Vol. 104, No. 1, 1982 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
January 15, 1982 Pages 58-65

N-ACETYL-D-GALACTOSAMINYL-β-(1→4)-D-GALACTOSE : A TERMINAL NON-REDUCING STRUCTURE IN HUMAN BLOOD GROUP Sd^a-ACTIVE TAMM-HORSFALL URINARY GLYCOPROTEIN

A.S.R. Donald, Cecilia P.C. Soh, Winifred M. Watkins and W.T.J. Morgan

Division of Immunochemical Genetics, MRC Clinical Research Centre, Watford Road, Harrow, Middlesex HA1 3UJ, U.K.

Received November 18, 1981

SUMMARY. Human Sd^a-active Tamm-Horsfall urinary glycoprotein labelled with galactose oxidase and tritiated sodium borohydride was found to contain both galactose and N-acetylgalactosamine as [^{3}H]-labelled terminal non-reducing sugars. Fragmentation of the macromolecule achieved by hydrazinolysis and acid hydrolysis was followed by fractionation of the degradation products by gel filtration, ion exchange and paper chromatography. A major product was a disaccharide which contained unlabelled galactose and [^{3}H]-labelled N-acetylgalactosamine. Sugar analysis, sodium borohydride reduction, methylation analysis and enzymic degradation enabled the structure N-acetyl-D-galactosaminyl-β-(1→4)-D-galactose to be assigned to the disaccharide.

INTRODUCTION

Human Tamm-Horsfall (T-H) urinary glycoprotein is a macromolecule composed of about 70% protein and 30% carbohydrate (see 1). The carbohydrate moiety of T-H preparations isolated from pooled urine was found to contain the sugars D-galactose, D-mannose, L-fucose, N-acetyl-D-glucosamine, N-acetyl-D-galactosamine and N-acetylneuraminic acid (1). Recently, however, it was shown that the inherited human blood group Sd^a character (see 2) is carried in urine on T-H glycoprotein (3) and there are two immunologically distinct forms of the glycoprotein, namely Sd(a+) and Sd(a-). Analyses of individual preparations isolated from the urine of blood group Sd(a+) and Sd(a-) donors revealed a consistent difference in the N-acetyl-D-galactosamine content (4). The value for this sugar in four Sd^a-active preparations fell within the range of 1-2% whereas four Sd^a-inactive preparations contained 0.2% or less. These findings were compatible with earlier indirect inferences that N-acetyl-D-galactosamine plays an important role in Sd^a blood group specificity on the red cell (5,6).

The carbohydrate chains in T-H glycoprotein are believed to be joined to peptide through N-acetylglucosamine-asparagine linkages (see 1) but their

0006-291X/82/010058-08$01.00/0

structures have not yet been elucidated. This paper reports the isolation of a terminal non-reducing structure, characterised as N-acetyl-D-galactosaminyl-β-(1→4)-D-galactose, from the degradation products of human Sd^a-active T-H glycoprotein.

MATERIALS AND METHODS

Anhydrous hydrazine and sodium borodeuteride were purchased from Aldrich Chemical Co. Ltd., Dorset, U.K. Sodium boro[^{3}H]hydride (8.7 Ci/mmol) was obtained from the Radiochemical Centre, Amersham, U.K. Bio-Gel P-4 and the ion exchange resins AG 1 X4 (200-400 mesh) and AG 50 X8 (200-400 mesh) were purchased from Bio-Rad Laboratories Ltd., Watford, U.K. Galactose oxidase (E.C. 1.1.3.9) and Jack bean β-N-acetylhexosaminidase (E.C. 3.2.1.30) were the products of Sigma Chemical Co. Ltd.

The A-active disaccharide, GalNAcα(1→3)Gal, was isolated as described by Cote and Morgan (7). Chemically synthesised GalNAcα(1→4)Gal-O-$(CH_2)_8CO_2CH_3$ was the gift of Dr D. Baker, Chembiomed Ltd., Alberta, Canada.

An α-N-acetylgalactosaminidase (E.C. 3.2.1.49) free from β-N-acetylgalactosaminidase and a preparation containing both α- and β-N-acetylgalactosaminidase activities were isolated from a crude extract of Trichomonas foetus (8) by chromatography on Sephacryl S-200 (Pharmacia Ltd.). The activities were measured with α- and β-p-nitrophenyl N-acetylgalactosaminides.

