Size heterogeneity of human cervical mucus glycoproteins

Studies performed with rate-zonal centrifugation and laser light-scattering

John K. SHEEHAN* and Ingemar CARLSTEDT††

*Department of Biological Sciences, University of Lancaster, Lancaster LA1 4YQ, U.K., and †Department of Physiological Chemistry 2, University of Lund, P.O.B. 94, S-221 00 Lund, Sweden

Mucus glycoproteins (mucins) from cervical pregnancy mucus were fractionated by using rate-zonal centrifugation in a gradient of guanidinium chloride. The distribution of the macromolecules, as assessed by using sialic acid determination, suggested the presence of three populations of different size. Individual fractions were subjected to laser light-scattering performed as total-intensity measurements as well as photon correlation spectroscopy. The results showed that points of inflexion were present in the distribution of both M_r and D_T (translational diffusion coefficient) and that the three populations have M_r values of approx. 24×10^6 , 16×10^6 and 6×10^6 respectively. The weight-average M_r for the whole distribution, as calculated from the values obtained for the individual fractions, was 13.6×10^6 , which is in good agreement with that found for the unfractionated material (11.1×10^6). Plots of $\log R_G$ (radius of gyration) and $\log(1/D_T)$ versus $\log M_r$ are in keeping with the macromolecules being linear flexible chains.

INTRODUCTION

The mucus glycoproteins (the mucins) are the gel-forming macromolecules of mucus. We have isolated cervical mucins from human pregnancy mucus (Carlstedt *et al.*, 1983a) and proposed a 'model' for their polymeric structure (Carlstedt *et al.*, 1983b). The hydrodynamic properties are consistent with the macromolecules (M_r 11.1×10^6) being linear flexible chains that occupy an expanded spheroidal domain in dilute solution (Sheehan & Carlstedt, 1984a,b). Pig gastric, pig submaxillary-gland and human respiratory mucins conform to this model, which has been corroborated by using electron microscopy (Harding *et al.*, 1983; Carlstedt & Sheehan, 1984; Rose *et al.*, 1984; Slayter *et al.*, 1984; Mikkelsen *et al.*, 1985; Shogren *et al.*, 1986; Sheehan *et al.*, 1986).

The polydispersity, in size, of mucus glycoproteins must be considered when a model of their polymeric structure is evaluated. The size distribution of ungraded mucins cannot generally be studied with gel chromatography, and the procedure most commonly used is sedimentation velocity in the analytical ultracentrifuge. The pronounced concentration-dependence of the sedimentation coefficient might, however, obscure the presence of more than one component, and an apparently unimodal distribution could well prove to be heterogeneous. The $g(s)$ function (apparent distribution of sedimentation coefficient) was used by Allen *et al.* (1984) as described by Hascall & Sadjara (1969) to study the polydispersity of pig gastric mucins, and species with s_{app} from approx. 10 to 60 S were observed. Creeth & Cooper (1984) revealed broad and 'skew' distributions of M_r for respiratory and gastric mucins with sedimentation-equilibrium experiments. The latter method has the advantage that the macromolecules can be studied at low concentrations, but has the drawback that very large species ($M_r > \text{approx. } 10 \times 10^6$) will be excluded in the analysis.

In the present work, rate-zonal centrifugation was used to disperse cervical mucins. Individual fractions were then analysed with laser light-scattering performed as absolute-intensity measurements and photon correlation spectroscopy. The results suggest that the macromolecules are not well described as a unimodal polydisperse distribution of M_r (Gibbons, 1963) because the distribution has points of inflexion corresponding to populations with M_r 6×10^6 , 16×10^6 and 24×10^6 .

EXPERIMENTAL

Materials

Sucrose (AnalaR) and guanidinium chloride (ultrapure grade) were bought from BDH Chemicals and Bethesda Research Laboratories respectively. Cervical mucins and fragments obtained after reduction ('subunits') as well as after subsequent trypsin digestion ('T-domains') were isolated as described in detail by Carlstedt *et al.* (1983a,b).

