Vol. 146, No. 3, 1987 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
August 14, 1987 Pages 1055-1063

MS and NMR ANALYSIS OF THE CROSS-REACTING DETERMINANT GLYCAN FROM TRYPANOSOMA BRUCEI BRUCEI MITat 1.6 VARIANT SPECIFIC GLYCOPROTEIN

Brigitte Schmitz, Roger A. Klein, and Imogen A. Duncan
Medical Research Council, Molteno Institute,
University of Cambridge, U.K.

Heinz Egge, Johannes Gunawan, and Jasna Peter-Katalinic
Institute for Physiological Chemistry, University of Bonn, F.R.G.

Ursula Dabrowski and Janusz Dabrowski
Max Planck Institute for Medical Research, Heidelberg, F.R.G.

Received June 16, 1987

SUMMARY: The cross-reacting determinant glycan from Trypanosoma brucei brucei MITat 1.6 is known to contain galactose, mannose and non-acetylated glucosamine [1,2,3]. The structural elucidation of this oligosaccharide has been impeded by an unusual non-glycosidic linkage to the peptide chain and a glycosidic linkage to inositol phosphate on either side of the oligosaccharide.

Using two different approaches for the isolation of the glycan, namely hydrolysis to give the oligosaccharide directly or pronase digestion to yield the glycan-containing C-terminal glycophosphopeptide, the structure of this glycan was elucidated by mass spectrometry and ^{1}H-NMR spectroscopy. There was evidence of heterogeneity in the glycan residue.

Antigenic variation in Trypanosoma brucei brucei is a well established phenomenon [4]. This haemoflagellate protozoan, pathogenic to cattle and used as a laboratory model for sleeping sickness in man, expresses a spectrum of immunologically distinct variant specific glycoproteins (VSGs) which cover the organism in a dense coat. As each antigenic variant develops in vivo, a new population of trypanosomes expresses this single antigenically homogeneous surface antigen. VSG exists in both a water-soluble form (sVSG) and a membrane-bound form (mfVSG), differing in a dimyristoyl diglyceride which is the anchor for mfVSG in the plasma membrane of the trypanosome, for all strains so far analyzed [5].

The VSG from T.b.brucei MITat 1.6 contains two glycan residues, as do other class I VSGs [6]. We have recently published structural details of the high-mannose glycan located some 50 amino

Abbreviations:
TLC, thin layer chromatography; CRD, cross reacting determinant; Hex, hexose; $GlcNH_2$, glucosamine; Man, mannose; Gal, galactose; VSG, variant specific glycoprotein; sVSG, soluble VSG; mfVSG, membrane form VSG; MITat, Molteno Institute antigenic type; DMSO, dimethyl sulphoxide; FAB-MS, fast atom bombardment mass spectrometry; GC-MS, gas chromatography mass spectrometry; NMR, nuclear magnetic resonance.

0006-291X/87 $1.50

acids from the C-terminus and N-glycosidically bound to an asparagine residue [7]. The other glycan moiety is attached to the C-terminus and forms the immunologically cross-reacting determinant (CRD) [1,8].

In this paper we report the structural elucidation of the cross-reacting determinant glycan from T.b.brucei MITat 1.6. The appropriate glycan-containing fraction was produced either by pronase digestion of VSG followed by affinity chromatography on ConA-Sepharose column, or by hydrolysis of the VSG using 60% HF, which is known to cleave phosphodiester bonds specifically under defined conditions [9,10], followed by purification on polyacrylamide columns and by TLC.

MATERIALS and METHODS

All solvents were of analytical reagent grade. HF (60%, technical grade) and LiOH (A.R.) were obtained from BDH; silica gel 60 TLC plates from Merck, Iatrobeads from Macherey and Nagel (Düren, FRG) and Biogel P2 and P4 resins from Biorad. Concanavalin A conjugated Sepharose was obtained from Pharmacia (Uppsala). Fast atom bombardment mass spectrometry (FAB-MS) was performed as described by Egge [11]. NMR spectroscopy was as described by Dabrowski [29-31]. Trypanosomes (T.b.brucei MITat 1.6) were prepared by passage through rats and obtained by aortic puncture, followed by separation from red and white cells by DE52 chromatography [12]. sVSG and mfVSG were purified from these trypanosomes as previously described [12,13].

