O-Glycosylation in *Aspergillus* glucoamylase

Conformation and role in binding

Gary WILLIAMSON,* Nigel J. BELSHAW* and Michael P. WILLIAMSON†

*AFRC Institute of Food Research, Norwich Laboratory, Norwich Research Park, Colney Lane, Norwich NR4 7UA, and †Department of Molecular Biology and Biotechnology, University of Sheffield, P.O. Box 594, Firth Court, Western Bank, Sheffield S10 2UH, U.K.

Functional peptides have been produced by proteolysis of glucoamylase (glucan 1,4- α -glucosidase; EC 3.2.1.3) from *Aspergillus niger* and purified by affinity chromatography, gel filtration and two ion-exchange-chromatography steps. The peptides correspond to residues 499–616 and 509–616 of the original glucoamylase molecule. Together with G1C (residues 471–616 from glucoamylase 1) [Belshaw & Williamson (1990) FEBS Lett. 269, 350–353], the three peptides all contain the C-terminal domain (residues 509–616) but, in addition, contain different proportions of the O-glycosylated region. The properties of these peptides have been compared to define the function of the O-linked oligosaccharides in this protein. The O-glycosylated region plays only a minor role in binding to hydrogen-bond ordered starch. The difference between the apparent free energy (ΔG) for binding between the non-glycosylated C-terminal domain (–26.0 kJ/mol) and the C-terminal domain containing the fully O-glycosylated region (–25.0 kJ/mol) is only 1.0 kJ/mol. Binding to β -cyclodextrin suggests that even this difference may reflect a small conformational change in the C-terminal domain rather than a direct effect of the O-linked sugars. The c.d. spectrum of the O-glycosylated region is deduced by comparison of the three peptides and is predominantly that of a random-coil structure. Two-dimensional n.m.r. spectra of glucoamylase and of the glycosylated peptide 499–616 show that the binding domain is more mobile than the catalytic domain and that its mobility is further increased on removal of the catalytic domain. The O-glycosylated region is more mobile still, and there is a marked increase in its mobility on removal of the catalytic domain. The O-glycosylated region in the intact protein can therefore be envisaged as a semi-rigid rod. The results show that a major function of O-glycosylation in glucoamylase 1 is to provide an extended peptide backbone and hence a fixed distance in linking the catalytic and binding domains. It does not in itself significantly increase the binding affinity for starch.

INTRODUCTION

Mono-, di- or tri-saccharides attached to serine or threonine residues in proteins are often clustered together to form heavily O-glycosylated regions. This structural motif is found amongst a diverse range of proteins, such as mammalian mucins (Slater & Coddington, 1973; Shogren *et al.*, 1989), mammalian low-density-lipoprotein receptor (Yamamoto *et al.*, 1984) and fungal carbohydases (Gunnarsson *et al.*, 1984; Knowles *et al.*, 1988). A suggested common function of this structural motif is to extend and rigidify the peptide backbone (Jentoft, 1990). In many proteins, O-glycosylation is within a relatively short region linking two protein domains (Jentoft, 1990).

Glucoamylase 1 (glucan 1,4- α -glucosidase, EC 3.2.1.3) from *Aspergillus niger* is an example of a protein containing two domains (catalytic domain, residues 1–470; C-terminal domain, 509–616) linked by a region (471–508) containing ~70% (w/w) carbohydrate (Svensson *et al.*, 1983, 1989; Evans *et al.*, 1990). This region consists of short oligosaccharides (predominantly mannose) O-linked to serine and threonine (Gunnarsson *et al.*, 1984). The covalently bound sugars in glucoamylase have been implicated in substrate binding (Hayashida *et al.*, 1989a,b) and in increasing thermostability (Shendy *et al.*, 1984).

Here we assess the role of glucoamylase O-glycosylation in substrate binding using proteolytically derived functional fragments of glucoamylase 1. We also demonstrate that the densely packed sugars impart rigidity to the peptide backbone and maintain it in a predominantly extended 'random-coil' structure.

EXPERIMENTAL

Materials

Glucoamylase 1, glucoamylase 2 and G1C (residues 471–616 from glucoamylase 1) were prepared as described previously (Belshaw & Williamson, 1990; Subbaramaiah & Sharma, 1988). *Bacillus licheniformis* proteinase type VIII, β -cyclodextrin, granular corn starch, α -mannosidase [from jack bean (*Canavalia ensiformis*)], 5,5'-dithiobis-(2-nitrobenzoic acid) and protein standards for chromatography were from Sigma Chemical Co. Proteinase type VIII (EC 3.4.21.4) has a broad specificity and is also known as subtilisin Carlsberg or subtilopeptidase A.

