Am J Hum Genet 38:38–44, 1986

Localization of the Human Procollagen α1(IV) Gene to Chromosome 13q34 by in Situ Hybridization

BEVERLY S. EMANUEL,[1] BEATRICE T. SELLINGER,[1] LORRAINE J. GUDAS,[2] AND JEANNE C. MYERS[3]

SUMMARY

Type IV (α1 and α2 chains) appears to be the only procollagen present in basement membranes. The structure of this protein is highly divergent from the interstitial and type V procollagens as exemplified by the interruptions in the Gly-X-Y region and unprocessed amino and carboxyl noncollagenous peptides. To expand our knowledge concerning the primary sequence of type IV and to investigate the factors influencing its unique distribution, we recently isolated cDNA clones coding for part of the human α1(IV) chain. To determine if the α1(IV) gene was cytologically linked to other procollagen genes that have been assigned to autosomes 17, 12, 7, and 2, overlapping clones covering 2.6 kilobases (kb) of the α1(IV) mRNA were used together for in situ hybridization to human metaphase chromosomes. Here, we show precise localization of α1(IV) at the telomere of 13q, thereby defining a fifth chromosome that contains members of this large and surprisingly dispersed multigene family.

INTRODUCTION

Basement membranes are widely distributed in the body; they provide support for epithelial and endothelial cells and serve as a filter between blood and

Received April 15, 1985.

This work was supported by grants AM20553, GM32592, GM07511, and P01-22427 from the National Institutes of Health.

[1] Children's Hospital of Philadelphia and Departments of Pediatrics and Human Genetics, University of Pennsylvania, Philadelphia, PA 19104.

[2] Dana Farber Cancer Institute, and Department of Pharmacology, Harvard Medical School, Boston, MA 02115.

[3] Connective Tissue Research Institute, and Departments of Medicine and Human Genetics, University of Pennsylvania, Philadelphia, PA 19104.

tissues (for reviews, [1–3]). Two high molecular weight glycoproteins, laminin and type IV procollagen, are the major constituents of these extracellular matrices. Type IV procollagen, consisting of α1 and α2 chains, has an unusual structure and a different tissue distribution as compared to the interstitial types I, II, and III and the pericellular type V [1–3]. Although related overall to the others by the presence of a central collagenous domain and noncollagenous propeptides, the type IV components are distinguished by interruptions in the Gly-X-Y repeating sequence and by retention of the terminal amino and carboxyl extensions after secretion [4–9]. Clarification of the type IV primary structure has come from two sources: isolation and sequencing of specific peptides or of cDNA clones that encode the protein [10–15]. We have been involved in the second approach since the recombinant molecules can also serve as probes for identifying α1(IV) mRNA in differentiated cell types, chromosomal mapping, and, ultimately, gene analysis. Here, we report on our continuing studies designed to further elucidate the organization of this multigene family in the human genome.

Recently, we characterized overlapping human α1(IV) clones coding for the COOH-terminus of the α-chain, the 3′ noncollagenous peptide, and the entire 3′ untranslated region ([15] and our unpublished data, 1985). Comparison of α1(IV) with the structurally related interstitial and α2(V) procollagens showed that although α1(IV) is extremely divergent there is limited homology in select areas, perhaps defining conserved functional domains [15]. To ascertain if α1(IV) was linked to the other procollagen genes encoding these proteins, we employed the chromosomal in situ hybridization technique that permits precise regional assignment of the locus. Previously, this procedure allowed us to determine that the α1(III) and α2(V) genes are located in the q24.3→q31 region of chromosome 2 [16]. Prior to this observation, no site with more than one procollagen gene had been described, since α1(I), α2(I), and α1(II) are dispersed on chromosomes 17, 7, and 12, respectively [17–20]. For the experiments presented here, we concomitantly used two α1(IV) clones that have in common 100 nucleotides and together code for 2.6 kb of the 6.5-kb α1(IV) mRNA [15]. Hybridization of these cDNAs to normal metaphase chromosomes sharply delineated the α1(IV) locus at band q34 of chromosome 13. This represents a fifth autosome with procollagen loci, suggesting that molecular contiguity does not play an important role in the expression of at least several members of this gene family.

