

Molecular Cloning of $\alpha 5(\text{IV})$ Collagen and Assignment of the Gene to the Region of the X Chromosome Containing the Alport Syndrome Locus

Jeanne C. Myers,* Tania A. Jones,† Eija-Riitta Pohjolainen,‡ Attia S. Kadri,*
Audrey D. Goddard,† Denise Sheer,† Ellen Solomon,† and Taina Pihlajaniemi‡

*Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, Philadelphia; †Imperial Cancer Research Fund, Lincoln's Inn Fields, London; and ‡Collagen Research Unit, Biocenter and Department of Medical Biochemistry, University of Oulu, Oulu, Finland

Summary

Type IV collagen is a major structural component of basement membranes. Four constituent polypeptides have been described and characterized to different degrees. Whereas the primary structure of the $\alpha 1(\text{IV})$ and $\alpha 2(\text{IV})$ chains has been completely established, only short protein sequences have been reported for the recently recognized $\alpha 3(\text{IV})$ and $\alpha 4(\text{IV})$ subunits. We have isolated overlapping human cDNA clones whose derived amino acid sequence is highly homologous to the $\alpha 1(\text{IV})$ and $\alpha 2(\text{IV})$ chains. However, these clones code for neither $\alpha 3(\text{IV})$ nor $\alpha 4(\text{IV})$, and thus this new polypeptide has been designated the $\alpha 5$ chain of type IV collagen. To determine whether the gene encoding the $\alpha 5(\text{IV})$ chain is syntenic with the contiguously arranged $\alpha 1(\text{IV})$ and $\alpha 2(\text{IV})$ genes at 13q34, the $\alpha 5(\text{IV})$ cloned DNA was hybridized to genomic DNA from somatic cell hybrids and to metaphase chromosomes. The results demonstrated that the $\alpha 5(\text{IV})$ collagen gene is located on the long arm of the X chromosome. Since 14 collagen genes have previously been assigned to nine autosomes, these data represent the first mapping of a collagen gene to the X chromosome. Most important, the $\alpha 5(\text{IV})$ gene has been sublocalized to bands Xq22→q23, which are in the same region known to contain the locus for the X-linked form of Alport syndrome. It is therefore possible that this severe dominantly inherited nephritis, manifested by splitting of the glomerular basement membrane, could be caused by mutations in the $\alpha 5(\text{IV})$ collagen gene.

Introduction

The collagens comprise a large family of triple-helical macromolecules primarily responsible for the structural and biological integrity of connective tissue (reviewed by Miller and Gay [1987]; Burgeson 1988). These proteins exhibit surprising heterogeneity in size, structure, and tissue distribution. Thirteen different types have already been formally classified, while others are at earlier stages of characterization. Reflecting the diversity in structural characteristics, collagen types I–XIII can

be subdivided into three categories essentially as described by Miller and Gay (1987). Group 1 molecules, types I, II, III, V, and XI, are the “classical” collagens formed by 95-kDa or greater polypeptide chains with continuous $(\text{Gly-X-Y})_n$ domains. Group 2 collagens, types IV, VI, VII, and XII, also contain large polypeptide chains, but the triple-helical domains are interrupted by nonhelical segments, while group 3 molecules, types VIII, IX, X, and XIII, are known as short-chain collagens. Some collagen types consist of three identical α chains (e.g., types II and III), while others consist of two or more genetically distinct but similar polypeptides (e.g., types I, IV, and V). Moreover, the molecular composition within a collagen type can vary considerably, as is found with types IV and V.

The polypeptide chains of collagen types I–XIII are encoded by a minimum of 24 genes, which are widely

Received December 20, 1989; revision received February 14, 1990.

Address for correspondence and reprints: Jeanne C. Myers, Ph.D., Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, Philadelphia, PA 19104-6059.