Purification steps

Proteinase digestion of glucoamylase 1 using *Bacillus licheniformis* proteinase type VIII and initial purification by affinity chromatography was performed as described previously (Belshaw & Williamson, 1990), except that the reaction was terminated before complete loss of glucoamylase 1. About 200 mg (calculated from original protein) was applied to a Mono Q (HR 10/10) column (Pharmacia-LKB technology) in 2 mM-Tris adjusted to pH 7.0 with 1 M-HCl (2 ml/min). Protein was eluted with a gradient of KCl, and the process was repeated 45 times. Pooled fractions (from the later of the two protein peaks which exhibited binding to granular starch but no hydrolytic activity) were loaded on to a Sephacryl S-200 (2.5 cm $\times$ 90 cm) column in 50 mM-sodium acetate, pH 5.0/0.1 M-NaCl (0.5 ml/min). Aliquots from gel filtration (0.25 ml, 6.5 mg/ml) were then re-

Abbreviations used: G1C, residues 471–616 from glucoamylase 1; G1C⁴⁹⁹, residues 499–616; G1C⁵⁰⁹, residues 509–616; s.e.c., size-exclusion chromatography; TOCSY, total correlation spectroscopy.

applied to the Mono Q column in 2 mM-Tris/HCl, pH 7.0, and the protein eluted with a shallow gradient of KCl.

Deglycosylation

G1C⁴⁹⁹ (residues 499–616 of glucoamylase 1) (0.5 mg) was incubated with α -mannosidase (4 units) at 35 °C in 0.25 ml of 0.05 M-sodium acetate (pH 4.5)/2 mM-zinc acetate. After the addition of 0.04 ml of 1 M-potassium phosphate, pH 7.0, the mixture was heated to 60 °C for 30 min. After cooling and centrifugation (10000 *g*, 10 min), the supernatant was applied to a column of Bio-Gel P-2 (10 cm $\times$ 1 cm) in water, and the fractions containing protein were pooled. Removal of α -mannosidase was confirmed by chromatography on a TSK G3000 SW column. The carbohydrate content decreased by 92 % after this treatment.

Estimation of M_r

A Tosoh TSK G3000 SW column (Anachem Ltd.) was equilibrated in 50 mM-sodium phosphate (pH 7.9)/0.3 M-NaCl. Samples (25 μ l) were injected (0.05 ml/min), and the elution position obtained by measuring peak A_{280} relative to standards of glucoamylase 1 and 2 (M_r 82000 and 70000 respectively; Svensson *et al.*, 1983) and of G1C (M_r 25000; Belshaw & Williamson, 1990).

Denaturing chromatography on the same column in 6 M-guanidinium chloride/0.2 mM-dithiothreitol (0.05 ml/min) employed the following standards in addition to the above: BSA (M_r 66000), ovalbumin (M_r 45000), trypsinogen (from bovine pancreas; M_r 24000) and α -lactalbumin (from bovine milk; M_r 14200). All standards were in 4–6 M-guanidinium chloride/0.2 mM dithiothreitol. The void volume in both cases was estimated with Blue Dextran.

SDS/PAGE (15 % acrylamide; 2.7 % cross-linked) (Laemmli, 1970) was run using the same standards as those used for non-denaturing size-exclusion chromatography (s.e.c.). Proteins were revealed by staining with Coomassie Blue.

Characterization of peptides

Amino acid composition was determined after hydrolysis in 6 M-HCl/0.1 mM-phenol at 115 °C for 15–17 h, using a Waters Pico Tag system with precolumn derivatization (phenyl isothiocyanate) and detection at 254 nm. Carbohydrate content was estimated by the phenol sulphuric acid method (Dubois *et al.*, 1956) relative to standards of D-mannose and glucoamylase 1. Protein concentration was estimated by using ϵ_{280} [30.7 mm⁻¹·cm⁻¹ for G1C, G1C⁴⁹⁹ and G1C⁵⁰⁹ (residues 509–616 of glucoamylase 1)] calculated from the tryptophan, tyrosine and cysteine content (Gill & von Hippel, 1989). In binding experiments this was used to calibrate the protein assay (Lowry *et al.*, 1951). The number of disulphide bonds was estimated with 5,5'-dithiobis-(2-nitrobenzoic acid) after reduction with NaBH₄ (Habeeb, 1972).

