Carbohydrate is linked through ethanolamine to the C-terminal amino acid of *Trypanosoma brucei* variant surface glycoprotein

Anthony A. HOLDER

Department of Molecular Biology, The Wellcome Research Laboratories, Langley Court, Beckenham, Kent BR3 4BS, U.K.

(Received 6 September 1982/Accepted 15 October 1982)

The C-terminal amino acid of the variant surface glycoprotein from *Trypanosoma brucei* is glycosylated. For two variant proteins that terminate in an aspartic acid and a serine residue respectively, it was shown that the sugar side chain is linked through ethanolamine to the α -carboxyl group of the amino acid.

Sequential expression of genes for variant surface glycoproteins of *Trypanosoma brucei* is responsible for the antigenic variation shown by the blood stage parasite (for review, see Cross *et al.*, 1980; Borst & Cross, 1982). The complete amino acid sequence (Allen *et al.*, 1982) and nucleotide sequence (Boothroyd *et al.*, 1982) of one variant protein have been determined, as well as partial nucleotide sequences of several other variants (Matthyssens *et al.*, 1981; Rice-Ficht *et al.*, 1981; Boothroyd *et al.*, 1981). Comparison of the amino acid sequence at the C-terminus of several variant proteins (Holder & Cross, 1981; Allen *et al.*, 1982) with the corresponding nucleotide sequence shows that the polypeptide is synthesized with a hydrophobic 'tail' peptide that is absent from the mature protein. In addition to N-linked glycosylation of internal asparagine residues, the C-terminal amino acid of several mature variant glycoproteins is glycosylated (Holder & Cross, 1981). This C-terminal oligosaccharide is responsible for the immunological cross reaction observed between different variant glycoproteins. It is shown here that in two variant proteins the C-terminal carbohydrate is attached to the polypeptide through ethanolamine, which is in an amide linkage with the α -carboxyl group of the terminal amino acid.

Materials and methods

Variant surface glycoproteins were purified from two clones of *Trypanosoma brucei* strain 427 essentially as described previously (Cross, 1975, 1977). Variant glycoprotein 121 was obtained from clone MIT at 1.6 (subclone 121) and variant glycoprotein 55 was obtained from clone MIT at 1.7 (subclone 55).

The glycosylated C-terminal amino acid from each of these glycoproteins was purified by ion-exchange chromatography and gel filtration after extensive proteolytic digestion of the protein (Holder & Cross, 1981). Variant glycoprotein 121 was

digested with Pronase (proteinase from *Streptomyces griseus*; Sigma Chemical Co., Poole, Dorset, U.K.). Variant glycoprotein 55 was reduced and carboxymethylated and then digested with trypsin. The tryptic peptides were fractionated by gel filtration on Bio-Gel P10 (Bio-Rad Laboratories, Watford, Herts., U.K.) and fractions containing carbohydrate were subsequently digested with Pronase. The glycosylated amino acids were assayed for their ability to inhibit immunological cross reaction between heterologous variant glycoproteins (Holder & Cross, 1981).

Amino acid analysis after hydrolysis in 6M-HCl for 20 h *in vacuo* at 110°C was performed by using a Rank-Hilger Chromaspek J180 amino acid analyser equipped with *o*-phthalaldehyde fluorimetric detection. For the analysis of ethanolamine a single buffer system at pH 9.45 was used. The buffer was prepared by dissolving 3 g of boric acid, 8.76 g of NaCl and 0.35 ml of Brij 35 in water and adjusting to pH 9.45 with NaOH and making up to a final volume of 1 litre. For end-group and sequence determination the dansyl (5-dimethylaminonaphthalene-1-sulphonyl)-Edman procedure described by Hartley (1970) and the direct isothiocyanate procedure of Chang *et al.* (1978) were used.

Results

The C-terminal glycosylated amino acids from variant glycoproteins 55 and 121 were purified after extensive proteolytic digestion of the proteins. After acid hydrolysis the amino acid analysis showed the presence of serine and glucosamine and aspartic acid and glucosamine for the two glycoproteins respectively. In addition, an unidentified compound was present that eluted as a narrow peak on top of the broad NH₃ peak at a basic pH. When the amino acid analyser column was eluted with a single buffer at pH 9.45 the unidentified compound eluted coincidentally with authentic ethanolamine. Quantitative amino acid analyses of the glycosylated amino

acids are shown in Table 1. The neutral sugar compositions showing the presence of mannose and galactose in these materials have been published previously (Holder & Cross, 1981). In each sample there was an equimolar amount of ethanolamine and amino acid present.

The glycosylated amino acids were analysed by dansylation both before and after hydrolysis with 6M-HCl at 110°C. When the material from glycoprotein 55 was dansylated and then hydrolysed only dansylserine was detected. If the material was first hydrolysed and then dansylated, three products, dansylserine, dansylglucosamine and dansyl-ethanolamine, were observed. When the material from glycoprotein 121 was dansylated and then hydrolysed only dansylaspartic acid was detected. After prior acid hydrolysis, dansylation gave dansyl-aspartic acid, dansylglucosamine and dansyl-ethanolamine. Each derivative was coincident with standard compounds in three chromatography systems.

