Jpn J Human Genet 38, 371–380, 1993

GENOMIC CLONING AND PARTIAL CHARACTERIZATION OF HUMAN CHYMOTRYPSINOGEN GENE

De-Xing HOU,[1,*] Kazuo OZAWA,[1] Naohiro TOMITA,[2] Yoshizane MAEDA,[3] Tsutomu HASHIGUCHI,[3] Kazushige YOKOYAMA,[1] and Eiichi SOEDA[1,**]

[1]*Gene Bank, Tsukuba Life Science Center, The Institute of Physical and Chemical Research (RIKEN), 3-1-1 Koyadai, Tsukuba, Ibaraki 305, Japan*
[2]*Department of Surgery II, Osaka University Medical School, Fukushima-ku, Osaka 553, Japan*
[3]*Department of Animal Science, Faculty of Agriculture, Kagoshima University, Kagoshima 890, Japan*

Summary Chymotrypsinogen is a principal precursor of pancreatic proteolytic enzymes. We previously isolated a cDNA clone for human prechymotrypsinogen from a human pancreatic cDNA library. In the present study, we used this cDNA sequences to isolate genomic DNA clones. Three overlapping cosmid clones spanning approximately 65-kb genomic sequences were isolated from a human cosmid library. The genomic DNA clones were characterized by restriction enzyme mapping and by hybridizing them to subfragments of the cDNA. The sequence tagged sites for human chymotrypsinogen gene were created by designing two oligonucleotides. Furthermore, the isolated genomic clones were confirmed to be localized on chromosome 16q23 by fluorescence *in situ* hybridization and G-banding analysis.

Key Words chymotrypsinogen gene, cosmid cloning, genomic analysis, chromosomal localization, sequence tagged site

INTRODUCTION

Chymotrypsin [EC 3.4.21.1] is a principal member of protein digestion enzymes. Its precursor, chymotrypsinogen, is synthesized in the acinar cells of the pancreas and secreted into small intestine. The chymotrypsinogen is activated there through

Received August 5, 1993; Accepted September 30, 1993.
*To whom correspondence should be addressed.
**To whom requests for reprints should be addressed.

tryptic cleavage and converted into active chymotrypsin(s) that hydrolyzes food proteins to peptides. The chymotrypsinogen B gene locus (CTRB) has been assigned to long arm of human chromosome 16 using a partial sequences of rat chymotrypsin B cDNA (pcXP33) as a probe to hybridize the human-rodent cell hybrids (Honey *et al.*, 1984), and has been further localized to chromosome 16q22.3-q23.2 with linkage analysis of haptoglobin (HP), tyrosine aminotransferase (TAT), and chymotrypsin B (CTRB) in 13 informative families (Westphal *et al.*, 1987). However, the genomic DNA clone for human chymotrypsinogen has not been isolated and genomic organization has not been characterized yet.

We previously isolated a cDNA clone encoding human prechymotrypsinogen from a human pancreatic cDNA library (Tomita *et al.*, 1989). In the present study, we employed the human prechymotrypsinogen cDNA to isolate genomic DNA clones from a human cosmid library. The candidate clones were further characterized by restriction enzyme mapping and hybridizing them to subfragments of the cDNA. Finally, the genomic DNA clones were confirmed to be localized on the human chromosome 16q23 by fluorescence *in situ* hybridization.

MATERIALS AND METHODS

Isolation of genomic clones for human chymotrypsinogen gene. A cosmid library was constructed from genomic DNA of HLA-homozygous, B-lymphoblastoid cell line AKIBA, which was partially digested with *Sau*3A and then ligated to *Bam*HI-digested cosmid vector pJB8 (Kawai *et al.*, 1989). Approximately 3×10^5 clones from this library were transferred to nylon membranes. The human prechymotrypsinogen cDNA containing the entire protein-coding sequences (Tomita *et al.*, 1989) was radiolabeled with [α-^{32}P]dCTP (3,000 Ci/mmol; Amersham Co., Japan) by the random primer labeling method (Feinberg and Vogelstein, 1983), and then hybridized to the cosmid filters according to standard methods (Maniatis *et al.*, 1989). The positive cosmid clones were purified by successive hybridization. The cosmid DNAs were extracted from purified candidate clones and then characterized by restriction enzyme mapping.