Rate-zonal centrifugation

Centrifugation tubes were loaded from the bottom with linear gradients of guanidinium chloride (6–8 M) by using a glass capillary held in position with a micromanipulator, which permitted its careful removal. Mucin solutions (100–200 μ l; concn. approx. 2 mg/ml) were layered on to the gradients with a Hamilton syringe. Centrifugation was performed in an MSE 6 $\times$ 5.5 ml swing-out rotor at 35000 rev./min for 2.5 h. After centrifugation, the tubes were emptied from the bottom and the individual fractions were dialysed into 4 M-guanidinium chloride by using a multichannel microdialysis cell (Bethesda Research Laboratories). After dialysis, each fraction was adjusted to 1 ml with the diffusate and the concentration of mucin was assessed with sialic acid analysis by using an automated version (Lohmander *et al.*, 1980) of the method described by

† To whom correspondence and reprint requests should be sent.

Jouradian *et al.* (1971). A solution with known concentration of the unfractionated mucins was used as a standard (Carlstedt *et al.*, 1983a). Light-scattering was then performed as described below.

Sucrose gradients (10–30%, w/v) in 0.2 M-NaCl were prepared in a similar way. The sample (100 μ l; concn. approx. 2 mg/ml) was dialysed into 0.2 M-NaCl in 10% (w/v) sucrose and layered on to the gradient and centrifuged as described above for the guanidinium chloride gradients. The distribution of mucins in these gradients was monitored at 280 or 206 nm with a u.v. detector, essentially as described by Franzén *et al.* (1982).

Light-scattering

A Malvern Molecular Analyser System 4300 (Precision Devices and Systems, Malvern, Worcs., U.K.) equipped with a 15 mW Model 124B He/Ne laser (Spectra Physics, St. Albans, Herts., U.K.) was used for both total intensity and quasi-elastic measurements (Sheehan & Carlstedt, 1984a). Clarification of solutions was performed by centrifugation at 15000 *g* for 1–2 h. Conventional round glass cells (Malvern) were used for measurements down to 20°.

Total-intensity measurements on fractions obtained from rate-zonal centrifugation were performed at $\theta = 20, 30, 40, 50$ and 60° and the apparent R_G (radius of gyration) and M_r were determined from, respectively, the slope of c/R_θ versus $\sin^2\theta/2$ and the value of c/R_θ extrapolated to $\theta = 0^\circ$ (see e.g. Tanford, 1961). A value for $\partial n/\partial c$ (refractive-index increment) of 0.120 ml/g was used for cervical mucins in 4 M-guanidinium chloride (Sheehan & Carlstedt, 1984b). The values of M_r were corrected for the concentration-dependence by using the second virial coefficient obtained for the unfractionated material ($B = 1.2 \times 10^{-4}$ mol $\cdot$ ml $\cdot$ g $^{-2}$). Diffusion coefficients were determined on the same solutions at $\theta = 20^\circ$ and $T = 25^\circ\text{C}$ and were corrected to $T = 20^\circ\text{C}$, but not for the angular dependence and for the effect of concentration. The values of D_T (translational diffusion coefficient) for 'whole' cervical mucins show no concentration-dependence within the range studied here (Sheehan & Carlstedt, 1984a).

Light-scattering measurements on mucin fragments (subunits and T-domains) in 4 M-guanidinium chloride were performed essentially as described by Sheehan & Carlstedt (1984a). T-domains were studied at 4.72, 3.15, 2.36 and 1.57 mg/ml and subunits at 1.72, 1.14, 0.86 and 0.57 mg/ml. The value used for $\partial n/\partial c$ in both cases was 0.122 ml/g. Diffusion coefficients were determined on the same solutions and the values of D_T were extrapolated to zero angle and zero concentration and corrected to 20°C .

RESULTS AND DISCUSSION

To avoid convection and to diminish broadening in rate-zonal centrifugation the sample is typically layered on to a solution of increasing density towards the bottom of the tube. Sucrose, which is commonly used in rate-zonal centrifugation, impairs subsequent carbohydrate analysis, whereas caesium salts rearrange in a gravitational field, making interpretation of the radial distribution as a size distribution difficult. Gradients of guanidinium chloride were used for these reasons and because the relevant physical data are available to evaluate the behaviour of cervical mucins in this solvent (Sheehan & Carlstedt, 1984a,b). Furthermore, the concentration of the macromolecules can be conveniently determined by sialic acid measurements in guanidinium chloride.