MATERIALS AND METHODS

In Situ Hybridization Probes

The probes used for the chromosomal mapping studies were two α1(IV) cDNA clones isolated from a normal fibroblast library (fig. 1). The 3′ clone, NB-3, contains a 1-kb insert coding for most of the 3′ untranslated region ([15] and our unpublished data, 1985). The 5′ clone, KK4, overlapping NB3 by about 100 nucleotides, contains a 1.7-kb insert coding for the 5′ portion of the untranslated region, the 229 residue C-terminal peptide, and 182 amino acids of the Gly-X-Y region including a four amino acid interruption [14, 15].

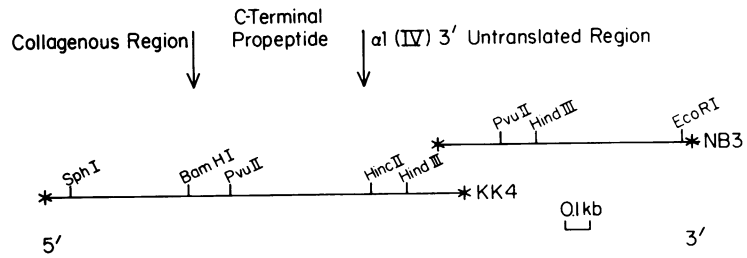

FIG. 1.—Restriction maps of human α1(IV) procollagen cDNA clones. The regions of the α1(IV) chain encoded by the cDNA clones are shown at *the top*. The 1-kb insert of NB3 corresponds to most of the 3′ untranslated region, and the 1.7-kb insert of KK4 codes for the COOH-terminus of the collagenous domain, the C-terminal propeptide, and part of the 3′ untranslated region. DNAs were inserted into the *Pst*I site of pBR322 (*asterisks*).

In Situ Hybridization to Metaphase Chromosomes

Metaphase chromosome spreads were prepared from peripheral blood lymphocytes of normal males having a 46,XY karyotype using standard techniques [21]. In situ hybridization and autoradiographic procedures were performed essentially as described [22]. Probe DNA was ^{3}H-labeled by nick-translation to a specific activity of 2×10^7 cpm/μg, separated from unincorporated nucleotides on a Sephadex G-50 column, and ethanol precipitated with salmon sperm carrier DNA. DNA was resuspended in a hybridization mixture consisting of 50% formamide/2 × SSC/10% dextran sulfate, pH 7.0, denatured for 5 min at 70°C, and added to the slides. The first experiment was carried out using the probe NB3 at a final concentration of 0.07 μg/ml. The three subsequent experiments were performed with equal microgram amounts of NB3 and KK4 totaling 0.035 or 0.07 μg/ml. After incubation for 18 hrs at 37°C, the slides were washed in 50% formamide, 2 × SSC at 39°C.

RESULTS

Pro α1(IV) cDNA Clones

Restriction maps of the two human procollagen α1 type IV cDNA clones are shown in figure 1. NB3 was identified using a mouse type IV clone coding for the 3′ part of the untranslated region (our unpublished data, 1985, and [11]), and the 5′ clone, KK4, was later isolated by hybridizing the cDNA library to NB3 [15]. DNA sequences and derived amino acids of the COOH-terminus of the collagenous region are reported in figure 2. Comparison of the residues with the published protein sequence obtained by peptide analysis [14] demonstrated that the clones encode part of the α1 chain of human type IV procollagen.

Chromosomal Localization of the Human α1(IV) Procollagen Gene

To determine the chromosomal location of the α1(IV) procollagen gene, four independent experiments were performed. In experiment 1, the first α1(IV) clone isolated, NB3, was used for in situ hybridization. Twenty-three metaphase spreads were examined in which 78 grains were on chromosomes. Of these, 22% were located on chromosome 13, and 15 of 17 grains were found in the q32→34 region. No other site had more than two grains. For the next three experiments, a 1.7-kb overlapping clone, KK4, was combined with the 1-kb clone NB3, which together code for 2.6 kb of the α1(IV) mRNA. The final DNA

```
PROTEIN  Gly Phe Asp Gly Ala Hyp Gly Gln Hyl Gly Glu Met Gly Pro Ala Gly Pro Thr Gly Pro Arg
CLONE   [Gly Phe Asp Gly Ala Pro Gly Gln Lys Gly Glu Met Gly Pro Ala Gly Pro Thr Gly Pro Arg]
        [GGG TTT GAC GGT GCC CCT GGC CAG AAA GGA GAG ATG GGA CCT GCC GGG CCT ACT GGT CCA AGA]
         866                                                                            886

PROTEIN  Gly Phe Hyp Gly Pro Hyp Gly Pro Asp Gly Leu Pro Gly Ser Met Gly Pro Hyp Gly Thr Pro
CLONE   [Gly Phe Pro Gly Pro Pro Gly Pro Asp Gly Leu Pro Gly Ser Met Gly Pro Pro Gly Thr Pro]
        [GGA TTT CCA GGT CCA CCA GGC CCC GAT GGG TTG CCA GGA TCC ATG GGG CCT CCA GGC ACC CCA]
         887                                                                            907
```

