Volume 282, number 2, 393–396 FEBS 09702 May 1991

ADONIS 001457939100409R

Cloning of human α1(X) collagen DNA and localization of the COL10A1 gene to the q21–q22 region of human chromosome 6

Suneel Apte[1], Marie-Genevieve Mattei[2] and Bjorn R. Olsen[1]

[1] *Department of Anatomy and Cellular Biology, Harvard Medical School, Boston, MA 02115, USA and* [2] *Hôpital d'Enfants, Groupe Hospitalier de la Timone, 13385 Marseille, Cedex 5, France*

Received 11 March 1991; revised version received 14 March 1991

With consensus primers based upon the nucleotide sequence of the chicken α1(X) collagen gene, we have used PCR with human genomic DNA as template to isolate a 289 bp fragment coding for part of the carboxyl non-triple helical domain of the human α1(X) gene. We have demonstrated the presence of the sequence of the PCR clone within the human genome by partial sequence analysis of a 1 kb *Hin*dIII genomic DNA fragment that hybridized with the PCR clone. Furthermore, using the PCR clone as a probe for in situ hybridization of human metaphase chromosome spreads, and for Southern analysis of a panel of human–hamster somatic cell hybrid DNAs, we have assigned the locus for the α1(X) gene to the q21-q22 region of human chromosome 6.

Collagen; Human type X; Nucleotide sequence; Chromosomal localization

1. INTRODUCTION

Type X collagen, a product of hypertrophic chondrocytes in regions of endochondral ossification, belongs to a distinct class of collagens designated short-chain collagens [1,2]. Molecules in this class contain subunits with a short-triple-helical domain of about 450 amino acid residues flanked by amino- and carboxyl-terminal globular domains. Their polypeptide subunits are encoded by genes with a unique exon structure; a distinct feature is the presence of a large exon that encodes the entire triple-helical and the carboxyl-terminal globular domains [3].

Three members of the short-chain collagen gene family have been identified and cloned. Two of these members, COL8A1 and COL8A2, encode polypeptides (α1(VIII) and α2(VIII), respectively), that form hetero- or homotrimeric type VIII collagen molecules synthesized by vascular and corneal endothelial cells [4–6]. The third member, COL10A1, codes for the polypeptide subunits (α1(X) collagen chains) of homotrimeric type X collagen molecules in hypertrophic cartilage [7].

Portions of the α1(VIII) and α2(VIII) collagen genes have been isolated from rabbit, mouse, and human libraries and the two genes have been localized to the human chromosomes 3 and 1 respectively [5,8]. For α1(X) collagen only the chicken and bovine genes have been reported [3,7,9]. The lack of human probes is unfortunate because the restricted expression of type X collagen suggests that it may play an important role during chondrocyte hypertrophy and endochondral ossification. Abnormalities in the structure or expression of type X collagen may therefore be associated with chondrodysplasias and other diseases of cartilage, such as osteoarthrosis. To provide the basis for molecular studies of the involvement of type X collagen in these conditions, we have isolated and sequenced a fragment of the human α1(X) collagen gene, and determined that the COL10A1 gene is located on the long arm of chromosome 6.

Correspondence address: B.R. Olsen, Department of Anatomy and Cellular Biology, Harvard Medical School, Boston, MA 02115, USA. Fax: (1) (617) 277 2732

2. MATERIALS AND METHODS

Human DNA was prepared from peripheral blood leukocytes by a standard procedure [10]. The DNA was used as template in the polymerase chain reaction (PCR) with the 5′ sense primer 5′ ACAGCAACACTATGATCC 3′ and 3′ antisense primer 5′ ACCTGAGAAAGAAGAATGAACATA 3′. The oligonucleotide primers were synthesized on a Milligen Cyclone DNA synthesizer. After deblocking with NH_4OH they were used in the PCR reaction without further purification.

The nucleotide sequences of the primers were selected based on the observation that the rabbit, mouse, and human α1(VIII), mouse and human α2(VIII), and chicken α1(X) chains contain highly conserved regions (both at the amino acid and nucleotide levels) within their non-triple-helical carboxyl-terminal globular (NC1) domains [5]. Two of these regions were therefore chosen for primer synthesis. Genomic DNA was used as template because all short-chain collagen genes isolated so far have no introns within the region that codes for the triple-helical and carboxyl-terminal globular domains.

