The effect of the carbohydrate moiety upon the size and conformation of human plasma galactoglycoprotein as judged by electron microscopy and circular dichroism

Structural studies of a glycoprotein after stepwise enzymic carbohydrate removal

Hildegard WATZLAWICK,* Mary T. WALSH,†‡§§§ Ingrid EHRHARD,* Henry S. SLAYTER,||¶ Heinz HAUPT,** H. Gerhard SCHWICK,** George W. JOURDIAN, †† Sumihiro HASE, †§ || || Karl SCHMID§ and Reinhard BROSSMER*

*Institut fur Biochemie II, University of Heidelberg, D-6900 Heidelberg, Federal Republic of Germany, tDepartment of Biophysics, tDepartment of Medicine and §Department of Biochemistry, Boston University School of Medicine, Boston, MA 02118, U.S.A., ||Dana Farber Cancer Institute and ||Department of Cellular and Molecular Physiology, Harvard Medical School, Boston, MA 02115, U.S.A., **Behringwerke A.G., D-3550 Marburg/Lahn, Federal Republic of Germany, and ††Rackham Arthritis Research Unit, University of Michigan Medical School, Ann Arbor, MI 48109, U.S.A.

Galactoglycoprotein is a unique human plasma protein $[76\%$ carbohydrate (23% N-acetylneuraminic acid, 20%) galactose, 3% mannose, 1% fucose and 29% N-acetylgalactosamine plus N-acetylglucosamine) and 24% polypeptide, ^a single polypeptide chain of about 200 amino acid residues that is high in serine and threonine content] [Schmid, Mao, Kimura, Hayashi & Binette (1980) J. Biol. Chem. 255, 3221-3226]. Highly purified exoglycosidases with well-defined specificities were used to prepare five derivatives of galactoglycoprotein in which sequential residues of N-acetylneuraminic acid, galactose, N-acetylglucosamine, ^a second galactose and N-acetylgalactosamine were removed with ⁸³ % of the total carbohydrate cleaved. C.d. shows that native galactoglycoprotein and all derivatives in aqueous buffer are predominantly random coil, suggesting that removal of a large number of electrostatic net charges, as well as the major portion of the carbohydrate moiety, does not alter the secondary structure of the polypeptide chain. Examination of the size and conformation of tungsten-shadowed galactoglycoprotein and asialo and agalacto derivatives by electron microscopy shows the size and conformation of all three preparations to be similar, with only minor differences in particle length and width.

INTRODUCTION

Human plasma galactoglycoprotein (galactogp) $(M, 81000)$ has an unusually high carbohydrate content (76%) , consisting of approximately a quarter of its weight of each of N-acetylneuraminic acid (NeuAc), neutral sugars, amino sugars and polypeptide (Schmid et al., 1980; Schwick & Haupt, 1984). Because of its high content of galactose (Gal) and N-acetylgalactosamine (GalNAc), it is not considered to be a typical plasma glycoprotein, which generally lacks GalNAc (Baenziger, 1984) and is high in mannose (Man) content. It is speculated that galactogp is derived from cells, since cell-surface glycoproteins are high in GalNAc content (Schmid et al., 1980). Thus galactogp might be responsible for certain receptor functions. The carbohydrate moiety of galactogp consists of approx. 40 0 glycosidic hexasaccharides, three O -glycosidic tetrasaccharides (for structural formulae see Scheme 1) and three biantennary Nglycans (Akiyama et al., 1984).

In the present study highly purified exoglycosidases were used to remove sequential carbohydrate residues from galactogp. The effect of their removal on the secondary structure, size and conformation of galactogp and its derivatives was assessed by c.d. and electron microscopy.

