Eur Biophys J (1997) 25: 477–480

ARTICLE

Kornelia Jumel · Fiona J.J. Fogg · David A. Hutton
J.P. Pearson · Adrian Allen · Stephen E. Harding

A polydisperse linear random coil model for the quaternary structure of pig colonic mucin

Received: 29 November 1996 / Accepted: 2 December 1996

Abstract The distribution of molecular weights for polymeric colonic mucus glycoprotein or "mucin" isolated and solubilised in the presence of protease inhibitors from pig colons is shown to be considerably greater than its "subunit" (thiol reduction product) and papain digested forms using the technique of size-exclusion chromatography coupled to multi-angle laser light scattering, and confirmed by sedimentation equilibrium measurements. The conformation of this mucin is probed by examining the molecular weight – intrinsic viscosity relationship in terms of the Mark-Houwink-Kuhn-Sakurada analysis for its polymeric (or "whole"), reduced and papain-digested forms: an exponent "*a*" of (1.1±0.1) is obtained indicating a linear random coil conformation consistent with other mucins. Size-exclusion chromatography coupled to multi-angle laser light scattering is shown to provide a relatively simple complementary technique to sedimentation equilibrium for the molecular weight distribution analysis of polydisperse materials.

Key words Mucin conformation · Molecular weight distribution

Introduction

Our understanding of the molecular structure of mucins – large molecular weight glycoproteins of up to 90% carbohydrate (existing as serine or threonine O-linked side chains from 3–30 residues length built around a polypeptide backbone and which give mucus in the tracheobronchial, reproductive and gastrointestinal tracts its characteristic protective properties) – has increased considerably over the last two decades with (i) the establishment of isolation protocols which preserve the molecular integrity of the mucin structure (Carlstedt et al. 1985) and (ii) the identification and partial sequencing of up to 9 human genes called "MUC genes" (Gendler and Spicer 1995) which code for the polypeptide backbone depending on its source. These genes include code for tandemly repeated sequences of amino acids, thought to be the principal regions of glycosylation of the polypeptide backbone. In the colon at least three of these genes are expressed: MUC2, 3 and 4 (Audie et al. 1993).

In a recently published paper (Fogg et al. 1996) we showed using both size-exclusion chromatography coupled to multi-angle laser light scattering and sedimentation equilibrium in the analytical ultracentrifuge that mucin isolated from pig colons and solubilised by mild homogenisation in the presence of protease inhibitors had a molecular size typical of other mucin glycoproteins [weight average $M_w=(5.5\pm0.3)\times10^6$ g/mol]. This mucin fraction contained 70% of the per-iodic acid Schiff stain (Mantle and Allen 1978) positive material in the adherent gel. Reduction by mercaptoethanol yielded a "subunit" molecular weight of $M_w=(2.1\pm0.5)\times10^6$ g/mol and digestion by papain yielded a form of $M_w=(0.60\pm0.04)\times10^6$ g/mol, suggesting a subunit composition similar to other mucins. In this short paper we (i) provide quantitative molecular weight *distribution* information on the mucin and its reduced (or "subunit") and papain digested (or "basic unit") forms and (ii) show from the (sedimentation equilibrium and light scattering) molecular weight versus intrinsic viscosity scaling relation that the colonic mucin in either its polymeric, reduced or digested forms adopts a "linear random coil" conformation proposed as a general model for mucins by Sheehan and Carlstedt (1989).

