Primary structures of bovine liver low molecular weight kininogen precursors and their two mRNAs

(bradykinin/recombinant DNA/DNA sequence/internal homology/RNA processing)

HIROYUKI NAWA*, NAOMI KITAMURA*, TADAAKI HIROSE†, MICHIKO ASA†, SEIICHI INAYAMA†, AND SHIGETADA NAKANISHI*

*Institute for Immunology, Kyoto University Faculty of Medicine, Kyoto 606, Japan; and †Pharmaceutical Institute, Keio University School of Medicine, Tokyo 180, Japan

Communicated by Motoo Kimura, October 5, 1982

ABSTRACT By using a mixture of synthetic oligodeoxyribonucleotides as a probe, cloned cDNA sequences specific for low molecular weight (LMW) kininogen have been isolated from a cDNA library of bovine liver mRNA sequences. Nucleotide sequence analyses of cloned cDNA inserts have revealed that bovine liver LMW kininogens are encoded by at least two very similar but distinct mRNAs. The corresponding amino acid sequences show that the LMW kininogen precursors of the two types, composed of 436 and 434 amino acid residues, both contain two internally homologous sequences in the amino-terminal portion between a signal peptide and a bradykinin moiety. The two mRNAs exhibit 15 nucleotide substitutions and 6 nucleotide deletions/additions in their protein-coding regions. The replacement of 13 amino acid residues and the deletions/additions of 2 amino acid residues in the two LMW kininogen precursors are all localized within the internally homologous regions, implying that these regions may be biologically significant in relation to the existence of two LMW kininogens. The nucleotide changes in the two mRNAs also occur in the limited portions that principally encode the internally homologous amino acid sequences. This suggests that the mRNAs are transcribed from the same gene to generate two LMW kininogen precursors differing only in the internally homologous sequences.

Kininogens are precursor proteins of bradykinin that are widely distributed in mammalian blood plasma (1). Bradykinin, which is released by limited proteolysis by kallikreins, has a number of pharmacological actions, including muscle contraction, hypotension, pain generation, and increase of vascular permeability (2–5). Mammalian plasma contains at least two distinct kininogens, designated low molecular weight (LMW) and high molecular weight (HMW) kininogens, which differ in molecular weight and in susceptibility to different kallikreins (1, 6). Another functional role of HMW kininogen, as a cofactor in contact activation of intrinsic coagulation and fibrinolysis, has been demonstrated (for reviews, see refs. 1, 7, and 8). Thus, bradykinin and kininogen play an important role in many pathophysiological conditions, especially in inflammation.

Both LMW and HMW kininogens are single-chain glycoproteins and carry the bradykinin moiety in the interior of the polypeptide chain bridged by a disulfide linkage (8–13). Both kininogens, thus, consist of three domains: an amino-terminal heavy chain, a bradykinin moiety, and a carboxyl-terminal light chain. Complete amino acid sequences of the light chains of both kininogens have been determined. They reveal that there are marked structural differences between the two kininogens (9–13). On the other hand, structural similarity in the heavy chains of the two kininogens has been postulated because of their similar amino acid compositions and their high degree of immunological crossreactivity (11). However, their primary structures remain to be determined.

A number of questions regarding the structures, functions, and biosynthesis of kininogens remain to be elucidated. The construction of bacterial plasmids containing the mRNA sequences for kininogens is a direct approach to analyzing the amino acid sequences, the mRNAs, and the genes for these proteins. We report here the cloning of DNA sequences complementary to the bovine liver mRNAs coding for LMW kininogens. Nucleotide sequence analyses of cloned cDNA inserts have revealed that bovine LMW kininogens are encoded by at least two very similar but distinct mRNAs. The intriguing primary structures of the LMW kininogen precursors and the characteristic nucleotide sequence relationship between the two mRNAs are reported.

