Carbohydrate Research, 160 (1987) 317–327
Elsevier Science Publishers B.V., Amsterdam – Printed in The Netherlands

LIGHT-SCATTERING STUDIES OF FRACTIONATED OVINE SUBMAXILLARY MUCINS

RANDAL L. SHOGREN[†], NEIL JENTOFT[†,‡], THOMAS A. GERKEN[†,‡], ALEXANDER M. JAMIESON[*], AND JOHN BLACKWELL[*]

Departments of Pediatrics[†], Biochemistry[‡], and Macromolecular Science[*], Case Western Reserve University, Cleveland, Ohio 44106 (U.S.A.)

(Received March 24th, 1986; accepted for publication in revised form, August 2nd, 1986)

ABSTRACT

Static and dynamic light-scattering studies of solutions of ovine submaxillary mucin (OSM) glycoproteins, fractionated by exclusion chromatography on Sephacryl S-1000, are reported. These experiments yielded information regarding the structure and conformation of the glycoprotein chain, in the form of weight-average molecular weights, M_w, z-average radius of gyration, $R_{g,z}$, and z-average of the inverse hydrodynamic radius, $\langle R_h^{-1} \rangle_z$. The values of $\langle R_h^{-1} \rangle_z$ are found to correlate very well with the S-1000 elution volume characteristics for four OSM fractions of different molecular weights. The structural parameters for these OSM fractions are, within experimental error, similar to those deduced for porcine submaxillary mucins (PSM) in earlier studies. The results suggest that, like PSM, the glycoprotein structure of OSM consists of linear chains constructed by covalently linking two or more elementary subunits together *via* disulfide bonds. In addition, the rigidity of the protein core of OSM is substantially greater than that observed for non-glycosylated-polypeptide random coils. Because $\langle R_h^{-1} \rangle_z$, and hence, elution volume depends only on the molecular weight of the mucin protein core, the M_w calibration obtained for OSM should be applicable to the chromatography of other mucin glycoproteins.

INTRODUCTION

Mucin glycoproteins constitute a family of biopolymers whose major chemical constituent is carbohydrate, and whose principal biological function is to impart viscoelastic, or viscosity-enhancing, character to gel-like, mucous secretions which perform protective functions in various biological systems[1–5]. The primary structure of mucins consists of a protein core, heavily substituted with oligosaccharides attached *via* glycosidic linkages to O-3 of serine and threonine residues. Three classes of mucin can be distinguished at this level of structure, based on the average length of the side chains. On average, submaxillary mucins have shorter oligosaccharide side chains (2–5 sugars) than cervical (5–10 sugars) or gastrointestinal

0008-6215/87/$ 03.50

mucins (10–15 sugars)[4,5].

In our earlier work, we focussed on the physicochemical characteristics of porcine submaxillary mucins (PSM). By applying light-scattering techniques[6–8], we established that, when appropriate precautions are taken to avoid mechanical shearing and proteolysis during purification, typical PSM preparations consist of polydisperse populations of molecules whose molecular weights are in the range $M_w = 1\text{–}20 \times 10^6$. Preparative chromatography on Sephacryl S-1000 is a convenient means for fractionating these species in order to obtain more-homogeneous preparations for structural analysis[9]. Based on studies of the molecular weight-dependence of the *z*-average radius of gyration, $R_{g,z}$, and the *z*-average inverse hydrodynamic radius, $\langle R_h^{-1} \rangle_z$, of PSM in 6M Gdn·HCl as solvent, we deduced that the mucin macromolecules exist as linear coils. Treatment of the high-molecular-weight fractions with 2-mercaptoethanol results in a decrease[6–8] in the molecular weight to $M_w = 1.8 \times 10^6$. Thus, the PSM glycoproteins appear to be synthesized by end-to-end linking of linear subunits, presumably *via* disulfide bonds.

