not an artifact due to data sets of differing resolution. The average of all thermal parameters for atoms of the protein and glycosyl chains is essentially the same in the two structures (Table 2).

Other than the regions mentioned above, the thermal parameters in both the pH 4 and 6 structures are uniformly low, with the vast majority of values below 30 Å². However, atoms of residue G of glycosyl chain 171 rise to and exceed 50 Å² in both structures. Atoms of residue I of glycosyl chain 395 exceed 50 Å² in the pH 4 structure. *B*-values of the mannosyl chain of residue 464 are above 40 Å² in the pH 6 structure and above 50 Å² in the pH 4 structure.

Of the residues modeled in alternative conformations, the disorder associated with Asp55 may be the most significant. Harris *et al.* (1993) suggest that Asp55 interacts with the 4 and 6-hydroxyls of the non-reducing end of substrates. In both the pH 4 and 6 structures Asp55 in one position of its side-chain has interactions with waters 696, 641 and 912, whereas in the other position Asp55 interacts with waters 641 and 637. The first position corresponds closely to the conformation of Asp55 in the complex of 1-deoxynojirimycin with glucoamylase (Harris *et al.*, 1993). As mentioned above, the disorder in Asp55 may reflect a mechanism whereby water molecules of the active site enter the solvent channel nearby and migrate to the surface of the enzyme.

(d) Crystal packing

Contacts between enzyme molecules in the crystal give rise to three protein/protein interfaces (Table 5). The predominant interface (Interface 3 of Table 5) involves extensively the *O*-glycosylated residues, Ser-mannoses 453, 455 and 459 and Thr-mannose 457. The participation of the *O*-glycosylated domain in a major lattice contact of the crystal infers that a homogeneous composition of the glycosyl chains for the residues above is an absolute requirement for the growth of this crystal form. Interface 1 of Table 5 is less extensive than interface 3, but nonetheless the interactions here are significant as they approach the outer regions of the active site. The presence of interface 1 may influence the binding of analogs of extended substrates to this crystal form.

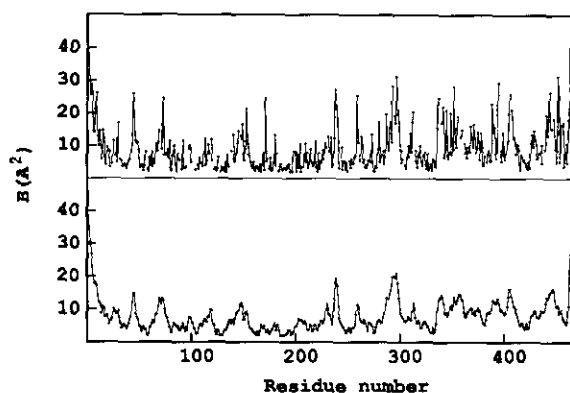


Figure 3. Averaged thermal parameters of the side-chains (top) and the main-chain (bottom) as a function of residue number for the pH 6 structure.

The superposition of the pH 4 and 6 structures, using the protocol of Honzatko (1986) revealed no significant differences in the conformation of the protein and attached glycosyl chains. The overall root-mean-square difference in the relative positions of α -carbon atoms is 0.11 Å, consistent with estimated uncertainty (0.1 to 0.2 Å) in coordinates obtained from the analysis due to Luzzati (1952). The highest root-mean square difference (2.6 Å) in the positions of atoms is for the side-chain of Thr-mannose 464 and probably reflects the true uncertainty in the position of that side-chain.

(e) Comparison of pH 4 and 6 structures

The absence of major conformational differences in the vicinity of the active site was confirmed by the calculation of a difference distance matrix (Liebman, 1982), employing 10 water molecules (waters 500, 646, 696, 815, 881, 894, 910, 911, 912 and 924) and 19 protein atoms (carbonyl oxygen 47, Tyr48 OH, Trp52 NE1, Arg54 NH2, Asp55 OD1 and OD2, Asp55 (alternative conformer) OD1 and OD2, Lys108 NZ, Trp120 NE1, Gln124 NE2, carbonyl oxygens 177 and 178, Glu179 OE1 and OE2, Glu180 OE2, Arg305 NH1, Tyr311 OH and Glu400 OE1). Nearly all possible distances between the selected protein atoms deviated by no more than 0.25 Å in the pH 6 structure relative to the pH 4 structure. The only exceptions were atoms OD1 and OD2 of the alternative conformer of Asp55 and NZ of Lys108. Coordinates for the atoms of the side-chain of Asp55 are subject to high uncertainty (0.2 Å), because they are associated with weak electron density. Lys108, however, is consistently closer to other residues of the active site in the pH 4 structure relative to the pH 6 structure, with the largest change in distance being a 0.48 Å decrease between NZ of Lys108 and OE2 of Glu180. The relative positions of the ten water molecules, listed above, differ generally by no more than 0.5 Å. The exception is water 881 which shows differences as high as 0.67 Å. However, water 881 is the only one of the ten listed, that does not hydrogen bond directly to the protein. Thus, a higher level of

Table 5

Direct and indirect polar contacts between atoms of glucoamylase-11(471) molecules in the crystal^a

First group	Bridging water ^b	Second group
Interface no. 1 ^c		
Ser356 OG		Thr-mannose 464 O
Thr360 O	water 1001	Asp25 OD1
Thr360 OG1		Asp25 OD1
Thr362 N		Ala24 O
Thr362 O	water 628	Ala32 O
Thr362 O	water 628	Gly31 N
Ser364 N	water 683	Gly31 O
Ser364 OG	water 1002	Gly31 O
Ser 366 O	water 745	Asp112 O
Ser366 OG		Asp112 OD1
Ser368 N		Asp33 OD1
Ser368 N	water 745	Asp112 O
Ser368 OG	water 745	Asp112 O
Ser368 OG		Glu113 OE2
Thr369 N		Asp33 OD1
Thr369 OG1	-	Asp33 OD2
Interface no. 2 ^d		
Asp144 O		Tyr312 N
Asp144 OD1	water 969	Asn313 OD1
Asn145 O	water 961	Tyr312 OH
Asn145 ND2	water 1003	Asn313 OD1
Asn145 ND2	water 1003	Asn313 ND2
Tyr147 OH		Asn45 OD1
Ser468 O	water 1153	Asn45 OD1 ^f
Interface no. 3 ^e		
Gly207 O	water 674	Ser-mannose 455 O2
Gly207 O	water 674	Ser99 O
Ser208 OG	water 799	Ser100 O
Ser208 OG	water 1262	Tyr168 OH ^f
Ser209 OG	water 651	Pro94 O
Ser209 OG		Ser100 O ^f
Ser211 N	water 651	Pro94 O
Ser211 OG		Pro94 O
Arg160 NH1	water 1006	Gln172 NE2
Arg160 NH2	water 1006	Gln172 NE2
Ser-mannose 453 O2	water 825	Ser119 OG
Ser-mannose 453 O3	water 825	Ser119 OG
Ser-mannose 453 O3		Ser119 OG
Ser-mannose 453 O5	water 825	Ser119 OG
Ser-mannose 455 N	water 792	Ser119 OG
Ser-mannose 455 O	water 605	Trp120 O
Ser-mannose 455 O	water 605	Arg122 NH2
Ser-mannose 455 O4		Tyr175 OH
Ser-mannose 455 O4	water 687	Arg122 O
Gly456 O	water 687	Arg122 O
Thr-mannose 457 O3		Asn182 ND2
Thr-mannose 457 O3		Thr173 OG1
Thr-mannose 457 O4		Thr173 O
Thr-mannose 457 O4		Thr173 OG1 ^f
Thr-mannose 457 O4	water 1006	Gln172 NE2
Thr-mannose 457 O6		Arg125 NH2 ^g
Thr-mannose 457 O6	water 1005	Arg125 NH1
Ser-mannose 459 O2		Thr173 N
Ser-mannose 459 O2		Gln172 NE2
Ser-mannose 459 O3	water 1004	Asn-glycosyl 171 O
Ser-mannose 459 O6		Gln172 NE2
Ser-mannose 459 O6		Gln172 NE2
Ser-mannose 459 O6	water 1007	Gln172 OE1

