Glycoconjugate Journal (1996) **13:** 833–837

The transcripts of the apomucin genes *MUC2, MUC4,* and *MUC5AC* are large and appear as distinct bands

DAN BAECKSTRÖM and GUNNAR C. HANSSON*

Department of Medical Biochemistry, Göteborg University, Medicinaregatan 9A, 413 90 Gothenburg, Sweden

Received 10 October 1995, revised 31 January 1996

RNA from four colorectal carcinoma cell lines was prepared and analysed in Northern blots using probes for the *MUC2*, *MUC4*, and *MUC5AC* mucin apoprotein genes. The sizes of the transcripts were very large, in the order of at least 12–16 kb. The presence of distinct bands is in contrast to earlier reports, where these transcripts showed extensive polydispersity. RNA from rat small intestine was also prepared and probed with cDNA for the rat *Muc2* mucin gene. This analysis also showed a large and discrete hybridizing band, indicating that apomucin mRNA of well-defined size can be obtained also from a tissue with high endogenous RNase activity.

Keywords: Northern blot, mRNA, mucin, colon carcinoma, intestine

Introduction

Mucins are a heterogeneous group of glycoproteins located at mucosal surfaces, where they are believed to be important for protection and lubrication. The larger mucins have a central core protein covered with oligosaccharide side chains accounting for more than 50% of the total molecular mass. Many mucins appear to be disulfide linked oligo- or polymers and the glycoproteins obtained by reduction have been called subunits [1]. Mucins have been purified according to many different protocols [2], several of which have not taken enough care to avoid mechanical or proteolytic degradation. However, even the most stringent purification conditions [1] have yielded mucin subunits that are polydisperse with respect to molecular size. The biochemical background to this phenomenon is not yet fully understood, although variable glycosylation has been believed to be the major reason. However, the possibility of a polydisperse apoprotein has not been excluded and the findings of polydisperse mRNAs in Northern blot analyses and variable number tandem repeat polymorphism in the coding sequences of mucin apoprotein genes [3, 4] have strengthened this proposal. Polydisperse transcripts have been observed for all mucin apoprotein genes cloned, except *MUC1*, and this polydispersity has been considered as one of the typical features of mucin mRNAs [5–7]. In the present investigation we have used Northern blots to analyse the transcripts of the mucin apoprotein genes *MUC2*, *MUC4*, and *MUC5AC* from four tumour cell lines, showing distinct bands with large sizes. Similar results were obtained by probing rat small intestine RNA with a rat apomucin cDNA.

*To whom correspondence should be addressed. Tel: +46 31 77 33488; Fax: +46 31 416108.

Materials and methods

Cell lines

The colon carcinoma cell lines COLO 205 (American Type Culture Collection no. CCL 222), LS174T (CCL 188), SW1116 (CCL 223), and LoVo (CCL 229) were grown in Iscove's medium (Gibco, Paisley, UK) with 10% foetal calf serum (Gibco) and supplemented as described [8]. LS174T and LoVo had been selected for homogenous expression of the carbohydrate antigen CA242 [9].

Probes

For the detection of the transcripts, probes were prepared from the following plasmids; *MUC2*, a plasmid containing the 836 bp SMUC41 tandem repeat insert [10]; rat *Muc2*, a plasmid containing the 705 bp VR-1A insert [11], *MUC5AC*, a plasmid containing the 3.5 kb NP3a insert [12]. The excised inserts were separated by agarose gel electrophoresis and isolated using QIAEX II (QIAGEN, Hilden, Germany) or Geneclean II (Bio 101, La Jolla, CA). For *MUC4* a synthetic 48 oligonucleotide GTCGGTGACAGGAAGAGGGGTGGCGTGACCTGTG-GATGCTGAGGAAGT probe based on [13] was used. The

MUC4 oligonucleotide probe was a gift kindly provided by Dr Sandra J. Gendler. The SMUC41-containing plasmid was a gift kindly provided by Drs Young S. Kim and James R. Gum. The NP3a plasmid was kindly provided by Dr Mary Rose. The inserts were labelled with α-(^{32}P)-CTP using the Megaprime random priming kit (Amersham, Amersham, UK) and the oligonucleotide probe with γ-(^{32}P)-ATP and T4 kinase (Pharmacia, Uppsala, Sweden).