K. Jumel · S.E. Harding (✉)
National Centre for Macromolecular Hydrodynamics,
University of Nottingham, Sutton Bonington, LE12 5RD, UK

F. J. J. Fogg · D. A. Hutton · J. P. Pearson · A. Allen
Department of Physiological Sciences, Medical School,
Framlington Place, The University,
Newcastle-upon-Tyne NE2 4HH, UK

Molecular weight distribution analysis

Isolation of the mucin, reduction by mercaptoethanol, digestion by papain and measurement of weight average mo-

lecular weights by size-exclusion chromatography coupled to multi-angle laser light scattering, and sedimentation equilibrium have been reported previously (Fogg et al. 1996). All measurements were performed in a Na^+/K^+ phosphate-chloride buffer (pH = 6.8, I = 0.3 M) containing 1 mM EDTA and 0.01% (w/v) sodium azide. Determinations of M_w were performed on three separate preparations (1, 2, and 3) for each case. Sedimentation equilibrium determinations on "sample 1" for each of the cases "polymeric", "reduced" and "digested" forms had also been performed and shown to be in good agreement with the light scattering results.

We can take advantage of the on-line fractionation and analysis ability of size-exclusion chromatography coupled to multi-angle laser light scattering method to determine the molecular weight distribution for the polymeric mucin and its treated derivatives. This procedure is outlined in Wyatt (1993): the molecular weight of each volume "slice" is determined using a "Debye" plot of R_θ/Kc, where R_θ is the Rayleigh ratio and K is the usual constant ($=2\,\pi^2\,n_0^2\,(\mathrm{d}n/\mathrm{d}c)^2/(N_A\,\lambda^4)$ with n_0 the refractive index of the solvent, $\mathrm{d}n/\mathrm{d}c$ the refractive index increment = 0.165 ml/g, N_A is Avogadro's number and λ the wavelength of the incident radiation (632.28 nm)). Up to 18 angles are recorded simultaneously and the extrapolation to zero angle is performed almost instantaneously. Because of the low concentration (~1.0 mg/ml) injected onto the size exclusion chromatography columns [two analytical PSS HemaBio columns (HemaBio linear and HemaBio 40) preceeded by a PSS Hema guard column to eliminate supramolecular aggregates] thermodynamic non-ideality can be reliably ignored. From calibration plots of elution volume, V_e, versus corresponding molecular weight for each value of V_e, it is trivial to convert the elution profiles into molecular weight distributions, and Fig. 1 shows these for the polymeric, reduced and digested forms. As expected, the polymeric form is clearly considerably more polydisperse than the reduced (subunit) and digested ("basic unit") forms.

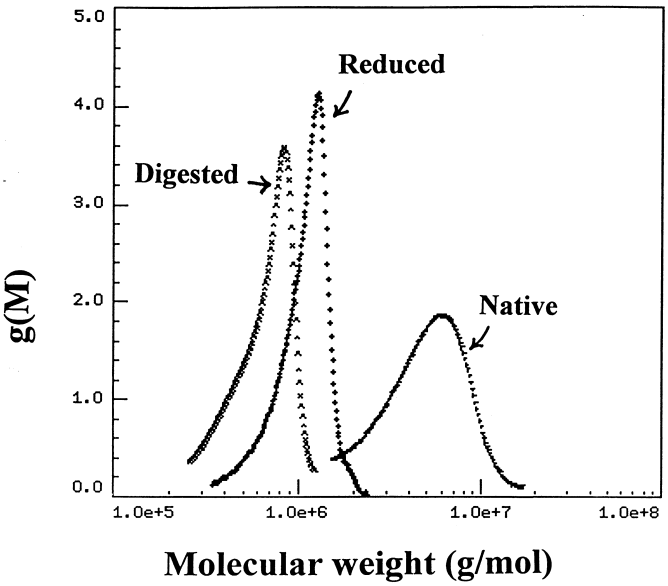

Fig. 1 Molecular weight distributions for polymeric, reduced and digested pig colonic mucin from of size-exclusion chromatography coupled to multi-angle laser light scattering analysis