Extraction of DNA and Southern blot analysis. A human-mouse hybrid A9-(3884)-1 [containing a human chromosome der(16),t(16;X)(q24;q26) as the only human chromosome, established by Dr. M. Oshimura, Tottori University] was provided from the Japanese Cancer Research Resources Bank (Tokyo). DNAs from human CGM-1, mouse A9 cells, and a human-mouse hybrid A9(3384)-1 were extracted as described previously (Maniatis *et al.*, 1989). High-molecular-weight DNAs were digested with 5 units of one of several restriction enzymes per microgram of DNA, as recommended by the manufacturer. Aliquots of 13 μg of each digested DNA were fractionated by electrophoresis in a 0.8% agarose gel and transferred to nylon membrane. The baked membrane was hybridized with ^{32}P-labeled cDNA probe.

Restriction enzyme mapping. A restriction enzyme map was constructed on the basis of three overlapping cosmid clones, and these clones were digested with restriction enzyme *Eco*RI and *Bam*HI. To determine the overlapping relationship among restriction fragments, the *Eco*RI-digested fragments were recovered with Geneclean II (Bio 101, Inc.) and then digested with *Bam*HI. The exon-containing fragments were detected by hybridizing them to the cDNA probe. The vector fragments in the clones were detected by hybridizing them to cosmid vector pJB8 digested with *Bam*HI.

Fluorescence in situ hybridization. The cosmid DNA from candidate clones was labeled with biotin-14-dATP (BRL) by nick-translation, as described previously (Lichter *et al.*, 1990). The labeled DNA was denatured at 75°C for 5 min, and preannealed with human Cot-1 DNA (BRL) at 37°C for 30 min. Chromosomal metaphase spreads were prepared by standard methods from cultures of phytohemagglutinin-stimulated normal human male lymphocytes that had previously been synchronized by treatment of 5-bromodeoxyuridine (5-BrdU, Sigma), essentially as described by Zabel *et al.* (1983). The chromosomal DNA on slide was denatured and hybridized with preannealed probe at 37°C for overnight. The slide was then incubated with goat-antibiotin (Vector Lab. Inc.), washed in 4× SSC, 0.05% Tween 20, and incubated again with second and third antibodies (anti-goat-IgG and anti-rabbit-IgG, Organon Teknika Co.) conjugated with fluorescent isothiocyanate (FITC). Finally, the slide was stained with propidium iodide and examined under a Nikon Optiphot microscope equipped with epifluorescence optics (UFX-IIA). Photographs were taken on Kodak Ektachrome 400 film. The G-banding was carried out by the Hoechst 33258-UV-Giemsa method (Zabel *et al.*, 1983) and photographs were taken in bright field with Minicopy HRII film (Fuji Film Co.).

Primer synthesis and PCR analysis. The oligonucleotides were synthesized with an Applied Biosystems 380B DNA synthesizer and subsequently purified by OPC columns. The PCR reactions were performed on a GeneAmp PCR System 9600 (Perkin-Elmer-Cetus, North Haven, CT), and the PCR reaction mixture contained 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% gelatin, 200 μM each dNTP, 100 ng each primer, and 2.5 U of *Taq* polymerase (Perkin-Elmer, Norwalk, CT) in a volume of 50 μl. DNA was initially denatured at 94°C for 2 min, and then 35 cycles of PCR were performed, each cycle consisting of a 20-s denaturation at 94°C, a 20-s annealing at 67°C, and 30-s extension at 72°C. Final extension was carried out at 72°C for 5 min. The amplified products were analyzed by electrophoresis in a 3% agarose gel.