The two primers were used for PCR with Thermus aquaticus (Taq) polymerase (United States Biochemical). The mixture of template (1 μg), primers (1 μM final concentration), dNTPs (200 μM final concentration) and reaction buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% gelatin) (reagents from United States Biochemical), was first heated to 95°C for 5 min and chilled on ice. 2.5 units of Taq polymerase (United States Biochemical) were added,

and thirty PCR cycles were performed with Programmable Dri-Block (Techne Inc., Princeton Junction, NJ). The conditions were: denaturation at 95°C for 1 min, annealing at 42°C for 1.5 min, extension at 65°C for 1 min, with a final extension at 65°C for 3.5 min. The PCR products were extracted once with phenol/chloroform, ethanol precipitated, and phosphorylated using *E. coli* T4 kinase (United States Biochemical) and ATP. They were subsequently methylated with *Eco*RI methylase (New England Biolabs), ligated to synthetic *Eco*RI linkers, digested with *Eco*RI and fractionated by electrophoresis through 2% agarose. The fraction containing ≈ 300 bp-long products (the expected size of the α1(X) product, based on comparison with the chicken α1(X) sequence) was eluted from the gel using the Geneclean kit and cloned into *Eco*RI-digested λgt10 (Stratagene). The resultant recombinants were packaged with 'Gigapack plus' packaging extracts (Stratagene) and plated onto *E. coli* strain C600 (Stratagene). This phage library was then screened by plaque hybridization, using duplicate filter lifts. As probe, a 180 bp *Pvu*II/*Pst*I fragment of the chick cDNA clone pYN3116 [7] was labelled with ^{32}P using a random-primed labelling kit (United States Biochemical). Filters were prehybridized for 2 h in 20% (v/v) formamide, 1 M NaCl, 0.05 M PIPES, 1% (w/v) SDS, 0.1% w/v sheared salmon sperm DNA and hybridized for 36 h with 10^5 cpm of labelled probe in a fresh change of hybridization fluid at 42°C. The nitrocellulose filters were washed at 42°C in 3×SSC, 0.5% (w/v) SDS and placed onto Kodak X-Omat R film for autoradiography. 10 000 plaques were screened. DNA from one positive clone, SAh10 was purified, and the 300 bp insert of SAh10 was excised with *Eco*RI, subcloned into pBluescript (Stratagene) to generate pSAh10, and into M13mp19 for the preparation of ssDNA for nucleotide sequencing.

Nucleotide sequence analysis of pSA10 was performed with the dideoxy nucleotide chain-termination technique [11]. Analysis of the sequence data was done using programs available through the National Institutes of Health-supported Molecular Biology Computer Research Resource of the Dana-Farber Cancer Institute/Harvard School of Public Health.

The insert of pSAh10 was also used as ^{32}P-labelled probe for Southern blot analysis of human genomic DNA digested with *Hin*dIII. A single hybridizing band in the *Hin*dIII digest of genomic DNA approximately 1 kb in size was eluted from an agarose gel and cloned into the *Xba*I site of λZAP (Stratagene), after the ends of vector and insert had been rendered compatible by partial fill-in using Klenow polymerase. Recombinants were packaged as described above, and the resultant library was plated onto XL1-blue host cells (Stratagene). Duplicate filter lifts were hybridized to the ^{32}P-labelled insert of pSAh10 overnight as described above and washed in 0.2×SSC, 0.5% SDS at 55°C. 15 000 plaques were screened, and the insert of one positive clone was rescued into pBluescript by in vivo excision to generate the recombinant plasmid clone pSAh10g. Nucleotide sequencing was performed as already described.

For determination of the chromosomal location of the α1(X) gene, in situ hybridization was carried out on chromosome preparations obtained from phytohemagglutinin-stimulated peripheral blood lymphocytes cultured for 72 h. 5-Bromo-deoxyuridine was added for the final 7 h of culture (60 μg/ml of medium), to ensure a posthybridization chromosomal banding of good quality. pSAh10, containing the insert cloned in pBluescript, was labelled with tritium by nick-translation to a specific activity of 1.7×10^8 dpm/μg. The radiolabelled probe was hybridized to metaphase spreads at a final concentration of 25 ng/ml of hybridization solution as described [12]. After coating with nuclear track emulsion (Kodak NTB2), the slides were exposed for 14 days at 4°C, then developed. To avoid slipping of silver grains during the banding procedure, the chromosome spreads were first stained with buffered Giemsa solution and metaphases photographed. R-banding was then performed by the fluorochrome-photolysis-Giemsa method and metaphases rephotographed before analysis.