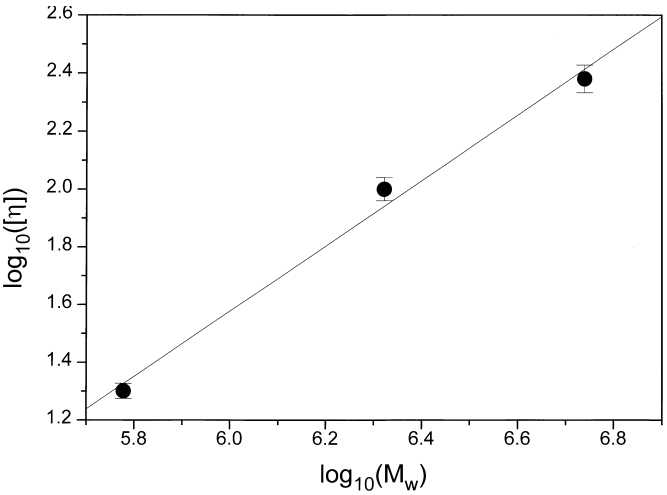

Fig. 2 Mark-Houwink-Kuhn-Sakurada [η] versus M_w plot for pig colonic mucin based on values for the native mucin, subunit and papain digested or "basic unit", following the lines of Sheehan and Carlstedt (1984) for human cervical mucin. The slope, $a = (1.1 \pm 0.1)$. The following "consensus" molecular weights were used, based on weight averages determined by sedimentation equilibrium and the size-exclusion chromatography coupled to multi-angle light scattering results: Native: $(5.5 \pm 0.3) \times 10^6$ g/mol; subunit $(2.1 \pm 0.5) \times 10^6$ g/mol and papain digested $(0.60 \pm 0.04) \times 10^6$ g/mol

Conformation analysis: linear random coil model

Silberberg and Mayer (1981) introduced the concept of a "mucin basic unit" of molecular weight ~530000 g/mol consisting of a heavily glycosylated backbone with one or two "naked" regions of peptide with no or little glycosylation at either end. This is essentially equivalent to the "T-domain" or "trypsin domain" (i.e. product left after trypsin digestion) introduced by Carlstedt and Sheehan to describe the structure of initially cervical and then pig and bronchial mucins (see e.g. Carlstedt and Sheehan 1984, 1989), but without most of the naked peptide. Similarly, Scawen and Allen (1977) described a protease resistant mucin unit. One or more of these basic units linked together forms the subunit which is the product remaining after reduction of the polymeric mucin by thiols. From Table 1 of Fogg et al. (1996) we can therefore infer that there are 3–4 basic units on average which make up a subunit and on average 3 subunits (or 9–12 basic units) which make up the polymeric mucin, but with polydispersity (cf Fig. 1) in terms of the numbers of basic units, the extent of glycosylation and the numbers of subunits.

The question which follows naturally next is: how are these units linked together? Earlier models based on pig gastric mucin (isolated with proteolytic inhibitors) suggested a windmill model, consistent with an overall spher-

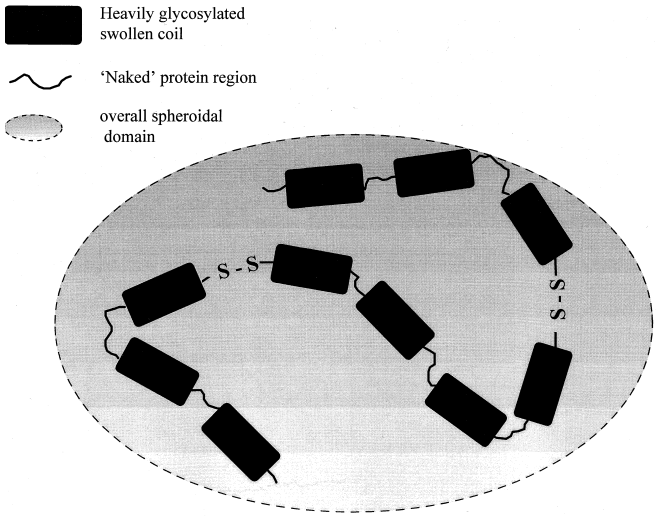

Fig. 3 Linear random coil model for pig colonic mucin. The *shaded area* shows the "effective" overall spheroidal volume of influence of the macromolecule

