Molecular cloning and expression patterns of mouse cartilage oligomeric matrix protein gene

J. R. Kipnes, L. Xu, F. Han, R. Rallapalli, S. Jimenez*, D. J. Hall, R. S. Tuan and Y. Li
Department of Orthopaedic Surgery and *Division of Rheumatology, Department of Medicine, Thomas Jefferson University, Philadelphia, PA 19107, U.S.A.

Summary

Objective: To develop transgenic mice harboring mutations in the COMP gene as animal models for pseudoachondroplasia (PSACH) and multiple epiphyseal dysplasia (MED), autosomal dominant disorders characterized by early onset osteoarthritis and epiphyseal abnormalities. As a first step in generating a mouse model for COMP mutations, we have cloned the cDNA of mouse COMP and examined its tissue expression pattern.

Design: Total mRNA was isolated from the skeletal tissues of newborn C57BL/6j mice and used as a template for oligo(dT) first-strand cDNA synthesis. The cDNA was used for PCR amplification of COMP using three oligonucleotide primer pairs designed from the published rat COMP cDNA sequence. Nested PCR was used to complete the sequence between the amplified fragments. The entire cDNA was sequenced and the expression pattern of the corresponding transcripts examined by Northern hybridizations.

Results: A full-length COMP cDNA was isolated. Analysis showed that the entire translated region of the mouse COMP gene is 2268 bp and the derived amino acid sequence shows 90% homology to human COMP. Of eight adult mouse non-cartilage tissues tested, COMP expression was detected only in testis.

Key words: Cartilage, COMP, Osteoarthritis, MED, PSACH.

Introduction

Cartilage oligomeric matrix protein (COMP) is an N-terminal disulfide-linked, 524 kd homopentameric glycoprotein found primarily surrounding chondrocytes in the cartilage extracellular matrix.1 COMP is a member of the thrombospondin family,2,3,4 whose members all contain a series of polypeptide domains, termed type II-repeats, that are homologous to epidermal growth factor (EGF), followed by a series of repeats, termed type III-repeats, that contain an EF-hand Ca ++-binding sequence.

Pseudoachondroplasia (PSACH) and multiple epiphyseal dysplasia (MED) are autosomal dominant diseases characterized primarily by short-limb dwarfism apparent in early childhood.5 PSACH is also associated with deformities of the epiphyseal centers in the knees, hands, shoulders, and hips, and with vertebral abnormalities that result in scoliosis. MED is a milder dysplasia of the epiphy- ses of peripheral joints. In both disorders, degeneration of the articular cartilage leads to the early onset of osteoarthritis. MED and PSACH are members of the same bone dysplasia family and have been mapped to a common region on chromosome 19 that corresponds to the human COMP gene locus.6–8 Subsequent studies have shown that many patients with PSACH and MED have mutations in the type III-repeats of the COMP gene.9–16

The Ca ++-binding type III-repeats of COMP have been suggested to play an important role in maintaining the integrity of the cartilage matrix, since many of the COMP gene mutations responsible for PSACH and MED are either deletions or substitutions of the conserved aspartate residues in the type III-repeats. Ca ++-binding activity is likely to be compromised in the mutated COMP, although the exact biological function of COMP and the contribution of the mutated COMP gene products to the chondrodysplasia seen in PSACH and MED patients are unknown.

To understand the mechanisms underlying the phenotype caused by COMP mutations, it is critical to generate transgenic mouse models to study how these mutations cause the structural alterations in cartilage, which lead to chondrodysplasia. As a first step to generating these mouse models, we report here the cloning of mouse COMP cDNA, analysis of its coding sequence, and its tissue expression pattern.
Fig. 1. (A) Map of COMP clones obtained using RT-PCR of mouse mRNA. The clones produced from primers derived from the published rat sequence are shown in gray. The clones produced from primers derived from the mouse sequence are shown in black. White represents the portions of the genomic clones, E27 and E55, which contain the 5′-end regulatory region and the 3′ end, respectively. (B) Nucleotide sequence and corresponding amino acid sequence of mouse COMP. The sites for potential N-linked glycosylation are boxed and the RGD sequence is underlined.
Materials and methods

CLONING OF MOUSE COMP cDNA BY REVERSE TRANSCRIPTION POLYMERASE CHAIN REACTION

Total mRNA was isolated from the skeleton of newborn C57BL/6j mice using the Rex Total RNA Extraction Kit (USB), and the PolyA Tract Kit (Promega) was used as a template for oligo(dT)-primed first-strand cDNA synthesis using the SUPERScript PreAmplication System (Gibco BRL). Following removal of the mRNA template with RNaseH, the cDNA was used for PCR amplification of mouse COMP using three oligonucleotide primer pairs (IDT), JK1 (Forward: ACGGTGATGGAATGTGACGC, Reverse: TGGCAGCCTGACCTGTGGTC), JK2 (Forward: CAACTCGATAAAAGACAAGT, Reverse: ATCTTGCTAATACCTGTGC), and JK3 (Forward: GCCTGGCTGTGGTTACAGGCG, Reverse: GGAGAAGGAGAACACCAGCCGG) designed from the published rat COMP cDNA sequence (GenBank No. X72914). The PCR mixture contained cDNA template, 1 μM of each primer, 0.2 mM of each dNTP (Perkin-Elmer), and Taq DNA Polymerase (Boehringer-Mannheim).