A. Isolation of the glycophosphopeptide by pronase digestion.

Pronase digestion was performed essentially according to Ferguson [5]. The digest of 60 mg of sVSG was clarified by centrifugation, concentrated to 1 ml and desalted by passage over Biogel P2 (50 cm x 2.5 cm) with water as eluent. Carbohydrate containing fractions were detected by the phenol-sulphuric acid procedure [14], pooled and lyophilized. N-acetylation with acetic anhydride was performed according to [15]. α-Galactosidase treatment of the glycophosphopeptide was performed according to the method of Hanfland et al. (28).

Separation of the N-acetylated glycophosphopeptide from the high mannose N-glycan was achieved on a 25 cm x 1.6 cm column of ConA-Sepharose as described by Dulaney [16]. The glycophosphopeptide was eluted by 75 ml of 0.1 M acetate buffer pH 5.0 together with peptides. Further elution with 75 ml of 0.1 M acetate buffer pH 5.0 plus 0.5 M NaCl and 0.2 M methyl α-mannoside yielded the high mannose glycopeptides. The fractions were desalted by passage over Biogel P2.

B. Isolation of the oligosaccharide by treatment with HF.

The hydrogen fluoride treatment was essentially performed as described by Fischer [9]. Hydrogen fluoride (60%) was pre-cooled to between -20 and -30 ^{0}C and then added to the VSG at 0 ^{0}C with efficient mixing (37.2 mmol HF (2 ml) per micromol VSG). All these operations were conducted in PTFE or polyethylene tubes. The sample was kept on ice for 14-16 hours before the HF was rapidly neutralized at 0 ^{0}C with saturated LiOH (7 ml LiOH per ml HF); the pH was adjusted with glacial acetic acid to pH 7. The precipitated LiF was removed by centrifugation and the supernatant concentrated to about 1 ml on a rotary evaporator (40 ^{0}C and 5 torr) before being desalted on a Biogel P2 column. Orcinol positive fractions were combined , freeze dried and then peracetylated with acetic anhydride/pyridine 1:1 by volume (1 ml) overnight at room temperature, followed by 1 hour at 60 ^{0}C. The peracetylated sample was dried under N_2 and chromatographed on TLC plates (Merck Kieselgel 60). Each plate was developed 3 times with benzene/ethanol 88/12 (v/v). The bands were identified by spraying the plates with water and were extracted with $CHCl_3$/ethanol 80/20 (v/v). CRD activity was measured using a standard radio-immunoassay procedure with affinity-column purified CRD-antibody [17] and ^{125}I-goat anti-rabbit IgG antibodies.

C. Permethylation of the oligosaccharide.

The oligosaccharide was permethylated in DMSO solution, after deacetylation, using sodium hydroxide and methyl iodide, based upon the method of Ciucanu and Kerek [18].

D. Methylation analysis.

The permethylated oligosaccharide was hydrolyzed overnight (16 h) in a sealed tube at 80 ^{0}C using 80% acetic acid/0.7N H_2SO_4. The oligosaccharide had been acetylated prior to hydrolysis with acetic anhydride (or a mixture of CD_3COOH and CH_3COOH) and pyridine 1:1 inorder to protect the

glucosamine residue. The hydrolyzate was reduced with sodium borohydride and then acetylated. The alditol acetates were analyzed by GC-MS using an LKB9000 mass spectrometer and 25 m fused silica columns containing either BP15 or BP1 (SGE) as the stationary phase (150-250 ^{o}C temperature programmed, at 3 ^{o}C per minute; linear gas velocity 40 cm/sec). The basis of this method is well documented [19-21]. The permethylated alditol acetates were identified by their gas chromatographic retention times compared to reference compounds.