^a All contacts listed are less than 3.2 Å in at least 1 of the 2 structures. If a contact less than 3.2 Å is present in one structure, then it is absent in the other structure if the distance exceeds 3.4 Å.

^b If no bridging water is listed, then the atom of the left-hand column is in direct contact with the corresponding atom of the right-hand column.

^c Entries in the right-hand column come into contact with

uncertainty would again be expected for its coordinates.

The only significant difference in the vicinity of the active site between the pH 4 and 6 structures is the thermal parameter for water 500, which is the putative catalytic water in the enzymic hydrolysis of maltooligosaccharides (Harris *et al.*, 1993). In the pH 6 structure water 500 has a thermal parameter of 15.7 Å². It makes a strong hydrogen to atom OE1 of Glu400 (donor-acceptor distance of 2.73 Å). Water 500 in the pH 4 structure has a slightly lower donor acceptor distance (2.62 Å) to atom OE1 of Glu400, but a much lower thermal parameter of 9.2 Å². Atom OE1 of Glu400 has essentially identical thermal parameters for both structures, 7.9 Å² in the pH 6 structure, and 7.0 Å² in the pH 4 structure. The data suggest that water 500 is associated more tightly with Glu400 in the pH 4 structure than in the pH 6 structure. The observation above with respect to the thermal values of water 500 defies a simple explanation, as presumably a significant increase in the fraction of protonated Glu400 would weaken its interaction with water 500. The tighter association water 500 with the enzyme at pH 4, the pH of optimum activity, is consistent, however, with the putative role of Glu400 in the removal of a proton from water 500 during catalysis (Harris *et al.*, 1993).

The absence of observable conformational differences between the pH 4 and 6 structures in the neighborhood of Glu179 is contrary to expectation, if Glu179 is indeed the catalytic acid of glucoamylase with a pK_a of 5.8. Glu179 is essential to catalysis (Sierks *et al.*, 1990) and it occupies a position with respect to 1-deoxynojirimycin that is consistent with its putative role as a catalytic acid (Sierks *et al.*, 1990; Harris *et al.*, 1993). Yet the environment of Glu179 is not what one would expect for an acid residue with an anomalous pK_a of 5.8 (Figure 4). Indeed, the closest ionized residue to Glu179 is Arg54. Atom OE1 of Glu179 is only 3.8 Å away from atom NH2 of Arg54. Furthermore, Lys108 is separated from Glu179 by a low dielectric barrier comprised of the side-chains of Trp52 and Trp120. The nearest electrostatic charges, then, should depress the pK_a of Glu179. In addition, the

those listed in the left-hand column after applying the operator $x+1/2$, $-y+1/2$, and $-z+1$ to their fractional coordinates. Entries in the left-hand column come into contact with those listed in the right-hand column after applying $x+1/2$, $-y+1/2$ and $-z+1$ to their fractional coordinates.

^d Entries in the right-hand column come into contact with those listed in the left-hand column after applying the operator x , y and $z+1$ to their fractional coordinates. Entries in the left-hand column come into contact with those listed in the right-hand column after applying x , y and $z-1$ to their fractional coordinates.

^e Entries in the right-hand column come into contact with those listed in the left-hand column after applying the operator $-x+1/2$, $-y$ and $z+1/2$ to their fractional coordinates. Entries in the left-hand column come into contact with those listed in the right-hand column after applying $-x+1/2$, $-y$ and $-z-1/2$ to their fractional coordinates.

^f Present only in the pH 6 structure.

^g Present only in the pH 4 structure.

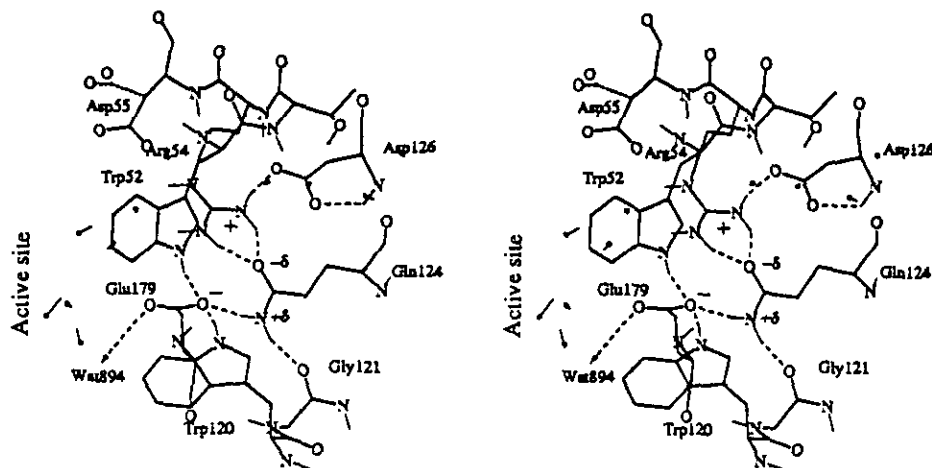


Figure 4. Stereo view of the environment of Glu179 of the pH 4 structure. Hydrogen bonds are presented as broken lines. The alternate conformation of Asp55 is given by dotted lines. Filled circles represent sites for water molecules in the pH 4 structure with lines indicating their displacement from equivalent sites of solvation in the pH 6 structure.

hydrogen bonds between atom OE1 of Glu179 and atoms NE2 of Gln124, NE1 of Trp52 and NE1 of Trp120 should polarize the carboxylate of Glu179 and thereby reduce the basicity of its OE2 atom. Thus, the crystallographic results infer an ionized state for Glu179 in the absence of substrate over the range of pH 4 to 6.

