Electron microscopy of cervical-mucus glycoproteins and fragments therefrom

The use of colloidal gold to make visible 'naked' protein regions

John K. SHEEHAN*† and Ingemar CARLSTEDT†

*Department of Biochemistry and Molecular Biology, University of Manchester, Stopford Building, Oxford Road, Manchester M13 9PT, U.K., and †Department of Physiological Chemistry 2, University of Lund, P.O. Box 94, S-221 00 Lund, Sweden

Subunits of human cervical-mucus glycoproteins obtained by reductive cleavage of whole mucins and high- M_r glycopeptides (T-domains) obtained after their trypsin digestion were studied with electron microscopy after spreading the macromolecules in a monolayer of benzyldimethylalkylammonium chloride. Subunits were observed as linear and apparently flexible particles, with number- and weight-average lengths of 390 nm and 460 nm respectively. T-domains randomly distributed on the grid have number- and weight-average lengths of 90 nm and 103 nm respectively, whereas when aligned (possibly stretched by flow) they were longer, with number-average and weight-average lengths of 150 nm and 170 nm respectively. Subunits complexed with gold appeared as segmented structures, with a distribution of inter-gold distances similar to the length distribution for the relaxed T-domains. The whole mucins had few binding sites for gold, suggesting that reduction exposes hydrophobic protein-rich regions with high affinity for gold. Most T-domains had a binding site at one end, indicating the presence of a residual protruding naked peptide region. We conclude that mucins are assembled from subunits joined end-to-end, and that each subunit consists of alternating oligosaccharide 'clusters' (approx. 100 nm) and naked peptide regions which have (after reduction) a high affinity for colloidal gold.

INTRODUCTION

Cervical mucus plays an important role in reproduction. The endocervix is filled with a mucous secretion, the properties of which vary during the ovulatory cycle. The highly visco-elastic luteal-phase mucus acts as a barrier to the spermatozoa, whereas the more abundant mid-cycle secretion is less visco-elastic and readily penetrated by the spermatozoa. For reviews, see Chantler & Elstein (1986) and Wolf & Litt (1986).

The macromolecules responsible for the rheological properties of mucus are large and complex glycoproteins referred to as the mucus glycoproteins or the mucins. Cervical mucins (M_r approx. 10×10^6) are composed of subunits (M_r 2×10^6) which are obtained after reduction of disulphide bonds. Trypsin digestion of subunits yields glycopeptides (T-domains) with M_r 380 000, which account for 80% of subunits (Carlstedt *et al.*, 1983a,b). Hydrodynamic studies show that the macromolecules are well described as linear flexible chains, which adopt a highly hydrated spheroidal shape in dilute solution (Sheehan & Carlstedt, 1984), and this view has been corroborated by using electron microscopy (Sheehan *et al.*, 1986). The architecture of cervical, respiratory, gastric and submaxillary mucins has been shown to be very similar (Carlstedt & Sheehan, 1984a,b; Rose *et al.*, 1984; Mikkelsen *et al.*, 1985; Sheehan *et al.*, 1986; Shogren *et al.*, 1986; Marianne *et al.*, 1987).

The T-domains correspond to stretches of the protein core that are highly substituted with oligosaccharides,

and these units are flanked by so-called 'naked' regions of protein (see, e.g., Carlstedt *et al.*, 1985). Consequently, carbohydrate-rich and protein-rich segments alternate along the mucin thread. It has been shown that colloidal gold has low affinity for glycosaminoglycans (Sheehan *et al.*, 1987) and thus may be a probe for localizing protein domains within a highly glycosylated structure. Here we have used electron microscopy to study further the mucin subunits and T-domains, and also introduced colloidal gold as a marker for the 'naked' stretches of protein within the macromolecules.