Paper chromatography was carried out on Whatman No. 40 paper in solvent a (ethyl acetate-pyridine-water, 2:1:1 by vol), solvent b (ethyl acetate-pyridine-water, 10:4:3 by vol), or solvent c (ethyl acetate-pyridine-acetic acid-water, 5:5:1:3 by vol).

Re-N-acetylation was carried out in 25% methanol by the addition of 1 M $NaHCO_3$ and acetic anhydride. Removal of salts was achieved by passage of solutions through columns of AG 1 (acetate) and AG 50 (H^+) resins.

Methods for the analysis of neutral sugars and their alcohols by g.l.c. and of hexosamines and hexosaminitols on an amino acid analyser are described elsewhere (9). The isolated disaccharide was methylated with methyl iodide in dimethylformamide with $BaO/Ba(OH)_2$ as catalyst (10). The methylated disaccharide was hydrolysed by the method of Yang and Hakomori (11).

T-H glycoprotein was isolated from the urine of an Sd(a+) individual (W.M.) by the procedure of Tamm and Horsfall (12). The glycoprotein recovered after precipitation three times with 0.58 M NaCl was dissolved in water, thoroughly dialysed against water, centrifuged at 30,000g for 20 min and dried from the frozen state. The dry material was sequentially extracted with 9:1, 2:1 and 1:2 mixtures of chloroform-methanol to remove lipid material, washed with methanol, and then dissolved in water, dialysed against water, centrifuged for 1 h at 100,000g and the clear colourless solution was dried from the frozen state.

A tritium label was introduced into the T-H glycoprotein by a modification of the procedure of Gahmberg (13). Sialic acid was removed from the glycoprotein (50 mg) by treatment for 1 h at 80° with 0.05 M H_2SO_4 and the treated preparation was exhaustively dialysed against water. Galactose oxidase (67.5 units) was added to the desialysed glycoprotein solution and the mixture was incubated at 37° for 2 h. Tritiated sodium borohydride (50 mCi in 0.1 ml 0.01 M NaOH) was added to the solution, followed after 30 min at room temperature by the addition of excess unlabelled sodium borohydride (50 mg). The resultant mixture was left to stand overnight at 4° and the solution was then thoroughly dialysed against water. The labelled glycoprotein was purified by chromatography on a column (90 x 1.5 cm) of Sepharose 4B. Elution with 0.01 M NaCl, and examination of the fractions for radioactivity, gave a single radioactive peak which corresponded with the peak recorded by absorbance at 280 nm.

RESULTS AND DISCUSSION

Tritium labelling of T-H glycoprotein. Desialysed Sd^a-active T-H glycoprotein oxidised with galactose oxidase and reduced with tritiated sodium borohydride was subjected to acid hydrolysis (1 M HCl for 16 h at 100°) followed by re-N-acetylation, desalting and chromatography in solvent b. Two radioactive spots were detected which co-chromatographed with galactose and N-acetylgalactosamine, respectively. Since galactose oxidase oxidises only terminal non-reducing galactose and N-acetylgalactosamine residues in glycoproteins (14) it follows that these sugars can both occupy terminal positions in the carbohydrate chains in the desialysed T-H glycoprotein. A procedure was therefore devised which would yield N-acetylgalactosamine-containing oligosaccharides.

Hydrazinolysis of the T-H glycoprotein. A solution of [^{3}H]-labelled T-H glycoprotein (8 mg, 65 x 10^6 c.p.m.) was mixed with a solution of non-radioactive T-H glycoprotein (500 mg) derived from the same Sd(a+) donor. The mixture was thoroughly dried and treated in a sealed tube with anhydrous hydrazine (6 ml) for 16 h at 105°. The hydrazinolysis products, after repeated evaporation with toluene to remove traces of hydrazine, were dried *in vacuo* over P_2O_5, dissolved in 5% acetic acid (5 ml) and fractionated on a column (80 x 2.2 cm) of Bio-Gel P-4. The eluate obtained with 5% acetic acid was divided into three fractions on the basis of the peaks detected by a refractive index monitor : Fraction I, elution volume 60-90 ml; Fraction II, elution volume 91-160 ml; Fraction III, elution volume 161-280 ml. Aliquots of these fractions were re-N-acetylated, desalted and analysed for N-acetylgalactosamine. Separate aliquots were counted for radioactivity. Fractions I and II, which together contained 84% of the radioactivity and 76% of the N-acetylgalactosamine in the starting material, were combined and evaporated to dryness.