FIG. 2.—Amino acid sequences derived from the α1(IV) procollagen cDNA clone. Shown are the nucleotides and derived amino acids from the cDNA clone KK4 and the human α1(IV) protein sequence reported by Babel and Glanville [14]. Residues 866-907 are the COOH-terminus of the α1(IV) 95-K collagenous fragment. Adjacent to this region is the 229 amino acid C-terminal noncollagenous peptide [15]. Sanger dideoxy DNA sequencing was carried out essentially as described by Messing [23].

concentration in experiment II was 0.035 μg/ml, and in experiments III and IV, 0.07 μg/ml. Consistent with the first results, the predominant site of hybridization in these three experiments was the long arm of chromosome 13 (13q) with 24%, 31%, and 25% of the total grains (table 1 and fig. 3). The chromosomal distribution of all grains from the four experiments is shown in figure 4. In each study, the number of grains localized to 13q32→13q34 represented 6–10 times the number on a chromosomal region of similar length. Since 14%, 20%, 23%, and 23% of the total grains in the four analyses were found at 13q terminus, the procollagen α1(IV) locus was more precisely assigned to 13q34.

DISCUSSION

Of the four closely related interstitial procollagens [24–31] encoded by genes on different chromosomes [16–20], α1(I), α2(I), and α1(III) are present in most tissues, while α1(II) has a specific location in cartilage [1, 3]. The one type V procollagen gene [α2(V)] [32] mapped so far is in the same region of chromosome 2 as α1(III), and whether this has an effect on their coordinate regulation remains to be determined [16]. These two genes in addition to α1(I), α2(I), and α1(IV) are expressed in normal cultured fibroblasts [15, 31, 32], but our recent experiments show that only α1(III), α2(V), and α1(IV) transcripts are found in human umbilical endothelial cells, with the latter species predominating (our unpublished data, 1985). Therefore, the studies presented here provide additional evidence that chromosomal distribution of a number of procollagen

TABLE 1

HYBRIDIZATION OF α1(IV) PROBES TO NORMAL METAPHASE CHROMOSOMES

Probe	Experiment no.	No. metaphases	No. grains	Grains on 13q
NB3	1	23	78	17/78 = 22%
NB3 + KK4 ...	2	100	189	46/189 = 24%
NB3 + KK4 ...	3	27	74	23/74 = 31%
NB3 + KK4 ...	4	75	180	46/180 = 25%
Total		225	521	132/521 = 25%