MATERIALS AND METHODS

Materials

All chemicals (analytical grade) were purchased from Merck (Darmstadt, Germany), Serva (Heidelberg, Germany) or Aldrich Chemical Co. (Milwaukee, WI, U.S.A.). Concanavalin A-Sepharose 4B, Sepharose 6B and CH-Sepharose 4B were from Pharmacia (Uppsala, Sweden), and Bio-Gel P-2 was from Bio-Rad (Muinchen, Germany). Galactogp was isolated from pooled normal human plasma (Schmid et al., 1980; Akiyama et al., 1984) with an additional step of solid-phase immuno-adsorption using rabbit IgG coupled to CNBr-activated Sepharose CL-4B. Since galactogp is not itself immunogenic, the rabbit antiserum contained antibodies to only the 'impurities', thus allowing removal of small amounts of proteinaceous contaminants.

Instrumentation

C.d. spectra were recorded at ambient temperature in a Cary 61 c.d. spectropolarimeter (Varian, Palo Alto, CA, U.S.A.) as described by Walsh et al. (1990). Methods used to evaluate the data (Greenfield & Fasman, 1969; Mao & Wallace, 1984) have also been described by Walsh et al. (1990). The carbohydrate moiety has been shown not to contribute significantly to the

Abbreviations used: galactogp, galactoglycoprotein; Fuc, fucose; Gal, galactose; GalNAc, N-acetylgalactosamine; GlcNAc, N-acetylglucosamine; Man, mannose; NeuAc, N-acetylneuraminic acid.

^{§§} To whom correspondence should be addressed, at Department of Biophysics, Boston University School of Medicine, ⁸⁰ East Concord Street, Boston, MA 02118-2394, U.S.A.

¹¹ ¹¹ Present address: Department of Chemistry, Osaka University College of Science, Osaka 560, Japan.

a-

was incubated for 48 h with 50 munits of Vibrio *1 *
1 \mathbf{I} D. I Is \mathcal{L} galactogp ellipticities, which were recorded from 250 to 200 nm (Schmid et al., 1978; Johnson, 1987; Walsh et al., 1990).

For the electron-microscopic studies, protein solutions (0.1 m/s/mol) in 0.1 M-ammonium acetate buffer, pH (7.0 cm) (0.1 mg/ml) in 0.1 M-ammonium acetate buffer, pH 7.0, containing 50 % (v/v) glycerol were sprayed on mica (Slayter *et al.*, 1984; Slayter, 1989). Protein was applied as an aerosol to freshly $(20.7, 0.01, 1.00)$, i fotom was applied as an acrosor to freship μ a very thin layer of the μ and μ is the contraction- $\frac{1}{2}$ a very thin layer of tungsten by electron-beam evaporation. nass inickness of the background film of tungsten was $\frac{3 \text{ ng/cm}^2}{2 \text{ Mg/cm}^2}$. This metal layer was coated with 2.5 nm of carbon. nicrographs were recorded in focus on a JEM TOU CA electron microscope at 100 kV and at a magnification of \times 40000–53000. High-resolution dark-field images were obtained from lightly shadowed specimens with the use of matched annular condenser and objective apertures and an exposure time of 0.35 s.

The h.p.l.c. system used for derivative purification has been described in detail by Walsh et al. (1990) and was used without modification.

T_{S}

The use of *Vibrio cholerae* sialidase, bovine testis β -galactosidase, Streptococcus pneumoniae β -galactosidase and jack-bean β -N-acetylglucosaminidase and their activities have been described in detail by Walsh et al. (1990). The enzymic procedures have been applied to the galactogp without modification. α -N-Acetylgalactosaminidase from bovine liver (6 units/mg of protein) was isolated with the aid of specific affinity matrices. All enzymes were free of other glycosidases and proteinases. All digestions were at 37 $^{\circ}$ C.

Hydrolysis conditions, analytical methods for specific sugars and all other protocols have been described in detail by Walsh $et \ al.$ (1990). The preparation of partially deglycosylated derivatives of galactogp is described and schematized in Scheme 1.