A further observation is that the PSM macromolecules undergo a self-association process in 0.1M NaCl *via* non-covalent interactions at the chain ends to give a linear aggregate[8]. Earlier observations[6] of branched PSM aggregate structures in 0.1M NaCl were proved[8] to be an anomalous species produced by lyophilization. Finally, when plotted *versus* the estimated molecular weight of the protein core, $M_{p,w}$, the $R_{g,z}$ and $\langle R_h^{-1} \rangle_z$ values for submaxillary[8], cervical[10], and gastrointestinal[11] mucins could be superposed on universal curves[8]:

$$R_{g,z} = 0.57\, M_{p,w}^{0.56}\ (\text{Å}) \tag{1}$$

and

$$1/\langle R_h^{-1} \rangle_z = 0.268\, M_{p,w}^{0.58}\ (\text{Å}) \tag{2}$$

This suggests that the conformational properties of the protein core in each of these mucin classes is the same, and is independent of the average lengths of the carbohydrate side-chains[8]. In comparison to non-glycosylated protein coils in 6M guanidine (Gdn)·HCl–0.1M 2-mercaptoethanol, the characteristic ratio for the mucin-core peptide is approximately three times as large, reflecting a substantial increase in conformational rigidity because of the side chains[8].

We now report results of a light-scattering study of four fractions of ovine submaxillary mucin (OSM) glycoprotein prepared by methodology similar to that developed for our earlier PSM investigation. It is of interest to compare the conformational properties of these two species in view of the fact that the average size of the carbohydrate side-chains of the OSM polymer (~2.0 sugars) is even smaller than that of PSM (2–3.5 sugars per side chain).

EXPERIMENTAL

Methods. — (a) Isolation and purification. High-molecular-weight, native OSM was purified from frozen sheep-submaxillary glands as previously described[12]. Protease inhibitors were added in the initial steps, to lessen proteolysis, and the purified mucin solution was stored for a short time at 3° prior to light-scattering analysis. This avoids freeze-drying procedures which, in our prior studies of PSM, were found[8] to cause changes in mucin structure.

(b) Chromatography. Solid guanidine hydrochloride was added to the purified OSM in water, to give a final concentration of 5M Gdn·HCl containing ~1 to 1.5 mg of OSM per mL. A portion (~400 mL) of this solution was loaded onto a column (9 × 80 cm) of Sephacryl S-1000, the column was eluted with 5M Gdn·HCl, pH 7 (Aldrich 95%, treated with activated charcoal, and ultrafiltered) at 200 mL per h, and 33-mL fractions were collected. Fractions were assayed for carbohydrate by the periodic acid–Schiff assay[13], and pooled as follows: pool I, fractions 56–78; pool II, fractions 79–98; pool III, fractions 99–118; and pool IV, fractions 119–130. The four pools were then extensively dialyzed against water, and concentrated to ~1 mg per mL in a Speed Vac evaporator (Savant Instruments). The pools were finally made to 5M Gdn·HCl (Pierce, Sequanal grade), 10mM phosphate buffer, pH 7, and were rerun on a column (120 × 1 cm) of Sephacryl S-1000 eluted with the same buffer. To the column was applied ~1.5 mL of OSM (~0.6 mg/mL); the flow rate was 5 mL per h, and 1.6-mL fractions were collected. Mucin was detected by measuring the u.v. absorbance of the fractions at 228 nm.

(c) Light-scattering analysis. These experiments were carried out by using methods described previously[6–8]. Briefly, solutions of OSM pools I–IV in water were made to 6M Gdn·HCl (Pierce), 10mM phosphate buffer, pH 6.7. Mucin concentrations were in the range of 50–200 μg per mL, based on residual weights obtained after the solutions were dialyzed against water and then lyophilized. The solutions were filtered through 5-μm Millipore filters, diluted with filtered buffer into light-scattering cells, and centrifuged for 1 h at 5000 r.p.m. immediately prior to light-scattering analysis. Less than 5% of the mucin was lost on filtration, as estimated by periodic acid–Schiff determination[13] of the filtered and unfiltered solutions.