© 1990 by The American Society of Human Genetics. All rights reserved. 0002-9297/90/4606-0003\$02.00

distributed throughout the genome. Fourteen genes have thus far been mapped and assigned to nine autosomes: 1, 2, 6, 7, 10, 12, 13, 17, and 21. Original speculation that genes under tight coordinate control might be clustered was replaced by suggestions of random dispersion after the first several collagen genes were mapped (reviewed by Myers and Emanuel [1987]). It is now clearly established that both extremes exist. The two genes coding for the type I collagen chains, $\alpha 1(I)$ and $\alpha 2(I)$, are segregated on chromosomes 17 and 7, respectively (Myers and Emanuel 1987), whereas the $\alpha 1(IV)$ and $\alpha 2(IV)$ genes are contiguously arranged (Pöschl et al. 1988; Soinin et al. 1988) at 13q terminus (Griffin et al. 1987; Killen et al. 1987; Boyd et al. 1988). Current data also indicate tight linkage at two other loci—the $\alpha 1(III)$ and $\alpha 2(V)$ genes at 2q24.3→q31 (Emanuel et al. 1985; Tsipouras et al. 1988) and the $\alpha 1(VI)$ and $\alpha 2(VI)$ genes at distal 21q (Weil et al. 1988). In contrast, the $\alpha 3(VI)$ gene is found on a separate chromosome, 2q37 (Weil et al. 1988), and two genes coding for type XI collagen, $\alpha 1(XI)$ and $\alpha 2(XI)$, are unlinked on the short arms of chromosomes 1 and 6 (Henry et al. 1988; Hanson et al. 1989; Kimura et al. 1989a). Additional genomic regions identified as containing single collagen loci are 6q, 10q, and 12q, to which the $\alpha 1(IX)$, $\alpha 1(XIII)$, and $\alpha 1(II)$ genes, respectively, have been assigned (Kimura et al. 1989b; Pajunen et al. 1989; Shows et al. 1989).

Recently, we isolated a unique human cDNA clone encoding a previously unknown collagen chain. The 3-kb nucleotide sequence corresponds solely to repeating Gly-X-Y triplets with interruptions identical in distribution to those in the basement membrane collagen $\alpha 1(IV)$ and $\alpha 2(IV)$ chains (Hostikka and Tryggvason 1988). Analysis of 3' overlapping clones encoding the junction sequence of the collagenous and COOH-terminal noncollagenous (NC1) domains surprisingly excluded the possibility that the clones coded for either one of the recently recognized $\alpha 3(IV)$ or $\alpha 4(IV)$ chains (Butkowski et al. 1987; Saus et al. 1988). The new collagen chain, which exhibits over 80% homology to $\alpha 1(IV)$ in the NC1 domain (Brinker et al. 1985; Pihlajaniemi et al. 1985), has therefore been designated the $\alpha 5$ chain of type IV collagen. To determine whether the $\alpha 5(IV)$ gene is syntenic with $\alpha 1(IV)$ and $\alpha 2(IV)$ at 13q34, we mapped the gene using somatic cell hybrids and by *in situ* hybridization. Our results revealed the first localization of a collagen gene to the X chromosome. Most important, the $\alpha 5(IV)$ gene has been cytologically assigned to the region which includes the locus for the X-linked form of Alport syndrome (Atkin

et al. 1988; Brunner et al. 1988; Flinter et al. 1989), a dominantly inherited progressive nephritis characterized by splitting of the glomerular basement membrane (Rumpelt 1980; Yoshikawa et al. 1981).

Material and Methods

Collagen $\alpha 5(IV)$ DNA Hybridization Probes

A detailed description of the human $\alpha 5(IV)$ collagen cDNA clones will be presented elsewhere. The clones were isolated from human placenta and umbilical vein endothelial cell λ gt11 cDNA libraries (Clontech Laboratories). The original cDNA clone, PF17, is 3 kb in length and contains three internal *EcoRI* sites. Two *EcoRI* subclones of PF17 were used for the chromosomal mapping studies—namely, a 0.68-kb 5' *EcoRI* fragment coding for the junction of the 7S domain and the NH₂-terminal part of the collagenous region, and the adjacent 1.2-kb *EcoRI* fragment coding for 409 residues of the (Gly-X-Y)_n region.