E. Methanolysis.

The methyl glycoside of glucosamine was prepared from the oligosaccharide according to published procedures (27). After acetylation with acetic anhydride-pyridine 1:1 the product was analyzed by GC-MS.

RESULTS and DISCUSSION

The oligosaccharide, which forms part of the CRD, is not, as in other glycoproteins, bound through a glycosidic linkage to the peptide chain, but via a bond at the non-reducing end of the oligosaccharide to ethanolamine, which in turn is amide-linked to the C-terminal aspartate or serine [22]. The question arises as to how the ethanolamine residue and the oligosaccharide are linked. As discussed by Ferguson [2], a second phosphodiester linkage would be quite likely. Based on this assumption we expected that treatment of the VSG with 60% HF would result in selective cleavage of the phosphodiester bonds on either side of the oligosaccharide, leaving the rest of the glycoprotein substantially intact [10]. Our results indicate rather clearly that an oligosaccharide species is produced by this procedure, which is the same for mfVSG and sVSG as demonstrated by FAB-MS (data not shown).

The FAB-mass spectrum of the peracetylated oligosaccharide mixture before TLC separation yielded a major molecular ion ($M+Na^{+}$) at m/z 2470 nominal mass, corresponding to an oligosaccharide containing six hexose units, equivalent to $C_{102}H_{137}NO_{67}Na$. Weaker molecular ion signals were observed at m/z 2182, 2758 and 3046 indicating the presence of oligosaccharide species containing 5, 7 and 8 hexose units respectively.

The mass spectrum of the main component purified by TLC is shown in Figure 1. Sequence ions were observed at m/z 331 for a terminal hexose and at m/z 619 and 907 indicating that three hexose

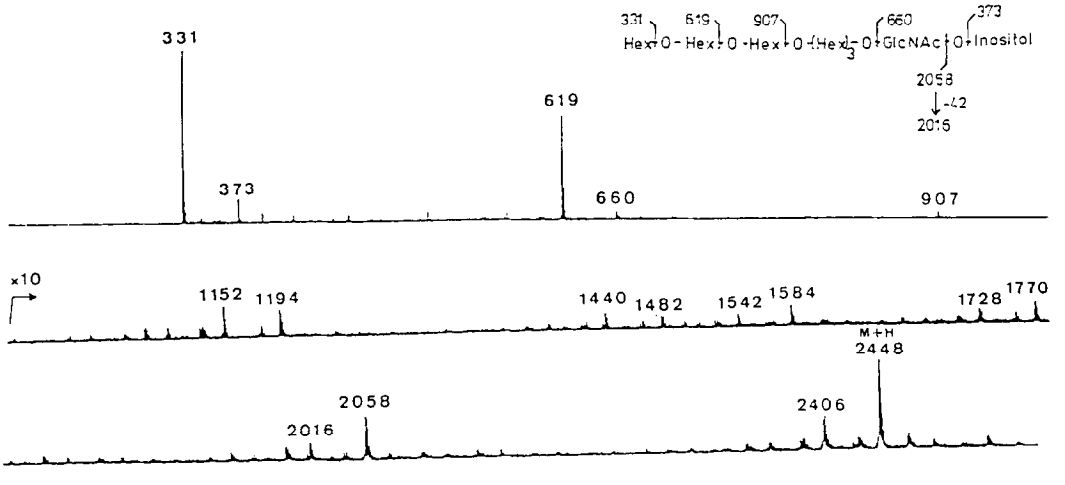

Figure 1. FAB-MS and fragmentation scheme of the peracetylated oligosaccharide Hex_6GlcNH_2-Inositol.