Light-scattering measurements were performed by using laser light of wavelength $\lambda = 6328$ Å and a BI 240 photogoniometer with BI 2020 autocorrelator[6–8] (Brookhaven Instruments Corp., Ronkonkoma, NY). Values for M_w, $R_{g,z}$, and the second osmotic virial coefficient, A_2, were determined from Zimm plots of average intensities of scattered light at scattering angles, θ, of 20–70°. The refractive index increment at constant chemical potential for OSM in 6M Gdn·HCl was found to be 0.13 mL per g, measured after exhaustive dialysis against solvent for one week. The intensity autocorrelation function of the scattered light was analyzed by using the method of cumulants in order to derive the first moment, Γ, and the second moment μ_2. Γ was measured as a function of OSM concentration

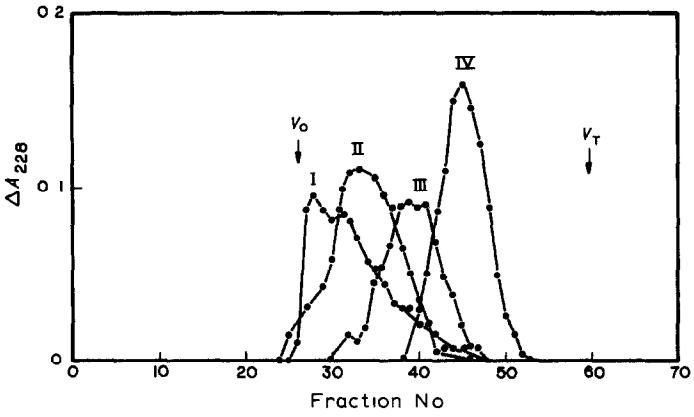

Fig. 1 Pools I–IV rechromatographed on Sephacryl S-1000 to confirm that the pools contained macromolecules whose sizes correspond to the elution characteristics of the preparative column. (Protein content determined by absorbance at 228 nm)

and scattering angle over the same range as already described. The concentration dependence of Γ was negligible over this concentration range, and hence, $D^0_{t,z}$ was determined by plotting $\Gamma/\mathbf{q}^2$ against $\mathbf{q}^2$ and extrapolating to $\mathbf{q} = 0$, where $\mathbf{q}$ is the scattering vector [$\mathbf{q} = (4\pi n/\lambda)\sin\theta/2$].

RESULTS

Fig. 1 shows the elution profiles of OSM Pools I–IV on Sephacryl S-1000 using 5M Gdn·HCl as solvent. The elution positions, V_e, of each pool confirm the absence of aggregation effects, demonstrating that these preparations contain macromolecules of various sizes corresponding to the elution characteristics of the preparative column. Table I summarizes the yield from the fractionation procedure, together with values of the distribution coefficient K_D [$= (V_e - V_0)/(V_T - V_0)$] determined from the analytical column, using λ and T4 phage DNA as markers for the void volume, V_0, and adenosine 5′-monophosphate as a marker for the total volume, V_T. Also shown are values of the ratio of the u.v. absorbance at 280 nm to that at 228 nm.

Fig. 2. is the Zimm plot representing angular and concentration extrapolations

TABLE I

YIELDS FROM FRACTIONATION, AND VALUES OF DISTRIBUTION COEFFICIENT

Pool	*Wt (mg)*	A_{280}/A_{228}	K_D
I	24	0.110	0 12
II	56	0 087	0.24
III	37	0.074	0 41
IV	12	0 070	0 56