Even if Glu179 is ionized rather than protonated in the absence of substrate, we still cannot exclude it from acting as a catalytic acid. When substrate binds to the active site, the aqueous environment around atom OE2 of Glu179 (there are seven water molecules within 5 Å of atom OE2) will be replaced by a medium with a lower dielectric constant. The negative charges of Asp55, Glu180 and Glu400, which lie across the solvent void of the unoccupied active site, should elevate the pK_a of Glu179. In addition, a true substrate of glucoamylase could evoke a conformational change, which disrupts one or more of the hydrogen bonds involving atom OE1 of Glu179. The disruption of hydrogen bonds to atom OE1 of Glu179 does not occur in the 1-deoxynorjirimycin complex (Harris *et al.*, 1983) or in the complex of glucoamylase with acarbose (unpublished results). However, neither 1-deoxynorjirimycin, which lacks the glycosidic linkage altogether, nor acarbose, which has a positively charged imino linkage in place of the natural glycosidic linkage, are faithful analogs of true substrates. Clearly, more studies are necessary in order to confirm the role of Glu179 in the mechanism of glucoamylase.

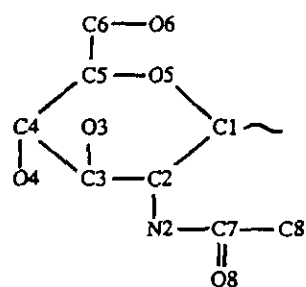
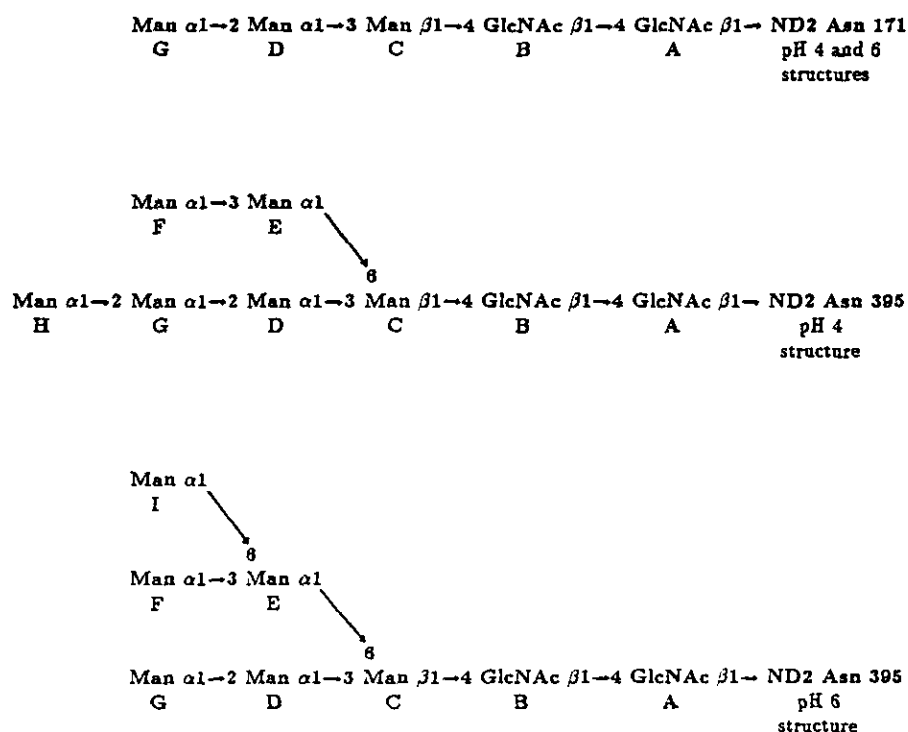
(f) N-linked glycosyl chains

Glucoamylase-II(471) has a total of 12 glycosyl chains, two of which are N-linked to asparagine residues and the remainder O-linked to serine or threonine. The N-linked glycosyl chains are the more complex of the two types of glycosylation. We observe electron density for five glycosyl residues in the chain linked to Asn171 and eight glycosyl residues in the chain linked to Asn395. The chemical

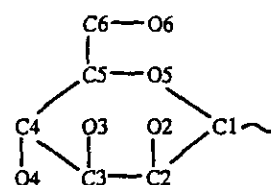
structures and nomenclature for residues and atoms of the N-glycosyl chains are presented in Figure 5. The two glycosyl chains are of the high-mannose type (Einsphar, 1989). Differences in the two chains occur after residue C (mannose). For the chain linked to Asn171, we observe "branching" only from the 3-hydroxyl position of residue C, whereas for the chain linked to Asn395, we observe branching from both the 3 and 6-hydroxyls of residue C. The 6-hydroxyl extension of Asn-glycosyl chain 171 may be present, but disordered, as the 6-hydroxyl group projects out into a solvent void of the crystal. Alternatively, that position may be glycosylated to only a limited extent. Weak density does appear in the vicinity of the 6-hydroxyl of residue C of glycosyl chain 171, but the density is too diffuse to allow an unambiguous interpretation.

The chemical composition of glycosyl chain 395 differs in the pH 4 and 6 structures. That segment of the glycosyl chain that branches from the 3-hydroxyl position of residue C has three α -1 $\rightarrow$ 3-linked mannosyl residues in the pH 4 structure, but only two in the pH 6 structure. (Residue H is absent in the pH 6 structure.) The glycosyl chain that branches from the 6-hydroxyl of residue C has two α -1 $\rightarrow$ 3-linked mannosyl residues in the pH 4 structure, but an additional mannosyl residue (residue I), α -1 $\rightarrow$ 6-linked to residue F, in the pH 6 structure. The additional residues in each of the two structures have weak density and have no interactions with the protein. The enzymes for the pH 4 and 6 structures come from separate preparations (see Material and Methods), this being the most likely explanation for the observed differences in the composition of the glycosyl chains at Asn395.

In spite of the differences in chemical composition, the conformation of the glycosyl chains at Asn171 and Asn395 are similar (Figure 6, Table 7). In particular residues A, B and C of both chains exhibit a significant hydrogen bonding pattern (Table 6) that fixes the relative orientations of these three sugars. The 3-hydroxyl of the first N-acetyl-D-glucosamine (residue A) hydrogen bonds to the



N-Acetylglucosamine (GlcNAc)



Mannose (Man)

Figure 5. Chemical structures of the *N*-glycosyl chains of glucoamylase, showing the convention used to label atoms of the mannosyl and *N*-acetylglucosamine residues.

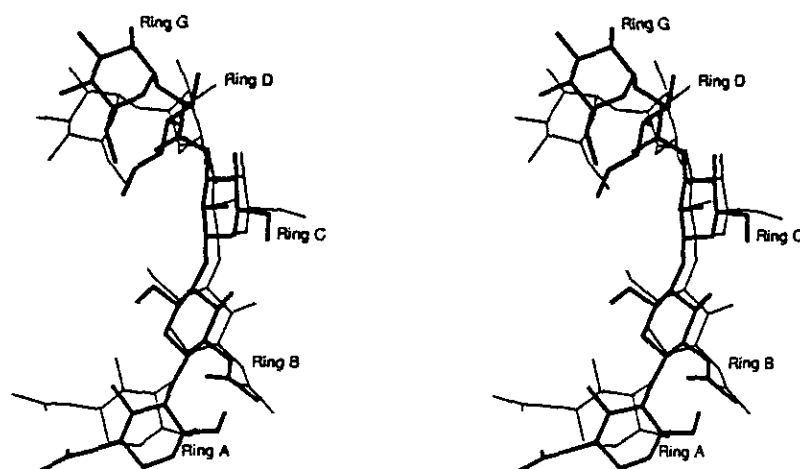


Figure 6. Superposition of the glycosyl chain of asparagine 171 (bold lines) onto the corresponding residues of the glycosyl chain of asparagine 395 of the pH 6 structure.