Assays

Hydrolysis of soluble starch by glucoamylase was measured by using 2,4-dinitrosalicylic acid (Rick & Stegbauer, 1974). Adsorption on to starch (Belshaw & Williamson, 1990) and binding of β -cyclodextrin (Belshaw & Williamson, 1991) were performed by published methods.

C.d.

Four averaged spectra were recorded on a Jasco J600 spectrophotometer at 20 °C at 10 nm/min with a 1 nm bandwidth. Samples (71 μ M) in 0.2 mM-sodium acetate, pH 4.5 (0.1 mM

pathlength) were run against buffer alone. The spectra were smoothed and the structural content analysed by the method of Provencher & Glockner (1981).

¹H n.m.r.

Data were collected on a Bruker AMX-500 spectrometer, using solutions of glucoamylase 1 (0.7 mM) and G1C⁴⁹⁹ (1.3 mM) in ²H₂O at pH 8.0, 30 °C. Two-dimensional pure-phase total correlation spectroscopy (TOCSY) spectra (Bazzo & Campbell, 1988) were acquired into 2048 real points using 220 incremental t_1 values, with spectral widths of 6000 Hz in F_2 and 5000 Hz in F_1 , 32 scans per increment. The residual solvent signal was suppressed by presaturation, and spin locking, of duration 32 ms, was achieved using the waltz-16 pulse sequence with a 4.5 kHz field. Data were processed using shifted-sine-bell window functions.

RESULTS

Production and purification of peptides

Fig. 1 shows the elution pattern on ion exchange after proteolysis and affinity chromatography. Peak C is unmodified glucoamylase, remaining after partial proteolysis. Peak B is G1C, and peak A is either a small amount of glucoamylase 2 or proteolytically cleaved glucoamylase 1 carried through from affinity chromatography. The fractions indicated by the bar were pooled. Gel filtration separated glucoamylase and G1C from the main protein fraction, which was eluted last. Fig. 2 shows resolution of this protein fraction into two major components. It is shown below that peak A is a peptide consisting of residues 499–616 from the original glucoamylase 1 (G1C⁴⁹⁹), whereas peak B consists of residues 509–616 (G1C⁵⁰⁹). Samples were dialysed against water, freeze-dried, and stored at –20 °C. The purity of each peptide was analysed by SDS/PAGE, and by s.e.c. under native and denaturing conditions. Both peptides, as well as G1C, gave fairly broad bands on SDS/PAGE, characteristic of glycoproteins (Segrest & Jackson, 1972). Despite this, G1C⁴⁹⁹ appeared to consist of only one component, whereas G1C⁵⁰⁹ exhibited a faint impurity at the same position as G1C⁴⁹⁹. By s.e.c., G1C⁴⁹⁹ gave one peak, whereas G1C⁵⁰⁹ was > 90 % pure with a small peak corresponding to G1C⁴⁹⁹.

Fig. 1. First ion-exchange chromatography of proteinase-treated glucoamylase

—, A_{280} ; ●, relative glucoamylase activity; ----, [KCl]. The elution positions of standards are shown by the arrows: A, glucoamylase 2; B, G1C; C, glucoamylase 1.

Fig. 2. Second ion-exchange chromatography on Mono Q

—, A_{280} ; ----, [KCl] (M).

Table 1. Characterization of G1C⁴⁹⁹ and G1C⁵⁰⁹

Parameter	Fragment ...	G1C ⁴⁹⁹	G1C ⁵⁰⁹
M_r			
Expected		16600	12500
S.e.c.		16400	10900
SDS/PAGE		16000	11500
S.e.c. in 6 M-GuHCl		17400	14000
Average		16600	12100
Carbohydrate content			
(%, w/w)			
Expected (approx.)		23	4
Obtained		19	1
ϵ (calculated)			
$\epsilon_{280}^{1\%}$ (mm ⁻¹ ·cm ⁻¹)		30.7	30.7
$\epsilon_{1\text{cm}, 280}^{1\%}$		5.5	4.1

Fig. 3. C-Terminal amino acid sequence of glucoamylase 1

The arrows indicate the cleavage positions by proteinase VIII. O-Glycosylated residues are indicated by asterisks.