RNA purification

RNA was prepared using the guanidinium thiocyanate method [14]. RNase free water (glass distilled or diethyl pyrocarbonate (DEPC) treated [15]) was used for all solutions. All pipettes were wide mouthed, all pipette tips used were cut, and care was taken to avoid shear forces when pipetting. Guanidinium thiocyanate (4 M) was prepared by dissolving 50 g in 10 ml of 1 M Tris-HCl, pH 7.6, adding water up to 100 ml, and filtering through Whatman 3MM paper (Whatman, Maidstone, UK). Mercaptoethanol was added to 1% before use. Guanidinium chloride (7.5 M) was prepared by dissolving 72 g of guanidinium chloride up to a volume of 100 ml after adjusting the pH to 7 with a few drops of 3 M NaOH. Dithiothreitol was added to 10 mM immediately before use.

For RNA preparation from tissue culture cells, the medium was decanted from four 175 cm^2 flasks containing confluent cells and the cells were lysed by rapid addition of a total of 20 ml of cold 4 M guanidinium thiocyanate. The cell lysate was then removed using a cell scraper (Nunc, Roskilde, Denmark).

For RNA preparation from rat small intestine, a rat (GOT-W inbred strain) was starved over night, anaesthetized with ether and killed by cervical dislocation. About 1/3 of the small intestine was removed, washed with cold 0.9% NaCl by attaching a pipette to one end of the intestine. The intestine was held on ice and cut open and the mucosa was quickly scraped off with a spatula into 20 ml of cold guanidium thiocyanate.

Cell or tissue lysates were homogenized on ice with four strokes in a Dounce all glass homogenizer using the loose pestle. One ml of a 10% lauryl sarcosinate (Sigma, St Louis, MO) solution was added. If the solution contained visible debris, the solution was centrifuged at 2500 rpm in 50 ml conical tubes. The supernatant was transferred to DEPC treated all teflon tubes (Nalge, Rochester, NY), and mixed with 1 ml 2 M KAc, pH 5.5 and 1.6 ml 1 M HAc. Subsequently, 15 ml ethanol (99.5%, room temperature) was added in portions with mixing between additions, and the tubes were placed in a −20 °C ethanol bath for at least 2 h before centrifugation at about 10 000 × **g** at +4 °C for 20 min. The supernatant was gently aspirated and the pellet was resuspended in 10 ml 7.5 M guanidinium chloride with 10 mM dithiothreitol added just before use. After addition of 0.5 ml 2 M KAc, pH 5.5, the tubes were mixed and 5 ml absolute ethanol was added before precipitation at −20 °C for at least 2 h followed by centrifugation at 10 000 × **g** for 20 min. The supernatant was removed using a pipette and the pellet resuspended as above in 7.5 M guanidinium chloride and the above step repeated. The obtained pellet was resuspended in 5 ml 20 mM EDTA, pH 7.0 and the solution was mixed with 10 ml chloroform:n-butanol 4:1 (by vol) added by a RNase free glass pipette. After shaking for 5 min and centrifugation for 10 min, the upper phase was transferred to a teflon tube and the interface washed twice with 4 ml 20 mM EDTA, pH 7.0. The upper phases were recovered after centrifugation and the combined upper phases were precipitated by the addition of 0.1 volume of 3 M NaAc, pH 7.0 and 2.5 volumes of ethanol followed by cooling to −20 °C for at least 2 h. The precipitated RNA was recovered by centrifugation at 10 000 × **g** for 20 min at +4 °C. The pellet was dissolved in RNase free water and the RNA was quantified by measurement of A_{260} and analysed by agarose gel electrophoresis in the presence of 6.5% formaldehyde. The RNA was always stored precipitated at −20 °C. The yield of RNA was 1.8 mg from the rat intestine, 4.4 mg from LS174T, 0.13 mg from LoVo, 1.43 mg from COLO 205, and 0.78 mg from SW1116.

Polyadenylated RNA was isolated using Dynabeads oligo-(dT)$_{25}$ (Dynal, Oslo, Norway). Total RNA (100 μg) in water was purified on 0.2 ml of resuspended Dynabeads according to the supplier's description. The amount of mRNA recovered was used for two slots in agarose gel electrophoresis.