MATERIALS AND METHODS

Cloning Procedures. Total RNA was extracted from a bovine liver as described (14) and poly(A)-containing RNA was isolated by subjecting the total RNA extracted to oligo(dT)-cellulose chromatography twice (15). A cDNA library was constructed by the method of Okayama and Berg (16), using 30 μg of poly(A)-containing RNA and 10 μg of the vector-primer DNA; avian myeloblastosis virus reverse transcriptase was provided by J. W. Beard. Escherichia coli χ1776 (17) or HB 101 (18) was used for transformation as described (19) and ampicillin-resistant transformants were screened by hybridization (20) at 36°C with a mixture of 16 oligodeoxyribonucleotides described in Results. The oligodeoxyribonucleotides were synthesized by the modified triester methods (21). Further details of the cloning procedures have been described previously (22). All of the cloning procedures were conducted in accordance with the guidelines for research involving recombinant DNA molecules issued by the Ministry of Education, Science and Culture of Japan.

Analytical Procedures. Restriction endonucleases were obtained from Takara Shuzo (Kyoto, Japan), New England BioLabs, and Bethesda Research Laboratories and were used under conditions described by the vendors. 5'-End-labeling of a mixture of oligodeoxyribonucleotides and restriction fragments was carried out as described (22, 23). Nick-translaction was performed with use of [α-32P]dCTP (22). DNA sequence analysis was carried out by the procedure of Maxam and Gilbert (24). Further details of the sequence analysis procedures have been described previously (23).

Abbreviations: LMW kininogen, low molecular weight kininogen; HMW kininogen, high molecular weight kininogen.
RESULTS
Isolation of cDNA Clones and DNA Sequence Analysis. Our approach for the isolation of cloned DNA sequences specific for kininogens was to screen a library of cDNA clones by hybridization with a mixture of oligodeoxynucleotides complementary to all possible coding sequences for a small portion of the known amino acid sequences of the proteins. As hybridization probes, we used a mixture of 16 oligodeoxynucleotides, 5'-C-C-Y-T-C-N-G-T-Y-T-T-C-A-T (Y = C or T and N = A, G, C, or T), which represents cDNA sequences predicted from the pentapeptide sequence, Met-Lys-Thr-Glu-Gly (excluding the third nucleotide residue of the glycine codon), present in the light chains of both bovine LMW and HMW kininogens (10, 12) (Fig. 1). Eighty-one hybridization-positive clones were isolated from about 50,000 transformants derived from the bovine liver cDNA library. Twelve of them were analyzed with several restriction endonucleases, either individually or in pairs. At least three clones shared several common restriction cleavage sites that corresponded well with those inferred from possible coding sequences for the known amino acid sequence of bovine LMW kininogen. Clone pKG13, which carried the largest cDNA insert of the three, was subjected to nucleotide sequence analysis according to the strategy indicated in Fig. 1.

Clone pKG13 carried a DNA segment corresponding to a large portion of LMW kininogen mRNA but did not contain the entire protein-coding region. Therefore, we rescreened the 81 above-mentioned cDNA clones and also newly screened about 100,000 transformants from the cDNA library by hybridization with the restriction fragments derived from the 5'-terminal portion of clone pKG13 (Fig. 1). Ten additional hybridization-positive clones were isolated and were analyzed with restriction endonucleases. Unexpectedly, several clones exhibited restriction cleavage patterns that could not be accounted for simply by the size differences of the inserts when compared with the insert of clone pKG13. Therefore, the 10 clones were analyzed in more detail and were classified into three groups according to their restriction maps. The first group, type I, which included four clones (e.g., pKG59) whose inserts varied in length from that of clone pKG13, carried cDNA inserts that showed identical restriction maps throughout the regions that overlapped with the insert of clone pKG13. The second group of four clones, type II (e.g., pKG146 and pKG254), also exhibited restriction maps essentially identical to the map of clone pKG13 but lacked some specific restriction sites—namely, Aat I (768), BstNI (781), Hpa II (792), and Dde I (842)—that were present in the cDNA insert of clone pKG13. These results suggested that the two types of cDNA inserts were derived from two similar but distinct mRNAs. Therefore, the largest cDNA inserts in each group of clones (pKG59 and pKG146), as well as another cDNA insert exhibiting the second type of restriction map (clone pKG254), were subjected to nucleotide sequence analysis as shown in Fig. 1. The third group, which included two clones, contained inserts that were almost twice as long as the insert in clone pKG13. Partial sequence analysis of a cDNA insert in the third group revealed that it encoded HMW kininogen (data not shown).

