Structure of murine enterokinase (enteropeptidase) and expression in small intestine during development

XIN YUAN, XINGLONG ZHENG, DESHUN LU, DEBORAH C. RUBIN, CHRISTOPHER Y. M. PUNG, AND J. EVAN SADLER
Howard Hughes Medical Institute, Departments of Medicine and Biochemistry and Molecular Biophysics, Barnes-Jewish Hospital, Washington University School of Medicine, St. Louis, Missouri 63110

Yuan, Xin, Xinglong Zheng, Deshun Lu, Deborah C. Rubin, Christopher Y. M. Pung, and J. Evan Sadler.
Structure of murine enterokinase (enteropeptidase) and expression in small intestine during development. Am. J. Physiol. 274 (Gastrointest. Liver Physiol. 37): G342–G349, 1998.—Enterokinase (enteropeptidase) is expressed only in proximal small intestine, where it initiates digestive enzyme activation by converting trypsinogen into trypsin. To investigate this restricted expression pattern, mouse enterokinase cDNA was cloned, and the distribution of enterokinase mRNA and enzymatic activity were determined in adult mice and during gestation. Analysis of enterokinase sequences showed that a mucinlike domain near the NH2 terminus is composed of repeated 15-amino acid Ser/Thr-rich motifs. By Northern blotting and trypsinogen activation assays, enterokinase mRNA and enzymatic activity were undetectable in stomach, abundant in duodenum, and decreased distally until they were undetectable in midjejunum, ileum, and colon. By in situ mRNA hybridization, enterokinase mRNA was localized to the enterocytes throughout the villus. Expression was not observed in goblet cells, Paneth cells, or Brunner’s glands. Enterokinase mRNA and enzymatic activity were not detected in the duodenum of fetal mice but were easily detected in the duodenum on postnatal days 2–6. Both enterokinase mRNA and enzymatic activity decreased to very low levels after day 7 but increased after weaning and reached a high level characteristic of adult life by day 60. Therefore, in mice, duodenal enterocytes are the major type of cells expressing enterokinase, which appears to be regulated at the level of mRNA abundance.

duodenum; trypsinogen activation; serine protease; pancreatic enzymes

ENTEROKINASE, also known as enteropeptidase, was discovered in Pavlov’s laboratory as an activity of small intestinal extracts that stimulated the digestive enzymes of pancreatic secretions (33). Subsequent studies showed that enterokinase is a protease that activates trypsinogen (21). The resultant trypsin then cleaves and activates the zymogens of other pancreatic digestive hydrolases. Inherited deficiency of enterokinase impairs the initiation of this digestive hydrolase cascade, thereby causing intestinal malabsorption that may be life threatening (13, 14).

The amino acid sequences of bovine (19, 23), human (18, 19), porcine (29), and rat (39) enterokinase were recently determined indirectly by cDNA cloning. Enterokinase appears to be synthesized as a single-chain serine protease zymogen with an NH2-terminal hydrophobic domain that could mediate membrane association. This mechanism of membrane localization has not been demonstrated directly, and other models could account for the observed distribution of enterokinase.

In all animal species studied, enterokinase is highest in duodenum and rapidly decreases aborally, becoming undetectable by the distal jejunum (27, 32). In humans (1) and rats (24), this distribution is established late in fetal life. Enterokinase is associated with the brush border of enterocytes and appears to be absent in other cell types, although enterokinase antigen was reported in some goblet cells (16, 26, 30). Substantial quantities of free enterokinase also occur in mucinous secretions of bovine (10) and porcine (28) small intestine, suggesting that enterokinase could be secreted by other cells and localized secondarily on enterocytes.

To address the distribution and developmental regulation of enterokinase, the expression of enterokinase mRNA and enzymatic activity in mouse intestine was characterized in fetal and adult life. The complete cDNA sequence encoding murine enterokinase was cloned and employed to identify the cells within the proximal small intestine that express enterokinase mRNA by in situ hybridization.