Northern blot analyses

Samples containing total RNA, freshly purified polyadenylated RNA or RNA size markers (Boehringer Mannheim, Mannheim, Germany) were mixed 1:1 (by vol) with the sample buffer (20 mM 3-(*N*-morpholino)-propanesulfonate, pH 7.0, 5 mM NaAc, 1 mM EDTA, 67% formamide, 25% formaldehyde). To this mixture (10–20 μl), 1 μl ethidium bromide (0.4 mg ml^{-1}) was added. After heating at 65 °C for 5 min, 3 μl of a solution containing 0.25% Bromphenol Blue and 30% Ficoll (Pharmacia) was added and the sample was loaded onto a 1% agarose gel containing 20 mM 3-(*N*-morpholino)-propanesulfonate, pH 7.0, 5 mM NaAc, 1 mM EDTA (1X MOPS) and 6.5% formaldehyde. The gel was run for 16–20 h at 25–30 V using 1X MOPS as running buffer.

The gel was subjected to capillary blotting overnight using Hybond-C (Amersham) nitrocellulose membranes prewetted with water and 20X SSC (20X SSC = 0.3 M sodium citrate, pH 7.0, 3 M NaCl). Subsequently, the membranes were air-dried and then baked at 80 °C over blue silica gel in a closed glass jar for 2 h. The baked membranes were wetted with 6X SSC and incubated at

42 °C for more than 3 h in prehybridization solution (5X Denhardt's solution [15], 4X SSC, 50 mM potassium phosphate, pH 7.4, 0.1% SDS, 50% formamide). The labelled probe was heated at 100 °C for 5 min, chilled on ice and then added to the prehybridization solution. The membranes were kept in this solution overnight at 42 °C, then washed at room temperature three times with 200 ml of 2X SSC, with 2X SSC, 0.1% SDS for 1 h, and with 0.1X SSC, 0.1% SDS for 1 h. After a final wash with 0.1X SSC, 0.1% SDS at 55 °C for 1 h, the membranes were exposed to Kodak X-OMAT S film at −80 °C for 1–7 days.

Results

Northern blot analyses were performed on mRNA from four colorectal carcinoma cell lines using probes to the *MUC2*, *MUC4*, and *MUC5AC* mucin genes. These three showed one or two distinct bands in each cell line as exemplified in Fig. 1A,B. The transcripts were large for all genes with an estimated size of more than 12 kb for *MUC2*, 15–16 kb for *MUC4*, and 0.5–1 kb less than *MUC2* for *MUC5AC* using RNA molecular size markers up to 7.4 kb as standards.

To show that it is also possible to obtain distinct bands from tissues with relatively large amounts of RNases we prepared RNA from rat small intestine and probed it with a rat *Muc2* probe (VR-1A) obtained during cloning the apoprotein of the 'insoluble' mucin complex of rat small intestine [11]. In this experiment, a comparison was also made between the hybridizations obtained with total RNA and with purified polyadenylated RNA. The results (Fig. 1C) show a distinct band in the size of more than 12 kb for both RNA and poly-A^+ RNA, although the total RNA blot is slightly distorted by the presence of rRNA.

Discussion

In this study, the expression and transcript sizes of three mucin apoprotein genes (*MUC2*, *MUC4*, and *MUC5AC*) in four different colorectal carcinoma cell lines (COLO 205, LS174T, LoVo, and SW1116) have been investigated. The expression of *MUC1* has previously been shown in the COLO 205 and SW1116 cell lines [16]. The COLO 205 and LoVo cell lines expressed only one of these mucin genes (*MUC1* and *MUC2*, respectively), whereas both LS174T (*MUC2* and *MUC5AC*) and SW1116 (*MUC2* and *MUC4*) expressed two. Size polymorphism, probably due to variable number of tandem repeats, was