The primary structures of LMW kininogen mRNAs were deduced from the nucleotide sequences determined for the two types of cloned cDNAs (clones pKG13 and pKG59, and clones pKG146 and pKG254). These results are summarized in Fig. 2. A comparison of the nucleotide sequence of the 1,530-base-pair insert (nucleotide residues −96 to 1,434) of clone pKG59 and that of the 1,562-base-pair insert (nucleotide residues −133 to 1,429) of clone pKG146 exhibited 15 nucleotide substitutions and 6 nucleotide deletions/additions in their overlapping nucleotides. Because several samples of each type of clone were found, it is reasonable to assume that they represent two different mRNAs coding for LMW kininogen. This assumption was

Fig. 1. Strategy of determining the sequences of the cDNA inserts in clones pKG13, pKG59, pKG146, and pKG254. For the isolation of clone pKG13, see the text. For the isolation of clone pKG59, those clones which hybridized with the labeled oligodeoxynucleotides were rescreened by hybridization at 60°C with the nick-translated 51P-labeled Eco R I fragment containing nucleotides 224–684 from clone pKG13. Clones pKG146 and pKG254 were isolated from the cDNA library by hybridization at 60°C with the 5'-32P-labeled Eco R I restriction fragment containing nucleotides 167–223 from clone pKG13. The restriction map displays only relevant restriction endonuclease sites, which are identified by numbers indicating the 5'-terminal nucleotide generated by cleavage; for the nucleotide numbers, see Fig. 2. Restriction sites absent from type II of the cDNA insert (clones pKG146 and pKG254) are indicated by asterisks. The sequence corresponding to the coding region is shown by the shadowed box; the solid box indicates the coding region for bradykinin. The sequences used as hybridization probes for selecting clones are indicated by the lines directly beneath the restriction map. The type of cDNA insert of each clone (see the text) is indicated in parentheses after the clone name. Gaps on the bars of clones pKG146 and pKG254 represent the six nucleotide deletions. The poly(dA)-poly(dT) tracts and the poly(dG)-poly(dC) tails are not included in the restriction map. The location of the terminus of pKG254 is approximate. The direction and extent of sequence determinations are shown by horizontal arrows under each clone used; the sites of 5'-end labeling are indicated by short vertical lines at the end of arrows. The slash marks at the end of arrows mean that the site of 5'-end labeling was located on the vector DNA.
Fig. 2. Primary structures of two LMW prekininogen mRNAs. The nucleotide sequences of the mRNAs were deduced from those of the cDNA inserts in the clones given in Fig. 1. According to the nucleotide sequence of LMW prekininogen I mRNA, nucleotide residues are numbered in the 5' to 3' direction, beginning with the first residue of the AUG triplet encoding the initiator methionine, and the nucleotides on the 3' side of residue 1 are indicated by negative numbers. The nucleotide differences of LMW prekininogen II mRNA when compared with LMW prekininogen I mRNA, as well as the nucleotide residues present in the 5' and 3' terminal regions in LMW prekininogen II mRNA, are displayed below the sequence of LMW prekininogen I mRNA; the absence of a nucleotide from the sequence of LMW prekininogen II mRNA indicates that the two mRNA sequences are the same; the presence of a colon indicates a deletion. Because there is a possibility that the 5' extremities of the two mRNAs are not included in these sequences, they are indicated by broken lines. The 3' terminal sequences shown are followed by a poly(A) tract connected with the vector DNA sequence, thus representing the complete sequences of these regions. The predicted amino acid sequence of LMW prekininogen I is displayed above the nucleotide sequence, and the amino acid residues are numbered beginning with the initiator methionine.
Fig. 3. Comparison of amino acid sequences in the two internally homologous regions. Amino acid sequences are described with the standard one-letter notation for amino acid residues. The replacements and deletions of amino acid residues in LMW prekininogen II are shown below the corresponding residues of the amino acid sequence of LMW prekininogen I; a hyphen indicates a deletion. Amino acid residues identical in two homologous regions of LMW prekininogen I are enclosed in boxes. The positions of cysteine residues discussed in the text are indicated by asterisks. For the amino acid numbers, see Fig. 2.