MATERIALS AND METHODS

Preparation of RNA. Total RNA was isolated from tissues of adult C57BL/6 mice (Jackson Laboratory) using RNAzol B (Biotech Laboratories, Houston, TX), according to the manufacturer’s recommended procedures. Briefly, fresh tissue samples were frozen and stored in liquid nitrogen until use. Tissues were thawed on ice in RNAzol solution and homogenized with a Polytron model PT3000 homogenizer (Brinkmann) for <30 s at 2,400 rpm. Chloroform (0.1 vol) was added, and samples were incubated for 15 min on ice. After centrifugation at 14,000 g for 20 min, the upper layer was collected and reextracted with RNAzol B and chloroform. RNA was precipitated with ethanol and resuspended in water. Poly(A)+ mRNA was prepared from total RNA of mouse duodenum by oligo(dt) chromatography (Clontech) (3).

Isolation and characterization of mouse enterokinase cDNA. A pair of degenerate 20-nucleotide (nt) oligonucleotide primers was synthesized based on the conserved nucleotide sequence surrounding the active serine residue and the NH2-terminal light-chain sequence for bovine and human enterokinase (19). A fragment of mouse enterokinase cDNA was synthesized with these primers by reverse transcription-polymerase chain reaction (RT-PCR) using mouse duodenum total RNA as the template. The 576-bp product was cloned into TA cloning vector (Invitrogen) to generate plasmid pMEK1. For cDNA library screening, an EcoRI fragment containing the insert of pMEK1 was radiolabeled with [32P]dCTP by random oligonucleotide priming (9).

A cDNA library was prepared from 2 µg of duodenal poly(A)+ RNA using a ZAP Express cDNA synthesis kit...
Spondyloptosis is a condition affecting the vertebrae, causing them to be out of alignment.

Gene expression was analyzed by real-time PCR using specific primer sets. The primer sequences are as follows:

- Forward primer: `5'-CTT GAC ACA TCG AAT GAT G-3'
- Reverse primer: `5'-TGC AGC ACT TAG CAT TAG CTT G-3'

Results showed a significant upregulation of expression in the experimental group compared to the control group.

In situ hybridization analysis was performed to further confirm the gene expression data. Hybridization was visualized using a fluorescent signal.

Immunohistochemical analysis was also conducted to localize the expression of the gene of interest. Positive staining was observed in the targeted tissue areas.

These findings suggest that the gene in question plays a crucial role in the development of spondyloptosis.

References:
Fig. 1. Amino acid sequence of murine proenterokinase. Translated composite cDNA sequence of mouse proenterokinase is aligned with sequences of rat, cow, human, and pig proenterokinase. Amino acid residues identical in all 4 sequences are boxed. Specific motifs are indicated by labeled brackets. These include proposed signal-anchor sequence and a potentially O-glycosylated mucinlike domain that is variable among different species. Four 21-residue repeats within mucinlike domain are indicated and labeled a–d above aligned sequences. Other conserved motifs include 2 low-density lipoprotein receptor (LDLR) domains, 2 complement component C1r or C1s (C1r/s) domains, a MAM domain (named for motifs found in Meprin, Xenopus laevis A5 protein, and protein tyrosine phosphatase µ), and a macrophage scavenger receptor (MSCR) domain. Cysteine residues are shown in boldface. Potential N-linked glycosylation sites are indicated by underlined boldface type. Positions of activation cleavage site (arrow) between heavy and light chains and serine protease His, Asp, and Ser active site residues (●) are indicated. A basic segment that is predicted to interact with acidic substrate residues is labeled (+++).
sequence of 792 nt and a poly(A) tail. The open reading frame encodes a protein of 1,069 amino acids that is 70–88% identical in sequence to bovine (19, 23), human (18, 19), porcine (29), or rat (39) enterokinase. A full-length murine enterokinase cDNA was assembled and expressed by transient transfection of 293T cells and COS-7 cells; the recombinant protein activated bovine trypsinogen and was cell associated, confirming that the cloned sequence encoded enterokinase.