Two overlapping 3' clones, PF6 and HE6, were subsequently isolated to allow sequence comparison with the 27 and 17 residues previously reported for the $\alpha 3(IV)$ and $\alpha 4(IV)$ collagen chains, respectively (Butkowski et al. 1987; Saus et al. 1988).

Cell Lines and Hybrids

All of the human and rodent lines, as well as the human × rodent hybrids CTP34B4, CTP412A2, Horp9.5, DT1.2.4, Dur4.3, F4SC13C112, Mog34A4, FG10, Mog1B/9, Sif4A31, Sir74ii, Twin19/D12, and 3W4C15, have been described elsewhere (VanHeyningen et al. 1975; Jones et al. 1976; Povey et al. 1980; Kielty et al. 1982; Edwards et al. 1985; Philips et al. 1985; Solomon et al. 1985; Wong et al. 1987).

Southern Blot Hybridization of $\alpha 5(IV)$ Clones to Rodent × Human Hybrids

DNA (7 μ g) from human and rodent parental cell lines and human × rodent hybrids was digested to completion with the restriction endonuclease *HindIII* or *BamHI*. The DNA samples were then fractionated by electrophoresis on 0.8% agarose gels and alkaline blotted onto Hybond N+ (Amersham International) membranes (Reed and Mann 1985) after depurination by a 10-min treatment in 0.25 N HCl.

The 0.68- or 1.2-kb $\alpha 5(IV)$ PF17 cDNA fragment was labeled with [α -³²P]dCTP by random-primer labeling (Feinberg and Vogelstein 1983) to a specific activity of 5×10^8 – 1×10^9 cpm/ μ g. Hybridization of

the filter-bound DNAs to 2×10^6 cpm/ml of the labeled probes was carried out in 0.6 M NaCl, 0.04 M sodium phosphate dibasic, 0.004 M EDTA, 0.5% low-fat dry milk, 1% SDS, 10% dextran sulfate, 50% formamide and 1 mg yeast RNA/ml at 42°C. After hybridization, filters were washed at 65°C in 0.1% SDS and $0.1 \times$ or $0.2 \times$ SSC ($1 \times$ SSC = 0.15 M NaCl, 0.015 M sodium citrate, pH 7.0).

In Situ Hybridization of $\alpha 5(\text{IV})$ Clones to Metaphase Chromosomes

Human lymphocytes obtained from normal female 46,XX blood were cultured with phytohemagglutinin for 72 h at 37°C (Watt and Stephens 1986). The cells were washed with fresh medium 16–17 h later and incubated in thymidine-rich (10^{-5} M) medium for an additional 6–7 h. Harvesting of the cells and all further procedures were carried out in subdued lighting.

In situ hybridization was performed essentially as described by Harper and Saunders (1981) and Zabel et al. (1983). DNA probes were labeled to a specific activity of 3×10^8 cpm/ μg with ^3H deoxynucleotides and oligonucleotide primers (Feinberg and Vogelstein 1983). Hybridizations were carried out for 18–22 h at 37°C in 50% formamide, 0.24 M sodium chloride, 0.03 M sodium citrate, 0.04 M sodium phosphate, 10% dextran sulphate, and 100 μg salmon sperm DNA/ml at pH 7.0. In one experiment, the 5' 0.68-kb *Eco*RI fragment of PF17 was present at a concentration of 0.1 $\mu\text{g}/\text{ml}$, and in the subsequent experiment, both the 5' fragment and the adjacent 1.2-kb fragment of PF17 were included at a concentration of 0.02 $\mu\text{g}/\text{ml}$. After incubation, slides were washed in 50% formamide, 0.3 M

sodium chloride, 0.03 M sodium citrate, pH 7.0 at 39°C, dehydrated, and immersed in Ilford K5 emulsion. Slides were developed after 1–2 wk and G-banded using a variation of the method reported by Wolff and Perry (1974). Slides were stained for 30 min in 0.5 μg of the fluorescent stain Hoechst 33258 per milliliter, exposed to long-wave ultraviolet light for 15 min, and stained with Wright's stain.