Table 6

Direct and indirect polar contacts involving atoms of the glycosyl chains of glucoamylase-II(471)^a

Glycosyl chain atom	Bridging water ^b	Contact residue and atom
Glycosyl chain of Asn171		
O3A [*]	—	Asn-glycosyl 171 O5B
O3A	water 663	Tyr223 OH
O4A [*]	water 923	Asn-glycosyl 171 O7B
O6A [*]	water 725	Asn-glycosyl 171 N2B
O6A	water 613	Gln219 O
O6A	water 613	Ala450 O
O7A	water 600	Trp170 O
O7A	water 600	Ser184 OG
O3B [*]	—	Asn-glycosyl 171 O5C
O4B [*]	water 972	Asn-glycosyl 171 O2C
O4B ^d	water 972	Ser226 OG
O2C [*]	—	Asn-glycosyl 171 O6D
O2C	water 972	Ser226 OG
O6D	—	Ser239 O
O3G	—	Ser239 O
O3G ^d	water 870	Arg241 O
O4G	—	Asp238 N
O4G ^d	water 1057	Asp238 OD1
Glycosyl chain of Asn395		
N2A	—	Ser397 OG
O3A [*]	—	Asn-glycosyl 395 O5B
O3A [*]	water 764	Asn-glycosyl 395 O6B
O3A ^d	water 850	Trp28 NE1
O3B [*]	—	Asn-glycosyl 385 O5C
O5B	water 713	Arg413 NE
O5B	water 713	Asp414 OD1
O6B ^d	water 850	Trp28 NE1
O7B [*]	—	Asn-glycosyl 395 O2F
O7B	water 684	Asp414 OD1
O4C [*]	water 760	Asn-glycosyl 395 O6G
O4C	water 760	Arg413 NH2
O4C	water 760	Thr43 OG1
O6D [*]	water 864	Asn-glycosyl 395 O6G
O6D	water 864	Trp28 O
O3E [*]	water 753	Asn-glycosyl 395 O4E
O3F [*]	water 746	Asn-glycosyl 395 O4F
O3F	water 746	Ser399 OG
O3F	water 746	Ser411 O
O3F	water 997	Glu409 O
O4F	—	Ser411 O
O4F	water 746	Ser399 OG
O4F	water 746	Ser411 O
O5F	—	Arg413 NH1
O6F	—	Ser42 OG
O3G	—	Ser30 O
O3G	water 808	Ser30 O
O4G	—	Ser30 O
O6G	—	Pro41 O
O6G	water 760	Thr43 OG1
O6G	water 760	Arg413 NH2
O6G	water 864	Trp28 O
Mannosyl chain of Ser443		
O2	water 685	Glu439 OE1
O2	water 685	Asp263 OD1
O2	water 685	Gly261 N
O2	—	Thr440 O
O6 ^c	water 909	Ser-mannose 444 O6 ^c
O6	water 909	Ser441 O
Mannosyl chain of Ser444		
O6 ^{c,e}	water 909	Ser-mannose 443 O6
O6 ^{c,e}	water 661	Ser441 O
O6 ^{c,e}	water 661	Ser-mannose 444 O

Table 6 (continued)

Glycosyl chain atom	Bridging water ^b	Contact residue and atom
Mannosyl chain of Thr452		
O3	water 839	Ser-mannose 453 O
Mannosyl chain of Ser453		
O2	water 869	Glu168 OE1
O2	water 825	Ser-mannose 455 O5
O3	water 825	Ser-mannose 455 O5
O3	—	Ser119 OG ^f
O6	water 943	Ala451 O
Mannosyl chain of Ser455		
O2	water 674	Gly207 O ^f
O2	water 674	Ser99 O
O3	water 632	Tyr458 N
O4	water 687	Gly456 O
O4	water 687	Arg122 O ^f
O4	—	Tyr175 OH ^f
O5	water 825	Ser-mannose 453 O3
O5	water 825	Ser119 OG ^f
Mannosyl chain of Thr457		
O3	—	Asn182 ND2 ^f
O3	—	Thr173 OG1 ^f
O4	water 1006	Arg160 NH1
O4	water 1006	Arg160 NH2
O4	water 1006	Gln172 NE2 ^f
O4	water 1006	Ser-mannose 459 O5
O4	—	Thr173 O ^f
O4 ^c	—	Thr173 OG1 ^f
O6	water 1005	Arg125 NH1 ^f
O6 ^d	—	Arg125 NH2 ^f
Mannosyl chain of Ser459		
O2	—	Gln172 NE2 ^f
O2	—	Thr173 N ^f
O3	water 1004	Asn-glycosyl 171 O ^f
O3	water 860	Thr-mannose 462 O5
O4	water 860	Thr-mannose 464 O5
O4 ^d	water 860	Thr-mannose 464 O6
O4	water 712	Thr-mannose 462 N
O4	water 712	Thr-mannose 462 O
O4 ^c	water 712	Thr-mannose 464 O3
O4	—	Ser-mannose 460 O
O6	water 1007	Gln172 OE1 ^f
O6	water 806	Arg160 NH1
O6	—	Gln172 NE2 ^f
Mannosyl chain of Ser460		
O2	water 733	Ile87 O
O2	— ^d	Ala86 O
O3	water 733	Ile87 O
O3	water 711	Val461 O
O4	water 988	Thr-mannose 462 O3
Mannosyl chain of Thr462		
O2	—	Ile460 O
O3	water 988	Ser-mannose 460 O4
O5	water 860	Ser-mannose 459 O3
O5	water 860	Ser-mannose 459 O4
O5 ^d	water 860	Thr-mannose 464 O6
O6 ^c	—	Thr-mannose 464 O4
O6 ^d	—	Thr-mannose 464 O6
Mannosyl chain of Thr464		
O3 ^c	water 712	Thr-mannose 462 O
O3 ^c	water 720	Thr-mannose 462 N
O3 ^c	water 720	Ser-mannose 459 O4
O4 ^c	—	Thr-mannose 462 O5
O4 ^c	—	Thr-mannose 462 O6
O4 ^d	water 957	Thr-mannose 462 O
O6 ^d	water 860	Ser-mannose 459 O4

Table 6 (continued)

Glycosyl chain atom	Bridging water ^b	Contact residue and atom
O6 ^d	water 860	Thr-mannose 462 O5
O6 ^d		Thr-mannose 462 O6

^a All contacts listed are less than 3.2 Å in at least 1 of the 2 structures. If a contact less than 3.2 Å is present in 1 structure, then it is absent in the other structure if the distance exceeds 3.4 Å.

^b If no bridging water is listed, then the atom of the left-hand column is in direct contact with the corresponding atom of the right-hand column.