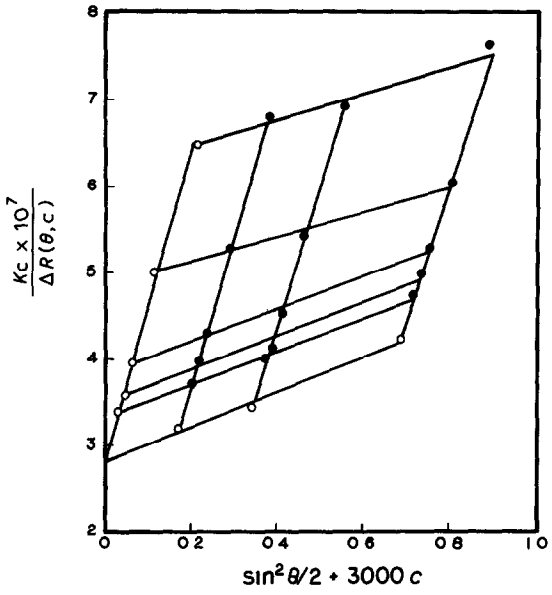

Fig. 2. Zimm plot of light-scattering intensities of OSM Pool II in 6M Gdn·HCl and 10mM phosphate buffer, pH 7.0 (This plot furnishes $M_w = 3\,6 \times 10^6$ $R_{g,z} = 1530$ Å, and $A_2 = 3.0 \times 10^{-4}$ cm^3.mol.g^2)

of the scattered intensity of OSM Pool II in 6M Gdn·HCl and 10mM phosphate buffer, pH 7.0. These plots, and similar extrapolations of the apparent *z*-average diffusion coefficient, provide values for $R_{g,z}$, $\langle R_h^{-1}\rangle_z$ and M_w.

Table II summarizes data for M_w, $R_{g,z}$, $\langle R_h^{-1}\rangle_z$, and μ_2/Γ^2. Also listed are values of the ratio $\rho = R_{g,z}\,\langle R_h^{-1}\rangle_z$ for each pool. It is noted that the M_w values for each fraction, together with the pooled weights of each fraction listed in Table I, lead to a calculated value $M_w = 3.2 \times 10^6$ for the entire sample. A further experiment was carried out by adding 10mM 2-mercaptoethanol to pool II. This resulted in a decrease of the molecular weight from $M_w = 3.6 \times 10^6$ to $M_w = 1.7 \times 10^6$, and of the radius of gyration from $R_{g,z} = 1530$ Å to $R_{g,z} = 1000$ Å.

DISCUSSION

In Figs. 3 and 4 are respectively plotted $\log R_{g,z}$ and $\log D^0_{t,z}$ *versus* $\log M_{p,w}$

TABLE II

STATIC AND DYNAMIC LIGHT-SCATTERING PARAMETERS OF OSM IN 6M Gdn·HCl, 10mM PO_4, AT pH 6 7

Pool	$M_w \times 10^{-6}$	$R_{g,z}$ (Å)	$1/\langle R_h^{-1}\rangle_z$ (Å)	ρ	μ_2/Γ^2
I	5.2 ±0 6	1530 ±150	1160 ±60	1.3	0 3–0.4
II	3 6 ±0.3	1530 ±150	830 ±46	1 8	0.2–0.3
III	2.2 ±0 2	1080 ±110	595 ±30	1 8	0.3–0.4
IV	0.78 ±0 1	920 ±90	450 ±25	2 0	0 5–0.6