Acid hydrolysis of hydrazinolysis products. Pooled Fractions I and II were hydrolysed with 0.5 M HCl (10 ml) for 6 h at 100°. Coloured decomposition products were removed by chromatography on a column (10 x 1.5 cm) of AG 50

(H^+) resin and elution with 2 M HCl. The eluate was freed from HCl on a column (15 x 3 cm) of AG 1 (acetate) resin, and, after concentration to 5 ml, was loaded onto a column (50 x 1 cm) of AG 50 (H^+) resin equilibrated with water. The column was washed with water (20 ml) to remove uncharged fragments (Fraction A, 28 x 10^6 c.p.m.) and then eluted with 0.33 M HCl. The fractions (3 ml) were monitored for radioactivity and for ninhydrin-positive material. A small radioactive peak (4.4 x 10^5 c.p.m.) was eluted in Fractions 16-21 (pooled to give Fraction B) and a larger radioactive peak (8.4 x 10^6 c.p.m.) was eluted in Fractions 22-29 (pooled to give Fraction C). Each of these radioactive peaks coincided with a ninhydrin-positive peak. Fraction B has not yet been further examined. Fraction C was freed from HCl, re-N-acetylated, desalted and examined by paper chromatography in solvent a. Radiochromatogram scanning revealed a strong radioactive peak with an $R_{Lactose}$ value 1.03 (C1) and a weaker peak (C2) which co-chromatographed with N-acetylgalactosamine. Staining with silver nitrate reagent showed the same two spots; no other components were detected. The two components were separated by preparative paper chromatography in solvent a and C1 was eluted from the paper. The eluate was passed through columns (1.5 x 0.3 cm) of AG 1 (acetate) and AG 50 (H^+) resins, evaporated to dryness and redissolved in water (1 ml).

Characterisation of Fraction C1. Chromatography of C1 in three different solvent mixtures revealed only one component when the papers were either scanned for radioactivity or stained with silver nitrate reagent. The compound had the following $R_{lactose}$ values : solvent a, 1.03; solvent b, 1.14; solvent c, 1.12. Hydrolysis of Fraction C1 with 2 M HCl for 2 h at 100°, followed by removal of acid and chromatography in solvent b, gave only one radioactive peak with the same mobility as galactosamine. Re-N-acetylation of the hydrolysis products followed by chromatography in solvent a again yielded one radioactive peak which now co-chromatographed with N-acetylgalactosamine. The [^{3}H]-labelled terminal non-reducing sugar in Fraction C1 was thus shown to be N-acetylgalactosamine.

Analysis of C1 for hexosamines gave 2.12 mg of N-acetylgalactosamine and less than 0.05 mg of N-acetylglucosamine. On g.l.c. analysis of the trimethyl-

silylated derivatives of C1, galactose (1.60 mg) was found to be the only other sugar present in the compound. Less than 0.1 mg of fucose, glucose or mannose were detected. Analysis for neutral sugars after reduction of C1 for 3 h with 1% $NaBH_4$ showed the complete loss of galactose and the formation of 1.41 mg of galactitol. Hence Fraction C1 contains N-acetylgalactosamine and D-galactose in the molar ratio of 1·08:1 and is therefore a disaccharide with galactose in the reducing position. The yield was 3.6 mg. This isolated fragment accounted for 20% of the N-acetylgalactosamine present in the parent glycoprotein.

The disaccharide was methylated, hydrolysed and the recovered neutral sugar methyl ethers were reduced with 1% $NaBD_4$ for 3 h. The alditol acetates on examination by g.l.c. gave a peak with a retention time corresponding to that given by a standard sample of 1,4,5-tri-O-acetyl-2,3,6-tri-O-methyl-galactitol. The identity of this methyl ether was confirmed by Dr A.M. Lawson (Clinical Research Centre, Harrow) who examined the alditol acetates by g.l.c.-mass spectrometry. The galactose residue is therefore substituted at the carbon-4 position. The hexosamine methyl ether fraction gave a single peak on the amino acid analyser with a retention time corresponding to that given by a standard preparation of 3,4,6-tri-O-methylgalactosamine. The disaccharide therefore has the structure GalNAc(1→4)Gal.