The insert of pSAh10 was also used as ^{32}P-labelled probe for Southern blot analysis of DNA isolated from a series of human-hamster somatic cell hybrids and digested with *Hin*dIII. The generation and characterization of these hybrids has been described elsewhere [13,14].

3. RESULTS AND DISCUSSION

The PCR-approach used here is based on the findings that (i) the short-chain collagen genes contain a large ex-

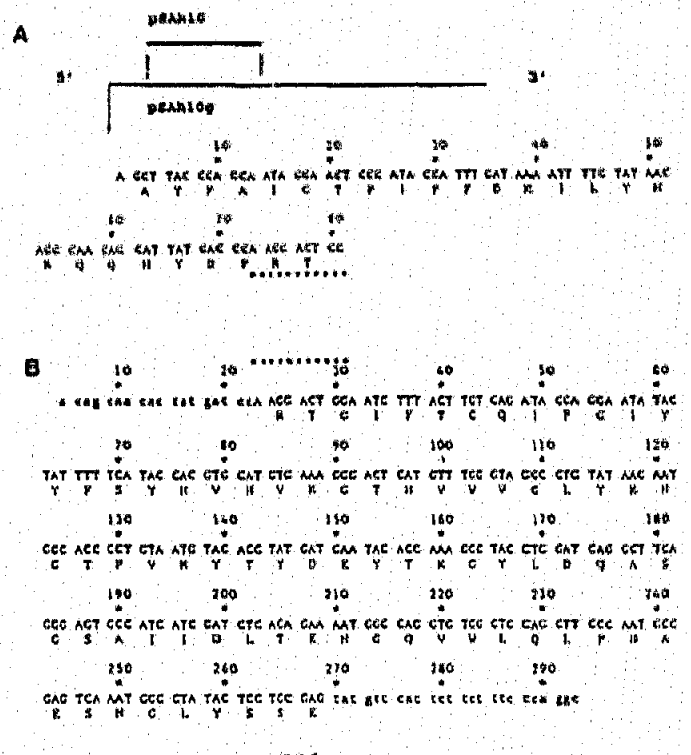

Fig. 1. (A) Diagram showing the position of pSAh10 relative to the genomic clone pSAh10g. Also shown is the nucleotide and derived amino acid sequence encoded by the 5′ end of pSAh10g. Note that this sequence overlaps and extends (in the 5′ direction) that of pSAh10. The overlap with pSAh10 is indicated by a series of dots. (B) Nucleotide sequence of pSAh10 and the deduced amino acid sequence, representing the carboxyl-half of the C-terminal domain of human α1(X) collagen. Sequences corresponding to the PCR primers are shown in lower case letters. The overlap with pSAh10g is indicated by a series of dots. (C) Southern-blot analysis of human genomic DNA digested with *Hin*dIII and probed with pSAh10. On the right are shown the positions of size markers.

on encoding the entire triple-helical and carboxyl-terminal globular domains, and (ii) the carboxyl 75% of the globular domain show a very high degree of sequence conservation, even at the nucleotide level [5]. Consequently, primers from within these conserved regions used for PCR with genomic DNA as template, should in principle allow isolation of fragments of all the genes that belong to this family. We have recently used this approach to isolate the mouse and human α2(VIII) collagen genes [5], and we report here the isolation of a fragment of the human α1(X) collagen gene.

The nucleotide sequence of the cloned PCR-product pSAh10 and the amino acid sequence for which it codes are shown in Fig. 1B. A comparison of the nucleotide sequence of pSAh10 (excluding the sequence of the primers) with those of the human α1(VIII) and α2(VIII) collagen genes [5,8], shows 61% and 64% identities. In contrast, the identity with the chicken α1(X) sequence at the nucleotide level is 76%. At the amino acid level, pSAh10 is 87% identical to chicken α1(X). Nucleotide and amino acid homologies with the recently published bovine α1(X) sequence [9] are even higher, being 83% and 92.7% respectively. This is much higher than the nucleotide identity between human α1(VIII) and human α2(VIII) in the same region, 73% [5,8]. We conclude therefore that pSAh10 represents a portion of the human α1(X) collagen gene, encoding the carboxyl-half of NC1.

