Volume 223, number 1, 74–76 FEB 05186 October 1987

Sequence of an intestinal cDNA encoding human motilin precursor

Y. Seino[+], K. Tanaka, J. Takeda*, H. Takahashi, T. Mitani, M. Kurono, T. Kayano, G. Koh*, H. Fukumoto*, H. Yano*, J. Fujita*, N. Inagaki*, Y. Yamada* and H. Imura*

[+] *Division of Metabolism and Clinical Nutrition, *Second Division, Department of Medicine, Kyoto University School of Medicine, Kyoto 606* and *Mie Research Laboratory, Sanwa Kagaku Kenkyusho Co., Ltd, Mie 511-04, Japan*

Received 19 August 1987

A cDNA clone encoding the human motilin precursor was isolated from an intestinal library using synthetic oligonucleotide probes. The predicted amino acid sequence indicates that the motilin precursor consists of 115 amino acids and includes a 25-residue N-terminal signal peptide followed by the 22-amino-acid motilin sequence and a long, 68-residue C-terminal peptide. The amino acid sequence of human motilin predicted from the cDNA sequence is indentical to its porcine counterpart, which has been determined by protein sequencing. Proteolytic processing of promotilin to motilin occurs at the sequence, Lys-Lys, this being the first reported instance of processing occurring at a pair of Lys residues. In other precursors it occurs at Lys-Arg, Arg-Arg, Arg, or very rarely Lys.

cDNA sequence; Motilin; (Human)

1. INTRODUCTION

Motilin is a unique gastrointestinal hormone, which plays an important role in regulating interdigestive gastrointestinal contraction [1]. Brown et al. [2,3] isolated porcine motilin from mucosa of the small intestine, and determined the sequence of this 22-amino-acid hormone. However, neither the sequence of human motilin nor that of the motilin precursor has been determined as yet. Here, we describe the isolation and sequencing of a cDNA clone encoding human motilin precursor.

Correspondence address: Y. Seino, Division of Metabolism and Clinical Nutrition, Kyoto University Hospital, Shogoin, Sakyo-ku, Kyoto 606, Japan

The nucleotide sequence presented has been submitted to the EMBL/GenBank database under the accession number Y00695

2. MATERIALS AND METHODS

2.1. *Oligonucleotide probe*

As the amino acid sequence of human motilin was unknown, an oligonucleotide probe (23-mer), carrying deoxyinosine residues at positions corresponding to ambiguous nucleotides [4], has been synthesized from the amino acid sequence of porcine motilin [3]; the sequence 3′-AA(A/G)CAIGGITA(A/G/T)AA(A/G)TGIAT(A/G)CC-5′ was complementary to the mRNA for the N-terminus of porcine motilin, Phe-Val-Pro-Ile-Phe-Thr-Tyr-Gly (excluding the third nucleotide of Gly).

2.2. *cDNA cloning and sequencing*

Human duodenum was obtained from a patient who had undergone a pancreatoduodenectomy for cancer of the pancreas. RNA was extracted [5] and poly(A)[+] RNA prepared by oligo(dT)-cellulose column chromatography [6]. A cDNA library was prepared according to Okayama and Berg [7] and screened for clones encoding human motilin, as described [8], except that the hybridization was

Published by Elsevier Science Publishers B.V. (Biomedical Division)

performed at 50°C. DNA sequencing was performed by the chain-termination method [9].

3. RESULTS AND DISCUSSION

3.1. *Isolation and sequencing of a cDNA clone encoding the human prepromotilin*

50000 ampicillin-resistant bacterial colonies were screened by hybridization with the oligonucleotide probe described above. Plasmid DNA, phMot-1, was prepared from the single bacterial clone that hybridized with the probe. The sequence of the 555 base pair (bp) insert (excluding GC tails and poly(A) tract) (fig.1) revealed a single open reading frame of 345 bp, containing the human motilin sequence. The 5′- and 3′-untranslated regions of this sequence are 55 and 155 bp long, respectively. There is a canonical polyadenylation signal [10], AATAAA (nucleotides 480–485), located 15 bp before the poly(A) tract.

```
          -50       -40       -30       -20       -10       -1
5'---- AGACAAGTAGAGAGACTCCTCCAGACCCACTCAGACCACGTGCACGCCCTCCAAG

  1       10        20        30        40        50        60
ATGGTATCCCGTAAGGCTGTGGCTGCTCTGCTGGTGGTGCATGTAGCTGCCATGCTGGCC
MetValSerArgLysAlaValAlaAlaLeuLeuValValHisValAlaAlaMetLeuAla
 /1                  10                                    20

         70        80        90        100       110       120
TCCCAGACGGAAGCCTTCGTCCCCATCTTCACCTATGGCGAACTCCAGAGGATGCAGGAA
SerGlnThrGluAlaPheValProIlePheThrTyrGlyGluLeuGlnArgMetGlnGlu
                    30                                    40

         130       140       150       160       170       180
AAGGAACGGAATAAAGGGCAAAAGAAATCCCTGAGTGTATGGCAGAGGTCTGGGGAGGAA
LysGluArgAsnLysGlyGlnLysLysSerLeuSerValTrpGlnArgSerGlyGluGlu
                    50                                    60

         190       200       210       220       230       240
GGTCCTGTAGACCCTGCGGAGCCCATCAGGGAAGAAGAAAACGAAATGATCAAGCTGACT
GlyProValAspProAlaGluProIleArgGluGluGluAsnGluMetIleLysLeuThr
                    70                                    80

         250       260       270       280       290       300
GCTCCTCTGGAAATTGGAATGAGGATGAACTCCAGACAGCTGGAAAAGTACCCGGCCACC
AlaProLeuGluIleGlyMetArgMetAsnSerArgGlnLeuGluLysTyrProAlaThr
                    90                                   100

         310       320       330       340       350       360
CTGGAAGGCTGCTGAGTGAGATGCTTCCCCAGCATGCAGCCAAGTGATGGCCACGCTGC
LeuGluGlyLeuLeuSerGluMetLeuProGlnHisAlaAlaLys***
                   110

         370       380       390       400       410       420
CGAGAAGGTGGACAGATTTGGGACGCCCCTCCTGCCCAAGTGAGGCCCTGGGAATTTACA

         430       440       450       460       470       480
GAGCCTGCCAGCTGGGCTTGGAAGGAAAACACCTTTCCAAAGCAAATTCCCCTCCAGCAA

         490       500
ATAAAGCATGAAATATACAG ------3'
```

