Communication

Cloning and Characterization of a Novel Human Lysyl Hydroxylase Isoform Highly Expressed in Pancreas and Muscle*

(Received for publication, December 31, 1996, and in revised form, January 10, 1997)

Minna Valtavaara, Hiinni Papponen, Anna-Maria Pirttilä, Kalervo Hiltunen, Heli Helander, and Raii Myllyläi

From the Biocenter and Department of Biochemistry, University of Oulu, FIN-90570 Oulu, Finland

We report the isolation and characterization of cdNA clones for a novel isofom of lysyl hydroxylase (lysyl hydroxylase 2), a posttranslational enzyme of collagen biosynthesis. The open reading frame predicted a protein of 737 amino acids, including an amino-terminal signal peptide. The amino acid sequence has overall similarity of over 75% to the lysyl hydroxylase (lysyl hydroxylase 1) characterized earlier. This similarity is even higher in the carboxyl-terminal end of the molecules. Lysyl hydroxylase 2 contains nine cysteine residues, which are conserved in lysyl hydroxylase 1. Furthermore, the conserved histidines and aspartate residues required for lysyl hydroxylase activity are present in the sequence. Northern analysis identified a transcript of 4.2 kilobases, which was highly expressed in pancreas and muscle tissues. Expression of cdNA in insect cells using a baculovirus vector yielded proteins with lysyl hydroxylase activity and an antiserum against a synthetic peptide of the deduced amino acid sequence recognized proteins with molecular weights of 88 and 97 kDa in homogenates of the transfected cells.

Collagens are the most abundant proteins in the human body, and they are found essentially in all tissues. To date, 19 different collagen types have been identified (1–4). The biosynthesis of collagens is characterized by several posttranslational modifications, one of which is hydroxylation of lysyl residues. Hydroxylysine occurs in the triplets within the helical region of the collagen molecule. Hydroxylysyl residues have an important role in the structure and stability of collagens. The hydroxy groups participate in the formation of intermolecular cross-links and serve as sites of attachment for carbohydrate units that are unique to collagens (1–4, 6).

Lysyl hydroxylase (EC 1.14.11.4) catalyzes the hydroxylation of lysyl residues in collagens (1). The enzyme requires Fe(II), 2-oxoglutarate, O2, and ascorbate in the reaction. The active enzyme is a homodimer consisting of subunits with a molecular weight of about 85,000 (1) and is constitutively expressed in a variety of tissues (7, 8). The complete cdNA-derived amino acid sequence has been reported for the enzyme from chick (9), human (10, 11), and rat (12). The enzyme has also been characterized at the genomic level (7), and the first mutations in the lysyl hydroxylase gene have been characterized in patients with type VI of the Ehlers-Danlos syndrome (EDSVI) (13–17).

We report here the cloning and characterization of a novel human lysyl hydroxylase (lysyl hydroxylase 2), which is highly expressed in pancreas and muscle tissues. The similarity in the amino acid sequence is about 75% when compared with the human and chicken lysyl hydroxylase characterized earlier (lysyl hydroxylase 1).

MATERIALS AND METHODS

PCR—PS1-PS2 fragment (PS1, 5'-TTTCTATTCTCTTGAGTC-3'; PS2, 5'-TCTAGTCAAGGCGAA-3'), sequences homologous to the lysyl hydroxylase cdNA sequence (obtained from the data base of expressed sequence tags, T11367, T18568) was generated by PCR from oligo(dT)-primed human kidney cdNA. The amplification was performed using Tth DNA polymerase (Amplitaq, Perkin-Elmer) in the presence of 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 1.5 mM MgCl2, and 0.01% (w/v) gelatin (94 °C, 2 min; 94 °C, 40 s; 43 °C, 40 s; 72 °C, 2 min; 30 cycles; 72 °C, 10 min). The subcloned fragment was sequenced and used as a probe for library screening.