^c Present only in the pH 6 structure.

^d Present only in the pH 4 structure.

^e Alternative position of an atom of a disordered side-chain.

^f Atom comes from a lattice contact molecule. See Table 5 for the pertinent symmetry transformation.

^g Intaresidue contact.

endocyclic oxygen (atom O5) of residue B. In addition, the 3-hydroxyl of residue B hydrogen bonds to the endocyclic oxygen (atom O5) of mannosyl C. Hydrogen bonds of the type mentioned above are observed in energy minimum conformations of disaccharides (Imberty *et al.*, 1990). In fact, the angles of Table 7 fall into energy allowed regions of ϕ, ψ plots (Imberty *et al.*, 1990) and, in most instances, correspond closely to conformations of minimum calculated energy. The root-mean-square difference in relative atomic positions after superposition of residues A, B and C of glycosyl chains 171 and 395 is 0.92 Å. The similarity in conformation of the two glycosyl chains is less beyond residue C. The entire glycosyl chain at Asn171 superimposes on that of Asn395 with a root-mean-square discrepancy of 1.9 Å. Apparently, the intrachain contacts largely determine the conformation of the glycosyl chain and perhaps the pentasaccharide core (residues A, B, C, D and E of glycosyl chain 395) will have a similar conformation irrespective of the protein to which it is attached.

Contacts between the *N*-glycosyl chains and the protein are limited to specific sugar residues. Glycosyl rings A, B, F and G contribute 25 of the 30 contacts to the protein, either by direct hydrogen bonding or by way of a bridging water molecule. Glycosyl residues F and G in particular curve back to the surface of the protein and interact strongly with side-chains and/or peptide linkages. The interactions involving residues F and G tether the *N*-glycosyl chain to the protein. This tethering effect may be responsible for the manifestation of ordered glycosyl chains at positions 171 and 395. Shaanan *et al.* (1991) report tethering of their *N*-linked glycosyl chain in the structure of a legume lectin. However, unlike the case of the lectin, which involves intermolecular hydrogen bonds, the *N*-linked glycosyl chains of glucoamylase are anchored entirely by intramolecular associations. Aside from the polar contacts listed in Table 6, the glycosyl chains at 171 and 395 cover surfaces of the protein that differ significantly in composition. Glycosyl chain 171 covers a hydrophobic groove that consists of the side-chains of Trp170, Tyr223,

Table 7

Torsion and three-center angles of glycosidic linkages for the *N*-glycosyl chains of glucoamylase^a

Linkage ^b	Glycosyl chain 171			Glycosyl chain 395			
	τ	ϕ	ψ	τ	ϕ	ψ	ω
Asn-A ^c	114	-92	180	116	-115	172	
A-B ^c	118	-81	-132	119	-79	-115	-
B-C ^c	113	-99	116	112	-93	-148	-
C-D ^c	118	94	170	110	84	133	-
D-G ^c	117	60	114	114	84	154	-
G-H ^{c,d}				115	97	137	-
C-E ^c				115	66	168	72
F-F ^c				113	78	139	-
E-I ^{c,e}				115	30	-163	-38

^a Values are averages of the pH 4 and 6 structures.

^b See Figure 5 for the definition of residues A to I.

^c Definitions of angles: for Asn-A, τ =C1A-ND2-CG, ϕ =O5A-C1A-ND2-CG, ψ =C1A-ND2-CG-CB; for A-B, τ =C1B-O4A-C4A, ϕ =O5B-C1B-O4A-C4A, ψ =C1B-O4A-C4A-C5A; for B-C, τ =C1C-O4B-C4B, ϕ =O5C-C1C-O4B-C4B, ψ =C1C-O4B-C4B-C5B; for C-D, τ =C1D-O3C-C3C, ϕ =O5D-C1D-O3C-C3C, ψ =C1D-O3C-C3C-C4C; for D-G, τ =C1G-O2D-C2D, ϕ =O5G-C1G-O2D-C2D, ψ =C1G-O2D-C2D-C3D; for G-H, τ =C1H-O2G-C2G, ϕ =O5H-C1H-O2G-C2G, ψ =C1H-O2G-C2G-C3G; for C-E, τ =C1E-O6C-C6C, ϕ =O5E-C1E-O6C-C6C, ψ =C1E-O6C-C6C-C5C, ω =O6C-C6C-C5C-O5C; for E-F, τ =C1F-O3E-C3E, ϕ =O5F-C1F-O3E-C3E, ψ =C1F-O3E-C3E-C4E; for E-I, τ =C1I-O6E-C6E, ϕ =O5I-C1I-O6E-C6E, ψ =C1I-O6E-C6E-C5E, ω =O6E-C6E-C5E-O5E.

^d Present only in the pH 4 structure.

^e Present only in the pH 6 structure.

Phe227, Leu234 and Phe237, whereas glycosyl chain 395 covers a surface with primarily hydrophilic side-chains (Ser397, Arg413 and Asp414). Tyr223 stacks with residue A of glycosyl chain 171, as has been observed in other protein-carbohydrate complexes (Quijcho, 1989).

Borne *et al.* (1990, 1992) observed hydrogen bonding networks between waters that act as bridging links between the glycan and the protein. For the two *N*-linked glycosyl chains of glucoamylase we observe two such networks. Waters 663, 736 and 796 mediate interactions between the protein and glycosyl chain 171 and waters 713, 864 and 865 mediate interactions between the protein and glycosyl chain 395.

(g) *O*-linked glycosyl chains

The *O*-linked glycosyl chains are all single mannose residues. We see no variation in the *O*-linked glycosylation of the two crystal structures. Additional glycosylation at the 2 and 6-hydroxyls, as observed by Gunnarsson *et al.* (1984) in their chemical analysis of *O*-glycosyl chains from glucoamylase-I from *A. niger*, are not observed in the crystal structures. As some *O*-glycosyl chains (residues 453, 455, 457 and 459) are major contributors to lattice contacts in the crystal (Table 5), enzyme molecules bearing more extensive glycosyl chains at these positions may not crystallize. Conceivably, larger *O*-glycosyl chains may be present only in the second half of the *O*-glycosylated domain, which is absent in our crystal structures.

