Volume 249, number 2, 248–252 FEB 07196 June 1989

Exon-intron organization, expression, and chromosomal localization of the human motilin gene

H. Yano, Y. Seino*, J. Fujita, Y. Yamada, N. Inagaki, J. Takeda*, G.I. Bell+, R.L. Eddy†, Y.-S. Fan†, M.G. Byers†, T.B. Shows† and H. Imura

*Second Division, Department of Internal Medicine and *Division of Metabolism and Clinical Nutrition, Kyoto University School of Medicine, Kyoto 606, Japan, +Howard Hughes Medical Institute and Departments of Biochemistry and Molecular Biology and of Medicine, The University of Chicago, Chicago, IL 60637* and *†Department of Human Genetics, Roswell Park Memorial Institute, New York State Department of Health, Buffalo, NY 14263, USA*

Received 20 March 1989

The human motilin gene has been isolated and characterized. The gene spans about 9 kilobase pairs (kb) and the 0.7 kb motilin mRNA is encoded by five exons. The 22-amino-acid motilin sequence is encoded by exons 2 and 3. The human motilin gene was mapped to the p21.2→p21.3 region of chromosome 6 by hybridization of the cloned cDNA to DNAs from a panel of reduced human-mouse somatic cell hybrids and by in situ hybridization to human prometaphase chromosomes. RNA blotting using RNA prepared from various regions of the human gastrointestinal tract revealed high levels of motilin mRNA in duodenum and lower levels in the antrum of the stomach; motilin mRNA could not be detected by this procedure in the esophagus, cardia of the stomach, descending colon or gallbladder.

Motilin gene; Chromosome 6; (Duodenum, Stomach, Human)

1. INTRODUCTION

Motilin is a 22-amino-acid polypeptide originally isolated from porcine intestine by Brown et al. [1,2]. Physiological studies suggest that it plays an important role in the regulation of interdigestive gastrointestinal motility and indirectly causes rhythmic contraction of duodenal and colonic smooth muscle [3,4]. The isolation of cDNAs encoding human and porcine motilin indicates that both are derived by proteolytic processing of precursors consisting of 115 and 119 amino acids, respectively [5,6]. Interestingly, a search of both protein and nucleic acid sequence data bases showed that the sequences of motilin and its precursor are unrelated to any other polypeptide hormones. This is in contrast to most other gastrointestinal hormones which are members of gene families. As a first step in studying the regulation of motilin gene expression, we have isolated and characterized the gene encoding this unique gastrointestinal hormone. We also have determined the chromosomal localization of the human motilin (designated *MLN*) gene and the distribution of motilin transcripts in several regions of the human gastrointestinal tract.

Correspondence address: H. Yano, Second Division, Department of Internal Medicine, Kyoto University School of Medicine, 54 Shogoin Kawahara-cho, Sakyo-ku, Kyoto 606, Japan

Abbreviation: MLN, motilin gene

The nucleotide sequence presented here has been submitted to the EMBL/GenBank database under accession no.Y07505

2. MATERIALS AND METHODS

2.1. *Isolation of the human motilin gene*

A human genomic library [7] was screened by hybridization with the human motilin cDNA clone phMot-1 [5] by procedures essentially as described by Maniatis et al. [8]. The exons were located using standard procedures and sequenced using the dideoxy chain-termination method [9].