units, each contributing 288 mass units, are linked together. Fragments which correspond to higher saccharides cannot be detected, because they could be hidden under the isotope peaks of the fragments belonging to the series which is produced by cleavage of the GlcNAc-inositol linkage and subsequent loss of 3-6 internal hexose units (m/z 2058, 1770, 1482 and 1194). The fragments which accompany this sequence, at masses 42 mass units lower, represent loss of a ketene group and are typical of peracetylated derivatives. The mass spectrum of the peracetylated Hex_5-GlcNAc-inositol is in agreement with the mass spectrum of the Hex_6-derivative. Fragments m/z 331, 619 and 907 produced from a terminal mono-, di- and trisaccharide indicate that in the case of a branched structure, one branch contains at least a trisaccharide. The mass spectrum of the permethylated Hex_5-GlcNAc-inositol was also analyzed; this mass spectrum confirmed the assignments for the peracetylated derivative. The permethylated oligosaccharide furnished $(M+H)^+$ at m/z 1530 and $(M+Na^+)^+$ at m/z 1552. Fragments corresponding to a terminal monosaccharide and disaccharide at m/z 219 and 423 were observed together with the ions resulting from the elimination of methanol (m/z 187, 391); however, no ion corresponding to a trisaccharide unit was observed. This is not altogether surprising since the decrease in intensity for mono- to disaccharide (m/z 219 → 423) would suggest that such a trisaccharide fragment would be of extremely low intensity; this effect is more pronounced for permethylated than for peracetylated derivatives (32). Two further series of ions are produced by cleavage of the GlcNAc-inositol linkage (m/z 1280, 1076, 872, 668 and 464) and by loss of internal hexoses from Hex_n-GlcNAc-inositol where n=0-3 (m/z 1090, 886, 682 and 478).

Consistent mass spectral data for the oligosaccharide and for the glycophosphopeptide, isolated after pronase digestion, were obtained. The peracetylated product produced a major molecular ion $(M-H)^+$ at m/z 2662 (nominal mass) corresponding to the presence of 6 hexose units with a molecular composition $C_{104}H_{142}N_3O_{73}P_2$. As for the oligosaccharide liberated by HF treatment, molecular species with 5, 7 or 8 hexose units were also detected with molecular ions at m/z 2374, 2950 and 3238. In the FAB-MS of the glycophosphopeptide fragmentations at m/z 331, 619, 907 and at low abundance m/z 1195, serve as indication for a branched structure containing up to four sugar residues at the non-reducing end, which is not linked to the group carrying the ethanolamine.

Mass spectrometric analysis of the glycophosphopeptide after α-galactosidase treatment shows a shift of the molecular ions corresponding to the loss of 2-3 galactose residues (m/z 2950 → 2662 → 2374 → 2086). This result provides evidence for galactose residues at the free non-reducing branch being α-glycosidically linked.

Methylation analysis confirmed the results of the FAB-MS, and demonstrated that the branching point is a mannose residue with linkages in position 1, 3 and 6 (Table 1). The permethylated alditol acetates of Hex_5-GlcNAc-Inositol show that there is a terminal galactose and mannose, and an internal galactose and mannose substituted in position 6 and 2 respectively.

Methanolysis of the oligosaccharides was performed to give information about the linkages of the glucosamine which could not be determined using the methylation procedure because of known difficulties (20). GC-MS of the methyl glycosides provided evidence for the glucosamine being substituted in the 4 position.

We have also evidence from ^{1}H-NMR analysis of the Hex_6-GlcNAc-inositol which confirms that: all galactose as well as mannose residues are in α-linkage; the branching point is a 1,3,6-substituted

Table 1. GC-MS analysis of the permethylated alditol acetates, or peracetylated methyl glycosides of Hex_5-$GlcNH_2$-Inositol

Substituted sugar	position of OCH_3	Linkage
Mannose	2,4	1,3,6
Mannose	3,4,6	1,2
Mannose	2,3,4,6	1
Galactose	2,3,4	1,6
Galactose	2,3,4,6	1
Glucosamine[1]	1,3,6	1,4

[1]: determined as the methyl glycoside; all others determined as the permethylated alditol acetates. The sugars were present at equimolar ratios, with an overall mannose to galactose ratio of 3:2.

mannose residue; the glucosamine is in a 1→4 linkage with the inositol; the branch-point mannose is 1→4 linked to the glucosamine; the terminal mannose is 1→3 linked to the branch-point mannose.