RESULTS

Isolation and identification of human chymotrypsinogen gene

The genomic clones for the human chymotrypsinogen gene were isolated from

a human cosmid library, using the human prechymotrypsinogen cDNA as a probe. Out of 3×10^5 cosmid clones, 22 clones gave positive hybridization signals. To determine whether the cosmid clones contain the entire chymotrypsinogen gene,

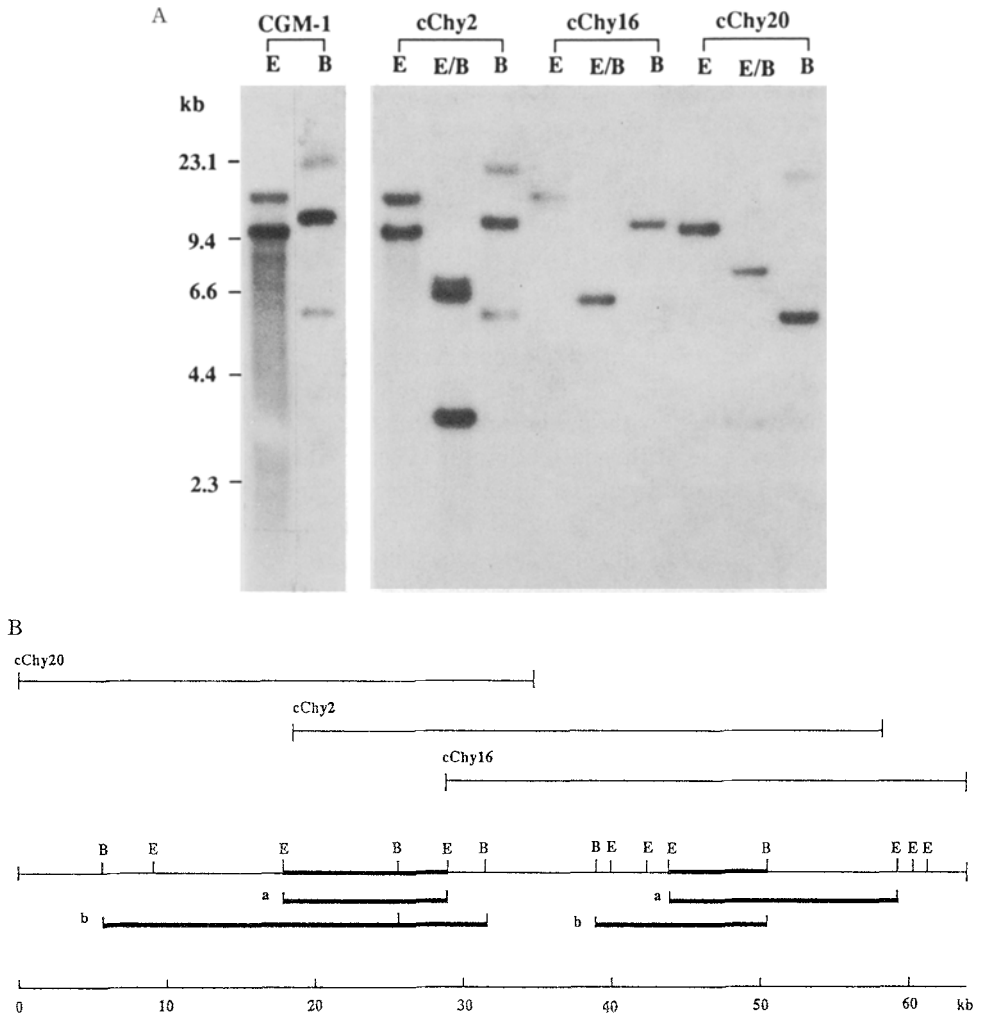

Fig. 1. (A) Southern blot analysis of DNAs from human CGM-1 (two lanes in left) and three cosmid clones (cChy2, cChy16, cChy20) isolated from the human cosmid library. The DNAs were digested with *Eco*RI (E), *Eco*RI/*Bam*HI (E/B) and *Bam*HI (B), and separated by electrophoresis in a 0.8% agarose gel. The DNA fragments were hybridized with the human prechymotrypsinogen cDNA probe. The positions of DNA size markers (*Hin*dIII-digested λDNA) are indicated on the left. (B) Restriction enzyme map and overlapping positions of the three cosmid clones. Exon-containing fragments in the map are represented by solid lines; a, *Eco*RI fragment; b, *Bam*HI fragment. The sequences covered by these cosmid inserts are indicated with kb at bottom.