Fig.1. Nucleotide and amino acid sequences of human prepromotilin. The region of the motilin sequence is boxed by straight line. The locations of the oligonucleotide probe used for screening (nucleotides 76–98) and putative poly(A) addition signal (nucleotides 480–485) are underlined.

3.2. *Human motilin precursor sequence*

The cDNA sequence indicates that human prepromotilin consists of 115 amino acids (M_r 12922) (fig.1). Hydropathic analysis [11] of the prepromotilin sequence reveals that the N-terminus is very hydrophobic, and since motilin is a secreted protein, it probably represents the signal peptide [12]. It is reasonable to assume that the signal peptide cleavage site follows Ala-25, which is adjacent to the N-terminus of motilin. Thus, the motilin precursor includes a 25-amino-acid signal peptide, followed by the 22-residue motilin sequence, and then a 68-amino-acid C-terminal peptide of unknown function.

The sequence of the 90-amino-acid promotilin polypeptide (M_r 10372) indicates that processing to motilin occurs via specific proteolytic cleavage at a pair of Lys residues adjacent to the C-terminus of the motilin sequence. Proteolytic processing most frequently occurs at the sequences Lys-Arg, Arg-Arg or Arg [13]. Processing at Lys occurs in the *Aplysia* FMRF-amide [14] and guinea pig pancreatic polypeptide (personal communication) precursors. The motilin precursor is the first known instance of a pair of Lys residues occurring at the cleavage site. Although there are no other pairs of basic amino acids in the promotilin sequence, a number of Arg and Lys residues are found in the C-terminal peptide sequence. It remains unknown whether processing might also occur at some of these sites so as to generate motilin-containing polypeptides retaining various portions of the C-terminal peptide.

3.3. *Comparison of motilin sequence*

The sequences of human and porcine motilin are identical. There are five amino acid differences between these sequences and that of canine motilin [15] at residues 7, 8, 12, 13 and 14 (fig.2).

A search of both protein and nucleic acid sequence data bases for sequences homologous to

```
          5         10        15        20
human   F V P I F T Y G E L Q R M Q E K E R N K G Q
pork    F V P I F T Y G E L Q R M Q E K E R N K G Q
dog     F V P I F T H S E L Q K I R E K E R N K G Q
```

Fig.2. A comparison of the sequences of motilin from human, pork and dog. Identical amino acids are boxed.

the human prepromotilin portion or cDNA sequence revealed no significant homologies. Thus, in contrast to most other gastrointestinal hormones, which belong to structurally related families, motilin and its precursor appear to be unrelated to previously described hormones.

The elucidation of the primary structure of human motilin and prepromotilin, as well as the availability of a cDNA clone encoding these proteins, now permits the examination of biosynthesis and physiological properties of this interesting hormone.

ACKNOWLEDGEMENT

We express our appreciation to Dr G.I. Bell for computer analysis for sequence homology and for helpful suggestions.

REFERENCES

[1] McIntosh, C.H.S. and Brown, J.C. (1980) in: Gastrointestinal Hormones (Glass, G.B.J. ed.) pp.233–244, Raven, New York.
[2] Brown, J.C., Cook, M.A. and Dryburgh, J.R. (1972) Gastroenterology 62, 401–404.
[3] Schubert, H. and Brown, J.C. (1974) Can. J. Biochem. 52, 7–8.
[4] Ohtsuka, E., Matsuki, S., Ikehara, M., Takahashi, Y. and Matsubara, K. (1985) J. Biol. Chem. 260, 2605–2608.
[5] Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, Y.J. (1979) Biochemistry 18, 5294–5299.
[6] Aviv, H. and Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 1408–1412.
[7] Okayama, H. and Berg, P. (1982) Mol. Cell. Biol. 2, 161–170.
[8] Takeda, J., Seino, Y., Tanaka, K.-I., Fukumoto, H., Kayano, T., Takahashi, H., Mitani, T., Kurono, M., Suzuki, T., Tobe, T. and Imura, H. (1987) Proc. Natl. Acad. Sci. USA, in press.
[9] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[10] Proudfoot, N.J. and Brownlee, G.G. (1976) Nature 263, 211–214.
[11] Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105–132.
[12] Steiner, D.F., Quinn, P.S., Chan, S.J., Marsh, J. and Tager, H.S. (1980) Ann. NY Acad. Sci. 342, 1–16.
[13] Schwartz, T.W., Wittels, B. and Tager, H.S. (1983) in: Proc. 8th Am. Peptide Symp., pp.229–238, Pierce, Rockford.
[14] Taussig, R. and Scheller, R.H. (1986) DNA 5, 453–461.
[15] Poitras, P., Reeve, J.R. jr, Hunkapiller, M.W., Hood, L.E. and Walsh, J.H. (1983) Regul. Peptides 5, 197–208.