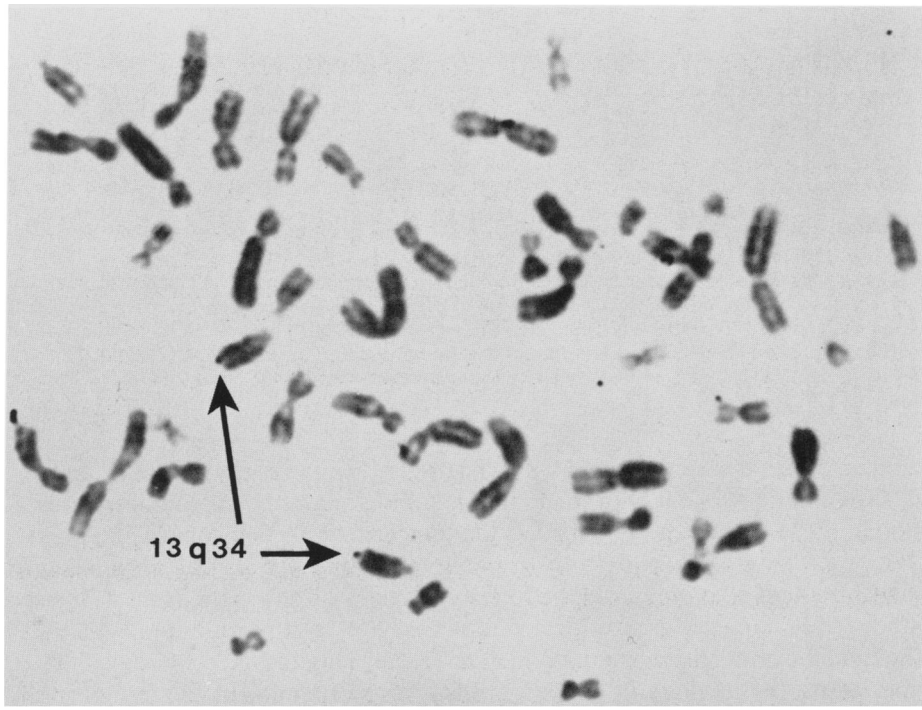

FIG. 3.—In situ hybridization of α1(IV) procollagen cDNA clones. The autoradiograph shown is representative of the pattern obtained from hybridization of the α1(IV) clones to normal metaphase chromosomes (46,XY). *Arrows* indicate grains on distal termini (q34) of both chromosome 13 homologs.

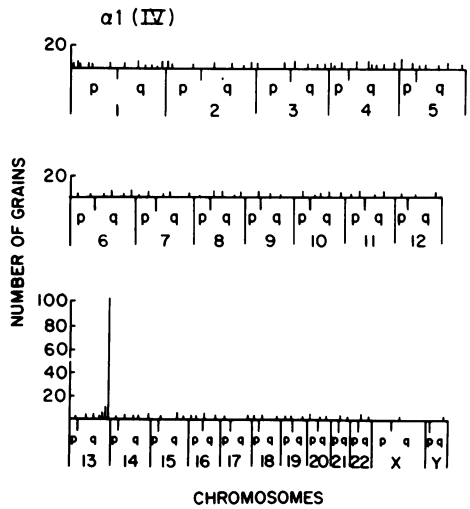

FIG. 4.—Histogram of grains distributed on metaphase chromosomes. The profile shows the grain localizations from the four independent in situ hybridization studies. The first experiment was carried out using the tritium-labeled probe NB3 and the three subsequent ones with both NB3 and the overlapping DNA probe KK4 (fig. 1). *Abscissa* represents the chromosomes in their relative size proportions, and *ordinate* shows the distribution and no. silver grains.

genes is not strictly correlated with either structural similarities of the polypeptide chains or expression of the mRNAs.

Sublocalization of α1(IV) to 13q34 shows this locus to be at a significant distance from the q14 region that is known to be involved in retinoblastoma [33, 34]. Carriers of the mutant allele are predisposed to develop this intraocular tumor of childhood. Investigators have generated a series of cloned anonymous DNA fragments from chromosome 13 that contain restriction fragment length polymorphisms in order to examine somatic genetic changes related to tumorigenesis. Further identification of such sites at the α1(IV) locus will not only be valuable for these studies, but also for establishing if linkage exists between inheritance of an α1(IV) allele and genetically transmitted diseases affecting basement membrane such as Alport syndrome and adult polycystic renal disease [2].

ACKNOWLEDGMENTS

We thank J. M. Brinker, H. R. Loidl, and S. Wang for DNA sequencing of the type IV clones; N. Kefalides and A. Martinez for helpful discussions; and M. Mason and T. White for typing the manuscript.