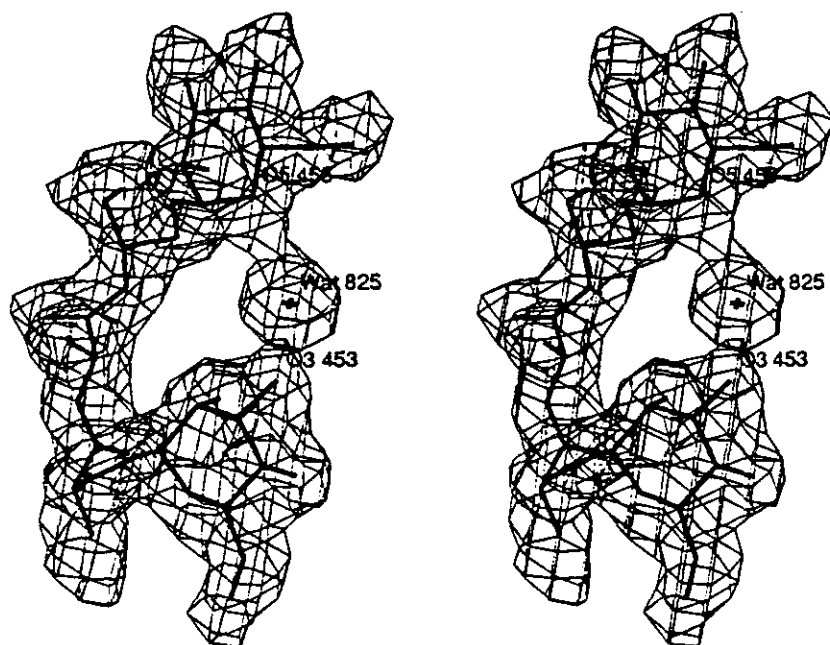


Figure 7. *O*-Glycosyl side-chains and associated electron density showing the interaction of the O4 atom of one mannosyl group with the endocyclic oxygen atom (O5) of its neighbor by way of a bridging water molecule.

Non-bonded interactions involving atoms of the *O*-linked mannosyl groups indicate several tendencies. If we consider only interactions between the mannosyl chains and the protein to which they are linked (lattice contacts excluded) and if we exclude Thr-mannose 464 from consideration due to the ambiguity of its conformation (see above), then the remaining nine mannosyl residues makes 4 direct contacts and 20 indirect contacts (through a single bridging water molecule) to the protein. Of these 24 interactions, 18 involve the atoms of peptide linkages, with the carbonyl group predominating (15 out of the 24 interactions). Excluding Thr-mannose 464 from consideration, eight contacts exist between the mannosyl chains of separate protein residues. All of the eight contacts involve bridging water molecules and in six of the eight interactions the water molecule makes a hydrogen bond to the endocyclic oxygen (atom O5) of the mannosyl group, the most common interaction being mannosyl O3 and/or O4 to a bridging water molecule to mannosyl O5 of a neighboring glycosyl chain (Figure 7).

The tendencies noted above for the interactions of the mannosyl chains arise out of specific conformational preferences exhibited by the collection of serine-mannose and Thr-mannose residues of glucoamylase-II(471) (Table 8). Taking the average of ϕ and ψ over the ten residues *O*-linked to mannose, one finds that eight of the ten residues have ϕ and ψ angles within $\pm 60^\circ$ of the average values. The outliers are Thr-mannose 464, which has weak density for the side-chain and Ser-mannose 443, which is the first of the *O*-glycosylated residues from the N terminus. Averages of ϕ and ψ using only eight *O*-linked residues (excluding residues 443 and 464) gives -125° and 144° , respectively. These

values are characteristic of an extended conformation of a polypeptide chain, similar to the one observed for the N-terminal half of the *O*-glycosylated domain of glucoamylase-II(471). The tendency toward an extended polypeptide is not surprising as such a conformation would bring about the maximum separation of the bulky *O*-linked mannosyl chains and thereby minimize steric clashes. Six of the eight residues used in the computation of average ϕ and ψ angles have χ^1 angles (defined by atoms N, CA, CB and OG or OG1) near 180° . When χ^1 is 180° , atom OG of serine (or OG1 of threonine) is near the nitrogen of the

Table 8
Selected torsion angles (in degrees) of *O*-glycosylated residues of glucoamylase-II(471)^a

Residue number	ϕ	ψ	χ^1	χ^2	χ^3
Ser443	-123	5	74	-100	-144
Ser444	-84	125	163	-174	-173
Thr452	-165	174	90	-175	-158
Ser453	-149	-173	-165	75	-170
Ser455	-94	130	180	144	-169
Thr457	-155	146	-170	138	-166
Ser459	-160	133	-171	179	-169
Ser460	-71	132	180	87	-171
Thr462	-120	123	-58	105	-166
Thr464 (pH 4)	-56	-60	-56	-166	-55
Thr464 (pH 6)	-69	-51	-27	153	-139
Typical values ^b	-125	144	180	60 to 180	-168

^a Entries for residues 442 to 461 are averages of the torsion angles from the pH 4 and 6 structures. The values for some of the torsion angles of residue 464 differ sufficiently to warrant a separate listing for the pH 4 and 6 structures. χ^1 is defined by atoms N, CA, CB and OG (or OG1); χ^2 is defined by atoms CA, CB, OG (or OG1) and C1 (of D-mannose); χ^3 is defined by atoms CB, OG (or OG1), C1 (of D-mannose) and C2 (of D-mannose).

^b See the text for an explanation of the values listed here.

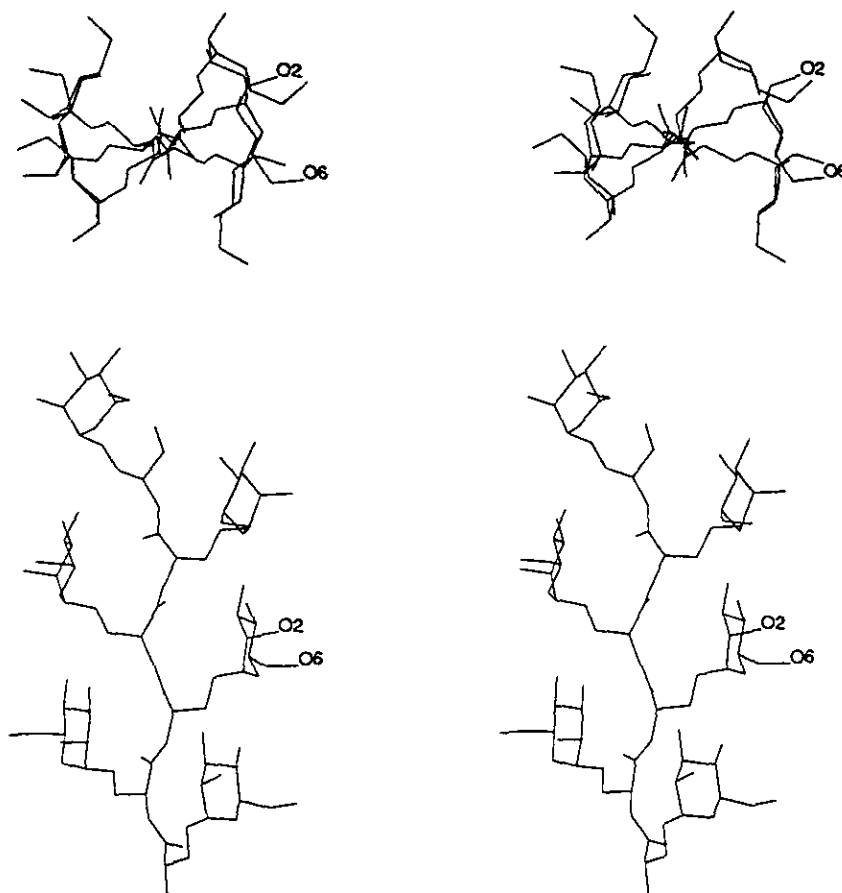


Figure 8. Model for the conformation of *O*-glycosylated polyserine viewed down its helix axis (top) and perpendicular to its axis (bottom). Labeled are atoms O2 and O6 of one mannosyl group.