Results

Identification of Human cDNA Clones Coding for the $\alpha 5(\text{IV})$ Collagen Chain

The cDNA clones PF17 and PF6 were isolated from a $\lambda\text{gt}11$ human placenta library while the 3' overlapping clone HE6 was later isolated from a $\lambda\text{gt}11$ human umbilical vein endothelial cell library (fig. 1). DNA sequencing revealed that these clones encoded a collagenous region highly homologous to the type IV polypeptides. By analogy to the $\alpha 1(\text{IV})$ collagen chain (Hostikka and Tryggvason 1988), the 1,013 amino acid residues derived from PF17 span residues 85–1080. PF6 begins at residue 895 and terminates at residue 1196. Clone HE6 overlaps PF6 by 221 nucleotides and encodes the remaining 244 residues of the collagenous region and 159 amino acids of the COOH-terminal noncollagenous domain, NC1 (our unpublished data).

Whereas the sequences of the human $\alpha 1(\text{IV})$ and $\alpha 2(\text{IV})$ collagen chains have been completely determined from DNA clones (Hostikka and Tryggvason 1988), only a few residues of the recently identified $\alpha 3(\text{IV})$ and $\alpha 4(\text{IV})$ chains have been established by protein anal-

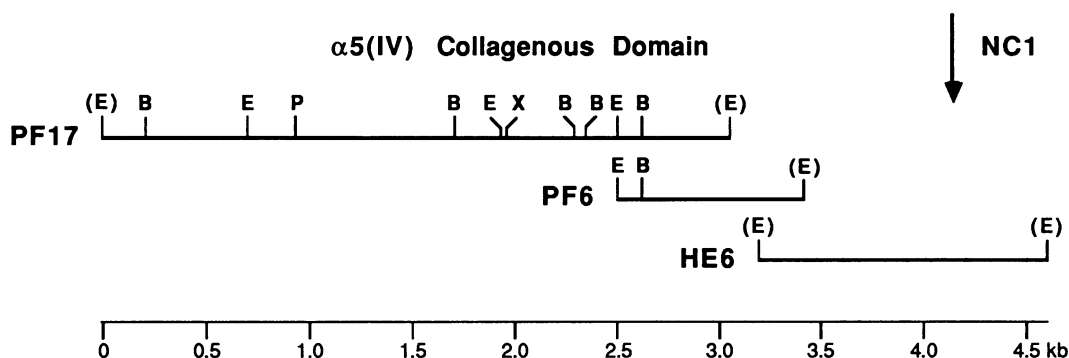


Figure 1 Cloned cDNAs encoding the human $\alpha 5(\text{IV})$ collagen chain. The arrow indicates the junction of the collagenous and COOH-terminal noncollagenous (NC1) domain. E = *Eco*RI restriction-endonuclease site, (E) = *Eco*RI linker sequence added during cloning. B = *Bam*HI restriction-endonuclease site; P = *Pst*I restriction-endonuclease site; X = *Xba*I restriction-endonuclease site. The nucleotide scale (kb) is shown at the bottom. The probes used in the mapping studies were the 5' 0.68-kb and the adjacent 1.2-kb *Eco*RI fragments of PF17.