EXPERIMENTAL

Materials

Guanidinium chloride (ultrapure grade) was obtained from Bethesda Research Laboratories. Benzyldimethylalkylammonium chloride (BAC) and colloidal gold (diameters 2–5 nm and 8–12 nm; prepared as described by Tschopp *et al.*, 1982) were gifts from, respectively, Dr. G. Klotz, Department of Virology, University of Ulm, Ulm, Germany, and from Dr. A. Ratcliffe, Kennedy Institute of Rheumatology, London, U.K. Mucins and fragments therefrom (subunits and T-domains) were prepared as described by Carlstedt *et al.* (1983a,b).

Electron microscopy

Carbon films were made by evaporation of carbon on to freshly cleaved mica. The films were floated off on

Abbreviation used: BAC, benzyldimethylalkylammonium chloride.

† To whom all correspondence should be addressed.

water and deposited on 600-mesh grids (Gilder Grids, Grantham, Lincs., U.K.). Specimen grids were prepared by a modification (Sheehan *et al.*, 1986) of the method described by Koller *et al.* (1969). In short, the molecules were spread in BAC monolayers on a hypophase of 50 mM-magnesium acetate, picked up on carbon-coated grids, stained with uranyl acetate and rotary-shadowed with platinum/tungsten. Stock solutions (approx. 2–3 mg/ml) of the macromolecules in 6 M-guanidinium chloride were diluted with 50 mM-magnesium acetate to final concentrations of 0.01–0.1 $\mu\text{g/ml}$, 0.05–1.0 $\mu\text{g/ml}$ and 1.0–5.0 $\mu\text{g/ml}$ for whole mucins, subunits and T-domains respectively.

Mucin-gold complexes were prepared as described by Sheehan *et al.* (1987). In short, stock solutions (5 μl ; 2–3 mg/ml) of mucins or fragments in 6 M-guanidinium chloride were diluted with 50 mM-magnesium acetate to a final concentration of about 100 $\mu\text{g/ml}$. This solution (50 μl) was then mixed with an equal volume of a suspension of colloidal gold in water (a large excess) and incubated for 30 min at room temperature. Unbound gold was removed by gel chromatography on a column (10 ml) of Sepharose CL-2B eluted with 50 mM-magnesium acetate and monitored with an LKB Uvicord S at 280 nm. The mucin-gold complexes were diluted with 50 mM-magnesium acetate and spread in BAC monolayers as described above.

Electron microscopy was performed with a JEOL 100CX scanning transmission instrument at 80 or 100 keV. Contour lengths were measured on photographically enlarged electron micrographs (final magnification 60 000 for subunits and 150 000 for T-domains) with a Joyce-Loebl Magiscan 2 image analyser. The macromolecules were traced with an 0.5 mm black ink pen, digitized, and lengths were measured by using a computer program based on that of Dixon & Tayler (1979). The precision is approx. 2%, and the overall accuracy of the contour-length measurements was estimated to be better than 10%.

RESULTS

Electron microscopy of cervical mucins shows that the macromolecules are linear flexible threads which are polydisperse (500–5000 nm) in length (Sheehan *et al.*, 1986). Subunits derived from these mucins also appear as polydisperse filamentous structures, which are irregular in shape, in keeping with them being flexible (Fig. 1a). The lengths of the particles range from 100 to 800 nm (Fig. 1b) and the size distribution is heterogeneous. The number-average (l_n) and weight-average (l_w) lengths are 390 nm and 460 nm respectively. The weight-average M_r for the subunits used in this study was 2.0×10^6 , and combining this value with that of l_w provides an estimate of the mass/length parameter (M_r/l_w) of 4300/nm.