Both the native and the peracetylated Hex_6-$GlcNH_2$-Inositol have been analysed with the aid of one-dimensional (1D) and two-dimensional (2D) correlated (COSY) ^{1}H-NMR spectroscopy at 500 MHz. For peracetylated sugars, glycosylation sites are readily determinable by the ca. 1 ppm relative upfield shift of the resonances of the protons at these sites [29-31]. The linkages shown in the partial formula of Hex_6-GlcNAc-Inositol given in Table 2 were established in this way. The connectivities substantiating the relevant assignements are indicated in the COSY spectrum (Fig. 2b). They have been additionally verified by analyzing a relayed coherence transfer spectrum (not shown) in which protons are correlated not only with their immediate neigbours (as in COSY), but also with the next

Table 2. Proton chemical shifts for a solution of the peracetylated Hex_6-GlcNAc-Inositol[a] in $CDCl_3$ at 303 K

	V	IV	III	II	I
Protons		→2Manα1	→ 6Manα1	→ 4GlcNAcα1	→ 4Inositol
	Manα1	→	3		
H-1	4.91	4.92	4.92	5.15	5.16
H-2	5.27	4.05	4.90	4.20	5.37
H-3	5.31	5.19	4.02	5.07	5.18
H-4	5.32	5.38	5.20	3.73	4.23
H-5	4.15	3.99	3.97	3.86	5.19
H-6	4.25	4.10	3.48	4.09	5.56
H-6'	4.25	4.27	3.77	4.56	
HN				6.08	

[a]: the assignments refer to the partial structure of the main component of this fraction (see text); the α-galactose residues, which are linked to Man-IV according to the MS data, have not been fully assigned.