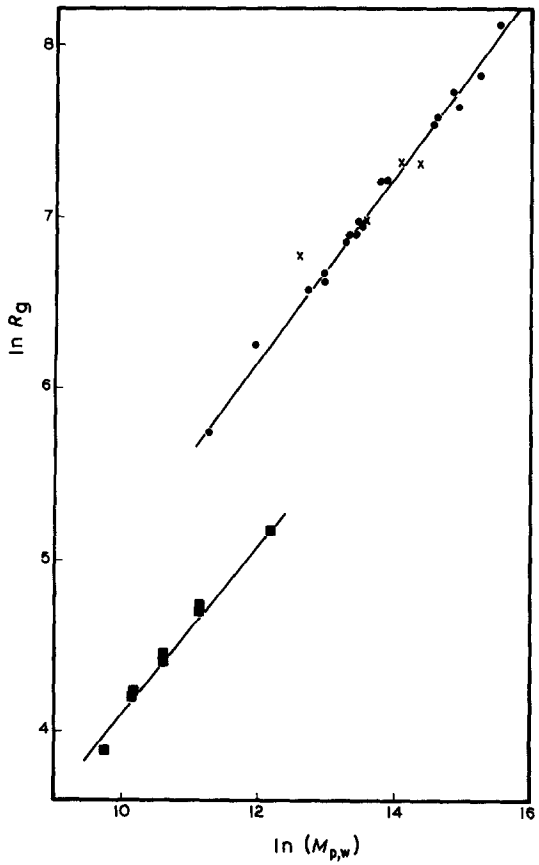

Fig 3 Natural logarithm of the *z*-average radius of gyration of mucin glycoproteins in aqueous solvents plotted *versus* the natural logarithm of the weight-average molecular weight of the protein backbone [Key: (×), OSM Pools I–IV in 6M Gdn·HCl–10mM phosphate, pH 7.0, (——) least-squares fit to collected data (●) for various mucin species, as described in ref 8, (■) data for non-glycosylated protein random coils (see ref 8)]

for OSM fractions I–IV. As in earlier work[8], $M_{p,w} = M_w \times$ (% protein), where, for OSM (ref. 14), % of protein = 37%. Also shown are the experimental relationships for $\log R_{g,z} - \log M_{p,w}$, and $\log D_{t,z} - \log M_{p,w}$ for mucins, summarized as Eqs. *1* and *2*. Clearly, when scaled against the estimated molecular weight of the protein core, the experimental data for OSM are reasonably consistent with the relationship previously derived. This suggests that the OSM glycoproteins are also linear coils whose core protein conformation is highly expanded compared to that of non-glycosylated polypeptide random coils, a conclusion that is similar to that deduced for other mucin species having longer carbohydrate chains. This deduction is supported quite well by experimental values of $\rho = R_{g,z} \langle R_h^{-1} \rangle_z = 1.8$–$2.0$ for OSM fractions II–IV. Such values are representative of polydisperse flexible coils in poor solvents at the non-draining limit. The comparatively low value $\rho = 1.3$ determined

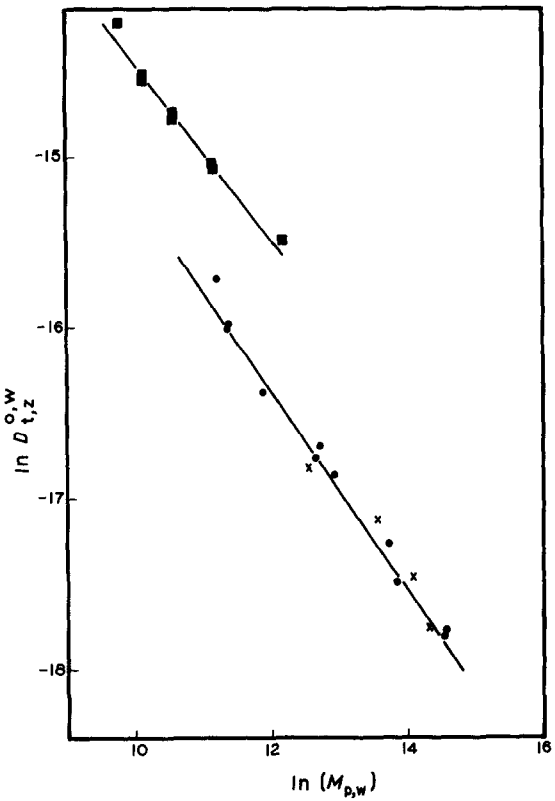

Fig 4 Natural logarithm of the *z*-average translational diffusion coefficient of mucin glycoproteins in aqueous solvents, extrapolated to zero angle and zero concentration, and corrected to the viscosity of water at 20°, plotted *versus* the natural logarithm of the weight-average molecular weight of the protein backbone [Key: (×), OSM Pools I–IV in 6M Gdn HCl–10mM phosphate, pH 7 0; (——) least-squares fit to collected data (●) for various mucin species, as described in ref 8, (■) data for non-glycosylated protein random coils (see ref 8)]

for pool I may reflect the presence of intramolecular cross-links or some other anomaly in the structure, composition, and/or molecular weight distribution of this pool. In this regard, it is noted that this pool appeared bimodal on rechromatography (see Fig. 1) and also showed a significantly higher ratio A_{280}/A_{228} than did pools II–IV, which may indicate contamination by non-mucin species.