RESULTS AND DISCUSSION

The use of the multi-step sequential enzymic degradation procedure for glycoprotein glycans offers two distinct advantages over other methodologies: (1) taking into account the established specificity of the applied pure glycosidases, structural data on such glycans can be obtained; (2) the use of partially de- $\frac{1}{2}$ relative to the compute $\frac{1}{2}$ and the contribution of $\frac{1}{2}$ and $\frac{1}{2}$ by the contribution of $\frac{1}{2}$ and $\frac{1}{2}$ by this may relate a defined structure to a biological activity.

On the basis of proposed structures of the given as of galactogp on the state of proposed en actually of the grypants of gametogy
A kivama et al. 1984) assuming full siglulation, galactogy would have 92 NeuAc residues or a NeuAc content of 33% . Thus, since 62 NeuAc residues are released by acid hydrolysis (Table 1), galactogp is sialylated to the extent of only 67 $\%$.

 $\frac{3}{2}$ $\frac{3}{2}$ Bovine testis β -galactosidase (Table 1, step 2a) hydrolyses cluding Gal β 1-3GlcNAc, Gal β 1-4GlcNAc and Gal β 1-3GalNAc (Distler & Jourdian, 1973). Substitution of GlcNAc for GalNAc, as in Gal β 1-3(GlcNAc β 1-6)GalNAc, produces a substrate that is not attacked by this enzyme, relevant to the enzymic cleavage of O -hexasaccharides of galactogp. This enzyme does not cleave any of the former three compounds if the monosaccharide to which Gal is linked is substituted with Fuc, such as in Gal β 1-4(Fuca1-2)GlcNAc, an important fact since asialo-galactogp has Le^a-factor activity, which is associated with the N -glycans (Schmid et al., 1980). Thus it was assumed that three branches of the biantennary N -glycans are associated with this activity. On the basis of the N-glycan and O-glycan structures of galactogp, the Gal residues accessible to bovine testis β -

Table 1. Monosaccharides liberated from galactogp by sequential enzymic degradation with different exoglycosidases

The bracketed text after each enzyme name indicates the specificity of the enzyme.

* For further information see the text.

Without 'branched' Gal β 1-3 of the O-hexasaccharides.

^I It was assumed that three Gal residues of the asialo-N-glycans are accessible.

§ The difference of one residue may be explained by the error of the methods.

|| 39 GalNAc residues of O-hexasaccharides and three GalNAc residues of O-tetrasaccharides.

Table 2. Carbohydrate composition of human plasma galactogp and its enzymically prepared derivatives

Fuc was not determined. Asialo-galactogp was used for quantification of neutral sugars since NeuAc interferes with their determination. Abbreviation: N.D., not determined.

* Taken from Schwick & Haupt (1984).

 \dagger *M*, values calculated from the amount of carbohydrate cleaved.

^t For explanation of the differences see the text.

§ Determined by acid hydrolysis (for protocol see Walsh et al., 1990).

galactosidase are only the Gal attached β 1-4 to the O-glycans of the 40 O-hexasaccharides and the Gal attached β 1-3 of the three O-tetrasaccharides and three of the two N-biantennary glycans.

Digestion of asialo-galactogp (Table 1, step 2a; Scheme 1, S2) released 41 Gal residues, representing 48% of the total Gal residues as determined by acid hydrolysis (86 total) or ⁸⁹ % of ⁴⁶ accessible Gal residues, in agreement with the previously described specificity of this enzyme. Digestion of asialo-galactogp with Streptococcus pneumoniae β -galactosidase released three fewer Gal residues (Table 1, step 2b) than did the bovine testis enzyme, in agreement with the structure of the three O -glycosidic tetrasaccharides containing Gal linked β 1-3 to GalNAc that cannot be cleaved by the Streptococcus pneumoniae enzyme.