The following oligonucleotide pairs and conditions were used in the amplifications to clone the 3' end of the novel cdNA. API (5'-CCATCAACTAGACTCTACATAGCCGC-3'), Marathon kit, Clontech) and PS22 (5'-GTACAATTGCTCTATTGAGTCA-3'), 94 °C, 10 min; 94 °C, 40 s; 55 °C, 40 s; 72 °C, 2 min; 30 cycles; 72 °C, 10 min. For cloning of the 5' end, the following oligonucleotide pairs were used: AP2 (5'-ACTACTGTTAGGGCCTGACCGGC-3'), Marathon kit, Clontech) and PS38 (5'-TTGGTAGAACATACGGTTGACATA-3'), 95 °C, 5 min; 95 °C, 40 s; 61 °C, 40 s; 72 °C, 2 min; 30 cycles; 72 °C, 5 min; PS39 (5'-CCATCTCCTCTTCCTCTCC-3') and oligo(dT) anchor primer (5'-GACCCGCTAGTGGTGCAGTCTTTTTTTTTTTTTTTTTTTTT-3'), 5' RACE kit, Boehringer Mannheim), 94 °C, 2 min; 94 °C, 20 s; 60 °C, 30 s; 72 °C, 2 min; 35 cycles; 72 °C, 10 min; PS41 (5'-TCATCCTCATCCTCTCCT-3') and PCR anchor primer (5'-GACCCGCTAGTGGTGCAGTCTTTTTTTTTTTTTTTTTTTTT-3'), 5' RACE kit, Boehringer Mannheim), 94 °C, 2 min; 94 °C, 20 s; 57 °C, 30 s; 72 °C, 2 min; 35 cycles; 72 °C, 10 min; PS46 (5'-CCTGCTCCGCCCAGACAGGATTCTC-3') and PCR anchor primer, 94 °C, 2 min; 94 °C, 20 s; 60 °C, 30 s; 72 °C, 2 min; 35 cycles; 72 °C, 10 min.

cDNA Cloning—Fetal and adult human kidney (agt11, Clontech) and adult human pancreas (agt10, Clontech) cdNA libraries were screened using standard protocols (18). Hybridization was carried out for 16 h in 5 × SSPE (1 × SSPE = 0.15 M NaCl, 0.01 M Na2HPO4, H2O; 0.001 M EDTA, pH 7.4), 5 × Denhardt's, 50% formamide, 0.01% SDS, 100 ng/ml denatured herring DNA at 42 °C. The filters were washed twice in 2 × SSC, 0.5% SDS and once in 1 × SSC, 0.1% SDS for 15 min at room temperature.

* This work was supported by grants from Sigrid Juselius Foundation and the Research Councils for Natural Sciences and Medicine in the Academy of Finland. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Dept. of Biochemistry, University of Oulu, Linnanmaa, FIN-90570 Oulu, Finland. Fax: 358-8-5531141.
The 3' and 5' end of the novel cDNA were obtained using the Marathon cDNA amplification kit (Clontech) and 5'/3' RACE kit (Boehringer Mannheim).

To clone the 3' end of noncoding region of the novel cDNA, double-stranded cDNA was prepared using poly(A) RNA from human kidney and oligo(dT) primer (Marathon kit) which contained two degenerate nucleotides following the poly(T) sequence. The Marathon cDNA Adapter was ligated to the blunt ended double-stranded cDNA and amplification was carried out by PCR using an adaptor specific primer and a gene-specific primer obtained from the sequence of clone 144.

To clone the 5' end of the novel cDNA, double-stranded cDNA was prepared using a sequence-specific primer. The Marathon cDNA adapter was ligated to the blunt ended double-stranded cDNA, and amplification was carried out using an adaptor-specific primer paired with a primer obtained from the sequence of clone 144. Using 5'/3' RACE kit, a single-stranded, tailed cDNA was converted to a double-stranded molecule by PCR using an oligo(dT) anchor primer paired with a sequence-specific oligonucleotide. Amplification was carried out using a PCR anchored primer paired with a sequence-specific primer.

Sequencing—DNA sequencing was performed by standard dideoxynucleotide sequencing using T7 sequencing kit (Pharmacia Biotech Inc.), D'Taq cycle sequencing kit (U.S. Biochemical Corp.) and PRISM™ AmpliTaq FS dye terminator cycle sequencing kit (Perkin-Elmer). Part of the sequencing was carried out manually and part using an Abi Prism 377 DNA sequencer (Perkin-Elmer). All clones were sequenced, despite the overlap in the sequence of the clones. The sequence of the ends of the molecule was confirmed by sequencing DNA fragments obtained by different oligonucleotide pairs in 5'/3' RACE.

Northern Blot Analysis—A human Multiple Northern blot containing poly(A) RNA from different tissues (Clontech) was hybridized for 16 h in 5X SSPE, 10X Denhardt's, 50% formamide, 2% SDS, 100 μg/ml denatured herring DNA at 42°C using radioactively labeled 144 clone as a probe. The blot was washed in 2X SSC, 0.05% SDS at room temperature for 40 min and then at 50°C for 40 min.