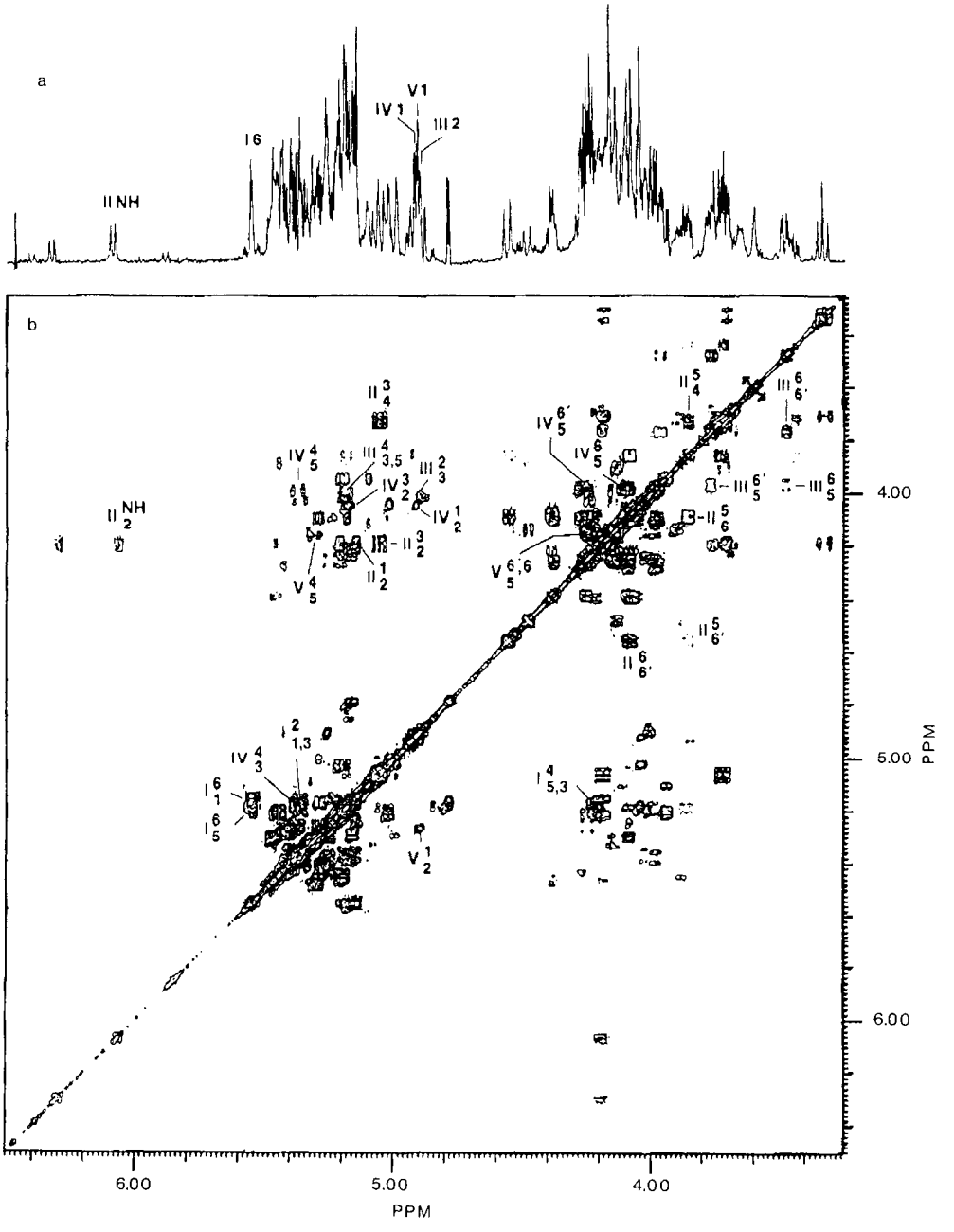

Figure 2a and 2b. Part of the 500 MHz ^{1}H-NMR spectra of a solution of peracetylated Hex_6-GlcNAc-Inositol in $CDCl_3$ at 303 K. (a) The resolution-enhanced 1D spectrum. Arabic numerals (or NH) denote proton positions in the given sugar residue specified by a Roman numeral (see formula in Table 2). (b) Contour plot of the COSY spectrum. The labelling of the correlation cross-peaks should be read as in the following example: IV^4_5 refers to Man-IV and indicates a coupling between H-4, whose chemical shift is defined by the F_2 coordinate (horizontal scale), and H-5, whose shift can be read on the F_1, vertical scale. To avoid crowding only one cross-peak of each pair is labelled, either above or below the diagonal.

but one neighbours, e.g., H-1 with H-2 and directly with H-3. These results are in accordance with the results of the methylation analysis for Hex_5-GlcNAc-Inositol (Table 1). Furthermore, the 1D and

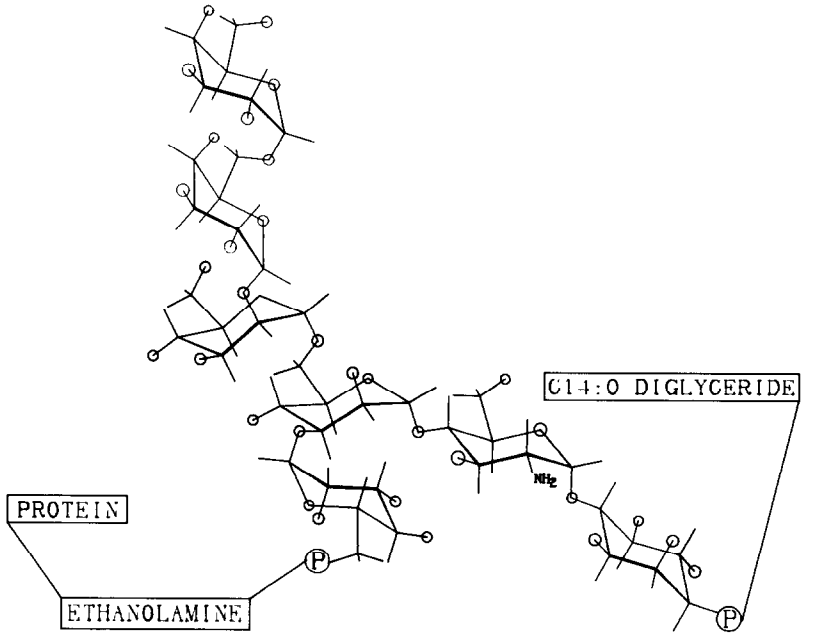

Figure 3. Proposed structure for the CRD-oligosaccharide: G = galactose; I = inositol; M = mannose; N = glucosamine. The structure is shown containing five hexoses.