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

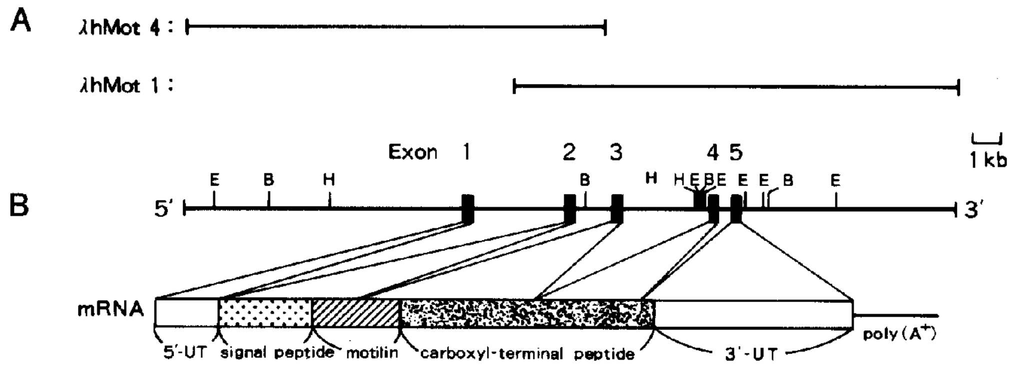

Fig.1. Structure of the human motilin gene. (A) Schematic representation of the extent of the inserts in λhMot-1 and -4. (B) Map of the motilin gene. The filled boxes indicate the positions of exons 1–5. The relationship of each exon to the mature mRNA is indicated. UT represents the untranslated region. The restriction sites are: E, *Eco*RI; H, *Hin*dIII; and B, *Bam*HI.

```
                                                         -301
                                          5'.......acgtcaccaatg
                                                         -241
tcatatatttaaggaacaaaaaaagaaaaggctttgttaaaatgacctctgagatgcagc
                                                         -181
tgagttttcagtggacgggagaatgccatctgggtggggctagcttccaccgaaccctg
                                                         -121
actgtcgctgttccttccaggaaagccctggaagcccatagcgtggctagccctgcctgg
                                                          -61
agtttccacgagcttcaagaatccaggctcccccctctgagggccccaaagctgtggtca
                                                           -1
aaggttaatgggctccaagggcagctcccagggttgggaggtatataagaacccgtcaga
1                            30                            60
TCAGCCGGACACCAGAAGACAAGCAGAGAGACTCCTCCAGACCCACTCAGACCACGTGCA
                       T
CGCCgtaagtagcc......Intron1 (≒2.7kb).........cattgtccagCTCCAAGATGGTA
                                                          MetVal
              90                             120
TCCCGTAAGGCTGTGGCTGCTCTGCTGGTGGTGCATGTAGCTGCCATGCTGGCCTCCCAG
SerArgLysAlaValAlaAlaLeuLeuValValHisValAlaAlaMetLeuAlaSerGln
          150                           180
ACGGAAGCCTTCGTCCCCATCTTCACCTATGGCGAACTCCAGAGGATGCAGgtaaagaac
ThrGluAlaPheValProIlePheThrTyrGlyGluLeuGlnArgMetGln
                                                         210
c......Intron2 (≒1.7kb).........ctgcccctagGAAAAGGAACGGAATAAAGGGCAAAA
                                          GluLysGluArgAsnLysGlyGlnLy
                  240                           270
GAAATCCCTGAGTGTATGGCAGAGGTCTGGGGAGGAAGGTCCTGTAGACCCTGCGGAGCC
sLysSerLeuSerValTrpGlnArgSerGlyGluGluGlyProValAspProAlaGluPr
                  300
CATCAGGGAAGAAGAAAACGAAATGATCAAGgtgagcagac.........Intron3 (≒3.5kb)
oIleArgGluGluGluAsnGluMetIleLys
                             330
.......tatttcacagCTGACTGCTCCTCTGGAAATTGGAATGAGGATGAACTCCAGACAGC
                LeuThrAlaProLeuGluIleGlyMetArgMetAsnSerArgGlnL
      360                           390
TGGAAAAGTACCCGGCCACCCTGGAAGGGCTGCTGAGTGAGATGCTTCCCCAGCATGgta
euGluLysTyrProAlaThrLeuGluGlyLeuLeuSerGluMetLeuProGlnHisA
                                                     420
cggggag......Intron4 (≒0.7kb).........ttttctccagCAGCCAAGTGATGGCCACGC
                                                 laAlaLys***
                  450                           480
TGGGGAGAAGGTGGACAGATTTGGGAGGCCCCTCCTGCCCAAGTGAGGCCCTGGGAATTT
                  510                           540
ACAGAGCCTGCCAGCTGGGCTTGGAAGGAAAACACCTTTCCAAAGCAAATTCCCCTCCAG
                  570                           600
CAAATAAAGCATGAAATATACAGaaatgactcttatttaatattttaattccaaaggcaa
                  630                           660
actcaggtcgcaagaaaaaggggggggaaagtaaagctcagagaaagttggccata.......3'
```

Fig.2. Partial nucleotide sequence of the human motilin gene. Exon sequences are shown in capital letters and intron and flanking sequences are in lower case letters. Nucleotides in the exons are numbered relative to the transcriptional start site. The approximate size of each intron is also indicated. A single base substitution relative to the cDNA sequence in the 5'-untranslated region is indicated. The motilin moiety, putative TATA motif and polyadenylation signal are noted.

2.2. *Primer-extension to map the 5′-end of motilin mRNA*

The transcriptional start site was determined by extension of the ^{32}P-labeled 20-mer, 5′-GCGTGCACGTGGTCTGAGTG-3′, complementary to nucleotides 44–63 of the 5′-untranslated region of motilin mRNA, using duodenal RNA and reverse transcriptase as described previously [10]. A sequencing ladder obtained using the same primer and the appropriate gene template was run in adjacent lanes to indicate the size of the extended product and the sequence at which extension terminated.