It was of interest to investigate whether our light-scattering experiments are consistent with the exclusion chromatography behavior of the OSM fractions on Sephacryl S-1000. We have tested this in two ways in Fig. 5. First, we compare values for the hydrodynamic radius with those of the distribution coefficient K_D according to the phenomenological equation of Ackers[15], as modified by Ohno *et al.*[16],

$$R_h = R_0 + \Delta R \, \mathrm{erfc}^{-1}(2K_D), \qquad (3)$$

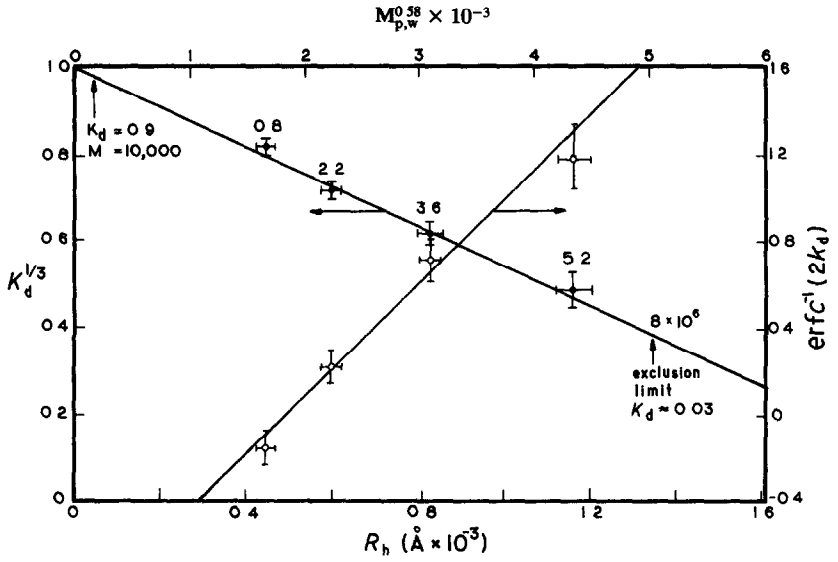

Fig. 5 The distribution coefficient values, K_D, for elution of OSM Pools I–IV on Sephacryl S-1000, compared with the hydrodynamic radii in 6M Gdn·HCl–10mM phosphate, pH 7.0 [Data plotted: (a) (○) as $\mathrm{erfc}^{-1}(2K_D)$ *versus* R_h, together with a least-squares fit to Eq. *3*, and (b) (●) as $K_D^{1/3}$ *versus* R_h, together with least-squares fit to Eq *4*]

where R_0 is the mean pore size of the gel matrix, ΔR is the standard deviation of the pore size distribution, which is assumed to be normal in form, and erfc^{-1} refers to the inverse error function complement. As may be seen, the results are consistent with Eq. *3*, yielding numerically reasonable values $R_0 = 500$ Å and $\Delta R = 500$ Å. An alternative and, perhaps, physically more realistic, model has been proposed by Porath[17] in which the exclusion characteristics of a solute molecule are determined based on its ability to penetrate a conical pore. The equation derived is of the form.

$$K_D^{1/3} = A - BR_h, \tag{4}$$

and is tested against K_D and R_h data for the OSM fractions in Fig. 5. Again, good correlation is found with values $A = 1$ and $B = 2200$ Å^{-1}. From the theory of Porath[17], a reasonable value of the base diameter of the pores, $D = 4400$ Å, was deduced.