The insert of pSAh10 hybridized to a single band in *Hin*dIII digested human DNA (Fig. 1C). Cloning of the human band in λZAP and screening of the resultant library resulted in the isolation of the clone pSAh10g. Partial sequence analysis of pSAh10g demonstrated that the sequence of pSAh10 was contained within it and that pSAh10g covers the 3' end of the human α1(X) collagen gene (Fig. 1C). Thus, the sequence we have cloned using PCR is represented in the human genome.

To determine the chromosomal location of the α1(X) gene, pSAh10 was used as probe for in situ hybridization of chromosome spreads. 200 metaphase cells were examined and 523 silver grains were found to be associated with chromosomes. Of these, 70 grains were associated with chromosome 6. The distribution of grains on this chromosome was not random; 61.4% (43/70 grains) mapped to the [q21-q22] region of the long arm, with a maximum in the 6q21 band (Fig. 2). Further support for the localization of the human α1(X) gene to the long arm of chromosome 6 came from Southern blot analysis of DNA isolated from a series of human-hamster somatic cell hybrids, the probe hybridizing only to those hybrids which contained chromosome 6q (data not shown).

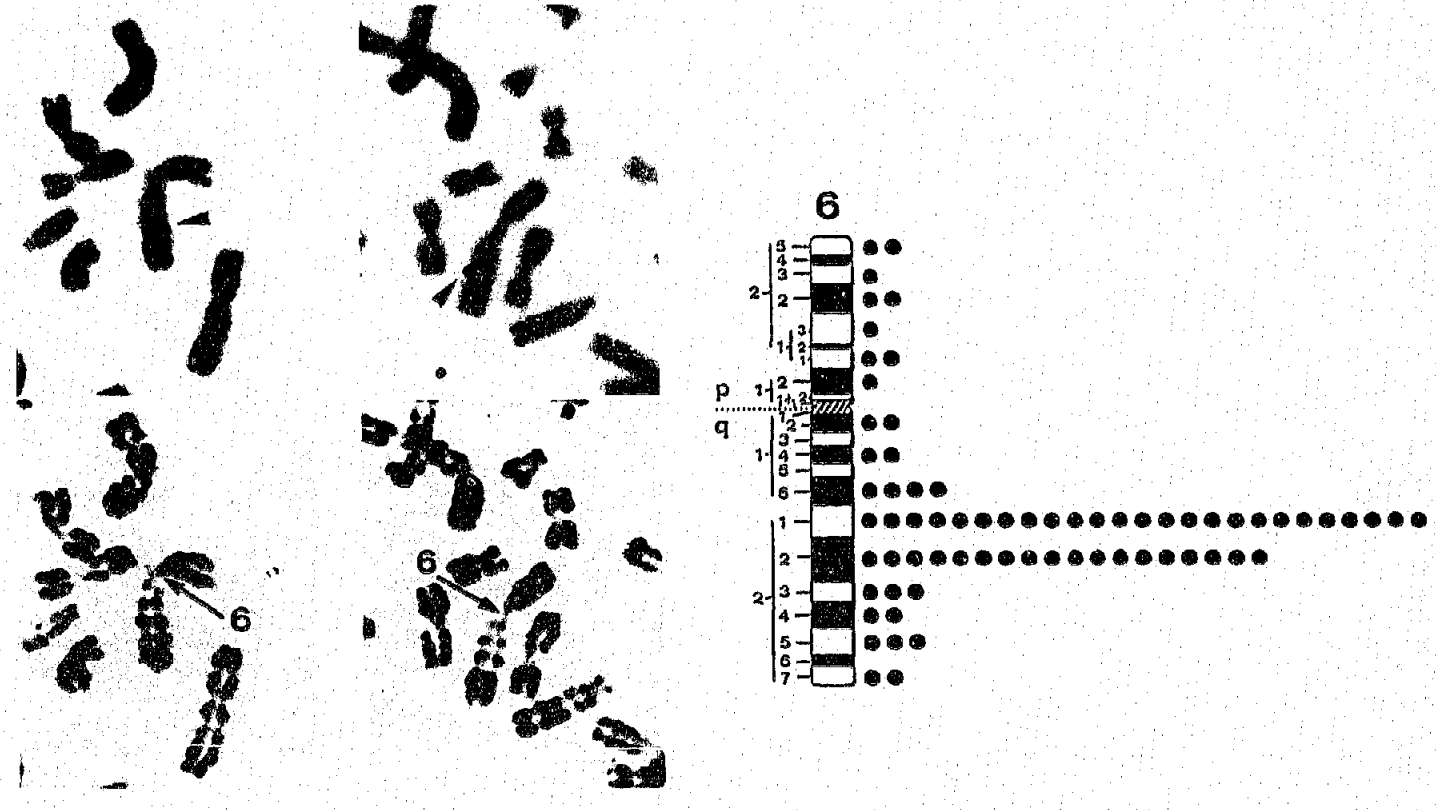

Fig 2. Ideogram demonstrating the distribution of silver grains localized on chromosome 6 by in situ hybridization.

On the basis of these findings we conclude that the human α1(X) collagen gene is located within the 6q21-6q22 bands of the human genome. Two other collagen genes have also been mapped to chromosome 6. The α2(XI) collagen gene, COL11A2, has been mapped to 6p212 region [15], while the α1(IX) gene COL9A1, has been mapped to 6q12-6q14 [16]. The three probes should be useful for linkage studies of heritable disorders of cartilage.

Acknowledgements: The authors are grateful to Dr. I. Nishimura and Dr. M. Warman for helpful discussion and technical assistance, and Drs. T. Khurana and G. Bruns for permitting us to use their hybrid panels. We are also grateful to Dr. K. von der Mark and E. Reichenberger for informing us that they have, independently of our efforts, cloned a different, but overlapping portion of human α1(X) by PCR. The study was supported by Research Grant AR36819 from the National Institutes of Health.