After coupling with dimethylaminoazobenzene isothiocyanate, followed by cleavage and conversion, only aspartic acid thiohydantoin was observed with the glycoprotein 121 material and serine thiohydantoin (together with its characteristic derivatives described by Chang *et al.*, 1978) was observed with the glycoprotein 55 material. For each sample a subsequent cycle produced no further thiohydantoin. This direct identification of the corresponding thiohydantoin demonstrated that the amino acid side chains were unmodified.

A single cycle of Edman degradation with phenyl isothiocyanate was performed on the two glycosylated amino acids, which exposed a new terminal amino group. After dansylation and acid hydrolysis dansylethanolamine was detected in both cases. If the acid hydrolysis was omitted after dansylation no derivative was observed. This indicates that the ethanolamine amino group is in an amide bond with the α -carboxy group of the amino acids and the ethanolamine hydroxy group is in a linkage, labile under strongly acidic conditions, presumably to the sugar moiety.

Table 1. Analysis of the C-terminal glycosylated amino acid from variant glycoproteins 55 and 121

The analyses were performed after hydrolysis in 6M-HCl for 20h. No correction for destruction or partial release of amino acids was made. Glucosamine values were corrected for 56% recovery under these conditions.

	Content (residues/mol)	
	Variant glycoprotein 55	Variant glycoprotein 121
Asp	—	1.00
Ser	1.00	—
Ethanolamine	1.10	1.01
Glucosamine	1.16	1.08

Discussion

Previously it had been shown that an oligosaccharide containing the neutral sugars mannose and galactose, together with glucosamine, was linked to the C-terminal amino acid of five variant surface glycoproteins (Holder & Cross, 1981). The results presented here are consistent with the oligosaccharide attached to the C-terminal serine of glycoprotein 55 and to the C-terminal aspartic acid of glycoprotein 121 through an ethanolamine group that is in an amide bond with the α -carboxy group of the amino acid.

The variant glycoprotein polypeptide is processed to remove a hydrophobic 'tail' peptide. Therefore, it seems likely that the ethanolamine-linked oligosaccharide is transferred to the α -carboxy group of the C-terminal residue exposed by proteolytic processing late in the biosynthesis of the glycoprotein. Alternatively the ethanolamine-linked sugar could have arisen by O-glycosylation of a serine residue with subsequent chain cleavage and decarboxylation of the terminal serine. Although all variant glycoproteins that are thought to terminate in aspartic acid have a serine as the first residue in the tail sequence, this is not true for the variants for which the mature protein terminates in serine (Boothroyd *et al.*, 1981, 1982; Matthyssens *et al.*, 1981; Rice-Ficht *et al.*, 1981). Therefore this sequence analysis would favour a mechanism based on transfer of the ethanolamine-linked oligosaccharide after removal of the tail peptide. An elucidation of the sequence of events in the biosynthesis and processing of this glycoprotein will be needed to resolve these alternatives.

I thank Dr. G. A. M. Cross and Ms. L. Davey for preparation of the glycoproteins and Dr. G. Allen for helpful discussions.

References

- Allen, G., Gurnett, L. P. & Cross, G. A. M. (1982) *J. Mol. Biol.* **157**, 527–546.
- Boothroyd, J. C., Paynter, C. A., Cross, G. A. M., Bernards, A. & Borst, P. (1981) *Nucleic Acids Res.* **9**, 4735–4743.
- Boothroyd, J. C., Paynter, C. A., Coleman, S. L. & Cross, G. A. M. (1982) *J. Mol. Biol.* **157**, 547–556.
- Borst, P. & Cross, G. A. M. (1982) *Cell* **29**, 291–303.
- Chang, J. Y., Brauer, D. & Wittmann-Liebold, B. (1978) *FEBS Lett.* **93**, 205–214.
- Cross, G. A. M. (1975) *Parasitology* **71**, 393–417.
- Cross, G. A. M. (1977) *Ann. Soc. Belge Med. Trop.* **57**, 389–399.
- Cross, G. A. M., Holder, A. A., Allen, G. & Boothroyd, J. C. (1980) *Am. J. Trop. Med. Hyg.* **29**, Suppl., 1027–1032.
- Hartley, B. S. (1970) *Biochem. J.* **119**, 805–822.
- Holder, A. A. & Cross, G. A. M. (1981) *Mol. Biochem. Parasitol.* **2**, 135–150.
- Matthyssens, G., Michiels, F., Hamers, R., Pays, E. & Steinert, M. (1981) *Nature (London)* **293**, 230–233.
- Rice-Ficht, A. C., Chen, K. K. & Donelson, J. E. (1981) *Nature (London)* **294**, 53–57.