REFERENCES

1. Bornstein P, Sage H: *Ann Rev Biochem* 49:957–1003, 1980
2. Martinez-Hernandez A, Amenta PS: *Lab Invest* 48:656–677, 1983
3. Miller EJ, in *Extracellular Matrix Biochemistry,* edited by Piez KA, Reddi AH, New York, Elsevier, 1984, pp 41–82
4. Minor RR, Clark CC, Strause EL, Koszalka TR, Brent RL, Kelfalides NA: *J Biol Chem* 251:1789–1794, 1976
5. Dehm P, Kefalides NA: *J Biol Chem* 253:6680–6686, 1978
6. Kresina TF, Miller EJ: *Biochemistry* 18:3089–3097, 1979
7. Crouch E, Sage H, Bornstein P: *Proc Natl Acad Sci USA* 77:745–749, 1980
8. Tryggvason K, Robey PG, Martin GR: *Biochemistry* 19:1284–1289, 1980
9. Schuppan D, Timpl R, Glanville RW: *Febs Lett* 115:297–300, 1980
10. Schuppan D, Glanville RW, Timpl R: *Eur J Biochem* 123:505–512, 1980
11. Wang SY, Gudas LJ: *Proc Natl Acad Sci USA* 80:5880–5884, 1983
12. Kurkinen M, Barlow DP, Helfman DM, Williams JG, Hogan BLM: *Nucleic Acids Res* 11:6199–6209, 1983
13. Schuppan D, Glanville RW, Timpl R, Dixit SN, Kang AH: *Biochem J* 220:227–233, 1984
14. Babel W, Glanville RW: *Eur J Biochem* 143:545–556, 1984
15. Brinker JM, Gudas LJ, Loidl HR, et al.: *Proc Natl Acad Sci USA* 82:3649–3653, 1985
16. Emanuel BS, Cannizzaro LA, Seyer JM, Myers JC: *Proc Natl Acad Sci USA* 82:3385–3389, 1985
17. Huerre C, Junien C, Weil D, et al.: *Proc Natl Acad Sci USA* 79:6627–6630, 1982
18. Henderson AS, Myers JC, Ramirez F: *Cytogenet Cell Genet* 36:586–587, 1983
19. Solomon E, Hiorns L, Dagleish R, Tolstoshev P, Crypstal R, Sykes B: *Cytogenet Cell Genet* 35:64–66, 1983
20. Strom CM, Eddy RL, Shows TB: *Somat Cell Mol Genet* 10:651–655, 1984
21. Moorhead PS, Nowell PC, Melman WJ, Battips DM, Hungerford DA: *Exp Cell Res* 20:613–616, 1960
22. Cannizzaro LA, Emanuel BS: *Cytogenet Cell Genet* 38:308–309, 1984
23. Messing J: *Methods Enzymol* 101:20–78, 1983

24. FULLER F, BOEDTKER H: *Biochemistry* 20:996–1006, 1981
25. BERNARD MP, MYERS JC, CHU M-L, RAMIREZ F, EIKENBERRY EF, PROCKOP DJ: *Biochemistry* 22:1139–1145, 1983
26. BERNARD MP, CHU M-L, MYERS JC, RAMIREZ F, EIKENBERRY EF, PROCKOP DJ: *Biochemistry* 22:5213–5223, 1983
27. YAMADA Y, KUHN K, DECROMBRUGGHE B: *Nucleic Acids Res* 11:2733–2744, 1983
28. SANDELL LJ, PRENTICE HL, KRAVIS D, UPHOLT WB: *J Biol Chem* 259:7826–7834, 1984
29. NINOMIYA Y, SHOWALTER A, VANDERREST M, SEIDAH NG, CHRETEIN M, OLSEN BR: *Biochemistry* 23:617–624, 1984
30. YOUNG MF, VOGELI G, NUNEZ AM, FERNANDEZ MP, SULLIVAN M, SOBEL ME: *Nucleic Acids Res* 12:4207–4228, 1984
31. LOIDL HR, BRINKER JM, MAY M, ET AL.: *Nucleic Acids Res* 12:9383–9394, 1984
32. MYERS JC, LOIDL HR, STOLLE CA, SEYER JM: *J Biol Chem* 260:5533–5541, 1985
33. SPARKES RS, MURPHREE AL, LINGUA RW, ET AL.: *Science* 219:971–973, 1983
34. CAVENEE WK, DRYJA TP, PHILLIPS RA, ET AL.: *Nature* 305:779–784, 1983

AN INTERNATIONAL CONFERENCE ON QUANTITATIVE GENETICS will be held May 31–June 5, 1987, in Raleigh, North Carolina. To be placed on the mailing list for future announcements or to seek further information about the conference, please write to: Dr. B. S. Weir, Conference Secretary, Department of Statistics, North Carolina State University, Box 8203, Raleigh, NC 27695-8203.