the three cosmid DNAs and total human DNA (CGM-1) were digested with *Eco*RI and *Bam*HI, separated in a 0.8% agarose gel and hybridized with the cDNA probe (Fig. 1A). The cosmid clone, cChy2, revealed all of fragments identical in size to those visible in total human DNA (lanes 1 and 2); the clones, cChy16 and cChy20, revealed a part of these fragments. A restriction enzyme map for human chymotrypsinogen gene was constructed on the basis of the three overlapping cosmid clones, and the sequences covered by these cosmid inserts stretched nearly 65 kb (Fig. 1B). These data indicate that human chymotrypsinogen gene was isolated completely with these cosmid clones.

Localization of the gene sequences within the genomic clone

As a means of localizing the gene sequences within the genomic clone, the exon-containing *Eco*RI fragments of 10-kb and 14-kb in cosmid clone cChy2 (Fig. 1A) were recovered and then digested with *Pst*I. The DNA fragments were subjected to Southern blot analysis using the entire cDNA as a probe. As shown in Fig. 2A, four exon-containing *Eco*RI-*Pst*I fragments were detected within the 10-kb *Eco*RI fragment (lane 2) while only two fragments of them were detected within the 14-kb *Eco*RI fragment (lane 1). To localize the cDNA sequences into the genomic DNA fragments, the cDNA was digested with *Pst*I and *Bal*I, and its subfragments were used as probes. The probe 1 was a 169-bp *Eco*RI-*Pst*I fragment of the cDNA which encompassed 16 bp of the 5′-untranslated region and the coding sequence for the first 51 amino acids of the human chymotrypsinogen. Probe 2 was a 226-bp *Pst*I-*Bal*I fragment encoding amino acids 52 to 127. Probe 3 was a 330-bp *Pst*I-*Eco*RI fragment containing the coding sequence for amino acids 173 to 263 and 61 bp of the 3′-untranslated region. Results from the Southern blot analysis, using these probes, are shown in Fig. 2B. Regarding the 10-kb *Eco*RI fragment, the probe 1 and 2 detected a 1.3-kb (lane 1) and a 1.2-kb (lane 2) *Eco*RI-*Pst*I fragment, respectively. The probe 3 detected two *Eco*RI-*Pst*I fragments of 0.7-kb and 0.5-kb (lane 3). Regarding the 14-kb *Eco*RI fragment, only the probe 3 detected two *Eco*RI-*Pst*I fragments (0.7-kb and 0.5-kb) (lane 4). These results can be summarized in Fig. 2C, the correspondence of the cDNA subfragments to genomic DNA indicates that the entire coding region of the chymotrypsinogen gene was encompassed within the 10-kb *Eco*RI fragment of the genomic clone. The additional 14-kb *Eco*RI fragment is at the downstream of this gene and shows homology to this gene. We also noted that the 1.2-kb *Eco*RI-*Pst*I fragment gave a signal that was more intense than other fragments. There would not be more than a single fragment comigrating in this band since it showed a clear single band in agarose gel with ethidium bromide staining (data not shown). It is possible that the 1.2-kb *Eco*RI-*Pst*I fragment contained most of the cDNA sequences while other fragments (1.3-kb, 0.7-kb, and 0.5-kb) contained the rest, therefore, they showed different intensity of signals under our highly stringent hybridization and washing conditions.