Release of GlcNAc from GlcNAc β 1-6(Gal β 1-3)GalNAcgalactogp (Scheme 1, S3) required lengthy incubations to release ⁹³ % of GlcNAc (Table 1, step 3; Scheme 1, S3) of approx. ⁴¹ cleavable GlcNAc residues.

Digestion of a-N-acetylglucosamino-galactogp with bovine testis β -galactosidase released 39 Gal residues (Table 1, step 4; Scheme 1, S4). The Gal β 1-3GalNAc linkage of the hexasaccharide of galactogp is cleaved only after removal of GlcNAc from Gal β 1-3(GlcNAc β 1-6)GalNAc. The cleavage of the

Fig. 1. C.d. spectra of galactogp and derivatives

Spectrum 1 (--), native galactogp; spectrum 2 (----), asialo $galactoop$; spectrum 3 (.....), agalacto-galactogp; spectrum 4 $(-\cdots)$, a-N-acetylgalactosamino-galactogp; spectrum 5 (\cdots a-N-acetylglucosamino-galactogp. A 1 cm quartz cell was used for measurements over the wavelength region 250-220 nm, whereas for measurements in the region $220-190$ nm a 0.1 cm quartz cell was used. On the day of the c.d. measurements freeze-dried protein samples were dissolved in 0.01 M-phosphate buffer, pH 7.0, containing 0.02% NaN₃. Protein was determined by the method of Lowry *et al.* (1951), with BSA as standard.

 $Gal β 1–3GalNAc bond, however, is relatively fast considering$ the close proximity of this bond to the polypeptide chain.

Theoretically, 40 α -linked GalNAc residues of the *O*-linked hexasaccharides and three GalNAc residues of the O-linked $tetrasaccharides$ on N -acetylgalactosamino-galactogp are accessible to bovine liver α -*N*-acetylgalactosaminidase (Table 1, step 5; Scheme 1, S5). After 72 h 40 residues were cleaved, although GalNAc residues are attached to serine and/or threonine residues of the polypeptide chain.

The carbohydrate compositions of galactogp and derivatives are presented in Table 2. In comparison with that calculated from the established structure (Scheme 1), a somewhat lower carbohydrate content was found for the asialo derivative. This result was expected, since the enzymic digestions would not yield complete release of the carbohydrate residues. For a-N-acetylgalactosamino-galactogp, which contains five GalNAc residues, a low content of Gal and GlcNAc is observed. The total amount of the carbohydrate cleaved by enzymic means was thus 83% .

The c.d. spectra of native galactogp and all enzymically prepared derivatives in aqueous buffer are similar to one another (Fig. 1), exhibiting a single negative minimum at 205 nm and molar ellipticity values [Θ] ranging from -9100 to -5000 degree · cm² · dmol⁻¹. Such spectra are characteristic of a protein that possesses predominantly random coil $(80-100\%)$, small amounts of β -structure (0-20%) and negligible content of α helical and β -turn regions. These results suggest that removal of up to 83% of the carbohydrate moiety of galactogp does not significantly alter its solution conformation or secondary struc- \mathbf{A}

Native galactogp and asialo-galactogp and agalacto-galactogp derivatives were compared by electron microscopy. In 0.1 M- ammonium acetate buffer a variable amorphous projected topography was observed, but when 50 $\%$ glycerol was added to the solvent the principal species observed was an extended particle (Fig. 2a). Such behaviour is consistent with that of a variety of mucins that form in measurable puddles in the absence of glycerol. The typical galactogp particle has an extended filamentous shape with an average length of 43 ± 15 nm. The distribution of particle lengths of the two derivatives was relatively broad (Figs. $3b$ and $3c$) and ranged from 30 to 70 nm. The particle width was approx. 4.0 nm. No significant qualitativeor quantitative differences in appearance or dimensions were observed for the preparations analysed, although the native particles tended to be straighter. The observed polydispersity in length is characteristic of flexible mucin-like molecules. and is characteristic of hexibic much material more differences.

