RESEARCH COMMUNICATION

Tissue-specific expression of a rat intestinal mucin-like peptide

Gongqiao XU, Dashou WANG, Ling Jun HUAN, Ernest CUTZ, Gordon G. FORSTNER and Janet F. FORSTNER*

Research Institute, The Hospital for Sick Children, Toronto, Ontario M5G 1X8, Canada

Expression of the gene for a rat intestinal mucin-like peptide (MLP) was studied by Northern-blot analyses of RNA prepared from a panel of rat tissues. Four probes (A–D) were constructed so as to span a 3.5 kb-long cDNA for rat MLP, and used for hybridization. Positive signals were obtained in intestine and colon, whereas lung, liver, stomach, submandibular gland and spleen were negative. The only transcript detected was approx. 9.5 kb in size. No mRNA splice variants were found. Hybridization *in situ* using probe B1, which corresponds to a cysteine-rich region near the C-terminus of MLP, confirmed that the gene for MLP is expressed by goblet cells of rat intestine and colon.

INTRODUCTION

Previous studies have shown that purified intestinal mucins from human, rat and other species give rise upon reduction to high-molecular-mass components which barely enter polyacrylamide gels during SDS/PAGE, as well as to a smaller glycoprotein (118 kDa) previously called the putative 'link' glycopeptide [1–3]. More recently we have sequenced a cDNA corresponding to the 'link' peptide of rat intestinal mucin and shown that it represents the C-terminal 689 amino acids of a large mucin-like peptide (MLP), which is estimated to contain a total of approx. 3000 amino acids [4,5]. On the N-terminal side of the cysteine-rich 'link' region we have identified two distinct regions both enriched in proline, threonine and serine, which are separated by a short cysteine-rich segment [5]. This pattern of alternating P,S,T-rich and cysteine-rich segments near the C-terminus has not been observed for other mammalian mucins cloned to date.

In the present study we sought further evidence that each of the distinct structural domains recognized in rat MLP is encoded by the same mRNA, in order to rule out the possibility that our sequence might have been generated by spurious recombination events during cDNA cloning. In addition, we have explored the possible existence of alternative splice variants and determined the tissue-specific expression of rat MLP by using Northern-blot analyses. Finally, we have localized, by hybridization *in situ*, the site of synthesis of MLP to intestinal and colonic goblet cells.

METHODS

Northern-blot analyses

Total RNA was isolated by the method of Han *et al.* [6] from a variety of rat tissues. RNA samples (6–12 µg of intestinal RNA, 25 µg of RNA from other tissues) were electrophoresed on 0.8% agarose/formaldehyde gels, soaked for 20 min in 50 mM-NaOH containing 10 mM-NaCl [7], blotted and probed as described previously [4]. Equivalent loading of RNA from non-intestinal tissues was judged visually by the equal intensity between blots of each lane, of the 28 S and the 18 S rRNA bands respectively. To prepare specific probes, DNA fragments were released by appropriate restriction enzymes (see the Results

section) from a 3.5 kb cDNA clone [4] encoding the C-terminus of rat intestinal MLP. Fragments were subcloned into plasmids pSP 72 or pGEM 4Z (Promega). ³²P-labelled cRNAs were synthesized with appropriate RNA polymerases by using [α -³²P]CTP (40 mCi/ml, > 800 Ci/mmol) (Amersham Canada Ltd., Oakville, Ont., Canada) according to the Promega Riboprobe protocol. All blots were stripped and re-probed with a rat actin probe (gift of Dr. J. Roummens, Division of Genetics, Hospital for Sick Children, Toronto) to confirm that mRNA was not degraded.

Hybridization *in situ*

Tissue and cell localizations of the mRNA for MLP were performed by the method of Harper & Marselle [8]. Briefly, tissue pieces from several organs of male Wistar rats were embedded in Tissue-Tek O.C.T. (Miles Scientific, Elkhart, IN, U.S.A.) and frozen in a solid-CO₂/ethanol bath. Sections (µm) were mounted on poly-L-lysine-coated slides, dried and fixed in 4% paraformaldehyde in phosphate-buffered saline (30 mM-sodium phosphate buffer, pH 7.4, in 140 mM-NaCl). Slides were immersed for 30 min in 0.1 M-triethanolamine, pH 8.0, acetylated with acetic anhydride (final concn. 0.25%, v/v), rinsed, dehydrated and dried as described previously [8]. Hybridization was performed for 17 h with a ³⁵S-labelled RNA probe (probe B1) made by transcription from an *Eco*R1-linearized pGEM 4Z vector bearing a 500 bp *Hind*III cDNA fragment of rat MLP in the presence of [α -³⁵S]CTP (40 mCi/ml, > 800 Ci/mmol) (Promega protocol). RNA in both antisense and sense orientations was synthesized from the same cDNA construct by appropriate RNA polymerases, and used as probes on adjacent tissue sections. Slides were then washed, dehydrated, air dried, incubated with RNAase solutions for 30 min and prepared for autoradiography and development as outlined previously [8]. Slides were exposed at 4 °C in the dark for 5 days, and examined with a Reichert-Jung Polyvar Photomicroscope. For light microscopy the slides were stained with haematoxylin and eosin.