Expression of cDNA in Insect Cells—Expression of the novel cDNA was carried out by baculovirus transfer vector (18) in the BAC-TO-BAC™ Expression system (Life Technologies, Inc.). PCR was used to generate two different constructs using human kidney cDNA as a template. Construct 1 contained the nucleotides from 1 to 2272 and construct 2 the nucleotides from 27 to 2272. The PCR products were confirmed by sequencing. Insect cells were harvested 48 or 72 h after infection according to protocol described for human lysyl hydroxylase (19).

Other Assays—Western blot analysis was carried out using a polyclonal antibody, produced in chicken, against a synthetic peptide of NPRTLKILIEQNRKI (amino acids 399–413 of lysyl hydroxylase 2). The homogenates of baculovirus-infected cells were fractionated in reducing conditions by SDS/10% polyacrylamide gel electrophoresis, blotted onto an Immobilon membrane (Millipore) and incubated with the antibodies against the synthetic peptide. Bound antibodies were visualized using the ECL detection system (Amersham Life Science, Inc.) and x-ray film (Eastman Kodak Co.). Lysyl hydroxylase activity was assayed by a method based on the hydroxylation-coupled decarboxylation of 2-oxo-[1-14C]glutarate (20), synthetic peptide ARGIKGIRGFSG or IKGIKGIKG was used as substrate.

RESULTS AND DISCUSSION

Isolation and Characterization of cDNA Clones—The first cDNA clone (PS1-PS2) was generated by amplification of oli-
TABLE I
Lysoyl hydroxylase activity in Nonidet P-40 supernatant of insect cells expressing the novel cDNA

| cDNA expressed          | Lysoyl hydroxylase activity<sup>a</sup> | dpm/plate<sup>b</sup> | %   |
|-------------------------|-----------------------------------------|-----------------------|-----|
| Noninfected cells       | 1217 ± 218                              | 100 ± 18              |     |
| Bacmid DNA infected control | 1429 ± 138                          | 117 ± 11              |     |
| Construct 1             | 2085 ± 502                              | 170 ± 41              |     |
| Construct 2             | 2866 ± 550                              | 235 ± 45              |     |

<sup>a</sup> Values are obtained using IKGIKGIKGK as a substrate for lysoyl hydroxylase. The data are given as mean ± S.D. of seven experiments for both constructs. Control values are based on five measurements for noninfected cells and two measurements for bacmid DNA-infected cells.

<sup>b</sup> One plate contained about 10 × 10<sup>6</sup> cells. All values are corrected further for minor differences in cell numbers of the plates.

Lysoyl hydroxylase activity in Nonidet P-40 supernatant of insect cells expressing the novel cDNA—

god(T)-primed human kidney cDNA using oligonucleotides PS1 and PS2 as primers (Fig. 1). The amplification yielded a fragment of about 300 nucleotides in length. Human kidney and pancreas cDNA libraries were then screened using PS1-PS2 clone as a probe. One positive recombinant was identified in the pancreas (λ4) and one in the kidney (λ6) library (Fig. 1). Clone λ44 was used as a probe to obtain four additional overlapping clones (Fig. 1, λ5, λ25, λ48, and λ86). The 5′ and 3′ ends of the novel cDNA were obtained using 5′ and 3′ RACE.

Nucleotide and Derived Amino Acid Sequences of the cDNA—The cDNA clones (Fig. 1) encode a polypeptide of 737 amino acids, which starts with a codon for methionine and contains seven potential attachment sites for asparagine-linked oligosaccharides, two of them have identical location in the sequence for human lysyl hydroxylase. The nine cysteine residues contained in the molecule. Examples of particularly variable regions are a 13-amino acid sequence (amino acids 283–295) and an 18-amino acid sequence (amino acids 340–357) with an 11% identity to lysyl hydroxylase. The 50 amino acids from amino terminus of the molecule have a 56% identity to human lysyl hydroxylase. The nine cysteine residues conserved in lysyl hydroxylase amino acid sequences are also conserved in the novel polypeptide. The amino acid sequence contains seven potential attachment sites for asparagine-linked oligosaccharides, two of them have identical location in the sequence for human lysyl hydroxylase.