REFERENCES

[1] Schmid, T.M. and Linsenmayer, T.F. (1987) in: Structure and Function of Collagen Types (R. Mayne and R.E. Burgeson, eds) pp. 223-259, Academic Press, New York.
[2] Yamaguchi, N., Mayne, R. and Ninomiya, Y. (1991) J. Biol. Chem. (in press).
[3] LuValle, P., Ninomiya, Y., Rosenblum, N. and Olsen, B.R. (1988) J. Biol. Chem. 263, 18378-18385.
[4] Yamaguchi, N., Benya, P.D., van der Rest, M. and Ninomiya, Y. (1989) J. Biol. Chem. 264, 16022-16029.
[5] Muragaki, Y., Jacenko, O., Apte, S., Mattei, M.-G., Ninomiya, Y. and Olsen, B.R. (1991) J. Biol. Chem. (in press).
[6] Mann, K., Jander, R., Korsching, E., Kühn, K. and Rauterberg, J. (1990) FEBS Lett. 273, 168-172.
[7] Ninomiya, Y., Gordon, M., van der Rest, M., Schmid, T., Linsenmayer, T. and Olsen, B.R. (1986) J. Biol. Chem. 261, 5041-5050.
[8] Muragaki, Y., Mattei, M.-G., Yamaguchi, N., Olsen, B.R., and Ninomiya, Y. (1991) Eur. J. Biol. (in press).
[9] Thomas, J.T., Kwan, A.P.L., Grant, M.E. and Boot-Handford, R.P. (1991) Biochem. J. 273, 141-148.
[10] Herrmann, B.G. and Frischauf, A.-M. (1987) Methods Enzymol. 152, 180-183.
[11] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
[12] Mattei, M.-G., Philips, N., Passage, E., Moisan, J.P., Mandel, J.L. and Mattei, J.F. (1985) Hum. Genet. 69, 268-271.
[13] Bruns, G.A.P., Mintz, B.J., Leary, A.C., Regina, V.M., and Gerald, P.S. (1979) Biochem. Genet. 17, 1031-1059.
[14] Simon, M., Phillips, M., Green, H., Shoh, H., Glatt, H., Bruns, G. and Latt, S.A. (1989) Am. J. Hum. Genet. 45, 910-916.
[15] Kimura, T., Cheah, K.S.E., Chan, S.D.H., Lui, V.C.H., Mattei, M.-G., van der Rest, M., Ono, K., Solomon, E., Ninomiya, Y. and Olsen, B.R. (1989) J. Biol. Chem. 264, 13910-13916.
[16] Kimura, T., Mattei, M.-G., Stevens, J.W., Goldring, M.B., Ninomiya, Y. and Olsen, B.R. (1989) Eur. J. Biochem. 179, 71-78.