RESULTS

To date we have determined the nucleotide sequence of 3.5 kb of rat intestinal mucin-like peptide, MLP [4,5]. The deduced

Abbreviation used: MLP, mucin-like peptide.

* To whom correspondence should be addressed.

Fig. 1. Schematic representation of the C-terminal 3.5 kb of rat intestinal MLP cDNA

The coding sequence is represented by a bar and the 3'-untranslated region by a line. The deduced peptide domains are marked with Roman numerals. P,S,T-rich domains are hatched. The arrowhead represents the starting point of the 'link' peptide region. Only restriction sites used in making probes are identified. Asterisks represent restriction-enzyme sites introduced for cloning. Black bars (below) indicate regions of the cDNA from which probes A, B, B1, C and D were derived. The MLP nucleotide numbering system has been changed from that shown initially [4] to accommodate a recently published [5] additional sequence of 852 nucleotides at the 5' end. The intersection of the new and the former sequence is indicated by a vertical broken line.

Fig. 2. Tissue-specific expression of rat MLP by Northern-blot analyses of total RNA from different rat tissues

1, Colon; 2, stomach; 3, liver; 4, lung; 5, submandibular gland; 6, spleen; 7, 7' and 7'', intestine. One blot (a) was hybridized with probe A, stripped, and re-probed with probe B (panel b), and a second blot (c) from a parallel gel was hybridized with probe C. All lanes carried the same amount (25 µg) of RNA, except for lanes 7, 7' and 7'', which contained 8, 12 and 6 µg of RNA respectively.

amino acid sequence can be divided into eight domains (Fig. 1): domains I–V are based on cysteine-alignment similarities with established domains in human von Willebrand factor [4]; domain VI, which contains tandem repeats, and domain VIII, which does not, are enriched in proline, serine and threonine; and domain VII is a short segment between them enriched in cysteine. Domains III and IV are particularly interesting because they contain segments having cysteine alignments similar to those in regions in the C-terminus of bovine and pig submaxillary mucin and frog integumentary mucin FIM-B1 [4,9]. However, domain I does not share cysteine alignment or sequence similarity with any other reported mucin except its human MLP counterpart [10].

Based on the different domains of MLP, five probes were constructed, as shown in Fig. 1, to cover the entire MLP sequence. Probe A was a 1.8 kb *Bam*HI/*Eco*RI fragment corresponding to nucleotides 331–2183. This long sequence covers half of the cysteine-rich domain I, the two short cysteine-rich domains V and VII, domain VI, which contains 4.5 tandem repeats rich in serine, threonine and proline, and most of the C-terminal end of the non-tandem repeat domain VIII, which is also enriched in proline, serine and threonine. The probe excludes

domains III and IV. Fig. 2(a) shows the results of Northern blots with probe A used on a panel of rat tissues. Intestinal and colonic RNA gave a positive signal at approx. 9.5 kb (or greater, since resolution is poor in this region of the gel), whereas stomach, liver, lung (including tracheal and bronchial epithelium), submandibular gland and spleen were negative. Even after long periods of exposure, no additional hybridization signals appeared in the other (non-intestinal) tissues. Probe B was a 500 bp *Hind*III fragment (nucleotides 1973–2456) corresponding to the C-terminal cysteine-rich half of domain I and having a 210 bp overlap with the 3' end of probe A. Probe B has been used previously to define the mRNA for MLP in rat intestine and human colonic cells [4,10]. Probe C was a 900 bp *Bam*HI/*Eco*RI fragment (nucleotides 2632–3529), encompassing most of domain III, all of domain IV and all of the 3' untranslated region. Figs. 2(b) and 2(c) show that the hybridization signals obtained with probes B and C were also confined to intestine and colon, with the dominant band again at approx. 9.5 kb. In Fig. 2(c) some smearing of the signal was observed (tracts 1, 7 and 7'). Although there was no evidence of RNA degradation, as judged from sharp 28 S and 18 S rRNA bands and intact hybridization signals with an actin probe (results not shown), we nevertheless