T-domains adopt various different conformations when spread for electron microscopy (Figs. 2a–2c). Randomly oriented particles (Fig. 2a) sometimes appear as 'kinked' rods and sometimes as 'blob'-like structures. In contrast, T-domains in parallel alignment over large areas of the grid, possibly under the influence of surface tension and/or flow forces, often appear as extended rods (Fig. 2c). In some cases 'blobs' are found on rod-like structures (Fig. 2b). The length distributions of randomly oriented (Fig. 2a) and aligned (Fig. 2c) molecules are shown in Figs. 3(a) and 3(b) respectively. The

number-average (l_n) and weight-average (l_w) lengths are 90 nm and 103 nm for the randomly distributed molecules and 150 nm and 170 nm for the orientated ones. Combining the two values for l_w with that for M_r (380 000) provides estimates of the mass/length parameter (M_r/l_w) of 3800/nm and 2200/nm for randomly oriented and extended molecules respectively.

Whole mucins labelled with colloidal gold are shown in Fig. 4. The polyvalent nature of the gold typically causes inter- and intra-chain associations, but occasional single molecules show a spacing between gold clusters of 500–700 nm (Fig. 4a). A more typical situation is seen in Fig. 4(b), where intra- and inter-chain cross-links make the delineation of individual molecules impossible. When subunits were treated with colloidal gold (Figs. 5a–5c) some molecules aggregated, as did the whole mucins, but the majority appeared as single species; however, we cannot rule out selective loss of aggregates of molecules in the column-chromatography step. The single molecules exhibited a segmental appearance with gold binding at the ends and at one to three internal sites, and had a somewhat shorter l_w value (about 350 nm) than subunits not treated with gold. The number of gold particles at each site depended on the size of the gold used, and typically one 'large' (12–15 nm; Fig. 5a) particle or one to five smaller ones (2–5 nm; Figs. 5b–5d) could be accommodated. One group of molecules, labelled only at the ends (Fig. 5b), formed a distinct subclass with an end-to-end distance of 150–180 nm. Subunits with affinity for gold at internal sites typically showed shorter (50–150 nm) inter-gold distances (Figs. 5c–5d). The number-average inter-gold distance is 93 nm, similar to the number-average length (90 nm) obtained for the randomly distributed T-domains.

T-domains complexed with gold are shown in Fig. 6. Most molecules only bind gold at one end and, when aligned, tend to project in the same direction from the gold particle, suggesting that the latter has a higher affinity for the carbon substrate than has the glycosylated region of the T-domain. The data suggest that in some cases more than one molecule may bind to the same gold particle, and it is thus possible that shorter T-domains are 'shielded' by longer ones. However, the length distribution (Fig. 7a) of gold-binding T-domains (with the assumption that gold is only bound at one end) is similar to that of the extended T-domains, with a value of l_n of 160 nm. This distribution accounts for at least 75% of the total number of particles. The size range (Fig. 7b) for T-domains with gold bound at internal sites (examples in Figs. 6b and 6c) is outside that for T-domains not complexed with gold (Fig. 3a), suggesting that the former are aggregates.

DISCUSSION

Electron microscopy has been used to study bronchial mucins (Rose *et al.*, 1984; Mikkelsen *et al.*, 1985; Sheehan *et al.*, 1986; Marianne *et al.*, 1987) as well as gastric and cervical mucins (Sheehan *et al.*, 1986). Although different techniques were used, very similar conclusions as to the nature of the molecules emerged from these studies, i.e. they are filamentous, flexible and very polydisperse in length. We have proposed that subunits, of $M_r (2-3) \times 10^6$, are the basic building blocks of cervical, respiratory and gastric mucins, and that the whole mucins are formed by their thiol-dependent end-to-end association. Each sub-

Fig. 1. Electron microscopy of subunits of human cervical mucins

The molecules were spread in a monolayer of BAC on a hypophase of 50 mM-magnesium acetate, picked up on carbon-coated grids, stained briefly with uranyl acetate and rotary-shadowed with platinum/tungsten. (a) The molecules appear as fairly 'stiff' but randomly disposed threads of variable lengths. (b) Histograms of the contour lengths of subunits from human cervical mucins. A total of 376 molecules was measured, and the frequencies of the number distribution (■) and weight distribution (▨) give l_n 390 nm and l_w 460 nm respectively.