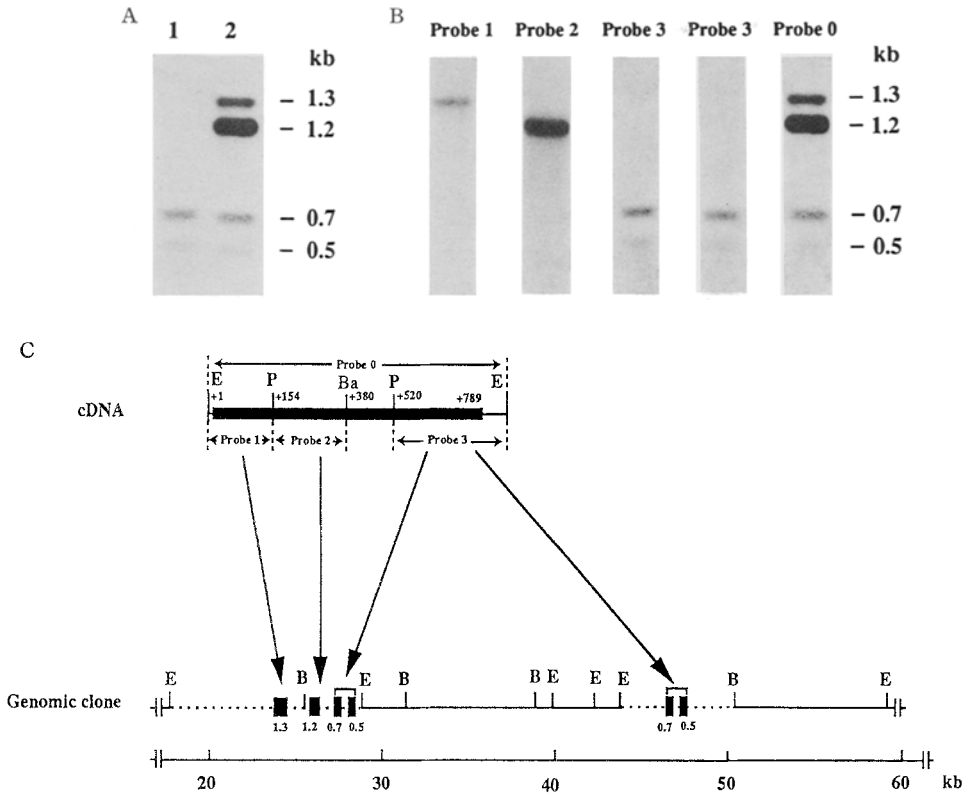

Fig. 2. Localization of the gene sequences within the genomic clone. (A) Southern blot analysis of exon-containing *Eco*RI fragments of 10-kb (lane 2) and 14-kb (lane 1). The *Eco*RI-digested cosmid DNA from cChy2 was fractionated by electrophoresis in a 1% agarose gel. Of which, 10-kb and 14-kb *Eco*RI fragments were recovered and secondly digested with *Pst*I. The *Eco*RI-*Pst*I DNA fragments were separated by electrophoresis in a 1.3% agarose gel and then hybridized with the entire cDNA probe. (B) Determination of corresponding fragments between the cDNA and genomic DNA. The *Eco*RI-*Pst*I DNA fragments from 10-kb (lanes 1–3 and 5) and 14-kb (lane 4) were hybridized with the cDNA subfragments (Probe 1–3) and the entire cDNA (Probe 0) as a control (lane 5). (C) Schematic representation of the correspondence between the cDNA subfragments and genomic DNA. The amino acid coding region is represented by solid bar. The numbers on the cDNA start from the first base of the initiation codon. The genomic clone was isolated from a human cosmid library in pJB8 cosmid vector, and the size of exon-containing *Eco*RI-*Pst*I fragments are indicated below the solid bar (kb). The dotted lines in the genomic DNA represent that the distances between exon and intron fragments were not determined.