2.3. *Gene mapping*

The chromosomal localization of the human *MLN* gene was determined by hybridization of ^{32}P-labeled phMot-1 to Southern blots of *Bam*HI-digested DNA from 31 human-mouse somatic cell hybrid cell lines involving 13 unrelated human cell lines and 4 mouse cell lines [11]. The hybrids had been previously characterized by chromosome analysis and the presence of human enzyme and DNA markers. The regional localization of the *MLN* gene was determined by in situ hybridization [12] of ^{3}H-labeled phMot-1 to normal human prometaphase lymphocyte chromosomes.

2.4. *RNA blotting*

RNA was isolated from human tissues using the guanidine isothiocyanate/cesium chloride procedure [7]. 20 μg of total RNA was denatured with glyoxal, electrophoresed on a 1% agarose gel and then blotted onto a nylon filter. The filter was hybridized with ^{32}P-labeled phMot-1.

3. RESULTS AND DISCUSSION

3.1. *Isolation and characterization of the human motilin gene*

Four of 1×10^6 phage of the human genomic library hybridized with the human motilin cDNA probe. Restriction mapping suggested that the inserts in three of these (λhMot-1, -2 and -3) were identical and overlapped with the insert in λhMot-4 (fig.1). The inserts of λhMot-1 and -4 were isolated and partially sequenced to obtain the exon-intron organization of the motilin gene. The five exons encoding human motilin mRNA span about 9 kb (figs 1 and 2). The nucleotide sequences of the protein coding and 3′-untranslated regions of the gene were identical to those of the cDNA; there is a single base difference in the 5′-untranslated region between these two sequences (fig.2). The first intron interrupts the gene region encoding the 5′-untranslated region of the mRNA. The second intron interrupts the region of the gene encoding the 22-amino-acid motilin sequence and the third and fourth introns are located in the region of the gene coding for the C-terminal peptide of unknown function (fig.2).

The human and porcine motilin precursors consist of 115 and 119 amino acids, respectively. The nucleotide sequences of the regions encoding these two proteins differ because of the apparent deletion in the human precursor sequence (or conversely an insertion in the porcine sequence) of a 12 nucleotide segment beginning 11 nucleotides upstream of the start of intron 4. The characterization of motilin genes from other species might provide a clue as to its origin. The 5′-end of motilin mRNA was identified by primer extension (fig.3) and was located 16 bases upstream from the 5′-end of the cDNA. Thus, the 5′-untranslated region of human motilin mRNA is 72 bases. There is a TATA motif 19 nucleotides upstream from the putative transcriptional initiation site; no other obvious transcriptional control signals occur in about

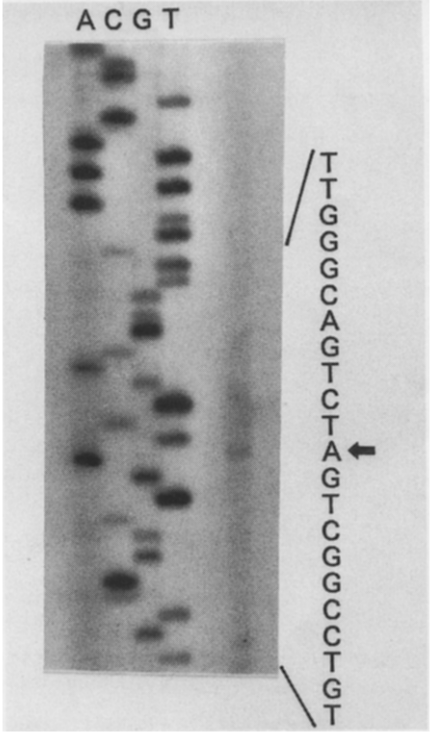

Fig.3. Primer extension analysis of the 5′-end of human motilin mRNA. The ^{32}P-labeled oligonucleotide was annealed to 3 μg of human duodenum poly(A)$^+$ RNA and then extended with reverse transcriptase. The right lane is the primer-extended cDNA. The sequence ladder obtained using the same oligonucleotide as a sequencing primer is indicated. The sequence around the end of the primer-extended cDNA is noted and is complementary to the mRNA sequence. The arrow indicates the putative transcription initiation site.