The carboxyl-terminal end of lysyl hydroxylase 1 is especially well conserved between chicken and human (10) having an identity of over 90% for the last 139 amino acids of the molecule. This finding suggests that this region of the molecule is important for the function of the enzyme. A search for conserved residues within the sequences of 2-oxoglutarate dioxygenases and a related dioxygenase, isopenicillin N synthase, suggested that two histidine-containing motifs in the carboxyl terminus, residues 656 and 708 in the human lysyl hydroxylase sequence, may function as ferrous binding ligands in the catalytic mechanism common to all 2-oxoglutarate dioxygenases (21, 22). Site-directed mutagenesis of prolyl 4-hydroxylase (23) and lysyl hydroxylase (19), the enzymes belonging to the group of 2-oxoglutarate-dependent dioxygenases, demonstrated the importance of these histidine residues in catalytic activity of the enzymes. Mutation analysis revealed three other functionally important histidines in the carboxy-terminal portion of lysyl hydroxylase polypeptide (19). It is remarkable that all these histidine residues are also conserved in the novel molecule (amino acids His-666, His-667, His-710, His-716, and His-718). Studies on human lysyl hydroxylase furthermore indicated that the mutation of asparagine residues in the carboxy-terminal region of the molecule cause dramatic inactivation of the enzyme (19); two of these residues (amino acids Asp-668 and Asp-684) can be also found in the carboxy-terminal region of the novel polypeptide. Asparagine-linked glycosylation has been shown to be a requirement for maximal lysyl hydroxylase activity (19, 24). One can speculate that the requirement holds true also for the novel protein, because (i) the sequence contains many putative glycosylation sites, (ii) two putative glycosylation sites are highly conserved; at least one of these is glycosylated in lysyl hydroxylase expressed in insect cells and both asparagines (corresponding to Asn-209 and Asn-696 of the novel sequence) are required for lysyl hydroxylase activity (19).

A Novel Polypeptide with Lysyl Hydroxylase Activity—To determine whether the novel polypeptide has lysyl hydroxylase activity, a recombinant baculovirus encoding for the novel cDNA was generated in insect cells (Table I). Two constructs were prepared for the experiments, one containing the whole coding region, including the signal peptide (construct 1) and the other containing the coding region but lacking the first nine amino acids of the signal peptide (construct 2). Recombinant proteins were harvested by homogenization of the cells in a buffer containing Nonidet P-40, and the insoluble pellet was further homogenized in a buffer containing glycerol (19). Activity measurements of the supernatants indicate that majority of lysyl hydroxylase activity was found in the Nonidet P-40 fraction. Both peptide substrates were hydroxylated by the recombinant proteins.
Immunological Studies—An antiserum was prepared against a synthetic peptide derived from the cDNA sequences of lysyl hydroxylase 2. The recombinant proteins produced by the constructs were analyzed by SDS-polyacrylamide gel electrophoresis and then immunostained using the peptide antibodies for detection. The peptide antibodies recognized protein bands with molecular weights of 88,000 and 97,000 in cells transfected with the constructs for lysyl hydroxylase 2 (not shown). The lower molecular weight band corresponds exactly to the size of an immunostained band found in kidney tissue and kidney cells (not shown). It is not known, however, if both molecular weight 88,000 and 97,000 proteins have lysyl hydroxylase activity. As indicated above, lysyl hydroxylase 2 contains many potential sites for glycosylation, and it is possible that the two forms on the immunoblot are due to variations in the glycosylation of proteins in insect cells.

Expression of the Novel Lysyl Hydroxylase in Various Tissues—Expression of the novel gene in various tissues was determined by Northern hybridization. A single band of 4.2 kilobases was obtained in the hybridization from heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas. The data indicate that the expression of the gene is highly regulated. A prominent hybridization signal was present in muscle tissue such as heart and placenta, as well as pancreas.

Conclusions—A novel isoform (lysyl hydroxylase 2) for lysyl hydroxylase has been described in this study. The molecule has a high identity to the lysyl hydroxylase characterized earlier (lysyl hydroxylase 1, Refs. 9–12), but the novel molecule clearly is a product of a different gene. This finding confirms the hypothesis of existence of isoforms for lysyl hydroxylase based on observations that hydroxylation of telopeptidyl lysyl residues may be under separate enzymic control distinct from that active toward lysyl residues within the helical regions of type I collagen (5, 25). The discovery of isoforms of lysyl hydroxylase is highly significant, because it may explain the great variation of lysine hydroxylation in collagen of different tissues in patients with EDSVI (26, 27), a disease resulting from the gene defect in lysyl hydroxylase 1. The isoforms are also in agreement with the finding that lysines in collagen types II, IV, and V in EDSVI patients were hydroxylated normally (27). Furthermore, the residual activity of EDSVI cells preferentially directed toward type IV collagen (28) suggests the presence of isoforms for lysyl hydroxylase. The presence of isoforms may also explain our previous observations that cells from an EDSVI patient producing a truncated form of lysyl hydroxylase 1, which lacks the highly conserved carboxyl-terminal portion of the molecule, nevertheless have detectable lysyl hydroxylase activity (13).

Acknowledgment—Dr. Timo Hautala, Department of Internal Medicine, University of Iowa, is acknowledged for his help in the computer search.