Characterization of peptides

Table 1 shows the M_r values for the peptides as estimated by three methods. Glucoamylase and its fragments run anomalously on SDS/PAGE and native s.e.c., and so the M_r of G1C⁴⁹⁹ and G1C⁵⁰⁹ were estimated from glucoamylase-derived standards in these methods. The expected carbohydrate content was calculated from the number of glycosylated residues (Fig. 3), each of which contains an average of two sugar residues (Gunnarsson *et al.*, 1984), although the distribution of O-glycosylation may vary between enzyme batches. The amino acid composition further supports the assignments of these peptides as residues 499–616 and 509–616 from the original glucoamylase. The assignments of 499 and 509 as the N-terminal amino acids in G1C⁴⁹⁹ and G1C⁵⁰⁹ respectively are tentative. It is possible, for example, that residue

Table 2. Amino acid composition of G1C⁴⁹⁹ and G1C⁵⁰⁹

No free cysteine was detected by titration with 5,5'-dithiobis-(2-nitrobenzoic acid) in 6 M-guanidinium chloride. The number of cysteine residues was estimated after NaBH₄ reduction. Abbreviation: ND, not determined. The amino acids serine and threonine are underestimated by about 10 and 5 % respectively by this method.

Amino acid	Composition (residues/molecule)			
	G1C ⁴⁹⁹		G1C ⁵⁰⁹	
	Expected	Obtained	Expected	Obtained
Asx	11	10.7	11	11.2
Glx	10	11	10	10.4
Ser	16	15.3	11	11.5
Gly	6	9.4	6	7.7
His	0	0	0	0
Arg	2	3.4	2	2.1
Thr	22	18.1	18	16.4
Ala	8	9.4	8	8.1
Pro	5	6	5	4.5
Tyr	6	4.1	6	5.6
Val	8	7.7	8	7.8
Met	0	0	0	0
Ile	5	4.7	5	4.6
Leu	6	5.8	6	5.8
Phe	3	2.8	3	2.6
Lys	3	3.1	2	1.6
Trp	4	ND	4	ND
$\frac{1}{2}$ -Cys	2	2	2	2

497 may be the N-terminal amino acid in G1C⁴⁹⁹. The alanine content is 9.4 residues/molecule of protein (Table 2), whereas a value of 8 is expected from the published sequence (Fig. 3). Consequently, it is possible that G1C⁴⁹⁹ is a mixture of G1C⁴⁹⁷, G1C⁴⁹⁸ and G1C⁴⁹⁹. On the other hand, since there are only two arginine residues in this part of the sequence, the amino acid analyses strongly suggest that the C-terminal amino acid is arginine, even though G1C⁴⁹⁹ apparently has one extra arginine residue. The properties of these two peptides, compared with G1C, are described below.

Binding to granular starch

G1C⁴⁹⁹ and G1C⁵⁰⁹ both bound to granular starch (Fig. 4). The relationship between the amount of protein adsorbed, q_{ad} , to the free protein in solution (q) is described by a Michaelis-Langmuir-type equation:

$$q_{ad}/q = Kp \cdot q_{max} / (1 + Kp \cdot q)$$

where Kp is a constant and q_{max} is the maximum amount of protein adsorbed by the starch (Kyriacou *et al.*, 1989). $1/Kp$ is an apparent dissociation constant of the protein from the ligand, and from the equation $\Delta G^{app} = -RT \ln Kp$, an apparent free-energy change can be calculated. For G1C⁴⁹⁹, $1/Kp = 16.8 \pm 0.9 \mu\text{M}$, and for G1C⁵⁰⁹, $Kp = 19.6 \pm 2.0 \mu\text{M}$. This compares with $12.7 \pm 0.5 \mu\text{M}$ for G1C (Belshaw & Williamson, 1990). The ΔG^{app} values calculated show decreasing free energy (i.e. increasing affinity) with an increasing number of glycosylated residues. The relationship appears to be linear, although there are only a limited number of points. From this (Fig. 4, inset) the contribution to binding of one O-glycosylated serine or threonine residue containing one disaccharide unit is 0.04 kJ/mol. Even considering that there are 28 O-glycosylated residues in G1C, their total contribution (~ 1 kJ/mol) is very small compared with that of G1C⁵⁰⁹ (~ 25 kJ/mol), which contains only two O-glycosylated residues.