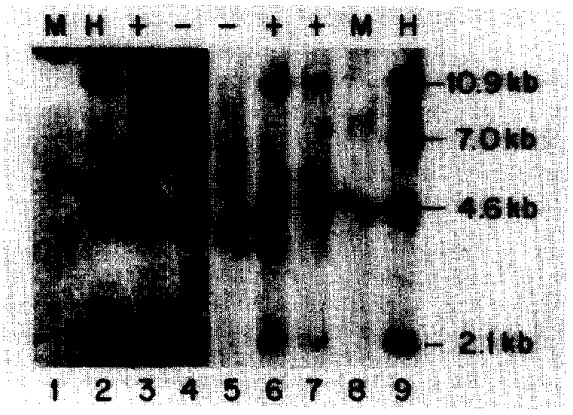

Fig.4. Hybridization of the human motilin cDNA probe to *Bam*HI-digested human-mouse cell hybrid DNAs. Lanes 1 and 8 contain mouse DNA and lanes 2 and 9, human DNA; note the polymorphic 7.0 kb fragment. Lanes 4 and 5 are digests from hybrid cell lines lacking human chromosome 6 and lanes 3, 6 and 7 are from cell lines having human chromosome 6.

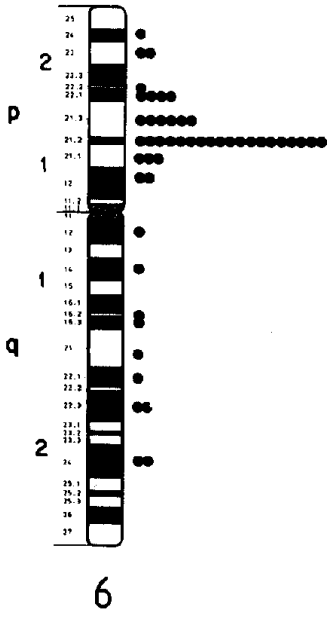

Fig.5. Ideogram of human chromosome 6 showing silver grain distribution after hybridization with the motilin cDNA. One hundred metaphase spreads were examined and 19.6% (48/245) of the grains were on chromosome 6 and 52.1% of these (25/48) were localized in the region p21.2 ⟶ p21.3. No other human chromosomes demonstrated a grain distribution above background.

300 bp of the 5′-flanking region of the motilin gene. The nucleotide sequence of the 3′-flanking region of the motilin gene contains a T-rich segment which could be involved in transcription termination/polyadenylation [13].

3.2. *Chromosomal localization of the human motilin gene*

The chromosomal assignment of the human *MLN* gene was determined from analysis of its segregation in a panel of reduced human-mouse somatic cell hybrids. The motilin cDNA probe hybridized to three human *Bam*HI fragments of 10.9, 4.6 and 2.1 kb (fig.4, lanes 2, 6 and 7); the probe also hybridized to an additional fragment of 7.0 kb in some DNA samples which most likely represents a restriction fragment length polymorphism (fig.4, lanes 3 and 9). The human cDNA probe hybridized very poorly to mouse DNA under the conditions used. The human-specific DNA fragments co-segregated with human chromosome 6 (table 1). In situ hybridization to human prometaphase chromosomes confirmed the assign-

Table 1

Segregation of the *MLN* gene with human chromosomes in *Bam*HI-digested human-mouse cell hybrid DNA

	Human chromosomes																						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X
Concordant no. of hybrids	15	23	17	21	17	31	17	16	12	18	22	13	16	16	20	13	15	21	23	18	17	17	20
Discordant no. of hybrids	13	7	11	10	14	0	13	15	16	12	8	18	15	15	11	18	14	10	8	13	14	13	7
MLN % discordancy	46	23	39	32	45	0	43	48	57	40	27	58	48	48	35	58	48	32	26	42	45	43	26

The scores were tabulated by the presence or absence of human DNA fragments in the different somatic cell hybrids. Concordant hybrids have either retained or lost the gene together with a specific chromosome. Discordant hybrids either retained the gene but not a specific chromosome, or the reverse. Percent discordancy indicates the degree of discordant segregation for a marker and a chromosome. A 0% discordancy is the basis for chromosome assignment

ment of the *MLN* gene to chromosome 6 and further localized it to the p21.2 ⟶ p21.3 region of the short arm of this chromosome (fig.5).

3.3. *Distribution of motilin mRNA*

Hybridization of the human motilin cDNA probe to Northern blots of RNA prepared from various regions of the human gastrointestinal tract revealed an abundant 0.7 kb motilin transcript in duodenum (fig.6). Lower levels of motilin mRNA were present in RNA prepared from the gastric antrum. The motilin mRNA levels in the esophagus, cardia of the stomach, descending colon and gallbladder were below the sensitivity of our RNA blotting assay. As discussed by Bond et al. [6], conflicting localizations of motilin have been reported using immunohistocytochemical procedures with some data suggesting widespread distribution of this peptide in the gastrointestinal tract and central nervous system. These localization studies should be repeated using the technique of hybridization histochemistry to identify specific cells containing motilin mRNA since nucleotide probes may be more specific than antibodies prepared against the peptide or fragments thereof. Nonetheless, the RNA blotting data presented here and in the report of Bond et al. [6] strongly suggest that motilin expression is restricted to the upper small intestine and gastric antrum.