The sequence of murine enterokinase is highly conserved compared with enterokinase from other species, although there are some potentially significant differences. All of the repeated motifs found in other mammalian enterokinase proteins are present in murine enterokinase (Fig. 1). These include the two low-density lipoprotein receptor repeats (37), two complement component C1r/s repeats (25), one MAM domain (5), and one macrophage scavenger receptor domain (11). The serine protease domain is also highly conserved, including a sequence of four basic amino acids (Fig. 1, residues 919–922) proposed to interact with the acidic propeptide of the trypsinogen substrate (19, 29). There are two insertions in the serine protease domain of murine enterokinase. Four extra amino acids are inserted at positions 850–853, in a predicted surface loop that could interact with substrate residues on the COOH-terminal side of the scissile bond. A single alanine residue is inserted at position 971, in a surface loop below the substrate binding cleft. Both of these insertion sites are relatively variable among serine proteases (19).

Some nonconserved features of the murine enterokinase heavy chain deserve comment. A potential N-myristoylation site is present at residue Gly2 in bovine, human and porcine enterokinase, and this glycine is replaced by Lys2 in murine enterokinase. The proposed signal-anchor segment of bovine, human, and porcine enterokinase contains a cysteine at position 32, and this is replaced by Ser32 in murine enterokinase. A more

Fig. 2. Ser/Thr-rich repeats in enterokinase mucin domains. Approximately 15 amino acid mucinlike segments of mouse, rat, cow, human, and pig enterokinase were lettered (a–d) as shown in alignment of Fig. 1 and aligned using program Megalign (DNASTAR). Amino acid numbers are indicated at right. Residues matching consensus sequence shown at bottom are boxed and shaded. Consensus N-glycosylation sites and a cysteine residue in murine enterokinase are indicated in boldface.

Fig. 3. Tissue distribution of mouse enterokinase mRNA and enzymatic activity. A: total cellular RNA was extracted from segments of mouse gastrointestinal tract and analyzed by Northern blotting. Membrane was hybridized sequentially with probes for mouse enterokinase (top), rat intestinal alkaline phosphatase (IAP), and human β-actin. B: enterokinase enzyme activity was assayed in detergent extracts of tissue segment as described under MATERIALS AND METHODS. Values are expressed as enterokinase (ng) relative to amount of total protein (mg) in extract.
striking difference is present in the Ser/Thr-rich domain between residues 193–222. This segment is 30 amino acids longer in murine enterokinase than in bovine or porcine enterokinase and 45 amino acids longer than in human enterokinase. The additional residue appears to constitute imperfect copies of a 15-residue motif (Fig. 2). This domain of murine enterokinase also contains two potential N-glycosylation sites and a nonconserved cysteine residue. Finally, the MAM domain of murine enterokinase contains four additional cysteines that are not present in other MAM domains (Fig. 1); at present, there is no information on MAM domain structure that could help predict whether these cysteines might form intradomain disulfide bonds.

Tissue and cellular expression of mouse enterokinase. The distribution of enterokinase expression in mouse intestine was determined by Northern blot (Fig. 3). Enterokinase mRNA was detected as a 4.4-kb band. The levels of both mRNA and enzymatic activity were highest in duodenum and decreased in the aborad direction, becoming undetectable by the midjejunum (Fig. 3). For comparison, the blots were stripped and rehybridized with a probe for intestinal alkaline phosphatase, another brush-border enzyme that was found to be distributed throughout the jejunum. Neither enterokinase nor alkaline phosphatase was expressed in the stomach, ileum, or colon. Enterokinase mRNA also was not detected in the pancreas, liver, testis, or brain of adult mice (data not shown).

Enterokinase expression along the villus-crypt axis was examined by in situ RNA hybridization with radiolabeled enterokinase single-stranded RNA probes (Fig. 4). When hybridized with the antisense probe, sections of duodenum demonstrated many autoradiographic grains over the villi. The crypt cells did not have overlying grains in a concentration significantly greater than background. Under higher bright-field magnification, in well-oriented 608-µm sections, the presence of grains was determined for 200 enterocytes and 50 goblet cells, and grains were seen only over the cytoplasm of enterocytes. No grains were detected over the rare Paneth cells, and no specific hybridization was detected at the level of Brunner’s glands. Tissue sections hybridized with the control sense single-stranded RNA probe did not exhibit hybridization signals.