Fig. 4. Adsorption on to granular starch

Various concentrations of GIC (●), GIC⁴⁹⁹ (■) and GIC⁵⁰⁹ (▲) were incubated with a suspension of granular starch (10 mg/ml) in 5 mM-sodium acetate, pH 4.5, at 4 °C until equilibrium was reached. The inset shows ΔG^{app} as a function of the number of *O*-glycosylated residues in each peptide.

The maximum amount of protein adsorbed, q_{max} , for GIC, GIC⁴⁹⁹ and GIC⁵⁰⁹ was 1.08 ± 0.02 , 1.58 ± 0.03 and 2.14 ± 0.11 nmol of protein/mg of starch respectively. This difference merely reflects a difference in M_r : for all three peptides, q_{max} , 27 μ g of peptide/mg of starch. Thus the length of the *O*-glycosylation does not affect the stoichiometry of adsorption on to starch, which from surface-area calculations is likely to be a monolayer of protein around a starch granule (S. G. Ring, unpublished work).

Binding of deglycosylated GIC⁴⁹⁹ to granular starch

The binding of deglycosylated GIC⁴⁹⁹ to granular starch was also measured (results not shown). The curve is almost a straight line up to 70 μ M-protein ($\sim 4 \times K_d$ for GIC⁴⁹⁹), and hence saturation is not seen. This suggests that the protein may aggregate. To test this, the protein was eluted from an s.e.c. column under non-denaturing conditions. This elution volume indicated an M_r of 13200 (expected M_r 12900), showing that the protein does not aggregate in solution. On the starch granule, therefore, the protein must associate to form a layer several protein molecules thick. This shows that the *O*-glycosylation helps to maintain the minimum binding stoichiometry, i.e. a protein monolayer.

Binding of β -cyclodextrin

GIC binds two molecules of β -cyclodextrin with a K_d of 1.7 μ M (Belshaw & Williamson, 1991). GIC⁴⁹⁹ and GIC⁵⁰⁹ were titrated with β -cyclodextrin to determine the effect of the *O*-glycosylated region brings about a small change in ΔG of 1.6 kJ/mol. The *O*-absorbance after addition of β -cyclodextrin to all three peptide exhibited peaks at 278, 286 and 294 nm. Table 3 shows the binding constants, the stoichiometry and ΔG calculated from these absorbance changes. The removal of the *O*-glycosylated region brings about a small change in ΔG of 1.6 kJ/mol. The *O*-glycosylation is unlikely to be involved in binding to β -cyclodextrin, which indicates that the $\Delta\Delta G$ is probably due to a small

Table 3. Binding of β -cyclodextrin to glucoamylase fragments

Enzyme fragment	K_d (μ M)	Stoichiometry (mol/mol)	ΔG (kJ/mol)	$\Delta\Delta G$ (kJ/mol)
GIC*	1.7	2.0	-32.3	-
GIC ⁴⁹⁹	3.0	1.7	-31.0	1.3
GIC ⁵⁰⁹	3.3	1.7	-30.7	1.6

* From Belshaw & Williamson (1991).

Fig. 5. C.d. spectrum of GIC⁴⁹⁹ (71 μ M) in 0.2 M-sodium acetate, pH 4.5

Table 4. Structure of glucoamylase fragments

The number of amino acid residues ($\pm$ S.D.) was estimated from c.d. spectra.

Enzyme fragment	No. of amino acids		
	α -Helix	β -Structure	'Random coil'
GIC	12 $\pm$ 1	77 $\pm$ 1	55 $\pm$ 1
GIC ⁴⁹⁹	12 $\pm$ 2	66 $\pm$ 2	39 $\pm$ 3
GIC ⁵⁰⁹	11 $\pm$ 2	68 $\pm$ 2	27 $\pm$ 3

conformational change in the protein which modifies the affinity for β -cyclodextrin. Since $\Delta\Delta G^{app}$ for binding to granular starch is of a comparable value (1.0 kJ/mol), it is concluded that the *O*-glycosylated region does not in itself play a significant role in binding to ligands, whether soluble or insoluble. Furthermore, the small observed effect is due solely to a small conformational change of the C-terminal domain after removal of the *O*-glycosylated region.

Conformation of the *O*-glycosylated peptide backbone by c.d.