Fig. 3. Detection of MLP mRNA in the small intestine by hybridization *in situ*

Sections were hybridized with the ^{35}S -labelled probe B1. Dark-field photomicrograph of a rat small-intestine section. The bar represents 300 μm .

wondered whether the observed smearing could be due to partial degradation of the very large mRNA for MLP. For this reason, we repeated the Northern-blot experiment. As shown in tract 7" the signal appeared as a discrete band at approx. 9.5 kb. Northern blots using Probe D, a 330 bp *EcoRI/BamHI* fragment corresponding to the *N*-terminus of the serine-, proline- and threonine-rich domain VIII, have been presented previously [5], and gave the same 9.5 kb signal and tissue specificity as seen with the other probes.

Taken altogether, these findings confirmed that all of the recognized domains of rat MLP are encoded by the same mRNA. The only transcript detected (~ 9.5 kb) was present in rat intestine and colon, but not in other tissues. True polydispersity of the mRNA signal in intestine and colon was not observed.

Localization of MLP by hybridization *in situ*

A 270 bp antisense ^{35}S -labelled *EcoRI/HindIII* cDNA probe B1 (Fig. 1) was used for hybridization *in situ* with tissue sections of rat organs known to produce mucins, i.e. stomach, intestine, colon and lung (including trachea and bronchi), as well as organs which do not produce mucins, i.e. brain, heart, kidney, spleen and liver. A comparable sense probe was used as a negative control. In a dark-field low-power view of a longitudinal section of the small intestine, a strong positive signal was localized in discrete foci throughout the epithelium of the villi and the crypts (Fig. 3). At higher magnification of sections counterstained with haematoxylin and eosin, the signal was localized in the basal cytoplasm of goblet cells (Fig. 4a). Neighbouring absorptive cells showed no specific signal. This was most easily seen in intestinal villi, where goblet cells could be distinguished by the clear space of their apical mucin-containing granule mass. Cross-sections of crypts, where goblet cells are relatively scarce, were useful in both intestinal (results not shown) and colonic (Fig. 4b) sections in confirming that goblet cells were distinctively labelled by the hybridization probe. No specific hybridization signal was seen in sections of other tissues tested, including stomach, lung, trachea, bronchi, liver, brain, kidney, heart, or spleen (results not shown).

DISCUSSION

The observations presented in this paper add strength to the notion that MLP represents a single mucin peptide produced in goblet cells of rat small and large intestine. Since mucin clones have been noted to have a tendency to recombine with unrelated sequences artefactually [11,12], we designed specific probes for MLP to span the full length of the 3.5 kb cDNA, and tested each probe in Northern blots of RNA samples. Each probe gave a common mRNA signal in intestine and colon at approx. 9.5 kb (or greater), suggesting that each region of MLP is encoded by the same mRNA species. Thus the pattern of multiple domains in MLP is not a result of DNA recombination events during cloning.

Apart from human intestinal MLP [10], no other reported mucins share significant sequence similarity with rat MLP. There is, however, a similarity in the cysteine alignments of domains III and IV with *C*-terminal regions of bovine and pig submaxillary mucins and FIM-B1 mucin [4]. One of our concerns in the present study was to exclude the possibility that alternative splice variants of the mRNA for MLP exist in the rat submandibular gland or other rat organs. Thus the probes for each MLP domain were tested in Northern-blot analyses against a panel of rat tissues. We found no evidence that selected domains of MLP (including III and IV) were expressed in other tissues, and thus no evidence for the existence of MLP splice variants.

Polydispersity in the size of mRNA has been a common finding in all reported Northern-blot analyses of mucins [13],

Fig. 4. Hybridization signal of MLP mRNA located in goblet cells

Bright-field micrograph of the lower half of one villus of rat small intestine (a) and a cross-section of crypts of rat colon (b), counterstained with haematoxylin and eosin. Bars represent 20 μm . Arrowheads point to typical goblet cells.

including a recently described rat intestinal mucin mRNA [12]. Polydispersity of rat MLP mRNA, however, was not a feature in the present or in our previous [4,5,10] experiments. In one blot (Fig. 2c) some smearing of the MLP signal was observed, but the results of a second experiment suggested that the smearing was probably a consequence of partial RNA degradation during electrophoresis.