The physical parameters characterizing the guanidinium chloride gradient are listed in Table 1. The expression $r(1 - \bar{v}\rho)/\eta_{\text{rel.}}$ (these terms are defined in Table 1) does not vary significantly with r , suggesting that the gradient is approximately isokinetic. This implies that an individual macromolecule would maintain a constant velocity during the experiment.

The distribution of mucins, as monitored by using sialic acid measurements, show three peaks of decreasing height from the top of the gradient (Fig. 1a). The cognate values for the weight-average M_r , z-average R_G and z-average D_T are shown in Figs. 1(b) and 1(c). The distribution of M_r increases from 6×10^6 to 35×10^6 . The three components occurring at M_r 6×10^6 , 16×10^6 and

Table 1. Characterization of the guanidinium chloride gradient

The gradient (6–8 M) was prepared as described in the Experimental section. The distance (r) from the centre of rotation to the concentrations of guanidinium chloride indicated in the Table were calculated by using the specifications on the rotor and the tubes supplied by the manufacturer. The cognate values for the density (ρ) and the relative viscosity ($\eta_{\text{rel.}}$) for guanidinium chloride were from Kawahara & Tanford (1966), whereas the values for the partial specific volumes ($\bar{v}$) were obtained by interpolation of data provided by Sheehan & Carlstedt (1984b). The latter values are given to three decimal places for the purpose of these calculations only and do not reflect the precision of the original data. The value of $r(1 - \bar{v}\rho)/\eta_{\text{rel.}}$ does not vary significantly with r , suggesting that the gradient is isokinetic for cervical mucins, i.e. the increase in the gravitational field with r is compensated for by the changes in ρ , $\eta_{\text{rel.}}$ and $\bar{v}$.

Concn. (M)	r (cm)	ρ (g/ml)	$\eta_{\text{rel.}}$	$\bar{v}$ (ml/g)	$r(1 - \bar{v}\rho)/\eta_{\text{rel.}}$ (cm)
6.00	6.10	1.145	1.620	0.668	0.91
6.25	6.68	1.150	1.685	0.669	0.91
6.50	7.26	1.156	1.770	0.670	0.93
6.75	7.83	1.162	1.845	0.671	0.94
7.00	8.41	1.168	1.925	0.672	0.94
7.25	8.99	1.174	2.030	0.673	0.93
7.50	9.57	1.180	2.125	0.674	0.92
7.75	10.14	1.186	2.230	0.674	0.91
8.00	10.72	1.192	2.400	0.675	0.87

Fig. 1. Distribution of cervical mucins after rate-zonal centrifugation in a gradient of guanidinium chloride

For full experimental details, see the text. (a) Distribution of the mucins assessed by using sialic acid measurements ($\bullet$); (b) weight-average M_r ($\bullet$) and the z-average radius of gyration (R_G ; $\circ$) determined with absolute-intensity light-scattering; (c) z-average translational diffusion coefficient (D_T ; $\bullet$) was determined with quasi-elastic light-scattering.

24×10^6 respectively. The z-average D_T , which is inversely related to M_r , decreases from $2.4 \times 10^{-8} \text{ cm}^2/\text{s}$ to $1.3 \times 10^{-8} \text{ cm}^2/\text{s}$ (4 M-guanidinium chloride; 20°C). By integrating over the whole distribution, a value for the weight-average M_r of 13.6×10^6 is obtained, which is in good agreement with that (11.1×10^6) obtained previously for the unfractionated material (Sheehan & Carlstedt, 1984a). The discrepancy could, for instance, be accounted for by a slight loss (approx. 4%) of the largest species during clarification by centrifugation when the unfractionated material was studied. On the other hand, to determine M_r from total-intensity light-scattering measurements requires that the absolute concentration of the sample is known. Consequently, determinations of M_r at both ends of the distribution are influenced by relatively large errors in the concentration measurements. The high values of M_r for fractions 3 and 4 may thus explain the slight overestimation of the average M_r .