oidal domain of the molecule as suggested by hydrodynamic measurements and also N-terminal analysis (Allen 1981). This provided a powerful starting point for future refinement and subsequent studies involving electron microscopy, (Harding et al. 1983a; Sheehan and Carlstedt 1984), the Wales van-Holde ratio of the sedimentation concentration dependence coefficient, k_s to the intrinsic viscosity, $[\eta]$ (Creeth and Knight 1967; Sheehan and Carlstedt 1984, 1989) and Mark-Houwink-Kuhn-Sakurada double logarithmic plots of $[\eta]$ versus M_w, the radius of gyration R_g versus M_w or translational diffusion coefficient D versus M_w all strongly suggested a linear molecule with the overall conformation of a random coil, which in turn occupied a spheroidal solvent domain. This "linear random coil model" is essentially the same as the "swollen coil array" model proposed earlier (Harding et al. 1983b) in which the individual units were also shown to have expanded coil like properties, consistent with the high proline content of mucins.

Since the intrinsic viscosities had also been measured for each of the colonic mucin forms (Fogg et al. 1996), following Sheehan and Carlstedt (1984) who performed the analogous plot for human cervical mucin we can perform a simple fit of $\log_{10}([\eta])$ versus $\log_{10}(M_w)$ (Fig. 2) to the Mark-Houwink-Kuhn-Sakurada scaling relation (see Mark 1938; Kuhn and Kuhn 1943, and Harding 1995 and references cited therein):

$$\log_{10}([\eta]) = \log_{10}(K') + a \cdot \log_{10}(M_w)$$

which gives $K' = 6.3 \times 10^{-6}$ and more significantly $a = (1.1 \pm 0.1)$. For a compact sphere, $a = 0$; for a random coil $a = 0.5 - 0.8$; for a rigid rod $a = 1.8$. Quite clearly both compact sphere and rigid rod forms can be ruled out and the most likely conformation is a slightly stiff linear random coil with the "basic units" and the subunits linked end- to end into the polymeric heterostructure. Figure 3 shows such a model based on the original scheme given for cervical mucin by Carlstedt et al. (1983), and where we have added the effective solvent domain.

Comment

It is worth commenting on the relative merits of the size-exclusion chromatography coupled to multi-angle laser light scattering procedure compared with the sedimentation equilibrium method for obtaining molecular weight distribution information. To obtain an equivalent molecular weight distribution using sedimentation equilibrium by coupling sedimentation equilibrium off-line with size-exclusion chromatography (Ball et al. 1988) would have required extensive work: fractionation and collection of samples from preparative size exclusion chromatograpphy followed by measurement by sedimentation equilibrium, and subsequent calibration of the column. Direct determination involving mathematical manipulation of the concentration distribution data is both time consuming, unwieldy (especially if non-ideality is taken into account – see Harding 1985) and has to be done by model fitting (log-normal, Gaussian, three species etc.): this illustrates the fact that size-exclusion chromatography coupled to multi-angle laser light scattering or "SEC/MALLS" method is really the method of choice for the determination of molecular weight distributions of polydispersed materials (of course it is a different matter for the characterization of self-associating or interacting systems where particular advantage can be made of the concentration distribution at sedimentation equilibrium). Sedimentation equilibrium is valuable in the context of polydisperse materials in the determination of the weight average molecular weight (and z-average molecular weight, if Schlieren optical records are used, or if Rayleigh or absorption optical records are of sufficient quality to be amenable to further analysis for the extraction of M_z). This latter thus provides independent conformation of a distribution determined from SEC/MALLS and its corresponding M_w as we have shown for the pig colonic mucin system.