The structure of the C-terminal domain containing no *O*-glycosylated region (512–616) was determined by running c.d. spectra of GIC⁵⁰⁹ (Fig. 5). The spectrum shows a low α -helix content, with a very large proportion of β -structure (Table 4), as predicted from the amino acid sequence (Svensson *et al.*, 1989). The addition of ten residues (nine of which are *O*-glycosylated) as in GIC⁴⁹⁹ gave no significant change in the number of residues with α or β structure, suggesting that the region 499–508 has a 'random coil' structure. The addition of a further 28 residues (17 of which are *O*-glycosylated) in GIC shows a small increase in β -structure, but a larger increase in 'random-coil' structure. These

Fig. 6. Two-dimensional TOCSY n.m.r. spectra of (a) glucoamylase 1 (0.7 mM) and (b) G1C⁴⁹⁹ (1.3 mM) in ²H₂O at pH 8.0 and at 35 °C

Only the aliphatic regions are shown.

results strongly suggest that the whole of the *O*-glycosylated region (471–509) has a 'random-coil' structure, possibly with a small amount of β -structure.

Mobility of different regions in the protein

Local mobility was estimated by studying ¹H n.m.r. spectra of G1C⁴⁹⁹. Initial estimates were made by inspection of linewidths in one-dimensional spectra and were confirmed and extended using two-dimensional spectra (Fig. 6), making use of the sensitivity of two-dimensional spectra to molecular mobility. The linewidths and intensities of the signals in the spectrum of G1C⁴⁹⁹ are as would be expected for a protein of total *M_r* ~ 17000. In

the spectrum of glucoamylase most of the observed signals are clearly derived from the starch-binding domain, because they correspond very closely with those seen in the G1C⁴⁹⁹ spectrum, although they are of significantly lower intensity, even taking into account the lower concentration of the protein used. The signals from the catalytic domain are broader and of lower intensity than those from the starch-binding domain, as would be expected for a globular domain of higher *M_r*. These results are consistent with a model in which the two domains are linked by a semi-flexible rod, which allows considerable amounts of independent motion to the two domains, but still constrains their relative positions (Gerken *et al.*, 1989).

A third class of signals is represented by those seen close to the diagonal around 3.8 p.p.m., and around $F_1 = 1.3$ p.p.m. + $F_2 = 4.2$ p.p.m. These signals correspond respectively to sugar protons and to threonine H^α – H^β cross-peaks, and originate from the *O*-glycosylated peptide region. They are sharper and more intense than most of the other signals, implying that the *O*-glycosylated region possesses considerable local flexibility. Comparison of Figs. 6(a) and 6(b) shows that these signals are much sharper in G1C⁴⁹⁹ than in glucoamylase. This provides a clear indication that the *O*-glycosylated region has restricted mobility in the intact protein. Taken together with the c.d. results, the n.m.r. studies indicate a semi-rigid extended conformation for the *O*-glycosylated region, which acts to hold the two domains apart, but allows them some mobility relative to each other.

DISCUSSION

Glucoamylase 1 hydrolyses granular starch much more rapidly than does glucoamylase 2 (Svensson *et al.*, 1982). This is attributable to an additional starch-binding domain present in glucoamylase 1 but not in glucoamylase 2. Hayashida's group (Hayashida *et al.*, 1989a,b) reported that the *O*-glycosylated region played an active part in breaking the hydrogen bonds between adjacent chains in granular starch ('water cluster dissociation model'). However, this model requires tight binding between the covalently bound mannose oligosaccharides and the glucose residues of starch involving several hydrogen bonds per sugar residue. The apparent free energy of the interaction between the entire *O*-glycosylated region and granular starch is only –1.0 kJ/mol, compared with –25.0 kJ/mol for G1C⁵⁰⁹. It is therefore thermodynamically unlikely that the *O*-mannose oligosaccharides could replace a significant number of hydrogen bonds in the starch, given such a weak interaction for this portion of the molecule. In addition, the difference between glucoamylase 1 and glucoamylase 2 is about 102 amino acids (Svensson *et al.*, 1983), and both enzymes possess all of the glycosylated region. If the latter were the starch-binding domain, then it would be expected that both glucoamylase 1 and 2 would bind to, and hydrolyse, granular starch to a similar extent. Furthermore, partial deglycosylation of glucoamylase 1 does not affect binding to granular starch (Hayashida & Yoshino, 1978; Shendy *et al.*, 1984), whereas chemical modification of Trp-590 and of Trp-616 in the C-terminal domain (absent in glucoamylase 2) completely abolished binding (Flor & Hayashida, 1983; Svensson *et al.*, 1986). Consequently, the *O*-glycosylated region plays only a small role, if any, in the binding and hydrolysis of hydrogen-bond-ordered granular starch.