peptide linkage, allowing at least for a favorable electrostatic interaction between N-H group of the peptide linkage and the lone pair electrons on OG (or OG1). In addition, at χ^1 of 180° bad steric contacts between the side-chain and atoms of the main-chain are at a minimum. The χ^3 angle of Table 7 (defined at atoms CB, OCG (or OG1), C1 of mannose and C2 of mannose) falls within $\pm 10^\circ$ of -170° . The torsion angle χ^2 is more variable, but with the values of ϕ , ψ , χ^1 and χ^3 , given above, χ^2 can adopt a value between 60 and 180° without serious steric conflicts. Figure 8 is an illustration of polyserine-*O*-mannose with the typical values listed in Table 7 for torsion angles and χ^2 set to 170° . The polymer has approximately a 20° right-handed twist between next-to-nearest neighbors. That twist is determined by the values of ϕ and ψ for the polypeptide chain and is similar in magnitude to the twist of an individual strand in a β -sheet. The 4-hydroxyl of one mannosyl chain is 3.8 Å from the endocyclic oxygen of the mannosyl chain of its next-to-nearest neighbor and thus, can participate by way of a bridging water molecule in a hydrogen bonding interaction. The structure depicted in Figure 8 has the 2-hydroxyl as well as the 6-hydroxyl projecting radially outward from the axis of the helix. These hydroxyl groups are available, then, for the attachment of additional glycosyl residues, consistent with the observation by

Gunnarsson *et al.* (1984) of α -1 $\rightarrow$ 2 and α -1 $\rightarrow$ 6 extensions of the *O*-glycosyl chain. The model of Figure 8 may be a good first approximation for the second half of the *O*-glycosylated domain. Thirty-three of the 42 residues of the C-terminal half of the *O*-glycosylated domain are either Ser-mannose or Thr-mannose in glucoamylase from *A. niger*. Although we anticipate a rigid structure for polyserine-*O*-mannose, in agreement with the analysis of NMR spectra by Williamson *et al.* (1992), the *O*-glycosylated domain of glucoamylase has glycine residues before and after the C-terminal half of the *O*-glycosylated domain. The glycine residues may represent points of articulation or conformational flexibility in the *O*-glycosylated domain that allows for some variation in the orientation of the starch binding domain with respect to the catalytic domain of glucoamylase I.

We thank Drs Arthur Arnone and Richard Ortega for collecting crystallographic data on crystals of native glucoamylase at pH 6. This work was supported by the National Science Foundation (Grant DMB-9018205), the United States Department of Agriculture (Grant 92-37500-8203), the National Institutes of Health (Grant 2-S07-RR07034-24) and the Iowa State University Biotechnology Program. Journal paper J-15499 of the Iowa Agriculture and Home Economics Experiment Station, Ames Iowa; Project No. 3159.

References

- Abdullah, M., Fleming, I. D., Taylor, P. M. & Whelan, W. J. (1963). Substrate specificity of the amyloglucosidase of *Aspergillus niger*. *Biochem. J.* **89**, 35P–36P.
- Aleshin, A., Golubev, A., Firsov, L. M. & Honzatko, R. B. (1992). Crystal structure of glucoamylase from *Aspergillus awamori* var. *X100* to 2.2-Å resolution. *J. Biol. Chem.* **267**, 19291–19298.
- Boel, E., Hjort, I., Svensson, B., Norris, F., Norris, K. E. & Fül, N. P. (1984). Glucoamylases G1 and G2 from *Aspergillus niger* are synthesized from two different but closely related mRNAs. *EMBO J.* **3**, 1097–1102.
- Bourne, Y., Rougé, P. & Cambillau, C. (1990). X-ray structure of a (α -man(1-3) β -man(1-4)glcNAc)-lectin complex at 2.1 Å resolution. *J. Biol. Chem.* **265**, 18161–18166.
- Bourne, Y., Rougé, P. & Cambillau, C. (1992). X-ray structure of a biantennary octasaccharide-lectin complex refined at 2.3 Å resolution. *J. Biol. Chem.* **267**, 197–203.
- Cambillau, C. & Horjales, E. (1987). TOM: a FRODO subpackage for protein-ligand fitting with interactive energy minimization. *J. Mol. Graphics* **5**, 174–177.
- Chothia, C., Levitt, M. & Richardson, D. (1978). Structure of proteins: packing of α -helices and pleated sheets. *Proc. Nat. Acad. Sci., U.S.A.* **74**, 4130–4134.
- Chu, S. S. C. & Jeffrey, G. A. (1968). The refinement of the crystal structures of β -D-glucose and cellobiose. *Acta Crystallogr. sect. B*, **24**, 830–838.
- Crawford, J. L., Lipscomb, W. N. & Schellman, C. G. (1973). The reverse turn as a polypeptide conformation in globular proteins. *Proc. Nat. Acad. Sci., U.S.A.* **70**, 538–542.
- Einsphar, H. M. (1989). Protein-carbohydrate interactions in biological systems. *Trans. Amer. Crystallogr. Assoc.* **25**, 1–22.
- Golubev, A. M., Neustroev, K. N., Aleshin, A. E. & Firsov, L. M. (1992). Crystallization and preliminary X-ray study of minor glucoamylase from *Aspergillus awamori* variant *X-100/D27*. *J. Mol. Biol.* **226**, 271–272.
- Gould, G. W. & Bell, G. I. (1990). Facilitative glucose transporters: an expanding family. *Trends Biochem. Sci.* **15**, 18–23.
- Gunnarsson, A., Svensson, B., Nilsson, B. & Svensson, S. (1984). Structural studies on the O-glycosidically linked carbohydrate chains of glucoamylase G1 from *Aspergillus niger*. *Eur. J. Biochem.* **145**, 463–467.
- Harris, E. M. S., Aleshin, A. E., Firsov, L. M. & Honzatko, R. B. (1993). Refined structure of the complex of 1-deoxynojirimycin with glucoamylase from *Aspergillus awamori* var. *X100* to 2.4-Å resolution. *Biochemistry*, **32**, 1618–1626.
- Hata, Y., Tsuchiya, K., Kitamoto, K., Gomi, K., Kumagai, C., Tamura, G. & Hara, S. (1991). Nucleotide sequence and expression of the glucoamylase-encoding gene *gla A* from *Aspergillus oryzae*. *Gene*, **108**, 145–150.
- Hayashida, S., Nomura, T., Yoshino, E. & Hongo, M. (1975). The formation and properties of subtilisin-modified glucoamylase. *Agric. Biol. Chem.* **40**, 141–146.
- Hayashida, S., Kuroda, K., Ohta, K., Kuhara, S., Fukuda, K. & Sakaki, Y. (1989a). Molecular cloning of the glucoamylase I gene of *Aspergillus awamori* var. *kawachi* for localization of the raw-starch-affinity site. *Agric. Biol. Chem.* **53**, 923–929.
- Hayashida, S., Nakahara, K., Kuroda, K., Miyata, T. & Iwanaga, S. (1989b). Structure of the raw-starch-affinity site on the *Aspergillus awamori* var. *kawachi* glucoamylase I molecule. *Agric. Biol. Chem.* **53**, 135–141.
- Hendrickson, W. A. & Konnert, J. H. (1980). Incorporation of stereochemical information into crystallographic refinement. In *Computing in Crystallography* (Diamond, R., Ramaseshan, S. & Venkutesan, K., eds), pp. 13.01–13.23, Indian Academy of Sciences, Bangalore.
- Hiromi, K., Takahashi, K., Hamauzu, Z. & Ono, S. (1966). Kinetic studies on gluc-amylase. III. The influence of pH on the rates of hydrolysis of maltose and panose. *J. Biochem.* **59**, 469–475.
- Hiromi, K., Ohnishi, M. & Tanaka, A. (1983). Subsite structure and ligand binding mechanism of glucoamylase. *Mol. Cell. Biochem.* **51**, 79–95.
- Honzatko, R. B. (1986). Automated calculation of coordinate transformations for the superposition of protein structures. *Acta Crystallogr. sect. A*, **42**, 172–178.
- Imberty, A., Gerber, S., Tran, V. & Pérez, S. (1990). Data bank of three-dimensional structures of disaccharides, a tool to build 3-D structures of oligosaccharides: Part I. Oligo-mannose type type N-glycans. *Glycoconjugate J.* **7**, 27–54.
- Juy, M., Amit, A. G., Alzari, P. M., Poljak, R. J., Caeyssens, M., Béguin, P. & Aubert, J.-P. (1992). Three-dimensional structure of a thermostable bacterial cellulase. *Nature (London)*, **357**, 89–91.
- Kraulis, J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950.
- Lewis, P. N., Mommay, F. A. & Scheraga, H. A. (1973). Chain reversals in proteins. *Biochem. Biophys. Acta*, **303**, 211–229.
- Liebman, M. N. (1982). Correlation of structure and function in biologically active small molecules and macromolecules by distance matrix partitioning. In *Molecular Structure and Biological Activity* (Griffin, J. & Duax, W. L., eds), pp. 193–212, Elsevier, New York.
- Longchambon, P. F., Avenel, D. & Neuman, A. (1976). Structure cristalline de l' α -D-mannopyranose. *Acta Crystallogr. sect. B*, **32**, 1822–1826.
- Luzzati, V. (1952). Traitement statistique des erreurs dans la détermination des structures cristallines. *Acta Crystallogr.* **5**, 802–810.
- Mueckler, M., Caruso, C., Baldwin, S. A., Panico, M., Blench, I., Morris, H. R., Allard, J. W., Lienhard, G. E. & Lodish, H. F. (1985). Sequence and structure of a human glucose transporter. *Science*, **229**, 941–945.
- Neustroev, K. N. & Firsov, L. M. (1990). Acid proteinase and the multiplicity of *Aspergillus awamori* glucoamylase forms. *Biokhimiia*, **55**, 776–785.
- Nikaido, H. & Saier, M. H., Jr (1992). Transport proteins in bacteria: common themes in their design. *Science*, **258**, 936–942.
- Nunberg, J. H., Meade, J. H., Cole, G., Lawyer, E. C., McCabe, P., Schweickart, V., Tal, R., Wittman, V. P., Flatgaard, J. E. & Innis, M. A. (1984). Molecular cloning and characterization of the glucoamylase gene of *Aspergillus awamori*. *Mol. Cell. Biol.* **4**, 2306–2315.
- Pazur, J. H. & Ando, T. (1960). The hydrolysis of glucosyl oligosaccharides with α -D-(1-4) and α -D-(1-6) bonds by fungal amyloglucosidase. *J. Biol. Chem.* **235**, 297–302.
- Quiocho, F. A. (1989). Protein-carbohydrate interactions:

- basic molecular features. *Pure Appl. Chem.* **61**, 1293–1306.
- Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7**, 95–99.
- Richardson, J. S. (1981). The anatomy and taxonomy of protein structure. *Advan. Protein Chem.* **34**, 167–339.
- Rose, G. D., Gierasch, L. M. & Smith, J. A. (1985). Turns in peptides and proteins. *Advan. Protein Chem.* **37**, 1–109.
- Saha, B. C. & Zeikus, J. G. (1989). Microbial glucoamylases: biochemical and biotechnological features. *Starch/Stärke*, **41**, 57–64.
- Savel'ev, A. N. & Firsov, L. M. (1982). Carboxyl groups in the active site of glucoamylase from *Aspergillus awamori*. *Biochimica*, **47**, 1618–1620.
- Shaanan, B., Halina, L. & Sharon, N. (1991). Structure of a legume lectin with an ordered N-linked carbohydrate in complex with lactose. *Science*, **254**, 862–866.
- Shibuya, I., Gomi, K., Iimura, Y., Takahashi, K., Tumura, G. & Hara, S. (1990). Molecular cloning of the glucoamylase gene of *Aspergillus shirousami* and its expression in *Aspergillus oryzae*. *Agric. Biol. Chem.* **54**, 1905–1914.
- Sierks, M. R., Ford, C., Reilly, P. J. & Svensson, B. (1990). Catalytic mechanism of fungal glucoamylase as defined by mutagenesis of Asp176, Glu179 and Glu180 in the enzyme from *Aspergillus awamori*. *Protein Eng.* **3**, 193–198.
- Svensson, B., Pedersen, T. G., Svendsen, I., Sakai, T. & Ottesen, M. (1982). Characterization of two forms of glucoamylase from *Aspergillus niger*. *Carlsberg Res. Commun.* **47**, 55–59.
- Svensson, B., Larsen, K., Svendsen, I. & Boel, E. (1983). The complete amino acid sequence of the glycoprotein, glucoamylase G1, from *Aspergillus niger*. *Carlsberg Res. Commun.* **48**, 529–544.
- Svensson, B., Larsen, K. & Gunnarsson, A. (1986). Characterization of a glucoamylase G2 from *Aspergillus niger*. *Eur. J. Biochem.* **154**, 497–502.
- Svensson, B., Jespersen, H., Sierks, M. & MacGregor, E. A. (1989). Sequence homology between putative raw-starch binding domains from different starch-degrading enzymes. *Biochem. J.* **264**, 309–311.
- Weill, C. E., Burch, R. J. & Van Dyk, J. W. (1954). An alpha-amylglucosidase that produces beta-glucose. *Cereal Chem.* **31**, 150–158.
- Williamson, G., Belshaw, N. J. & Williamson, M. P. (1992). O-Glycosylation in *Aspergillus* glucoamylase. *Biochem. J.* **282**, 423–428.
- Wilmot, C. M. & Thornton, J. M. (1989). Analysis and prediction of the different types of β -turn in proteins. *J. Mol. Biol.* **203**, 221–232.

Edited by R. Huber

(Received 23 September 1993; accepted 24 January 1994)