Developmental expression of mouse enterokinase. Enterokinase expression in the duodenum during embryonic and postnatal development was evaluated by Northern blotting and enzymatic assays (Fig. 5). Enterokinase mRNA and activity were not detected before

Fig. 4. In situ mRNA hybridization analysis of mouse enterokinase in duodenal tissues. Sections of adult mouse duodenum were hybridized with 35S-labeled RNA probes for enterokinase (magnification: ×125, A and B; ×160, C and D). A: antisense probe, 10- to 12-µm section. B: same section as in A with bright-field illumination. C: antisense probe, 6- to 8-µm section. D: control sense probe, 6- to 8-µm section.
birth. Enterokinase activity was highly expressed within the first 24 h of life, at a time when enterokinase mRNA was barely detectable. Enterokinase mRNA levels peaked during postnatal days 2–4, declined to very low levels from postnatal days 6 to 12, became easily detected by Northern blotting by postnatal day 21, and reached a high level of expression characteristic of adult life by postnatal day 60. The more sensitive RT-PCR method demonstrated the presence of enterokinase mRNA on postnatal days 0 and 8–12 (data not shown). Enterokinase activity tended to follow enterokinase mRNA levels but appeared to be more sensitive; activity was always readily detected in the postnatal period, and an increase in activity was observed on postnatal day 12 before the increase in mRNA expression was apparent by Northern blotting on postnatal day 21. The timing of this later increase was somewhat variable; another cohort of mice showed a rise in enterokinase mRNA to the adult level by postnatal day 16 (data not shown). In contrast, mRNA for intestinal alkaline phosphatase could be detected on postnatal day 0, and the level remained relatively constant throughout postnatal life (Fig. 5A).

DISCUSSION

Several features of the murine enterokinase sequence provide new insights into the structure-function relationships of this protease. The NH2-terminal amino acid sequences of bovine, human, and porcine enterokinase appear to meet the substrate specificity requirements for N-myristoyltransferase (35), suggesting that Gly2 may be myristoylated. However, Gly2 is not conserved in murine or rat enterokinase (39), indicating that myristoylation is not generally necessary for enterokinase biosynthesis or membrane localization. Enterokinase from several species was reported to have a segment of variable length, beginning at amino acid residue ~166, which is rich in Ser and Thr residues (18, 19, 29, 39) and may be O-glycosylated (29). One bovine enterokinase clone lacked the entire segment, which would be consistent with additional length polymorphism or alternative splicing (19). Examination of the longer murine enterokinase sequence (Fig. 1) indicates that, in all species, this segment is composed of tandem repeats of an ~15-amino acid motif (Fig. 2). Similar repeats are characteristic of heavily O-glycosylated epithelial mucins. For example, human small intestinal mucin MUC2 contains ~100 repeats of a 23-amino acid sequence, and this tandem array apparently is not divided by introns within the MUC2 gene (38). The frog integumentary mucin FIM-B.1 contains at least 22 repeats of an 11-residue sequence, but each of these repeats is encoded by a separate exon and alternative splicing yields a polydisperse family of FIM-B.1 mRNA species (34). The enterokinase mucin-type repeats are distinctive because some contain consensus sites for N-glycosylation, and one murine enterokinase repeat contains a cysteine (Fig. 2). The function of the variable mucin domain in enterokinase is not known, although extensive O-linked glycosylation may protect enterokinase from proteolysis.

The cell type in intestine that produces enterokinase has been controversial. Both enterokinase activity and antigen are associated with the brush border of enterocytes in the duodenum and proximal jejunum (16, 26, 30), although enterokinase antigen was reported in occasional goblet cells (30). By Northern blotting, enterokinase mRNA appears to be restricted to the small intestine (18). This pattern would be consistent with the synthesis of enterokinase by enterocytes and lo-
calization therein as an integral protein of the microvillus membrane. However, substantial quantities of free enterokinase appear to occur in the mucinous secretions of bovine (10) and porcine (28) small intestine. Furthermore, purified porcine enterokinase was found to lack the predicted NH$_2$-terminal transmembrane domain, presumably due to proteolytic processing (29), suggesting that enterokinase could be synthesized and secreted by other cell types and localized secondarily in the brush border of enterocytes. In that case, the observed proximal-distal gradient of enterokinase in the small intestine would also not necessarily reflect the pattern of synthesis, and localization of enterokinase in jejunum, for example, might be due to transport from the duodenum. However, studies of mRNA localization indicate that the sites of enterokinase synthesis and activity in vivo do correspond. Enterokinase mRNA and enzymatic activity are present in duodenum and proximal jejunum (Fig. 3), and the results of in situ mRNA hybridization (Fig. 4) appear to exclude cells other than enterocytes as major sources of enterokinase.