The starch-binding domain has been suggested by Svensson *et al.* (1989) to be the C-terminal 100 amino acids. This domain of *A. niger* glucoamylase 1 has five regions with sequences conserved in the C- or N-terminal domain of several different types of microbial amylase (Svensson *et al.*, 1989). The tight binding of G1C⁵⁰⁹, with almost no *O*-glycosylation, shows that assignment of this domain as the granular-starch-binding domain is correct.

Additionally, in the light of this information it is also clear that this domain is also responsible for binding soluble ligands, such as β -cyclodextrin (Belshaw & Williamson, 1991). The difference in β -cyclodextrin binding constants for G1C, G1C⁴⁹⁹ and G1C⁵⁰⁹ is probably due to a small conformational change in the binding domain brought about by the *O*-glycosylated region.

The general molecular architecture (catalytic domain, *O*-glycosylated linker, binding domain) is also seen in many cellobiohydrolases (Rouvinen *et al.*, 1990). Light-scattering studies have shown a long molecule consistent with a rigid and extended conformation of the *O*-glycosylated region. Another variation is found in sucrase-isomaltase, a membrane-bound protein. By electron microscopy it appears to consist of a catalytic domain, an *O*-glycosylated extended linker region and a membrane-anchoring domain (Cowell *et al.*, 1986).

The conformation of the *O*-glycosylated region of ovine submaxillary mucin has been studied in detail by light-scattering (Shogren *et al.*, 1989) and by ¹³C n.m.r. (Gerken *et al.*, 1989). The sugars are mainly sialic acid-GalNAc disaccharides. Studies on its asialo- and fully deglycosylated derivatives have shown that steric interactions between the *O*-linked sugar and the adjacent amino acid gave rise to a chain stiffening. The fully glycosylated derivative was a highly extended, relatively inflexible, random coil, which collapsed on removal of the carbohydrate. Our results on glucoamylase demonstrate that the short mannose oligosaccharides (α 1,2- or α 1,6-linked, with some α 1,6-linked glucose attached to mannose) also confer on the peptide an extended structure in which the individual mobility both of the sugars and of the peptide backbone is reduced. Since the arrangement of sugars is the same as that in mucin (densely packed, one to three sugar molecules per amino acid, *O*-linked to Ser/Thr) but the chemical nature of sugars is very different, the similarities in conformation suggest that the type of *O*-linked sugar variety is relatively unimportant in producing an extended polypeptide chain.

Thus the structure of glucoamylase I is an elongated one, with the catalytic domain linked to the binding domain via a linker region of considerable internal mobility, but of a generally extended and linear nature. The length per amino acid for a densely packed *O*-glycosylated sequence is ~ 0.25 nm, giving a total length of ~ 10 nm. Assuming that the catalytic domain and the binding domain are globular proteins, their dimensions would be ~ 6 nm and ~ 2.5 nm (diameter) respectively, and in intact glucoamylase I, these proteins would be separated by up to 10 nm. Whether communication between the two domains is facilitated by the *O*-glycosylated linker remains to be seen.

Other functions for *O*-glycosylation in glucoamylase have been suggested. The glucoamylase secreted by yeasts, although of much higher M_r , contains a Thr- and Ser-rich *O*-glycosylated region which directs the glucoamylase to the cell envelope and facilitates secretion (Yamashita, 1989). There is also a decrease in the thermostability of glucoamylase I after deglycosylation (Shendy *et al.*, 1984). These observations, together with the data reported here, suggest several functions of the *O*-glycosylated region. These are: (1) a linker maintaining an extended and semi-rigid structure; (2) prevention of aggregation and ensuring stoichiometric binding to a starch surface; (3) protection against proteolysis; and possibly (4) increasing thermostability and (5) directing the enzyme for secretion. Although this supports the suggestion (Jentoft, 1990) that the major function of *O*-glycosyl-

ation is to induce the peptide core to adopt a relatively stiff and extended conformation, in glucoamylase at least, the *O*-glycosylation also plays other important roles.

We thank Mr. Paul Brown for running the amino acid analysis, Ms. Linda Smith for help in collecting c.d. spectra and Dr. David Hornby for discussions. The work was funded by the Agriculture and Food Research Council (G.W., N.J.B.) and Science and Engineering Research Council (M.P.W.).