PCR analysis of human chymotrypsinogen gene

To produce the sequence tagged site for human chymotrypsinogen gene, we used its cDNA sequences to design the oligonucleotides that can specifically amplify

the putative exon-containing region of the chymotrypsinogen gene, guided in part by the exon-intron structure of the rat chymotrypsin B gene (Bell *et al.*, 1984). Different primer sets were tested with human and rodent genomic DNA templates, and a primer set that amplified a single human-specific band was selected. As shown in Fig. 3, these two unique primers (sense primer: 5′-CTGGTCTGCCAA-AAGGATGGAGCC-3′, and antisense primer: 5′-GTTGGCAGCCAGGATCTTC-TGCAC-3′) designed from the 3′-end of the coding region in the cDNA and flanking the active-site Ser-195, gave an expected 140-bp band in total human DNA (lane 1), the cDNA (lane 2) and the genomic cosmid DNA (lane 3) while no amplified fragment was detected in mouse (lane 4), hamster (lane 5), and vector DNAs (lanes 6 and 7).

Confirmation of human chymotrypsinogen gene on 16q23

To confirm the chromosomal region of the genomic DNA clone isolated by us, we performed fluorescence *in situ* hybridization and G-banding analysis. The DNA from three overlapping cosmids was labeled with biotin-14-dATP and hybridized to metaphase spreads of chromosome. Simultaneously, a biotinylated α-satellite DNA that is specific to D16Z2 region of chromosome 16 was added as a marker for distinguishing chromosome 16. Figure 4A displays the two pairs of yellow fluorescent signals to the cosmid probe on both sister chromatids (large arrows). Comparison between fluorescent signals and the G-banding pattern of the chromosome indicated that signals were consistently localized on the long arm of chromosome 16, 16q23 (Fig. 4, B and C). This result was confirmed by analysis of 16 cells (data not shown). The fluorescent signals were observed on both chromatids of chromosome 16 in all of 16 cells. No other chromosome region showed signals on

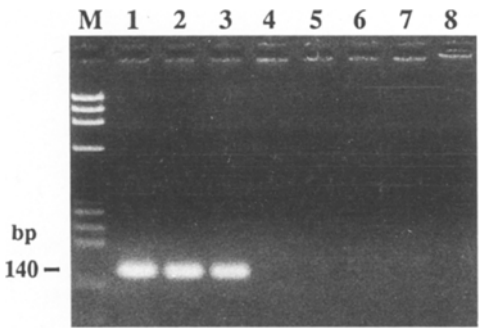

Fig. 3. The PCR analysis of DNAs from human CGM-1 (lane 1), the cDNA (lane 2), the genomic clone (lane 3), mouse (lane 4), hamster (lane 5), plasmid vector puc-18 (lane 6), cosmid vector pJB8 (lane 7), and no DNA as a blank (lane 8) with the designed two oligonucleotides (see MATERIALS AND METHODS). The PCR product was analyzed by electrophoresis in a 3% agarose gel. The expected PCR product is indicated on the left. The molecular weight standard (M) is ϕX174 DNA digested with *Hae*III.