the COSY spectrum of the native (non-acetylated) sample have shown that the Hex_6 moiety consists of three α-Man and three α-Gal residues. The α-anomeric configuration of Gal residues follows directly from the small $^3J_{1,2}$ coupling constants visible in the spectrum. As for mannoses, $^3J_{1,2}$ is small for both configurations due to the equatorial position of H-2. However, it was shown (ref. in [31]) that H-1 → H-2 magnetization transfer in COSY spectra is much more effective for α- than for β-mannoses due to the larger coupling constant in the case of the former ($^3J_{1,2}$ ~ 2 Hz for α- and <1Hz for β). In fact, the β-$Man\frac{1}{2}$ connectivity could only be observed with large samples [31]. Since the COSY spectrum recorded here for a ca. 1 mg sample exhibited three $Man\frac{1}{2}$ cross-peaks as strong as the majority of other cross-peaks, it was concluded that all three Man residues are α-anomers.

It should be added that the peracetylated Hex_6-GlcNAc-Inositol fraction was heterogenous to some extent, as seen from the four N-H signals of relative 53:24:12:11 intensities (Fig. 2a). The above analysis refers to the main component.

The structure for the Hex_5-component which is consistent with our present data and those already published (2,23), is shown in Figure 3. This structure also incorporates the findings of Ferguson et al. (24) as quoted by Low (25), namely that: (i) the phosphodiester bond is part of a terminal 6-phospho-mannose residue; (ii) there is a core structure containing Man_2GlcNH_2; (iii) this core structure is modified by 0-8 galactose residues; (iv) the inositol residue is linked glycosidically to the glucosamine through the 4-position.

In contrast to Ferguson [24] we found there to be three mannose residues, rather than just two. However, the NMR data and the methylation analysis showed unequivocally that this was so. There can be no confusion in the methylation analysis between galactose and mannose as the retention times for the galactose and mannose derivatives are very different [19].

Radioimmunoassay for CRD activity demonstrated that the CRD titre of the purified oligosaccharide after HF hydrolysis was reduced almost to zero. Mass spectral analysis yields molecular weights for the glycophosphopeptide and the oligosaccharide which agree, so that degradation of the oligosaccharide by treatment with HF can be excluded. Therefore we would like to suggest that one, or both phosphate groups may be involved in the maintenance of some structural feature which is essential for CRD activity. Treatment with α-galactosidase did not reduce CRD activity, suggesting that at least not all of the galactose residues are necessary for CRD activity, although it has been shown by Cross [26] that a variant containing no galactose exhibits a much reduced CRD activity. The heterogeneity found for this and other variants (1) may include differences in the substitution and number of galactose as well as mannose residues; however, there must be a structural feature which is common to all since these glycoproteins are recognized by the anti-CRD antibody.

ACKNOWLEDGEMENTS

We should like to acknowledge support from the WHO/World Bank/UNDP Special Programme for Research and Training in Tropical Diseases, the European Community (01.104) (R.A.K. and H.E.) and the Fritz Thyssen Stiftung (U.D.). We should also like to acknowledge expert technical assistance from Salih Eresh.