Fig. 2. Electron microscopy of T-domains of human cervical mucins

The molecules were spread in a monolayer of BAC on a hypophase of 50 mM-magnesium acetate, picked up on carbon-coated grids, stained briefly with uranyl acetate and rotary-shadowed with platinum/tungsten. (a) The molecules appear as randomly disposed 'kinked' particles and blobs, (b) rod-like structures with occasional blobs or (c) rod-like particles with a parallel orientation over wide areas of the grid, suggesting that they have been influenced by flow.

Fig. 3. Histograms of the contour lengths of T-domains

(a) Length distributions (▨, number frequency; ■, weight frequency) of 190 randomly distributed molecules gave values for l_n and l_w of 90 nm and 103 nm respectively. (b) Length distributions (symbols as in a) of 100 aligned and possibly stretched molecules gave values for l_n and l_w of 150 nm and 170 nm respectively.

Fig. 4. Electron microscopy of human cervical mucins complexed with colloidal gold

The mucins were treated with colloidal gold, spread in a monolayer of BAC on a hypophase of 50 mM-magnesium acetate, picked up on carbon-coated grids, stained briefly with uranyl acetate and rotary-shadowed with platinum/tungsten. (a) A single mucin with gold (2–5 nm) clustered in tight groups of one to three particles 500–700 nm apart. (b) Typically the gold caused intra-chain cross-linking and the formation of large aggregates.

Fig. 5. Electron micrographs of subunits from cervical mucins complexed with colloidal gold

The molecules were treated with colloidal gold, spread in a monolayer of BAC on a hypophase of 50 mM-magnesium acetate, picked up on carbon-coated grids and stained with uranyl acetate. (a) Molecules labelled with 'large' (12–15 nm) gold typically appeared as segmented structures, with single gold particles occurring at the ends and at one to three internal sites. (b) Some molecules, typically 150–180 nm in length, only bound gold (two to four particles, diameter 2–5 nm) at the ends. (c) and (d) Longer subunits treated with 'small' (diameter 2–5 nm) gold particles usually contained two to five segments flanked by single or clustered (three to five) gold particles.

Fig. 6. Electron microscopy of T-domains complexed with colloidal gold

The molecules were treated with gold, spread in a monolayer of BAC on a hypophase of 50 mM-magnesium acetate, picked up on carbon-coated grids, stained briefly with uranyl acetate and rotary-shadowed with platinum/tungsten. (a) Most of the molecules (at least 75 %) bind one or more gold particles close to one end. (b) and (c) When gold is bound at both ends, the molecules (5–10 %) are frequently aggregated to form particles longer than those present in the distribution of the untreated ones.

Fig. 7. Histograms of T-domains complexed with colloidal gold

Measurements were made on 130 molecules, assuming that the gold particles were bound at one end. (a) The length distribution gives a value for l_n of 160 nm, in good agreement with that obtained for the extended T-domains without gold. (b) Some molecules could be interpreted as binding gold at other sites than at the end, and the length distribution of such particles (40 were measured) yielded a value of l_n of 300 nm. This number is beyond the size range of untreated T-domains, and indicates that such particles consist of two or more glycopeptides attached at their ends.

unit contains, on average, four or five oligosaccharide 'clusters' flanked by 'naked' regions of protein. These 'clusters' are obtained as high- M_r glycopeptides (T-domains) after trypsin digestion of the subunits (Carlstedt *et al.*, 1983b). The subunits, as seen by electron microscopy, are linear and polydisperse, with a distinct heterogeneity in their size distribution. This finding was not expected on the basis of data obtained with gel chromatography (Carlstedt *et al.*, 1983b), but it is possible that the latter technique does not offer the same resolution.