Enterokinase expression during fetal and postnatal development has previously been studied only by enzymatic activity assays. In the rat, enterokinase activity appeared at embryonic day 20 and rapidly increased to the adult level by postnatal day 2 (24). In the mouse, we found a burst of enterokinase activity in the duodenum on postnatal day 0 and lower activity during the following 3 wk of life, after which enterokinase increased to the adult level (Fig. 5B); this pattern is consistent with a previous report (2). The appearance of enterokinase activity at the time of birth coincides with a major step in the maturation of small intestinal epithelium, the development of crypts and restriction of cell proliferation to stem cells within them (7, 17). The later increase corresponds approximately to the time of weaning and also is accompanied by functional changes in the intestinal epithelium to accommodate the adult diet. A similar pattern has been reported for the activities of several other brush-border hydrolases, including sucrase-isomaltase and maltase (15). The mechanisms that regulate the expression of these genes are not fully characterized.

Enterokinase mRNA appeared to parallel these changes in enzyme activity, indicating that enterokinase is regulated at least by mRNA abundance, possibly at the level of transcription. However, the correlation of mRNA and activity was imperfect. For example, substantial enterokinase activity was present at times when mRNA levels were low or undetectable by Northern blotting, especially on postnatal day 0 and between postnatal days 8 and 12 (Fig. 5). This discrepancy may reflect a higher sensitivity of enterokinase activity assays compared with Northern blotting, since enterokinase mRNA was detected easily at all of these time points by RT-PCR. Alternatively, enterokinase may also be regulated at a posttranscriptional level such as translation or posttranslational modification.

The mRNA and enzyme activity of enterokinase restricted mainly in duodenum and proximal jejunum suggests that enterokinase gene expression was controlled by a promoter whose properties could provide insight into the regional specialization of enterocytes during development. In addition, enterokinase is regulated dramatically in mature duodenum by pancreatic secretions. An isolated loop of dog intestine lost enterokinase activity after several hours, and addition of pancreatic juice restored enterokinase within 30 min; this stimulatory activity of pancreatic juice was destroyed by boiling (33). Similar results were obtained in a mouse model of exocrine pancreatic deficiency: enterokinase activity was nearly absent from the duodenum of CBA/J-eple mice but was induced to normal levels by feeding trypsinogen (22). The dependence of enterokinase activity on contact with pancreatic secretions also has been demonstrated in rat (31) and guinea pig models (6). Additional study will be required to determine whether luminal contents regulate enterokinase at the level of transcription, mRNA stability or translation, or zymogen activation.

The authors thank Dr. Qingyu Wu (Berlex Biosciences) and Dr. Yasunori Kitamoto (Kumamoto University) for helpful discussions; Drs. David H. Alpers and Michael J. Engle (Washington University) for the intestinal alkaline phosphatase probe; Cathy Gapart, Jill Robey, Teresa Tolley, and Qingme Xie (Washington University) for assistance with in situ mRNA hybridization assays; and Lisa Westfield for animal breeding.

These studies were supported in part by National Institute of Diabetes and Digestive and Kidney Diseases Grants DK-50053 (to J. E. Sadler) and DK-46122 (to D. C. Rubin).

Address for reprint requests: J. E. Sadler, Howard Hughes Medical Institute, Washington University School of Medicine, 660 South Euclid Ave., Box 8022, St. Louis, MO 63110.

Received 15 October 1996; accepted in final form 21 October 1997.

REFERENCES