Assignment of Amino Acid Sequences of LMW Kininogen Precursors. The amino acid sequences of two bovine LMW prekininogens, derived from the cDNA sequences, are given in Fig. 2. The sequence of nucleotide residues 1,126 to 1,305 in LMW prekininogen I mRNA (residues 1,200 to 1,299 in LMW prekininogen II mRNA) corresponded precisely to the amino acid sequence of 60 amino acid residues that was previously reported for the carboxyl-terminal portion, including the bradykinin moiety, of bovine LMW kininogen (12). However, an additional codon specifying arginine is present at the carboxyl end of the two LMW prekininogens, and this amino acid is followed by the translational termination codon UAG. Thus, the 3' noncoding regions of LMW prekininogen I and II mRNAs are 126 and 127 nucleotides long (excluding the poly(A) tract), respectively; the additional nucleotide residue of C was found at the end of the 3' noncoding region of LMW prekininogen II mRNA. The sequence sequence A-A-U-A-A-A commonly found near the 3' end of eukaryotic mRNAs (25) is present in LMW prekininogen mRNAs either 18 or 19 nucleotide residues upstream from the poly(A) tract.

The translational initiation site was assigned to the methionine codon AUG at positions 1 to 3 because this is the first AUG triplet that appears downstream from the nonsense codon UAA (residues -6 to -4) found on the reading frame corresponding to the amino acid sequences of the bradykinin and the light chain of LMW kininogen. This assignment seems plausible because the sizes of the cDNA sequences in clones pKG59 and pKG146 are both close to those of the LMW prekininogen mRNAs (approximately 1,700 nucleotides) estimated by RNA blot hybridization analysis (data not shown). The sequence of the first 32 amino acid residues starting with the putative methionine exhibits a feature characteristic of the signal peptide present at the amino-terminal end of secretory proteins (26): it contains many hydrophobic amino acids (11 nonpolar residues including 6 leucines). The signal peptide generally terminates in a residue with a neutral side chain such as alanine, glycine, or serine (27). The amino-terminal amino acid residue of bovine LMW kininogen was reported to be masked, and the first two amino-terminal amino acid residues of bovine HMW kininogen were reported to be <Glu-Glu (<Glu = pyrogulatamic acid) (13). If the masked amino acid residue of LMW kininogens was <Glu, as is the case for HMW kininogen, a possible site for cleavage of the signal peptide of LMW prekininogens seems to be between Ser-22 and Gin-23. Assuming that the cysteine residue at position 10 is eliminated with the signal peptide, the heavy chain and the light chain of both LMW kininogens would contain 17 cysteine residues and one residue, respectively. This finding is consistent with the results obtained by structural analysis of LMW kininogen, which indicate that the two polypeptide chains are linked by a single disulfide bond and all remaining cysteine residues form disulfide bonds within the heavy chain (11).

All nucleotide changes in the overlapping regions of the two LMW prekininogen mRNAs are located in the protein-coding regions, and the 15 nucleotide substitutions result in the replacement of 13 amino acid residues. LMW prekininogens I and II are composed of 436 and 434 amino acid residues, respectively, having calculated molecular weights of 48,428 and 48,149, respectively. A total of either 357 or 355 amino acid residues of the heavy chain and the composition of the heavy chain agree well with the number and the composition of the residues reported from amino acid analysis of the purified heavy chain of LMW kininogen (11).

Inspection of the deduced amino acid sequences of the heavy chains reveals that they contain two internally homologous sequences whose outlines are defined by the distribution of the cysteine residues, as shown in Fig. 3. In the case of LMW prekininogen I, six cysteine residues in the two homologous sequences are present at exactly equivalent positions. Moreover, the homologous regions of 122 amino acids each (amino acid residues 103–224 and 225–346) have 40 positions of identical amino acid residues, corresponding to about 33% identity. It is interesting to note that all amino acid differences in the two LMW prekininogens are located within the internally homologous regions. The amino acid sequence of LMW prekininogen II can also be aligned in the same way, as long as a single gap of two amino acid residues is left open between the second and third cysteine residues. LMW prekininogen II thus exhibits a high degree of internal sequence homology, which is remarkably similar to the sequence homology of LMW prekininogen I. This observation suggests that the internally homologous regions of both LMW prekininogens evolved as a consequence of a series of gene duplications of a primordial gene.