The values given above rely on the assumption that sialic acid is uniformly distributed with size in the mucin populations. The distribution of mucins in the guanidinium chloride gradient was therefore compared with those obtained in sucrose gradients analysed at 280 and

Fig. 2. Distribution of cervical mucins after rate-zonal centrifugation in a sucrose gradient

The gradients [10–30% (w/v) sucrose] were prepared and the samples centrifuged as described in the Experimental section. The sample was monitored at 280 nm (a) and 206 nm (b). The macromolecules appear as a skew distribution similar to that obtained in the guanidinium chloride gradient but with no components resolved.

206 nm (Figs. 2a and 2b). A similar skew distribution was observed at both wavelengths, although different components were not resolved. Absorption at 280 nm is mainly due to the protein part of the macromolecules, whereas the oligosaccharides make a significant contribution at 206 nm. Since very similar distributions are obtained with sialic acid and with absorption at 280 nm as well as 206 nm, we conclude that sialic acid is indeed monitoring the true concentration of the mucins here. Furthermore, if this were not the case, the weight-average M_r obtained by integrating the whole distribution is not likely to compare well with that obtained for the unfractionated material. Individual fractions from the guanidinium chloride gradient were finally re-run in sucrose gradients and monitored at 206 nm to confirm that a true separation of species with different migration rates had been achieved (results not shown).

As discussed above, the data in Table 1 suggest that the guanidinium chloride gradient is isokinetic for cervical mucins under these conditions. The distribution of a polydisperse sample in such a gradient will reflect the initial distribution of sedimentation coefficients (see, e.g., Young, 1984). We have previously shown that the relationship between the sedimentation coefficient and

Fig. 3. Plot of $\log \Delta r$ versus $\log M_r$ for cervical mucins after rate-zonal centrifugation in a gradient of guanidinium chloride

The values of M_r are from Fig. 1(b) and Δr is the radial distance travelled by the molecules from the meniscus assuming a uniform cross-section of the tubes. The slope (α) obtained by linear regression with the use of least-squares analysis is 0.55. The broken line indicates how the weight-average M_r is obtained by interpolation at the position corresponding to the second moment of the distribution.

M_r for cervical mucins and fragments thereof can be expressed as $s = AM^\alpha$ in 6 M-guanidinium chloride (Sheehan & Carlstedt, 1984a). Therefore we expect a linear relationship between the logarithm of the radial distance travelled by the molecules in the gradient and $\log M_r$. From such a plot (Fig. 3) a value for the slope (α) of 0.55 is obtained, in agreement with that (0.46) found previously (Sheehan & Carlstedt, 1984a). By interpolation (Fig. 3) a value for the weight-average M_r of 12.4×10^6 is obtained at the radial position corresponding to the second moment of the distribution, close to that determined by using total-intensity light-scattering.

Plots of $\log R_G$ and of $\log (1/D_T)$ versus $\log M_r$ for the fractionated mucins are shown in Fig. 4(a) and Fig. 4(b) respectively. Data on the unfractionated whole mucins as well as for mucin fragments (subunits and T-domains) are included in the plots. A linear least-squares fit to the data in Fig. 4(a), omitting the 'last' two points, for which the errors are likely to be large, provides a straight line with a slope (α) of 0.56. However, a slight downward curvature can be discerned with increasing values of M_r , and the initial slope (T-domains $\rightarrow$ subunits) is 0.62, whereas that for the fractionated 'whole' mucins is 0.51. Thus it is possible that there is slight dependence of α on M_r .

A linear least-squares fit to all data in Fig. 4(b) yields a straight line with a slope (α) of 0.53, whereas the initial slope (T-domains $\rightarrow$ subunits) is 0.62 and that obtained for the fractionated material is 0.44, again suggesting a deviation from a straight line. The values of D_T were measured at 20° and were not extrapolated to zero angle. Although previous experiences (Sheehan & Carlstedt, 1984a) suggest that there is only a small angular