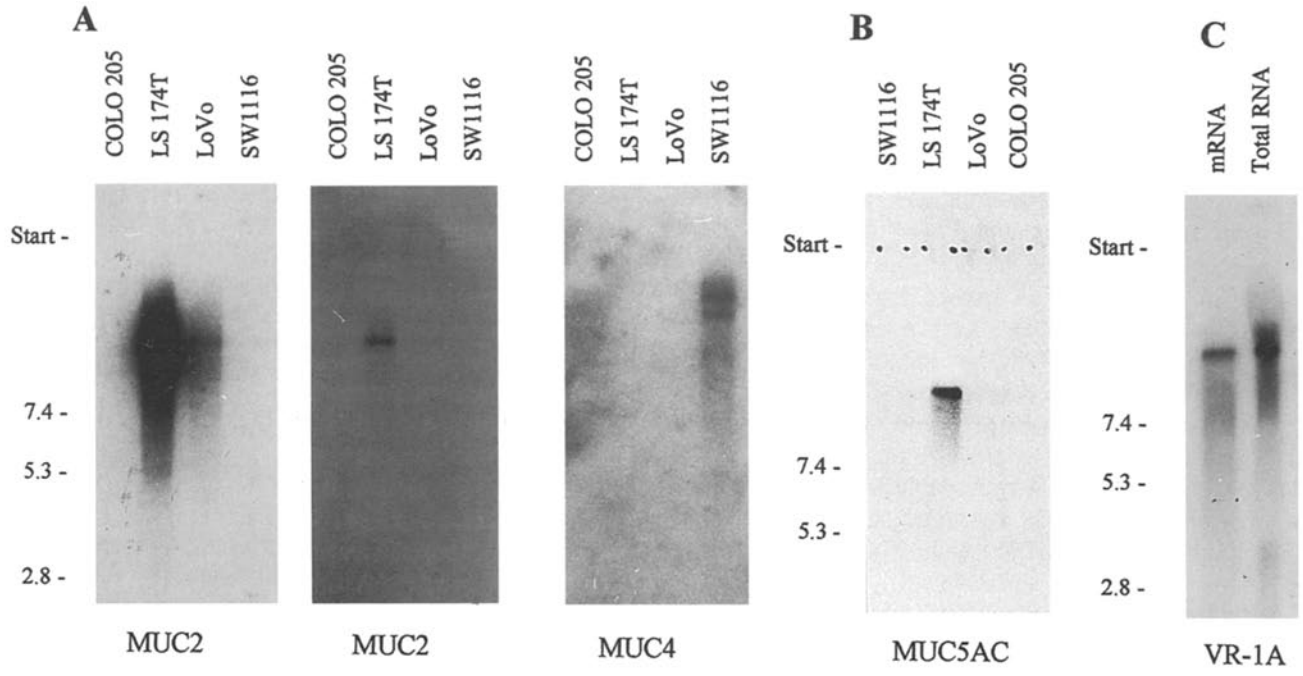

Figure 1. Northern blot analyses of mRNA from four cell lines (A, B) and of mRNA and total RNA from the epithelial cells of rat small intestine (C). mRNA corresponding to 50 μg RNA before polyA purification has been applied to each lane, except for 25 μg of total RNA in C. The blots have been stained with the probes indicated. The membrane probed with the *MUC2* probe was exposed for different periods to show both a strong and a weak band. The gel analysing MUC5AC was electrophoresed for a longer time. The numbers given are the sizes of the RNA markers.

observed in the *MUC1* transcripts of COLO 205 [16] and in the *MUC4* transcripts of SW1116. Due to the different lengths of the probes, a comparison of the intensities is not appropriate. However, the expression of *MUC2* in the LS174T cell line is probably substantially higher than any other of these mucins.

The *MUC2*, *MUC4*, and *MUC5AC* probes detected very large transcripts where the estimated sizes were more than 12 kb. These estimations are very uncertain, as they are far out of range of the RNA size markers used (1.6–7.4 kb) and 12 kb is probably underestimated. From the sequence analysis of *MUC2* cDNA [17] it has been shown that the translation product of a typical *MUC2* allele with 100 tandem repeats has a transcript size of at least 15 kb. Thus, if the number of tandem repeats in the *MUC2* alleles of LS174T and LoVo are in the normal range, *MUC4* should be at least 17–18 kb and *MUC5AC* around 14 kb. The size of the rat *Muc2* transcript from the strain used is about 0.5 kb smaller than that of the human *MUC2* mRNAs in LS174T.

In all the RNA blot experiments presented here the hybridized transcripts appear as discrete bands. The results of the analysis of rat small intestine RNA shows that a large, apparently undegraded apomucin transcript can be obtained also from a tissue with high endogenous RNase activity. RNA blots of mucins published by Khatri *et al.* and Lesuffleur *et al.* also show discrete bands [18, 19]. In contrast, most published RNA blots probing for transcripts of apomucin genes (with the exception of *MUC1*) show a high degree of polydispersity in the hybridized bands [5–7, 10, 13, 20]. The reasons for this phenomenon have been discussed only briefly, but it has been suggested that apomucin gene transcripts are intrinsically polydisperse [5–7]. The results presented here are difficult to reconcile with such theories. It is also difficult to explain the generation of the very polydisperse mRNAs observed in the previously published blots with any known mechanism except degradation.