The bradykinin sequence is boxed with a solid line. The replacements and deletions of amino acid residues in LMW prekininogen II are shown above the corresponding residues in the amino acid sequence of LMW prekininogen I; a line indicates a deletion. Note that, due to these deletions, the amino acid numbers from the deletions for LMW prekininogen II differ by two, and the nucleotide residue numbers differ by six, from those numbers for LMW prekininogen I.
DISCUSSION

The internal duplications of the amino acid sequences in the heavy chains raise interesting questions concerning the biological role of the characteristic structure of the duplicated amino acid sequences. Although no information on the linkage pattern of the disulfide bridges in the heavy chain is available, the extensive sequence homology of the internally homologous regions, as well as the equivalent locations of their cysteine residues of six each, may result in similar secondary and tertiary structures formed by each internally homologous sequence. Thus, the homologous regions may serve as the structural and functional domain in the heavy chain. It is to be noted in this context that there is structural similarity between kininogens and two zymogens, prothrombin and plasminogen. Release of bradykinin from the kinogens, as well as conversion of the two zymogens in their active forms, occurs as a result of limited proteolysis with the same type of enzymes, namely serine proteases (29). Furthermore, the amino-terminal portions of prothrombin and plasminogen contain two and five internally homologous structures, respectively, called kringle, which are formed by three disulfide bonds located at equivalent positions (29, 30). It has been shown that the kringle structures provide binding sites for specific structures with which the particular proteolytic enzyme interacts during activation or, in the case of plasminogen, perhaps after activation so that its proteolytic effect is prevented from spreading indiscriminately (29–31). Thus, it is tempting to speculate that the internally homologous structures in LMW kininogens are involved in a specific interaction with kalikrein that facilitates the limited and highly specific proteolytic activity characteristic of kalikrein. Furthermore, the fact that all amino acid differences are localized within the internally homologous regions in LMW prekininogens suggests that these regions are biologically significant in relation to the existence of two LMW prekininogens. It has recently been reported that there are several kalikrein-like proteases that have an extensive sequence homology with one another (32). Thus, it is an intriguing hypothesis that the homologous regions of both LMW kininogens define interactions of the LMW kininogens with specific but different types of kalikrein.

Because there was previously no explicit report concerning the heterogeneity in the protein structure of LMW kininogen, the finding of the two mRNAs encoding slightly different LMW prekininogen sequences was unexpected. This finding introduces a new question regarding the gene(s) for LMW prekininogens. Because the mRNA was extracted from one bovine liver, the different sequences observed represent the mRNA species derived from a single individual. We noted that all but one of the nucleotide changes in the overlapping regions of the two mRNAs are located within the sequences encoding the internally homologous amino acid sequences. This variability is a marked contrast to the nucleotide sequence homology in other portions of the two mRNAs. For example, the nucleotide sequences of the 499 residues specifying the region that encodes the carboxyl-terminal portion of the protein and the 3′ noncoding region are exactly identical. Therefore, if the two mRNAs are transcribed either from allelic forms of the gene or from two different genes, it must be postulated that some type of evolutionary mechanism exists whereby nucleotide changes are distributed in only certain portions of the gene(s). An alternative possibility is that the two LMW prekininogen mRNAs are transcribed from a single gene. Under a gene organization in which the dissimilar middle portions of the two mRNA sequences are specified by separate genomic segments within the same gene, different splicing events would ultimately generate these two overlapping but not identical mRNAs. This mechanism could be very useful to produce two LMW prekininogens that must differ, for some physiological reason, only in the internally homologous regions.

We thank Drs. S. Iwanaga, H. Kato, and S. Numa for helpful discussions. This investigation was supported in part by research grants from the Ministry of Education, Science and Culture of Japan, the Institute of Physical and Chemical Research, and the Mitsubishi Foundation.