Fig. 4. Plot of $\log R_G$ (a) and $\log (1/D_T)$ (b) versus $\log M_r$

(a) ● Points are the data for the fractionated whole mucins from Fig. 1(b), whereas ●, ■ and ▲ points are respectively the values obtained for unfractionated whole mucins, subunits and T-domains in 4 M-guanidinium chloride. A value for the slope (α) of 0.56 was obtained by fitting all the data to a straight line with the use of least-squares analysis. (b) ● Points are the data for the fractionated whole mucins from Figs. 1(c), and ●, ■ and ▲ points are respectively the values obtained for unfractionated whole mucins, subunits and T-domains in 4 M-guanidinium chloride. A value for the slope (α) of 0.53 was obtained by fitting all the data to a straight line with the use of least-squares analysis.

dependence of D_T below 20° , it is possible that this behaviour could, in part, explain the decrease in α with increasing size. On the other hand, it is well known from studies on DNA that the slope of e.g. $\log [\eta]$ versus $\log M_r$ decreases with increasing values of M_r (see, e.g., Godfrey & Eisenberg, 1976). This behaviour has been interpreted as a gradual change from 'rod-like' to

'worm-like' behaviour with increasing size. In analogy, it is conceivable that the largest species of the whole mucins have configurational possibilities that are not available to, in particular, the mucin fragments.

General discussion

Polydispersity is a characteristic feature of mucus glycoproteins, with respect to both size and buoyant density. However, in this investigation the presence of populations with different mobilities in the gradient shows that heterogeneity could well be superimposed on polydispersity.

Proposed models for the macromolecular architecture of mucins must consider their broad size distribution. For a linear structure this property is well accounted for by a variation in the contour length. In contrast, the 'windmill' model proposed by Allen (1983) for pig gastric mucins could accommodate polydispersity only to the extent that is compatible with the size distribution of the four subunits. We find it unlikely that such a model could accommodate the degree of polydispersity that these authors have shown to be present in gastric mucins (Allen *et al.*, 1984). Furthermore, recent data obtained by using electron microscopy show that gastric mucins are well described as a polydisperse population of linear flexible chains (Sheehan *et al.*, 1986).

The relationship between R_G and M_r for a series of similar polymers can be expressed as $R_G = AM_r^\alpha$ where α is a shape-dependent parameter. The value (0.56) obtained here from a plot of $\log R_G$ versus $\log M_r$ for the fractionated whole cervical mucins and the fragments (subunits and T-domains) is in keeping with the macromolecules being linear flexible chains as proposed previously (Carlstedt *et al.*, 1983b; Sheehan & Carlstedt, 1984a; Sheehan *et al.*, 1986). The value of α is in excellent agreement with that (0.55) obtained from data for unfractionated whole cervical mucins and the same fragments (Carlstedt *et al.*, 1983b) as well as with that (0.56) for whole gastric, respiratory and cervical mucins and their subunits in 6 M-guanidinium chloride (Carlstedt & Sheehan, 1984). In a recent paper, Shogren *et al.* (1986) also report a value of $\alpha = 0.56$ from data for a number of preparations of pig submaxillary-gland mucins differing in size and subunits thereof. In addition, these authors showed that $\ln R_G$ for mucins from many different sources could be fitted to the same straight line when plotted against $\ln M_r$ for the *protein core* of the macromolecules, i.e. when correction was made for the differences in protein content. The excellent agreement between the various studies strongly suggests that the mucins investigated indeed have a common macromolecular architecture and thus that cervical, respiratory, gastric and submaxillary mucus glycoproteins all are well described as linear flexible chains.

Values of D_T can be related to M_r in a similar way, and a plot of $\log(1/D_T)$ versus $\log M_r$ for the fractionated whole cervical mucins and the fragments provides a slope (α) of 0.53. This value is in reasonable agreement with that (0.64) obtained previously for the unfractionated whole cervical mucins and their fragments in 6 M-guanidinium chloride (Sheehan & Carlstedt, 1984a) and that (0.58) found by Shogren *et al.* (1986) for pig submaxillary-gland mucins. The differences could possibly be explained by difficulties in correcting completely for the angular dependence of D_T and/or by a slight

dependence of α with M_r , as discussed above. In the latter case, the value of α would depend on the size range studied. However, the various sets of data support the proposed model, although more work might be needed to refine measurements of D_T for large mucins with photon correlation spectroscopy.