The T-domains adopt a variety of shapes on the grid, which suggests that they are conformationally versatile. They may appear as compact 'blob'-like structures, as well as kinked or extended rods, in keeping with a jointed and highly flexible structure. T-domains contain approx. 12% of protein (Carlstedt *et al.*, 1983b), and, by assuming an average M_r value of 110 for the amino acids, it is estimated that the polypeptide core contains, on average, about 400 amino acids. This corresponds to an extension per amino acid of 0.42 nm which, within experimental error, would correspond to that of a completely extended polypeptide. Hydrodynamic studies of cervical T-domains (Sheehan & Carlstedt, 1984) and of related mucin fragments (Creeth & Knight, 1965) show that these structures in solution behave as random coils within a highly hydrated spheroidal solvent domain. The extended conformation observed here is thus likely to be an experimental artefact induced by mechanical forces operating on the grid during deposition of the molecules. However, the possibility that the highly glycosylated regions of the mucins in mucus can be stretched when subjected to mechanical forces may help to explain, at the molecular level, some of the physical properties of the gel.

The value (3800/nm) of M_r/l_w for the 'relaxed' T-domains is similar (within errors) to that for subunits (4300/nm), suggesting that this is the conformation for T-domains within these structures. Previous work (Sheehan *et al.*, 1986) on whole cervical mucins yielded an estimate for M_r/l_w of 6000/nm by electron microscopy, and there is thus a significant discrepancy in this parameter between subunits and whole mucins. However, the average length for the whole mucins might be underestimated, owing to difficulties in tracing the largest species in the distribution, which are thus excluded from the analysis.

Colloidal gold forms stable complexes with hydrophobic domains on many globular proteins (Geoghegan & Ackerman, 1977). In contrast, this reagent does not appear to bind significantly to hydrophilic structures such as carbohydrates (Sheehan *et al.*, 1987). When combined with T-domains, colloidal gold finds a site at one end on most (more than 75%) of the particles. Some species (about 10%) remain unlabelled, whereas a few bind gold at both ends or at internal sites. These observations suggest that most of the molecules have a peptide portion protruding at one end of the highly glycosylated region of the protein core, i.e. the T-domains are 'polarized'. As with the unlabelled T-domains those complexed with gold usually showed evidence of alignment, adopting the fully extended conformation. The observation that aligned T-domains always project in the same direction from the gold particles (Fig. 6a) suggests a mechanism by which the molecules may become fully extended. Thus, if the gold-binding, presumably hydrophobic, region of the T-domain is strongly bound to the carbon substrate, the more weakly attached carbohydrate-rich part of the molecule may be pulled and stretched by the flow fields generated during the transfer and washing procedures.

In subunits, the colloidal gold was localized at the ends, and in numerous cases at internal sites 50–170 nm apart. The average inter-gold binding distance (93 nm) is significantly shorter than the average length of the extended T-domains, but compares well with that of the randomly dispersed ones. The increased number of gold-binding sites in the subunits as compared with the whole mucins may be explained by the unfolding of hydrophobic regions after reduction of intra-molecular disulphide bonds, and the inter-gold site distance is in agreement with subunits containing 'naked' protein regions, which may bind gold, flanking oligosaccharide 'clusters' with no affinity for this 'probe'. Subunits were seldom observed as elongated molecules. This is consistent with there being significant hydrophobic peptide regions at both ends, as well as at internal sites, which attach the molecule firmly to the carbon substrate and prevent elongation during transfer to the grid.

In contrast, colloidal gold binds relatively infrequently to the intact mucins, suggesting that many 'naked' protein regions within the native molecule may be folded into structures that do not expose a large enough hydrophobic surface to bind gold.