Recently we have shown that the C-terminal region of rat MLP has a human intestinal cDNA counterpart (400 bp), which shares a high degree of similarity in nucleotide and deduced amino acid sequence [10]. Northern blots and chromosomal mapping suggested that the human MLP gene is probably equivalent to the gene for human intestinal MUC-2. Since MUC-2 is expressed in the lungs of humans with chronic bronchitis [14] and rats with lung damage due to SO₂ [15], we might expect that rat MLP could also be expressed in rat lung. This was not the case for the normal rat lung used in the present study, but MLP expression in diseased rat lung tissue has not yet been tested. It is possible, of course, that both species-specific as well as tissue-specific expression of mucin genes occur and that parallelism in the tissue distribution of rat MLP and human MUC-2 will not be found.

A rat intestinal mucin clone (called RMUC 176) has been reported recently by Gum *et al.* [12], and shown to encode a long stretch of threonine-rich tandem repeats (each 6 amino acids) followed by a cysteine-rich segment of 92 amino acids. The C-terminal sequence of the mucin peptide was not obtained. The mRNA for this gene was shown to be polydisperse, but to give a major signal at approx. 9 kb in Northern blots probed with RMUC 176. The mucin was expressed in rat intestine and colon. No species cross-hybridization was found by the investigators using human intestinal MUC-2 or MUC-3 or rat intestinal RMUC 176 tandem repeat probes in Northern blots of both rat and human tissues. Since our rat MLP shows sequence similarity to human MLP, and the latter exhibits similarities to human MUC-2 (similar-size transcript, similar chromosome localization), we speculate that rat MLP and RMUC 176 may represent two different rat intestinal mucin genes. Definitive confirmation, however, will require further sequencing to extend

RMUC 176 in a 3' direction and/or to extend rat MLP in a 5' direction to look for overlapping sequence.

Financial support was provided by the Medical Research Council of Canada and the Canadian Cystic Fibrosis Foundation. Assistance with photography was provided by Mike Starr and Amy Mercer-Connolly, Department of Pathology, Hospital for Sick Children, Toronto.

REFERENCES

1. Fahim, R. E. F., Specian, R. D., Forstner, G. G. & Forstner, J. F. (1987) *Biochem. J.* **243**, 631–640
2. Mantle, M., Forstner, G. G. & Forstner, J. F. (1984) *Biochem. J.* **217**, 159–167
3. Robertson, A. M., Mantle, M., Fahim, R. E. F., Specian, R., Bennick, A. & Kawagishi, S. (1989) *Biochem. J.* **261**, 637–647
4. Xu, G., Huan, L. J., Khatri, I. A., Wang, D., Bennick, A., Fahim, R. E. F., Forstner, G. G. & Forstner, J. F. (1992) *J. Biol. Chem.* **267**, 5401–5407
5. Huan, L. J., Xu, G., Forstner, G. & Forstner, J. (1992) *Biochim. Biophys. Acta*, in the press
6. Han, J. H., Stratowa, C. & Rutter, W. J. (1987) *Biochemistry* **26**, 1617–1625
7. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn., pp. 743–748, Cold Spring Harbor Laboratory, Cold Spring Harbor
8. Harper, M. E. & Marselle, L. M. (1987) *Methods Enzymol.* **151**, 539–551
9. Eckhardt, A. E., Timpote, C. S., Abernethy, T. L., Zhao, Y. & Hill, R. L. (1991) *J. Biol. Chem.* **266**, 9678–9687
10. Xu, G., Huan, L., Khatri, I., Sajjan, U. S., McCool, D., Wang, D., Jones, C., Forstner, G. & Forstner, J. (1992) *Biochem. Biophys. Res. Commun.* **183**, 821–828
11. Bhargava, A. K., Weitach, J. T., Davidson, E. A. & Bhavanandan, V. P. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6798–6802
12. Gum, J. R., Hicks, J. W., Lagace, R. E., Byrd, J. C., Toribara, N. E., Siddiki, B., Fearney, F. J., Lamport, D. T. A. & Kim, Y. S. (1991) *J. Biol. Chem.* **266**, 22733–22738
13. Strous, G. J. & Dekker, J. (1992) *Crit. Rev. Biochem. Mol. Biol.* **27**, 57–92
14. Jany, B. H., Gallup, M. W., Yan, P. S., Gum, J. R., Kim, Y. S. & Basbaum, C. B. (1991) *J. Clin. Invest.* **87**, 77–82
15. Basbaum, C., Gallup, M., Gum, J., Kim, Y. & Jany, B. (1990) *Biorheology* **27**, 485–489.