REFERENCES

1. Holder, A.A. (1985) Current Topics in Microbiology and Immunology 117, 57-74.
2. Ferguson, M.A.J., Low, M.G. and Cross, G.A.M. (1985) J. Biol. Chem. 260, 14547-14555.
3. Strang, A.M., Williams, J.M., Ferguson, M.A.J., Holder, A.A., Allen, A.K. (1986) Biochem. J. 234, 481-484.
4. Boothroyd, J.C. (1985) Ann. Rev. Microbiol. 39, 475-502.
5. Ferguson, M.A.J., Haldar, K. and Cross, G.A.M. (1985) J. Biol. Chem. 260, 4963-4968.
6. Turner, M.J. (1985) Ann. Inst. Pasteur/Immunol. 136C, 41-49.
7. Egge, H., Gunawan, J., Peter-Katalinic, J., Schmitz, B., Duncan, I.A., and Klein, R.A. (1986), in "Topics in Lipid Research - from Structural Elucidation to Biological Function", (Klein, R.A. and Schmitz, B., eds.), pp. 142-150, Royal Society of Chemistry, London.
8. Holder, A.A. and Cross, G.A.M. (1981) Mol. Biochem. Parasitol. 2, 135-150.
9. Fischer, W., Ishizuka, I. Landgraf, H.R. and Herrmann, J. (1973) Biochim. Biophys. Acta 296, 527-545.
10. Toon, P., Brown, P.E. and Baddiley, J. (1972) Biochem. J. 127, 399-409.
11. Egge, H. and Peter-Katalinic, J. (1985) in "Mass Spectrometry in the Health and Life Sciences", (Burlingame, A.L. and Castagnoli, N., eds.), Analytical Chemistry Symposia Series 24, 401-424.
12. Cross, G.A.M. (1984) J. Cell. Biochem. 24, 79-90.
13. Jackson, D.G. Owen, M.J. and Voorheis, H.P. (1985) Biochem. J. 230, 195-202.
14. Dubois, M., Gilles, K.A., Hamilton, J.K., Rebers, P.A. and Smith, F. (1956) Anal. Biochem. 28, 350-356.
15. Takasaki, S., Mizuochi, T. and Kobata, A. (1982) Methods in Enzymology 83, 263-269.
16. Dulaney, J.T. (1979) Anal. Biochem. 99, 254-267.
17. de Almeida, M.L.C. and Turner, M.J. (1983) Nature Lond. 302, 349-352..
18. Ciucanu, I. and Kerek, S. (1984) Carbohydr. Res. 131, 209-217.
19. Lindberg, B. (1972) Methods in Enzymol. 28, 178-195.
20. Lindberg, B. and Lönngren, J. (1978) Methods in Enzymol. 50, 3-33.
21. McNeil, M., Darvill, A.G., Åman, P., Franzén, L-E. and Albersheim, P. (1982) Methods in Enzymol. 83, 3-45.
22. Holder, A.A. (1983) Biochem. J. 209, 261-262.
23. Schmitz, B., Klein, R.A., Egge, H. and Peter-Katalinic, J. (1986) Mol. Biochem. Parasitol. 20, 191-197.

24. Ferguson, M.A.J., Dwek, R.A., Homans, S.W. and Rademacher, T.W. (1987) in NATO/ASI Volume: Host-Parasite Cellular and Molecular Interactions in Protozoal Infections (Chang, K-P., ed.), Springer Verlag, Heidelberg, in press.
25. Low, M.G. (1987) Biochem. J. 244, 1-13.
26. Cross, G.A.M. (1979) Nature Lond. 277, 310-312.
27. Fournet, B., Strecker, G., Leroy, J., Montreuil, J. (1981) Anal. Biochem. 116, 489-502.
28. Hanfland, P., Egge, H., Dabrowski, U., Kuhn, S., Roelcke, D. and Dabrowski, J. (1981) Biochem. 20, 5310-5315.
29. Dabrowski, U., Egge, H. and Dabrowski, J. (1983) Arch. Biochem. Biophys. 224, 254-260.
30. Dabrowski, J., Egge, H. and Dabrowski, U. (1983) Carbohydr. Res. 114, 1-9.
31. Dabrowski, J. (1987) in "Two Dimensional NMR Spectroscopy for Chemists and Biologists (Croasmun, W.R. and Carlson, M.L., eds.), Verlag Chemie International, Deerfield Beach, Florida.
32. Egge, H., Michalski, J.C. and Strecker, G. (1982) Arch. Biochem. Biophys. 223, 318-326.