REFERENCES
1. Kivirikko, K. I., Myllyla, R., and Pihlajaniemi, T. (1991) in Post-translational Modifications of Proteins (Crabbe, M. J. C., and Harding, J., eds) pp. 1–51, CRC Press, Boca Raton, FL
2. Kielty, C. M., Hopkinson, I., and Grant, M. E. (1993) in Connective Tissue and Its Heritable Disorders (Royce, P. M., and Steinmann, B., eds) pp. 103–147, Wiley-Liss, Inc., New York
3. Kivirikko, K. I. (1993) Annu. Med. 25, 113–126
4. Prockop, D. J., and Kivirikko, K. I. (1985) Annu. Rev. Biochem. 64, 403–434
5. Barnes, M. J., Constable, B. J., Morton, L. F., and Royce, P. M. (1974) Biochem. J. 139, 461–468
6. Robins, S. P. (1982) in Collagen in Health and Disease (Weiss, M. I. V., and Jayson, M. I. V., eds) pp. 169–178, Churchill Livingstone, Edinburgh, UK
7. Heikkinen, J., Hautala, T., Kivirikko, K. I., and Myllyla, R. (1994) Genomics 24, 464–471
8. Yeowell, H. N., Ha, V., Clark, L. C., Marshall, M. K., and Pinnell, S. R. (1994) J. Invest. Dermatol. 102, 382–384
9. Myllyla, R., Pihlajaniemi, T., Pajunen, L., Turpeenniemi-Hujanen, T., and Kivirikko, K. I. (1991) J. Biol. Chem. 266, 2805–2810
10. Hautala, T., Byers, M. G., Eddy, R. L., Shows, T. B., Kivirikko, K. I., and Myllyla, R. (1992) Genomics 13, 62–69
11. Yeowell, H. N., Ha, V., Walker, L. C., Murad, S., and Pinnell, S. R. (1992) J. Invest. Dermatol. 99, 864–869
12. Amstrong, L. C., and Last, J. A. (1995) Biochim. Biophys. Acta 1264, 93–102
13. Hyland, J., Ala-Kokko, L., Royce, P., Steinmann, B., Kivirikko, K. I., and Myllyla, R. (1992) Nat. Genet. 2, 228–231
14. Hautala, T., Heikkinen, J., Kivirikko, K. I., and Myllyla, R. (1993) Genomics 15, 399–404
15. Pousi, B., Hautala, T., Heikkinen, J., Kivirikko, K. I., and Myllyla, R. (1994) Am. J. Hum. Genet. 55, 899–906
16. Ha, V. T., Marshall, M. K., Elias, L. J., Pinnell, S. R., and Yeowell, H. N. (1994) J. Clin. Invest. 93, 1716–1721
17. Heikkinen, J., Toppinen, T., Yeowell, H., Krieg, T., Steinmann, B., Kivirikko, K. I., and Myllyla, R. (1997) Am. J. Hum. Genet. 60, 48–56
18. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., and Struhl, K. (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York
19. Pirskanen, A., Kaimio, A.-M., Myllyla, R., and Kivirikko, K. I. (1996) J. Biol. Chem. 271, 9398–9402
20. Kivirikko, K. I., and Myllyla, R. (1982) Methods Enzymol. 82, 245–304
21. Myllyla, R., Gunzler, V., Kivirikko, K. I., and Kasaka, D. (1992) Biochem. J. 286, 923–927
22. Roach, P. L., Clifton, I. J., Fulop, V., Harlos, K., Barton, G. J., Hadju, J., Andersson, I., Schofield, C. J., and Baldwin, J. E. (1995) Nature 375, 700–704
23. Lamberg, A., Pihlajaniemi, T., and Kivirikko, K. I. (1995) J. Biol. Chem. 270, 9926–9931
24. Myllyla, R., Pajunen, L., and Kivirikko, K. I. (1988) Biochem. J. 253, 489–496
25. Royce, P. M., and Barnes, M. J. (1988) Biochim. Biophys. Acta 930, 475–480
26. Pinnell, S. R., Krane, S. M., Kenzora, J. E., and Glimcher, M. D. (1972) N. Engl. J. Med. 286, 1013–1020
27. Ihme, A., Krieg, T., Nerlich, A., Feldmann, U., Rauterberg, J., Glanville, R. W., Edel, G., and Muller, P. K. (1984) J. Invest. Dermatol. 83, 161–165
28. Risteli, L., Risteli, J., Ihme, A., Krieg, T., and Muller, P. (1980) Biochem. Biophys. Res. Commun. 96, 1778–1784