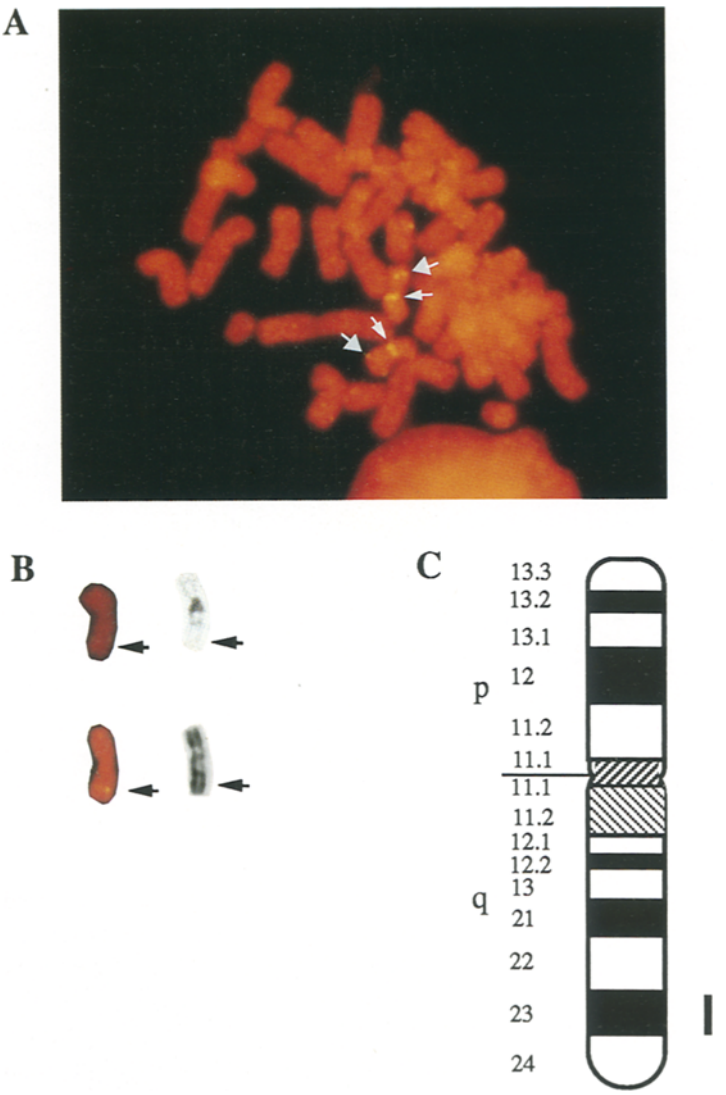

Fig. 4. Chromosomal mapping of human chymotrypsinogen gene by fluorescence *in situ* hybridization and G-banding analysis. (A) The metaphase spreads after hybridization with the biotinylated cosmid DNA and a α-satellite DNA that is specific to centromere of chromosome 16 for distinguishing chromosome 16, and detection with FITC. The large arrows indicate the cosmid DNA signals, and the small arrows indicate the α-satellite DNA signals. (B) Examples of two chromosome 16 from two different metaphase spreads. G-banding after fluorescence detection indicates that signals were consistently localized to the band 16q23. (C) A G-banded idiogram of human chromosome 16 depicting the localization of human chymotrypsinogen gene to 16q23 (the bar beside the idiogram).

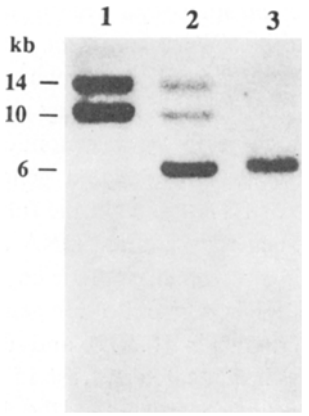

Fig. 5. Hybridization of human prechymotrypsinogen cDNA probe to human (lane 1), hybrid A9(3884)-1 (lane 2), and mouse A9 DNAs (lane 3) which were digested with *Eco*RI and separated by electrophoresis in a 0.8% agarose gel. Each lane contained 13 μg DNA. The positions of the hybridization bands are indicated on the left.

both chromatids. Additionally, a Southern blot analysis of hybrid A9(3884)-1, which contains a human chromosome der(16),t(16;X)(q24,q26) as the only human chromosome, also confirmed that the human chymotrypsinogen gene was located on chromosome 16 (Fig. 5). From these data, we conclude that the isolated genomic DNA encompassing human chymotrypsinogen gene is located on chromosome 16q23.

DISCUSSION

We have isolated the genomic DNA clones that encompassed human chymotrypsinogen gene, using human prechymotrypsinogen cDNA as a probe. The evidence come from that (i) the isolated genomic DNA clones revealed all of fragments identical in size to those obtained by hybridizing the cDNA probe to total human genomic DNA, and (ii) the restriction map showed excellent correspondence between genomic DNA and the cDNA subfragments. Furthermore, we have designed two oligonucleotides that can specifically detect human chymotrypsinogen gene by PCR. The primers will be useful as a sequence tagged site for further analysis of the human chymotrypsinogen gene.