In conclusion, electron microscopy of cervical subunits and T-domains gives further support to the structural model for these mucins originally proposed on the basis of hydrodynamic data (Carlstedt *et al.*, 1983b; Sheehan & Carlstedt, 1984). The molecules were envisaged as a linear array of two to ten subunits, each consisting of three to five oligosaccharide-rich domains flanked by

'naked' protein regions. This segmental structure is supported by, in particular, the binding pattern of colloidal gold to the subunits. The few binding sites for colloidal gold in the whole mucins as compared with the subunits suggest that these 'naked' segments, when within the whole mucins, are not available for binding gold. A probable explanation may be that they are folded and stabilized by intra-molecular disulphide bonds. There is evidence that the T-domains may adopt either a fully extended or a more compact conformation. Transitions between such states may underly some mechanical properties of mucus, and thus warrant further investigation.

This investigation was supported by grants from the Medical Faculty of Lund, Sweden, the Swedish Medical Research Council (project 7902), the Swedish Tobacco Company, Greta and Johan Kocks Stiftelser, Trelleborg, Sweden, the National Swedish Board for Technical Development, the Sunnerdahl Handikappfond, Sweden, the Swedish Society for Cystic Fibrosis, the Swedish National Association against Heart and Chest Diseases, the Cystic Fibrosis Research Trust, U.K., and the Wellcome Trust, U.K. Thanks are also due to Dr. David Thornton for help in the preparation of the manuscript.

REFERENCES

- Carlstedt, I. & Sheehan, J. K. (1984*a*) *Ciba Found. Symp.* **109**, 157–172
- Carlstedt, I. & Sheehan, J. K. (1984*b*) *Biochem. Soc. Trans.* **12**, 615–617
- Carlstedt, I., Lindgren, H., Sheehan, J. K., Ulmsten, U. & Wingerup, L. (1983*a*) *Biochem. J.* **211**, 13–22
- Carlstedt, I., Lindgren, H. & Sheehan, J. K. (1983*b*) *Biochem. J.* **213**, 427–435
- Carlstedt, I., Sheehan, J. K., Corfield, A. & Gallagher, J. T. (1985) *Essays Biochem.* **40**, 40–76
- Chantler, E. & Elstein, M. (1986) *Semin. Reprod. Endocrinol.* **4**, 333–342
- Creeth, J. M. & Knight, C. G. (1965) *Biochim. Biophys. Acta* **102**, 549–558
- Dixon, R. N. & Tayler, C. J. (1979) in *Machine-Aided Image Analysis* (Gardner, W. E., ed.), pp. 178–185, Institute of Physics (Conference Series no. 44), Bristol
- Geoghegan, W. O. & Ackerman, G. A. (1977) *J. Histochem. Cytochem.* **25**, 1187–1200
- Koller, T., Harford, A. G., Lee, Y. K. & Beer, M. (1969) *Micron* **1**, 110–118
- Marianne, T., Perini, J.-M., Lafitte, J.-J., Houdret, N., Pruvot, F.-R., Lamblin, G., Slayter, H. S. & Roussel, P. (1987) *Biochem. J.* **248**, 189–195
- Mikkelsen, A., Stokke, B. T., Christensen, B. E. & Elgsaeter, A. (1985) *Biopolymers* **24**, 1683–1704
- Rose, M. C., Voter, A., Brown, C. F. & Kaufman, B. (1984) *Biochem. J.* **222**, 371–377
- Sheehan, J. K. & Carlstedt, I. (1984) *Biochem. J.* **217**, 93–101
- Sheehan, J. K., Oates, K. & Carlstedt, I. (1986) *Biochem. J.* **239**, 147–153
- Sheehan, J. K., Ratcliffe, A., Oates, K. & Hardingham, T. E. (1987) *Biochem. J.* **247**, 267–276
- Shogren, R. L., Jamieson, A. M., Blackwell, J. & Jentoft, N. (1986) *Biopolymers* **25**, 1505–1517
- Tschopp, J., Podack, E. & Muller-Eberhard, H. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 7474–7478
- Wolf, D. P. & Litt, M. (1986) *Semin. Reprod. Endocrinol.* **4**, 323–332

Received 30 May 1989/28 July 1989; accepted 10 August 1989