 α Fig. 2(a) with Fig. 2(b) and 2(c) α Fig. 2(c) galactographs. of Fig. $2(a)$ with Figs. $2(b)$ and $2(c)$. In Fig. $2(a)$ galactogp appears to be generally extended and quite homogeneous, whereas asialo-galactogp (Fig. $2b$) shows many similar particles, but with more variability in their conformation, of which many more appear to be bent or collapsed than for the native material. A significant fraction (approx. 50%) of the asialo-galactogp and agalacto-galactogp populations tended to collapse into an amorphous mass from which a contour length could not be determined. The number of collapsed particles was greater for asialo-galactogp. Consequently the average lengths were similar, but the peaks were somewhat different.

Assuming that a single polypeptide chain of galactogp contains 200 amino acid residues, and using Pauling's value of 0.364 nm per residue, the maximum extension per peptide bond for a fully stretched polypeptide chain, approx. 70 nm is calculated for the completely extended length of the particle (Slayter et al., 1984). Thus there is close agreement between the maximum theoretical length and the observed maximum length.

In galactogp, on the average, approximately every fourth amino acid residue possesses a carbohydrate side chain. Thus there is less steric hindrance to folding than in other carbohydraterich mucins (Swann et al., 1985), even though the distribution of glycosidic side chains is homogeneous. Since the mean length (approx. 40 nm) is only a little more than half that of the maximum length possible (70 nm) , it is concluded that this particle is essentially linear with an accordion-like flexibility that is due to the lack of secondary structure. In this respect galactogp resembles a number of other glycoproteins investigated by electron microscopy, including plasma I antigen (77 $\%$ carbohydrate), bronchial mucins (70% carbohydrate), epiglycanin (77 $\%$ carbohydrate) and lubricin (60 $\%$ carbohydrate) (Swann et al., 1985; Slayter, 1989).

In an effort to relate the conformation found by electron microscopy for the galactogp to the conformations in solution, calculations were made of the sedimentation coefficient (Slayter, 1976) for both a flexible-coil model and a rod model by using the average radius taken from the data shown in Figs. $2(a)$ and $3(a)$. The result of this calculation indicates a sedimentation coefficient of 5.9 S for the flexible-coil model and 4.2 S for the rod model. The micrographs suggest a model with both characteristics. The small deviation of the calculated value from the published value of 4.6 S (Schmid et al., 1980) is quite consistent with the configurations observed by electron microscopy. These results provide reasonable confirmation that the same macromolecular entity exists in solution as is observed in the micrographs.

Furthermore, the c.d. spectra for these three preparations indicate that they are very similar in their secondary structures. i.e. possess predominantly random-coil conformation, which would be expected to be stabilized sterically by glycosidic side chains terminating in charged sialic acid residues.

1991

Fig. 2. Dark-field electron micrographs of lightly tungsten-coated galactogp and selected derivatives

(a) Native galactogp; (b) asialo-galactogp; (c) agalacto-galactogp. Arrows indicate typical particles. For the magnification see the Materials and methods section.

Fig. 3. Histograms showing distributions of lengths of native and modified galactogp

(a) Native galactogp; (b) asialo-galactogp; (c) agalacto-galactogp.

a series of consecutively deglycosylated derivatives of galactogp were prepared in a controlled manner in which the native state of the polypeptide chain remains intact. The native galactogp and derivatives were examined by c.d. to assess the effect of removal of sequential carbohydrate residues on their solution conformation and secondary structure. Removal of about ⁸³ % of the carbohydrate moiety does not alter the secondary structure of galactogp, which exists predominantly in a random coil. β_{2} -Glycoprotein I, with which a similar study has been conducted (Walsh *et al.*, 1990), showed that only after removal of 96 $\%$ of the carbohydrate was its secondary structure changed. Electron microscopy was used to assess the size and gross conformation of the native galactogp and its asialo and agalacto derivatives. All three preparations possessed an extended filamentous shape, an average length of approx. 43 nm and width of approx. 4 nm, similar to flexible mucin-like molecules. As to the biological significance of such glycoprotein derivatives, recent studies with human granulocyte colony-stimulating factor (Oh-eda et al., 1990) and thyrotropin (Thotakura et al., 1990) and the earlier investigation of α_1 -acid glycoprotein (Schmid *et al.*, 1978; Bennett & Schmid, 1980) demonstrated that, following stepwise enzymic degradation of the carbohydrate moieties, the monosaccharides that are associated with a specific biological function can be identified.