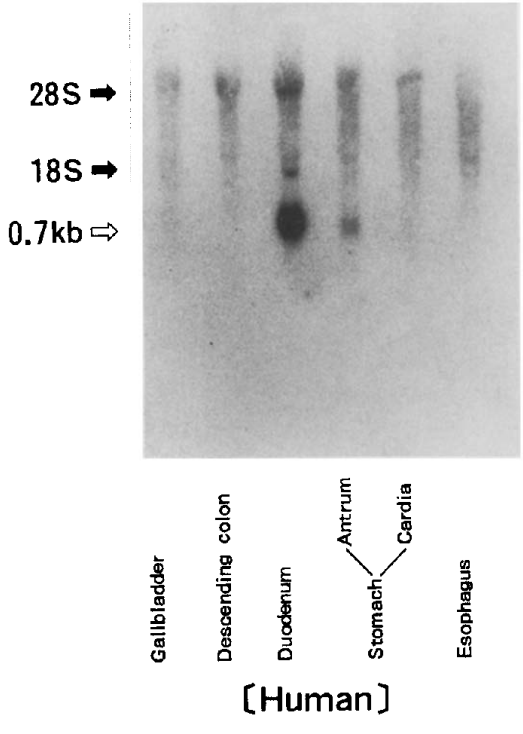

Fig.6. Tissue distribution of motilin mRNA. The positions of 18 S and 28 S ribosomal RNAs are indicated.

Acknowledgements: We are very grateful to Dr S. Ishii (Riken), Dr H. Fukumoto and Dr G. Koh (Kyoto University) for valuable advice in these studies. We also express our appreciation to T. Mitani (Sanwa Kagaku Kenkyusho) for preparation of the synthetic oligonucleotide used for primer extension analysis, and to Ms H. Tachikawa for her expert assistance in preparing this manuscript. This research was supported by Grants in Aid for Scientific Research 63440042 from the Ministry of Education, Science, and Culture, Japan (to H.I.), in part by Uehara Memorial Foundation and Grant-in-Aid from the Mochida Memorial Foundation for Medical and Pharmaceutical Research (to Y.S.), and by National Institutes of Health Grants GM-2054 and HD-05196 (to T.B.S.).

REFERENCES

[1] Brown, J.C., Cook, M.A. and Dryburgh, J.R. (1972) Gastroenterology 62, 401–404.
[2] Schubert, H. and Brown, J.C. (1974) Can. J. Biochem. 52, 7–8.
[3] Itoh, Z., Honda, R., Hiwatashi, K., Takeuchi, S., Aizawa, I., Takayanagi, R. and Couch, E.F. (1976) Scand. J. Gastroenterol. 11 (suppl.39), 93–110.
[4] Rennie, J.A., Christofides, N.D., Bloom, S.R. and Johnson, A.G. (1979) Gut 20, A912.
[5] Seino, Y., Tanaka, K., Takeda, J., Takahashi, H., Mitani, T., Kurono, M., Kayano, T., Koh, G., Fukumoto, H., Yano, H., Fujita, J., Inagaki, N., Yamada, Y. and Imura, H. (1987) FEBS Lett. 223, 74–76.
[6] Bond, C.T., Nilaver, G., Godfray, B., Zimmerman, E.A. and Adelman, J.P. (1988) Mol. Endocrinol. 2, 175–180.
[7] Lawn, R.M., Fritsch, E.F., Parker, R.C., Blake, G. and Maniatis, T. (1978) Cell 15, 1157–1174.
[8] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[9] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[10] Fukumoto, H., Seino, S., Imura, H., Seino, Y. and Bell, G.I. (1988) Diabetes 37, 657–661.
[11] Shows, T., Eddy, R., Harley, L., Byers, M., Henry, M., Fujita, T., Matsui, H. and Taniguchi, T. (1984) Somatic Cell Mol. Genet. 10, 315–318.
[12] Zabel, B.U., Naylor, S.L., Sakaguchi, A.Y., Bell, G.I. and Shows, T.B. (1983) Proc. Natl. Acad. Sci. USA 80, 6932–6936.
[13] McDevitt, M.A., Hart, R.P., Wong, W.W. and Nevins, J.R. (1986) EMBO J. 5, 2907–2913.