Using the molecular-weight calibration of the S-1000 column shown in Fig. 5, which corresponds to a combination of Eqs. *2* and *4*, and setting $1/\langle R_h^{-1}\rangle_z \equiv R_h$, it is possible to calculate the mass distribution of molecular weights $M\rho(M)$, of fractions I–IV from the mass distribution of K_D values $\rho(K_D)$, because

$$\rho(K_D)\Delta K_D = M_p\rho(M_p)\Delta M_p = M\rho(M)\Delta M, \tag{5}$$

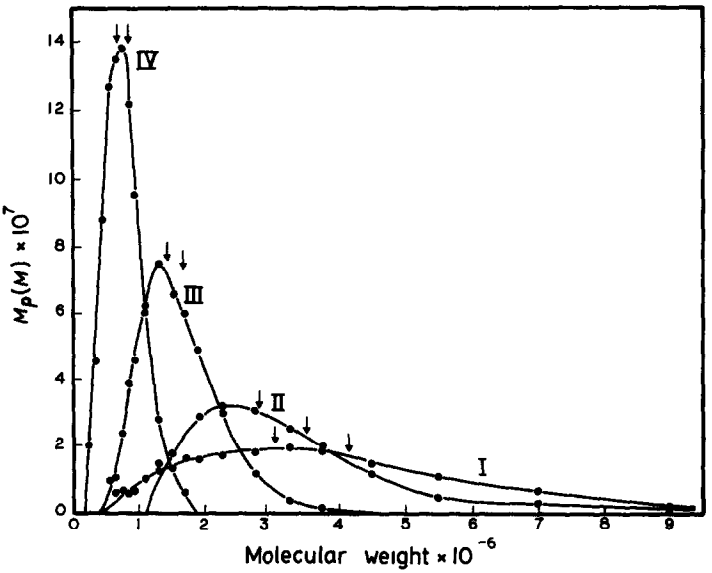

Fig 6 The mass distribution of molecular weights for OSM fractions I–IV, calculated from the Sephacryl S-1000 separation profiles in Fig. 1, using the molecular weight calibration for K_D *versus* $M_{p,w}$ shown in Fig. 5 [The arrows indicate location of calculated values of M_n and M_w for each pool: Pool IV, M_n = 680,000, M_w = 790,000; Pool III, $M_n = 1\ 44 \times 10^6$, M_w 1 66 × 10^6; Pool II, $M_n = 2\ 8 \times 10^6$, $M_w = 3.5 \times 10^6$, Pool I, $M_n = 3\ 0 \times 10^6$, $M_w = 4.1 \times 10^6$.]

where $\rho(M_p)$ and $\rho(M)$ are the number distributions of core protein molecular weight and mucin molecular weight, respectively. Hence,

$$M\rho(M) = M_p\rho(M_p)\frac{dM_p}{dM} = \rho(K_D)\frac{dK_D}{dM_p}\frac{dM_p}{dM} \tag{6}$$

$$= \rho(K_D)(1 - aBc^bM^b)\ 3abBc^bM^{b-1}, \tag{7}$$

where $K_D = (1 - BR_h)^3$ and $R_h = aM_p^b$ with $M_p = cM$ are used. Fig. 6 shows the results of such computations using $a = 0.268$, $b = 0.58$, $c = 0.37$, and $B = 2200$ Å^{-1}. Calculated values of M_w and M_n for each OSM pool are also shown. Interestingly, this analysis, which neglects effects of column spreading, generates values of $M_w/M_n < 1.36$, indicating that the fractions are of comparatively narrow molecular-weight distribution.

Because the correlation between K_D and R_h is universal for all polymers on a given gel, Fig. 5 can be used to obtain R_h for any polymer, once K_D is determined. The OSM fractions are, however, polydisperse, and the hydrodynamic radius measured by dynamic light-scattering is weighted to the largest species (although not as strongly as $R_{g,z}$), so that the R_h for a monodisperse sample having the same K_D would be slightly smaller[18]. It may also be noted that, as $1/\langle R_h^{-1}\rangle_z \propto M_{p,w}^{0.58}$ for all mucins, regardless of the size of the carbohydrate side-chain, it follows that it

should be possible to obtain an estimate of M_w for any mucin from the measured K_D on Sephacryl S-1000 when the protein-to-carbohydrate ratio is known.