Although the chymotrypsinogen B gene locus (CTRB) has been localized in chromosome 16q22.3-q23.2 using rat cDNA as a probe and with linkage analysis (Honey *et al.*, 1984; Westphal *et al.*, 1987), the human genome may carries more than one genes for chymotrypsinogens since two different forms of chymotrypsinogen have been reported to be present in human pancreas (Caro *et al.*, 1975). To confirm the chromosomal region of the genomic DNA clones isolated by us, we performed fluorescence *in situ* hybridization and G-banding analysis. The isolated genomic

clones were mapped into the chromosomal region within CTRB gene locus that was localized by linkage analysis (Westphal *et al.*, 1987). We have furthermore compared the sequence homology between human prechymotrypsinogen cDNA and rat chymotrypsin B cDNA, and found that they showed 83.6% homology (data not shown). These results suggest that the genomic DNA isolated by us is, in fact, a genomic gene for CTRB. The isolated genomic DNA will be useful to analyze the genomic structure of CTRB gene in the future.

The correspondence between the genomic DNA and the cDNA subfragments indicates that the entire coding region of human chymotrypsinogen gene was encompassed within the 10-kb *Eco*RI fragment of the genomic clone. Simultaneously, additional two *Eco*RI-*Pst*I fragments (0.5-kb and 0.7-kb) were generated from the 14-kb *Eco*RI fragment which is located at least 15 kb apart at the 3′-side of the 10-kb fragment (Fig. 2). It is unknown whether the 14-kb *Eco*RI fragment represents another chymotrypsinogen gene or other related gene that shows partial sequence homology to the chymotrypsinogen gene. In the future study, we intend to determine their sequences to demonstrate their relationships.

Acknowledgments We thank Ms. S. Okano for secretarial assistance. This work was supported by the human genome research program of RIKEN and the Science and Technology Agency of Japan.

REFERENCES

Bell GI, Quinto C, Quiroga M, Valenzuela P, Craik CS, Rutter WJ (1984): Isolation and sequence of a rat chymotrypsin B gene. J Biol Chem **259**: 14265–14270

Caro AD, Figarella C, Guy O (1975): The two human chymotrypsinogens purification and characterization. Biochim Biophys Acta **379**: 431–443

Feinberg AP, Vogelstein B (1983): A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem **132**: 6–13

Honey NK, Sakaguchi AY, Quinto C, MacDonald RJ, Bell GI, Craik C, Rutter WJ, Naylor SL (1984): Chromosomal assignments of human genes for serine proteases trypsin, chymotrypsin B, and elastase. Somat Cell Mol Genet **10**: 369–376

Kawai J, Ando A, Sato T, Nakatsuji T, Tsuji K, Inoko H (1989): Analysis of gene structure and antigen determinants of DR2 antigen using DR gene transfer into mouse L cells. J Immunol **142**: 312–317

Lichter P, Chang-Tang CJ, Call K, Hermanson G, Evans GA, Housman D, Ward DC (1990): High-resolution mapping of human chromosome 11 by *in situ* hybridization with cosmid clones. Science **247**: 64–69

Maniatis T, Fritsch EF, Sambrook J (1989): Molecular Cloning: A Laboratory Manual (2nd ed.), Cold Spring Harbor Laboratory Press, New York

Tomita N, Izumoto Y, Horii A, Doi S, Yokouchi H, Ogawa M, Mori T, Matsubara K (1989): Molecular cloning and nucleotide sequence of human pancreatic prechymotrypsinogen cDNA. Biochem Biophys Res Commun **158**: 569–575

Westphal E, Burmeister M, Wienker TF, Lehrach H, Bender K, Scherer G (1987): Tyrosine aminotransferase and chymotrypsinogen B are linked to haptoglobin on human chromosome 16q: comparison of genetic and physical distances. Genomics **1**: 313–319

Zabel BU, Naylor SL, Sakaguchi AY, Bell GI, Shows TB (1983): High resolution of human genes for amylase, proopiomelanocortin, somatostatin, and a DNA fragment (D3S1) by *in situ* hybridization. Proc Natl Acad Sci USA **80**: 6932–6936