Instead, we propose that intact apomucin transcripts from a given allele generally are monodisperse, but that they are very sensitive to degradation. As the researchers that report polydisperse apomucin mRNAs also find that control transcripts such as β-actin or GAPDH mRNA appear as sharp bands, it is reasonable to suggest that the large sizes of the apomucin messages are a critical factor, making them exceptionally vulnerable to degradation from RNase activity and shear forces.

The RNA extraction method used by us was originally designed for obtaining RNA from tissues rich in RNases [14], and strongly denaturing agents are present during most of the preparation. In the guanidinium thiocyanate/cesium chloride protocol used in the studies where polydisperse transcripts were found, the denaturing conditions are lost early. This difference in extraction method, together with the steps taken here to avoid shear forces may explain the discrepancy between our results and those previously presented. Other parts of the experimental procedures may also be important and therefore the methods used here, although far from novel, have been described in some detail. Enrichment of polyadenylated RNA does not seem to be critical for obtaining discrete bands, although some improvement is achieved by removing the interference of rRNA (Fig. 1C). We conclude that the pronounced polydispersity of mature mucin molecules is primarily caused by a highly variable glycosylation [2] and not by variable transcript sizes.

Acknowledgements

We are indebted to the skillful technical assistance of Mrs Helena Alstermark and Ms Malin Ek. The work was supported by grants from the Swedish Medical Research Council (Nos. 7491 and 10443), the Swedish Cancer Foundation (No. 3280) and the IngaBritt and Arne Lundbergs Stiftelse.

References

1. Carlstedt I, Sheehan JK, Corfield AP, Gallagher JT (1985) *Essays Biochem* **20**: 40–76.
2. Strous GJ, Dekker J (1992) *Crit Rev Biochem Mol Biol* **27**: 57–92.
3. Swallow DM, Gendler S, Griffiths B, Corney G, Taylor-Papadimitriou J, Bramwell ME (1987) *Nature* **328**: 82–84.
4. Toribara NW, Gum JR, Culhane PJ, Lagace RE, Hicks JW, Petersen GM, Kim YS (1991) *J Clin Invest* **88**: 1005–13.
5. Eckhardt AE, Timple CS, Abernethy JL, Zhao Y, Hill RL (1991) *J Biol Chem* **266**: 9678–86.
6. Tsuda T, Gallup M, Jany B, Gum J, Kim Y, Basbaum C (1993) *Biochem Biophys Res Commun* **195**: 363–73.
7. Verma M, Davidson EC (1993) *Proc Natl Acad Sci USA* **90**: 7144–48.
8. Larson LN, Johansson C, Lindholm L, Holmgren J (1988) *Int J Cancer* **42**: 877–82.
9. Johansson C, Nilsson O, Baeckström D, Jansson E-L, Lindholm L (1991) *Tumor Biol* **12**: 159–70.
10. Gum JR, Byrd JC, Hicks JW, Toribara NW, Lamport DTA, Kim YS (1989) *J Biol Chem* **264**: 6480–87.
11. Hansson GC, Baeckström D, Carlstedt I, Klinga-Levan K (1994) *Biochem Biophys Res Commun* **198**: 181–90.
12. Meerzaman D, Charles P, Daskal E, Polymeropoulos MH, Martin BM, Rose MC (1994) *J Biol Chem* **269**: 12932–39.
13. Porchet N, Cong NV, Dufosse J, Audie JP, Guyonnet-Duperat V, Gross MS, Denis C, Degand P, Bernheim A, Aubert JP (1991) *Biochem Biophys Res Commun* **175**: 414–22.
14. MacDonald RJ, Swift GH, Przybyla AE, Chirgwin JM (1987) *Methods Enzymol* **152**: 219–27.
15. Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning. A Laboratory Manual* 2nd edition Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
16. Baeckström D, Nilsson O, Price MR, Lindholm L, Hansson GC (1993) *Cancer Res* **53**: 755–61.

17. Gum JR, Hicks JW, Toribara NW, Siddiki B, Kim YS (1994) *J Biol Chem* **269**: 2440–46.
18. Khatri IA, Forstner GG, Forstner JF (1993) *Biochem J* **294**: 391–99.
19. Lesuffleur T, Porchet N, Aubert J-P, Swallow D, Gum JR, Kim YS, Real FX, Zweibaum A (1993) *J Cell Sci* **106**: 771–83.
20. Hollingsworth MA, Strawhecker JM, Caffrey TC, Mack DR (1994) *Int J Cancer* **57**: 198–203.