References

Allen A (1981) Structure and function of mucus. In Johnson LR (ed) Physiology of the gastrointestinal tract. Raven Press, New York, pp 617–639

Audie JP, Janin A, Porchet MC, Copin B, Gosselin B, Aubert JP (1993) Expression of mucin genes in respiratory, digestive and reproductive tracts ascertained by *in situ* hybridisation. J Histochem Cytochem 41:1479–1485

Ball A, Harding SE, Mitchell JR (1988) Combined low-speed sedimentation equilibrium/gel permeation chromatography approach to molecular weight distribution analysis. Application to a sodium alginate. Int J Biol Macromol 10:259–264

Carlstedt I, Lindgren H, Sheehan JK (1983) The macromolecular structure of human cervical-mucus glycoprotein. Biochem J 213: 427–435

Carlstedt I, Sheehan JK (1984) Is the macromolecular architecture of cervical, respiratory and gastric mucins the same? Biochem Soc Trans 12:615–617

Carlstedt I, Sheehan JK (1989) Models for the macromolecular structure of mucus glycoproteins. In: Harding SE, Rowe AJ (eds) Dynamic properties of biomolecular assemblies. Royal Society of Chemistry, Cambridge, UK, pp 256–275

Carlstedt I, Sheehan JK, Corfield AP, Gallagher JT (1985) Mucus glycoproteins: a gel of a problem. Essays Biochem 20:40–76

Creeth JM, Knight CG (1967) The macromolecular properties of blood group substances. Sedimentation velocity and viscosity measurements. Biochem J 105:1135–1145

Fogg FJJ, Hutton DA, Jumel K, Pearson JP, Harding SE, Allen A (1996) Characterization of pig colonic mucins. Biochem J 316: 937–942

Gendler SJ, Spicer AP (1995) Epithelial mucin genes. Ann Rev Physiol 57:607–634

Harding SE, Rowe AJ, Creeth JM (1983a) Further evidence for a flexible and highly expanded spheroidal model for mucus glycoproteins in solution. Biochem J 209:893–896

Harding SE, Creeth JM, Rowe AJ (1983b) Modelling the conformation of mucus glycoproteins in solution. In: Chester A, Heinegard A, Lundblad A, Svensson S (eds). Proc 7th Int Glycoconjugates Conf Olsson-Reklambyra, pp 558–559

Harding SE (1985) The representation of equilibrium distributions for non-ideal polydisperse systems in the analytical ultracentrifuge. Application to mucus glycoproteins. Biophys J 47:247–250

Harding SE (1995) On the hydrodynamic analysis of macromolecular conformation. Biophys Chem 55:69–93

Kuhn W, Kuhn H (1943) Die Frage nach der Aufrollung von Fadenmolekülen in strömenden Lösungen. Helv Chim Acta 26: 1394–1465

Mantle M, Allen A (1978) A colourimetric assay for glycoproteins based on the periodic/Schiff stain. Biochem Soc Trans 6:607

Mark H (1938) Der feste Körper. Hirzel, Leipzig, p 103

Scawen M, Allen A (1977) The action of proteolyic enzymes on the glycoprotein from pig gastric mucus. Biochem J 163:363–368

Sheehan JK, Carlstedt I (1984) Hydrodynamic properties of human cervical-mucus glycoproteins in 6M guanidinium chloride. Biochem J 217:93–101

Sheehan JK, Carlstedt I (1989) Models for the macromolecular structure of mucus glycoproteins. In: Harding SE, Rowe AJ (eds) Dynamic properties of biomolecular assemblies. Royal Society of Chemistry, Cambridge, UK, pp 256–275

Silberberg A, Meyer FA (1982) Structure and function of mucus. In: Chantler EN, Elder JB, Elstein M (eds) Mucus in health and disease II. Plenum, New York, pp 53–74

Wyatt PJ (1993) Light scattering and the absolute characterization of macromolecules. Anal Chim Acta 272:1–40