Anomeric linkage of the N-acetylgalactosamine residue. The disaccharide had an optical rotation $[\alpha]_D^{38}$ +69° (c 0.28, water)(measured by Miss Lathway, National Institute for Medical Research, Mill Hill). This value is slightly higher than the rotation ($[\alpha]_D^{28}$ +55.5°, c 1, water) recorded for a synthetic crystalline specimen of GalNAcβ(1→4)Gal (15). However, by analogy with the α-linked A-active disaccharide, GalNAcα(1→3)Gal, for which $[\alpha]_D$ values of +150° (7), +154° (16) and +201° (17) have been recorded, a considerably higher positive rotation would have been expected for a disaccharide with an α-anomeric linkage.

Hydrolysis with α- and β-N-acetylgalactosaminidases confirmed that the anomeric linkage was β. A preparation of α-N-acetylgalactosaminidase from

T. foetus failed to hydrolyse C1 but released N-acetylgalactosamine from GalNAcα(1→4)Gal-O-$(CH_2)_8CO_2CH_3$, and GalNAcα(1→3)Gal. A T. foetus preparation containing a mixture of α- and β-N-acetylgalactosaminidases hydrolysed C1 as well as the two α-linked disaccharides. Jack bean β-N-acetylhexosaminidase had no action on the α-linked disaccharides but released N-acetylgalactosamine from C1. This disaccharide therefore has the structure :

N-acetyl-D-galactosaminyl-β-(1→4)-D-galactose.

Serological activity of C1. The fact that C1 occurs as a non-reducing end group in the carbohydrate chains of Sd(a+) T-H glycoprotein, and that N-acetylgalactosamine is believed to be involved in Sd^a specificity, makes it highly probable that this structure forms part of the Sd^a determinant in the intact glycoprotein. However, in serological inhibition tests with two different specimens of human anti-Sd^a serum no inhibition of agglutination was detectable when the disaccharide was tested at concentrations up to 20 mg per ml. The agglutination of Sd(a+) cells by Dolichos biflorus lectin was inhibited to the same extent by C1 (conc. 2.5 mg per ml) as by GalNAcα(1→3)Gal and GalNAcα(1→4)-Gal-O-$(CH_2)_8CO_2CH_3$. The failure of the disaccharide to inhibit the human anti-Sd^a reagents may arise from the fact that the combining site of the antibody requires a structure larger than a disaccharide, a situation which occurs in the P blood group system where P_1 agglutinins in human serum appear to be strongly orientated to the conformation of a trisaccharide and to have little or no affinity for the terminal disaccharide (18). Alternatively in the intact T-H macromolecule either of the sugars in the disaccharide may have an additional substituent which forms part of the determinant or the disaccharide structure may interact with other parts of the molecule to take up a specific Sd^a conformation.

General comments. The lectin from Dolichos biflorus reacts with the blood group A determinant (19) and is generally considered to be specific for α-linked N-acetylgalactosamine (20). It is therefore somewhat surprising to find that Sd^a-active molecules, which combine with Dolichos biflorus, contain terminal non-reducing β-linked N-acetylgalactosamine residues. Fractions I

and II from the Bio-Gel column, on which the preliminary separation of the hydrazinolysis products was carried out, contained 76% of the N-acetylgalactosamine present in the starting material, and as judged by paper chromatography, apart from monosaccharides, the disaccharide C1 was the major labelled acid hydrolysis product. It can therefore be assumed that the majority of the carbohydrate chains in the T-H glycoprotein which terminate in N-acetylgalactosamine contain this sugar in β-(1→4)-linkage to galactose and that in this instance the Dolichos reagent is combining with β-linked N-acetylgalactosaminyl residues.