We thank Cynthia Bates, Rosilyn Ford, Martina Rohnacher and Birgit Hobl for excellent technical assistance. The expertise of Anne M. Plunkett and Margaret Gibbons in the preparation of this manuscript is gratefully acknowledged. This work was supported by the Deutsche Forschungsgemeinschaft SFB-136 to H. W. and R. B. and grants (HL-26335 to M.T.W., GM-14237 to H.S.S., AR-10531 to G.W.J. and GM-10374 to K. S.) from the National Institutes of Health, U.S. Public Health Service.

REFERENCES

- Akiyama, K., Simons, E. R., Bernasconi, P., Schmid, K., van Halbeek, H., Vliegenthart, J. F. G., Haupt, H. & Schwick, H. G. (1984) J. Biol. Chem. 259, 7151-7154
- Baenziger, J. U. (1984) in The Plasma Proteins, 2nd edn. (Putnam, F. W., ed.), vol. 4, pp. 272-315, Academic Press, New York
- Bennett, M. & Schmid, K. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 6109-6113
- Distler, J. J. & Jourdian, G. W. (1973) J. Biol. Chem. 248, 6772-6780
- Greenfield, N. & Fasman, G. D. (1969) Biochemistry 10, 4108-4116
- Johnson, W. C. (1987) Adv. Carbohydr. Chem. Biochem. 45, 73-93
- Lowry, 0. H., Rosebrough, N. H., Farr, A. L. & Randall, R. J. (1951) J. Biol. Chem. 193, 265-275
- Mao, D. & Wallace, B. A. (1984) Biochemistry 23, 2667-2673
- Oh-eda, M., Hasegawa, M., Hattori, K., Kuboniwa, H., Kojima, T., Orita, T., Tomonou, K., Yamazaki, T. & Ochi, N. (1990) J. Biol. Chem. 265, 11432-11435
- Schmid, K., Burlingame, R. W., Paulson, J. C. & Hill, R. L. (1978) Fed. Proc. Fed. Am. Soc. Exp. Biol. 37, 1298
- Schmid, K., Mao, S. K. Y., Kimura, A., Hayashi, S. & Binette, J. P. (1980) J. Biol. Chem. 255, 3221-3226

Schwick, H. G. & Haupt, H. (1984) in The Plasma Proteins, 2nd edn. (Putnam, F. W., ed.), vol. 4, pp. 167-220, Academic Press, New York Slayter, H. S. (1976) Ultramicroscopy 1, 341-357

- Slayter, H. S. (1989) Methods Enzymol. 169, 326-335
- Slayter, H. S., Lamblin, G., LeTreut, A., Galabert, C., Noudret, N.,
- Degand, P. & Roussel, P. (1984) Eur. J. Biochem. 142, 209-218
- Swann, D. A., Silver, F. H., Slayter, H. S., Stafford, W. & Shore, E. (1985) Biochem. J. 225, 195-201
- Thotakura, N. R., LiCalzi, L. & Weintraub, B. D. (1990) J. Biol. Chem. 265, 11527-11534
- Walsh, M. T., Watzlawick, H., Putnam, F. W., Schmid, K. & Brossmer, R. (1990) Biochemistry 29, 6250-6257

Received 23 October 1990/5 February 1991; accepted 13 February 1991