Finally, it is noted that the molecular weights listed in Table II for these OSM fractions (M_w = 780,000–5.2 × 10^6) are smaller than those observed for PSM fractions of similar preparative history, which were also obtained by elution from Sephacryl S-1000 (M_w = 4 × 10^6–12 × 10^6). This comparison has been made for only one preparation of OSM, however, and thus may not reflect a real difference. Also, the molecular weight of the subunits derived by cleaving disulfide bonds in the OSM pool II (M_w = 1.7 × 10^6) is similar to that determined for the corresponding subunits obtained by similar treatment of PSM fractions[6–8] (M_w = 1.8 × 10^6). Thus, the possibility exists that the average number of covalently linked subunits in the fresh OSM preparations may be smaller than that of those present in comparable OSM samples. This could indicate either a lower degree of polymerization in the biosynthesized material, or a greater ease of enzymic or hydrolytic degradation, in OSM.

In summary, light-scattering measurements of M_w, $R_{g,z}$, and $\langle R_h^{-1}\rangle_z$ are presented for four ovine submaxillary mucin glycoprotein fractions prepared by Sephacryl S-1000 exclusion chromatography. The values of $\langle R_h^{-1}\rangle_z$ measured for each fraction are consistent with their elution characteristics based on chromatography theory. The experimental data suggest that, like PSM, the OSM glycoprotein is a linear, oligomeric species formed by covalently linking linear subunits *via* end-to-end disulfide bonds. Like PSM, the rigidity of the protein core of OSM is substantially larger than that observed for non-glycosylated polypeptide random coils.

ACKNOWLEDGMENTS

This work was supported by NIH grants AM 27651, AM 33365, and HL 32227, and by grants from the Rainbow Chapter of the Cystic Fibrosis Foundation.

REFERENCES

1 I CARLSTEDT, J K SHEEHAN, A P CORFIELD, AND J T GALLAGHER, *Essays Biochem*, 20 (1985) 40–76.
2 A ALLEN, *Trends Biochem. Sci.*, 8 (1983) 169–173
3 A. E BELL, A ALLEN, E R MORRIS, AND S. B. ROSS-MURPHY *Int J Biol Macromol*, 6 (1984) 309–315
4 E F HOUNSELL AND T. FEIZI, *Med. Biol*, 60 (1982) 237–254
5 L A TABAK, M J. LEVINE, I D MANDEL, AND S A ELLISON, *J Oral Pathol*, 11 (1982) 1–17
6 R. L SHOGREN, A M JAMIESON, J BLACKWELL, P W CHENG, D G DEARBORN, AND T F BOAT, *Biopolymers*, 22 (1983) 1657–1675.
7 R L SHOGREN, A M. JAMIESON, J BLACKWELL, AND N JENTOFT, *J Biol Chem.*, 259 (1984) 14,657–14,662
8 R L SHOGREN, A M. JAMIESON, J BLACKWELL, AND N JENTOFT, *Biopolymers*, 25 (1986) 1505–1517
9 N JENTOFT, R L SHOGREN, A M JAMIESON, J BLACKWELL, AND J E. JENTOFT, manuscript in preparation
10 J. K SHEEHAN AND I CARLSTEDT, *Biochem J*, 217 (1984) 93–101

11 D SNARY, A. ALLEN, AND R. H. PAIN, *Biochem. J*, 141 (1974) 641–646
12 T A GERKEN AND D. G. DEARBORN, *Biochemistry*, 23 (1984) 1485–1497
13 M MANTLE AND A. ALLEN, *Biochem Soc. Trans.*, 6 (1978) 607–609
14 H D HILL, JR, J A REYNOLDS, AND R. L. HILL *J Biol. Chem*, 252 (1977) 3791.
15 G K ACKERS, *J. Biol Chem.*, 242 (1967) 3237–3238.
16 H OHNO, J BLACKWELL, A M JAMIESON, D A. CARRINO, AND A. I. CAPLAN, *Biochem. J.*, 235 (1986) 553–557
17 J PORATH, *Pure Appl. Chem.*, 6 (1963) 233–244
18 P PATTERSON AND A M JAMIESON, *Macromolecules*, 18 (1985) 266–272