The direct analysis of fractions obtained from, for example, rate-zonal centrifugation or gel chromatography with modern laser light-scattering techniques is illustrated here and, in particular, the application of quasi-elastic light-scattering could certainly become very useful in this area. Because the concentration-dependence of D_T is negligible within the concentration range used, a value for an apparent D_T can, with appropriate calibration, be transformed into a value of M_r without measuring the absolute concentration. With this approach the size distribution can be examined for molecules too large to be studied by equilibrium methods in the analytical ultracentrifuge or where the concentration-dependence of the sedimentation coefficient makes use of the $g(s)$ function unreliable.

We thank Ms. Birgitta Jönsson for the artwork and Ms. Inger Carlsson for secretarial aid. This investigation was supported by grants from the Swedish Medical Research Council (project no. 7902), the Medical Faculty of Lund, Sweden, the Swedish Tobacco Company, Greta och Johan Kocks Stiftelser, Trelleborg, Sweden, The National Swedish Board for Technical Development, The Sunnerdahl Handikappsfond, Sweden, The Swedish Society for Cystic Fibrosis and the Cystic Fibrosis Research Trust, U.K.

REFERENCES

- Allen, A. (1983) *Trends Biochem. Sci.* **8**, 169–173
- Allen, A., Hutton, D. A., Mantle, D. & Pain, R. H. (1984) *Biochem. Soc. Trans.* **12**, 612–615
- Carlstedt, I. & Sheehan, J. K. (1984) *Biochem. Soc. Trans.* **12**, 615–617
- Carlstedt, I., Lindgren, H., Sheehan, J. K., Ulmsten, U. & Wingerup, L. (1983a) *Biochem. J.* **211**, 13–22
- Carlstedt, I., Lindgren, H. & Sheehan, J. K. (1983b) *Biochem. J.* **213**, 427–435
- Creeth, J. M. & Cooper, B. (1984) *Biochem. Soc. Trans.* **12**, 618–621
- Franzén, A., Björnsson, S. & Heinegård, D. (1982) *Anal. Biochem.* **120** 38–45
- Gibbons, R. A. (1963) *Nature (London)* **200**, 665–666
- Godfrey, J. E. & Eisenberg, H. (1976) *Biophys. Chem.* **5**, 301–318
- Harding, S. E., Rowe, A. J. & Creeth, J. M. (1983) *Biochem. J.* **209**, 893–896
- Hascall, V. C. & Sadjera, S. W. (1969) *J. Biol. Chem.* **244**, 2384–2396
- Jourdian, G. W., Dean, L. & Roseman, S. (1971) *J. Biol. Chem.* **246**, 430–435
- Kawahara, K. & Tanford, C. (1966) *J. Biol. Chem.* **241**, 3228–3232
- Lohmander, L. S., De Luca, S., Nilsson, B., Hascall, V. C., Caputo, C. B., Kimura, J. H. & Heinegård, D. (1980) *J. Biol. Chem.* **255**, 6084–6091
- Mikkelsen, A., Stokke, B. T., Christensen, B. E. & Elgsaeter, A. (1985) *Biopolymers* **24**, 1683–1704
- Rose, M. C., Voter, W. A., Brown, C. F. & Kaufman, B. (1984) *Biochem. J.* **222**, 371–377
- Sheehan, J. K. & Carlstedt, I. (1984a) *Biochem. J.* **217**, 93–101
- Sheehan, J. K. & Carlstedt, I. (1984b) *Biochem. J.* **221**, 499–504
- Sheehan, J. K., Oates, K. & Carlstedt, I. (1986) *Biochem. J.* **239**, 147–153

Shogren, R. L., Jamieson, A. M., Blackwell, J. & Jentoft, N. (1986) *Biopolymers* **25**, 1505–1517
Slayter, H. S., Lamblin, G., LeTreut, A., Galabert, C., Houdret, N., Degand, P. & Roussel, P. (1984) *Eur. J. Biochem.* **142**, 209–218

Tanford, C. (1961) *Physical Chemistry of Macromolecules*, pp. 297–316, John Wiley and Sons, New York
Young, B. D. (1984) in *Centrifugation – a Practical Approach*, 2nd edn. (Rickwood, D., ed.), pp. 127–159, IRL Press, Oxford

Received 17 November 1986/20 February 1987; accepted 15 April 1987