REFERENCES

- Bazzo, R. & Campbell, I. D. (1988) *J. Magn. Reson.* **76**, 358–361
- Belshaw, N. J. & Williamson, G. (1990) *FEBS Lett.* **269**, 350–353
- Belshaw, N. J. & Williamson, G. (1991) *Biochim. Biophys. Acta* **1078**, 117–120
- Cowell, E. M., Tranum-Jensen, J., Sjostrom, H. & Noren, O. (1986) *Biochem. J.* **237**, 455–461
- Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A. & Smith, F. (1956) *Anal. Chem.* **28**, 350–356
- Evans, R., Ford, C., Sierks, M., Nikolov, Z. & Svensson, B. (1990) *Gene* **91**, 131–134
- Flor, P. Q. & Hayashida, S. (1983) *Appl. Environ. Microbiol.* **45**, 905–912
- Gerken, T. A., Butenhof, K. J. & Shogren, R. (1989) *Biochemistry* **28**, 5536–5543
- Gill, S. C. & von Hippel, P. H. (1989) *Anal. Biochem.* **182**, 319–326
- Gunnarsson, A., Svensson, B., Nilsson, B. & Svensson, S. (1984) *Eur. J. Biochem.* **145**, 463–468
- Habeeb, A. F. S. A. (1972) *Methods Enzymol.* **25**, 457–464
- Hayashida, S. & Yoshino, E. (1978) *Agric. Biol. Chem.* **42**, 927–934
- Hayashida, S., Nakahara, K., Kanlayakrit, W., Hara, T. & Teramoto, Y. (1989a) *Agric. Biol. Chem.* **53**, 143–150
- Hayashida, S., Kuroda, K., Ohta, K., Kuhara, S., Fukada, K. & Sakaki, Y. (1989b) *Agric. Biol. Chem.* **53**, 923–929
- Jentoft, N. (1990) *Trends Biol. Sci.* **15**, 291–294
- Knowles, J., Teeri, T. T., Lehtovaara, P., Penttila, M. & Saloheimo, M. (1988) *Biochemistry and Genetics of Cellulose Degradation*, Academic Press, New York
- Kyriacou, A., Neufeld, R. J. & Mackenzie, C. R. (1989) *Biotechnol. Bioeng.* **33**, 631–637
- Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685
- Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951) *J. Biol. Chem.* **193**, 265–275
- Provencher, S. W. & Glockner, J. (1981) *Biochemistry* **20**, 33–37
- Rick, W. & Stegbauer, H. P. (1974) in *Methods of Enzymatic Analysis*, 2nd edn. (Bergmeyer, H. U., ed.), vol. 2, pp. 885–890, Verlag-Chemie, Weinheim
- Rouvinen, J., Bergfors, T., Teeri, T., Knowles, J. K. C. & Jones, T. A. (1990) *Science* **249**, 380–385
- Segrest, J. P. & Jackson, R. L. (1972) *Methods Enzymol.* **28**, 54–63
- Shendy, B. C., Rao, A. G. A. & Rao, M. R. R. (1984) *J. Biosci.* **6**, 601–612
- Shogren, R., Gerken, T. A. & Jentoft, N. (1989) *Biochemistry* **28**, 5525–5536
- Slater, H. S. & Codington, J. F. (1973) *J. Biol. Chem.* **248**, 3405–3410
- Subbaramaiah, K. & Sharma, R. (1988) *Starch* **40**, 182–185
- Svensson, B., Pedersen, T. G., Svendsen, I., Sakai, T. & Ottersen, M. (1982) *Carlsberg Res. Commun.* **47**, 55–69
- Svensson, B., Larson, K., Svendsen, I. & Boel, E. (1983) *Carlsberg Res. Commun.* **48**, 529–541
- Svensson, B., Clarke, A. T. & Svendsen, I. (1986) *Carlsberg Res. Commun.* **51**, 61–74
- Svensson, B., Hespersen, H., Sierks, M. R. & MacGregor, E. A. (1989) *Biochem. J.* **264**, 309–311
- Yamamoto, T., Davis, C. G., Brown, M. S., Schneider, W. J., Casey, M. L., Goldstein, J. L. & Russell, D. W. (1984) *Cell (Cambridge, Mass.)* **39**, 27–38
- Yamashita, I. (1989) *Agric. Biol. Chem.* **53**, 483–496