The RT-PCR generated three fragments of mouse COMP cDNA, with sequence gaps between the clones. To recover the two gaps, we used nested PCR with four additional primer pairs (IDT), JK4 (Forward: AACTCG GACAGTGATAACCTGG, Reverse: ATGGCACGCTGAAAGCTGGAG)/JK5 (Forward: AAGATACACACCTGGATG GCC, Reverse: GTCTTGTGAGCAAAGATGAGG) and JK6 (Forward: TGCCAGCTGTAATAGTGRT, Reverse: ATCCCCAACCACCACCATAC/CJ7 (Forward: ATCAA TACCAGCCCTGGCTTC, Reverse: TCACTCTGCTAACGTTGGG) derived from the sequence of the three mouse cDNA clones. All PCR products were purified using Centricon-100 (Centricon) before sequence determination.

The 242-nucleotide sequence at the 5’-end of the translated region, the 482 nucleotide sequence upstream from the ATG start codon, and the 61-nucleotide sequence at the 3’-end of the translated region of mouse COMP were derived using a mouse COMP genomic clone. This clone was cut with Eco RI and subcloned in two pieces, E27, derived using a mouse COMP genomic clone. This clone containing the 5’-end of the trans- each strand cDNA synthesis (sense: ACGGT GATGGAATGTGACGC; antisense: CGTCGGTCGAGTG CATACCTGTGCTTC, Reverse: TCACTCTGCTAACGTTGGG) derived from the sequence of the three mouse cDNA clones. All PCR products were purified using Centricon-100 (Centricon) before sequence determination.

The 242-nucleotide sequence at the 5’-end of the translated region, the 482 nucleotide sequence upstream from the ATG start codon, and the 61-nucleotide sequence at the 3’-end of the translated region of mouse COMP were derived using a mouse COMP genomic clone. This clone was cut with Eco RI and subcloned in two pieces, E27, containing the 5’-end of the clone and E55, containing the 3’-end of the clone. This genomic clone was initially identified from a mouse Phage Artificial Chromosome (PAC) library (Research Genetics) by PCR using primers designed from the published rat sequence (sense: ACGGT GATGGAATGTGACGC; antisense: CGTCGGTCGAGTG CATACCTGTGCTTC, Reverse: TCACTCTGCTAACGTTGGG) derived from the sequence of the three mouse cDNA clones. All PCR products were purified using Centricon-100 (Centricon) before sequence determination.

All DNA sequencing was performed by the institutional sequencing facility using the dideoxynucleotide method.

NORTHERN HYBRIDIZATION ANALYSIS OF COMP EXPRESSION

Total RNA was extracted from adult mouse cartilage and a multi-tissue blot (Clontech; poly-A+ RNA from heart, brain, spleen, lung, liver, skeletal muscle, kidney, and testis) was hybridized with a random-primed [32P]-labeled, 700 bp fragment of mouse COMP cDNA (bp 1464–2097) generated by nested PCR using JK4 and JK5 primers.

Results and discussion

NUCLEOTIDE AND AMINO ACID SEQUENCE OF MOUSE COMP

The five cDNA fragments generated by PCR and the partial DNA fragments of E27 and E55 shown in Fig. 1A were used to determine the complete nucleotide sequence of mouse COMP cDNA, including 482 nucleotides upstream of the ATG start codon (Fig. 1B). The translated sequence of mouse COMP is 2268 bp corresponding to 756 amino acids (Fig. 1B).

The mouse and human (GenBank No. L32137) COMP coding sequences are similar in size, with the former being six nucleotides shorter, and in nucleotide homology (86%) and amino acid homology (90%). Two types of repeating sequences, type II-repeats and type III-repeats, are found within mouse COMP. There are four type II-repeats with 86% nucleotide identity to those in human COMP, and conservation of all 24 of the cysteine residues required for proper folding. There are eight contiguous Ca2+-binding type III-repeats spanning from Gly-265 to Ala-522. The mouse and human COMP type III-repeats have 91% homology over the entire domain with many conserved aspartate-residues, suggesting the importance of the Ca2+-binding region to the function of COMP. Also conserved are two cysteine residues in each type III repeat, suggesting their structural importance to the function and activity of the molecule. Thus, a mutation in one of these residues may be responsible for the PSACH and MED phenotype. Interestingly, the third type III repeat of mouse COMP has an Arg-Gly-Asp (RGD) integrin-binding motif, which is also present in human COMP.

There is also conservation of N-linked glycosylation sites between mouse and human COMP. The three potential sites for N-linked glycosylation at Asp-119, -142, and -740 of mouse COMP are the same as those found in the human COMP sequence.

TISSUE EXPRESSION PATTERNS OF MOUSE COMP

Northern analysis of eight adult non-cartilage tissues, heart, brain, spleen, lung, liver, skeletal muscle, kidney, and testis, revealed an intense 2.6 kb transcript, comparable in size to that found in cartilage (Fig. 2). Expression of COMP in the testis has not been previously described and there is no documentation of testicular or reproductive abnormalities in MED and PSACH patients. Whether a high level of tissue-specific COMP expression is related to testicular function remains to be studied.

In conclusion, availability of the mouse COMP cDNA should enable us to generate transgenic mice harboring mutant COMP. Given the structural homology between
human and mouse COMP, these animals should serve as useful experimental models to examine the functional role of COMP and the mechanisms by which COMP mutations result in MED and PSACH phenotypes.

Acknowledgments
This study was supported by NIH grant AR39740.

References