The structure GalNAcβ(1→4)Gal has not previously been reported in glycoproteins but is well known as a constituent of glycosphingolipids of the ganglio series (21). The disaccharide occurs as an internal structure in the major brain ganglioside G_{M1}(αNeuAc-GgOse$_4$Cer) and was first isolated from this source by Kuhn and Wiegandt (22). G_{M2}(αNeuAc-GgOse$_3$Cer), the biosynthetic precursor of G_{M1}, terminates with the disaccharide but this ganglioside is present in very small amounts in normal tissues (21). It occurs in increased amounts in the brain in Tay-Sachs disease (23), where a degradative enzyme is missing, and on the cell surface of transformed fibroblasts where it is considered to arise from the blocked biosynthesis of G_{M1} (24). G_{M1} has also been isolated from human kidney (25), and it is therefore interesting to speculate on the relationship between the β-N-acetylgalactosaminyltransferase involved in the biosynthesis of the ganglioside and that concerned with the formation of the disaccharide end-group in the T-H urinary glycoprotein. In the ganglioside G_{M2} the subterminal galactose residue is substituted with sialic acid and since this sugar is a constituent of T-H glycoprotein (1) it is possible that in the intact macromolecule this same trisaccharide structure occurs. In attempts to resolve these questions further work is in progress to isolate larger N-acetylgalactosamine-containing oligosaccharides from the T-H glycoprotein.

ACKNOWLEDGEMENTS. We thank Dr A. Lawson for the g.l.c.-mass spectrometry analysis and Miss Lathway for determining the optical rotation. We are most grateful to Dr D. Baker for the gift of the compound GalNAcα(1→4)Gal-O-$(CH_2)_8$-CO_2CH_3.

REFERENCES

1. Fletcher A.P. (1972) Glycoproteins (Ed. A. Gottschalk) p.p. 892-908, Elsevier, Amsterdam.
2. Race R.R. and Sanger R. (1975) Blood Groups in Man, 6th edition, p.p. 395-405, Blackwell, Oxford.
3. Morgan W.T.J., Soh C. and Watkins W.M. (1979) Glycoconjugates (Eds. R. Schauer, P. Boer, E. Buddecke, M.F. Kramer, J.F.G. Vliegenthart and H. Wiegandt) p.p. 582-583, Thieme, Stuttgart.
4. Soh C.P.C., Morgan W.T.J., Watkins W.M. and Donald A.S.R. (1980) Biochem. Biophys. Res. Commun., 93, 1132-1139.
5. Sanger R., Gavin J., Tippett P., Teesdale P. and Eldon K. (1971) Lancet, i, 1130.
6. Bird G.W.G. and Wingham J. (1971) Vox Sang., 20, 55-61.
7. Côté R. and Morgan W.T.J. (1956) Nature, London, 178, 1171-1172.
8. Watkins W.M. (1959) Biochem. J., 71, 261-274.
9. Donald A.S.R. (1981) Eur. J. Biochem. (In press)
10. Kuhn R., Baer H.H. and Seelinger A. (1958) Liebigs Ann., 611, 236-241.
11. Yang H. and Hakomori S-I. (1971) J. Biol. Chem., 246, 1192-1200.
12. Tamm I. and Horsfall F.L. (1950) Proc. Soc. Exp. Biol. Med., 74, 108-114.
13. Gahmberg C.G. (1978) Methods Enzymol., 50, 204-206.
14. Morell A.G. and Ashwell G. (1972) Methods Enzymol., 28, 205-211.
15. Shapiro D. and Acher A.J. (1970) J. Org. Chem., 35, 229-231.
16. Etzler M.E., Anderson B., Beychok S., Gruezo F., Lloyd K.O., Richardson N.G. and Kabat E.A. (1970) Arch. Biochem. Biophys., 141, 588-601.
17. Yosizawa Z. (1962) J. Biochem. (Tokyo) 51, 1-11.
18. Watkins W.M. and Morgan W.T.J. (1976) J. Immunogenetics, 3, 15-27.
19. Bird G.W.G. (1952) Nature, London, 170, 674.
20. Sharon N. and Lis H. (1972) Science, 177, 949-959.
21. Sweeley C.C. and Siddiqui B. (1978) The Glycoconjugates, Vol.1, p.p.459-540, Academic Press, New York.
22. Kuhn R. and Wiegandt H. (1963) Chem. Ber., 96, 866-880.
23. Svennerholm L. (1962) Biochem. Biophys. Res. Commun., 9, 436-441.
24. Rosenfelder G., Young W.W. and Hakomori S-I. (1977) Cancer Research, 37, 1333-1339.
25. Rauvala H. (1976) J. Biol. Chem., 251, 7517-7520.