A																
	Collagenous Residues															
$\alpha 5(\text{IV})$	Gly	Pro	Asp	Gly	Leu	Gln	Gly	Pro	Pro	Gly	Pro	Pro	Gly	Thr	Ser	
$\alpha 1(\text{IV})$	-	-	-	-	-	Pro	-	Ser	Met	-	-	-	-	-	Pro	↓
$\alpha 2(\text{IV})$	-	Arg	Pro	-	Ser	Pro	-	Leu	-	-	Met	-	-	Arg	-	↓
	Non-Collagenous (NC1) Residues															
$\alpha 5(\text{IV})$	Ser	Val	Ala	His	Gly	Phe	Leu	Ile	Thr	Arg	His	Ser	Gln	Thr	Thr	
$\alpha 1(\text{IV})$	-	-	Asp	-	-	-	-	Val	-	-	-	-	-	-	Ile	
$\alpha 2(\text{IV})$	Δ	-	Ser	Ile	-	Tyr	-	Leu	Val	Lys	-	-	-	-	Asp	
B																
	Collagenous Residues															
$\alpha 5(\text{IV})$	Gly	Pro	Asp	Gly	Leu	Gln	Gly	Pro	Pro	Gly	Pro	Pro	Gly	Thr	Ser	↓
$\alpha 3(\text{IV})$	-	Leu	Xaa	-	Lys	Pro	-	Asp	Thr	-	-	-	↓Ala	Ala	Gly	↓
$\alpha 4(\text{IV})$							-	-	-	-	Phe	Δ	-	↓Pro	Gly	↓
	Non-Collagenous (NC1) Residues															
$\alpha 5(\text{IV})$	Ser	Val	Ala	His	Gly	Phe	Leu	Ile	Thr	Arg	His	Ser	Gln	Thr	Thr	
$\alpha 3(\text{IV})$	Ala	-	Met	Arg	-	-	Val	Phe	-	-	-	Tyr	-	-	-	
$\alpha 4(\text{IV})$	Δ	Tyr	Leu	Ser	-	-	-	Leu	Val	Leu	-	-	-	-	-	

Figure 2 Comparison of the human $\alpha 5(\text{IV})$ collagen/noncollagen junction with those of the human $\alpha 1(\text{IV})$ and $\alpha 2(\text{IV})$ chains and bovine $\alpha 3(\text{IV})$ and $\alpha 4(\text{IV})$ chains. *A*, Sequences of the COOH-terminal end of the collagenous domain (upper three rows) and the beginning of the NC1 domain (lower three rows) in $\alpha 5(\text{IV})$, $\alpha 1(\text{IV})$, and $\alpha 2(\text{IV})$. In each group of three rows, the first row shows the human $\alpha 5(\text{IV})$ sequence, the second row the human $\alpha 1(\text{IV})$ sequence, and the third row the human $\alpha 2(\text{IV})$ sequence. *B*, Sequences at the junction of the two domains in $\alpha 5(\text{IV})$, $\alpha 3(\text{IV})$, and $\alpha 4(\text{IV})$. In each group of three rows, the first row shows the human $\alpha 5(\text{IV})$ sequence, the second row the bovine $\alpha 3(\text{IV})$ sequence, and the third row the bovine $\alpha 4(\text{IV})$ sequence. The junctions of the two domains are indicated by arrows. Dashes designate amino acids in the $\alpha 1(\text{IV})$, $\alpha 2(\text{IV})$, $\alpha 3(\text{IV})$, and $\alpha 4(\text{IV})$ chains that are identical with the $\alpha 5(\text{IV})$ residues. Gaps of one amino acid (Δ) have been introduced in the $\alpha 2(\text{IV})$ and $\alpha 4(\text{IV})$ chains for maximal alignment of the different chains. The X denotes an unidentified amino acid in the $\alpha 3(\text{IV})$ chain.

ysis (Butkowski et al. 1987; Saus et al. (1988). The 27 residues of $\alpha 3(\text{IV})$ and 17 residues of $\alpha 4(\text{IV})$, corresponding to the end of the collagenous region and the beginning of the NC1 domain, show only 48% and 42% homology, respectively, to the HE6-derived sequence (fig. 2). Although a similar low degree of homology is found when the HE6-derived sequence is aligned with the $\alpha 2(\text{IV})$ junction sequence (Griffin et al. 1987; Myers et al. 1987) there is a 79% identity between HE6 and $\alpha 1(\text{IV})$ (Brinker et al. 1985; Pihlajaniemi et al. 1985) (fig. 2). It is therefore suggested that the polypeptide encoded by clones PF17, PF6, and HE6 should be designated the $\alpha 5$ chain of type IV collagen.

Southern Blot Hybridization of the $\alpha 5(\text{IV})$ Collagen cDNA Clone to Human \times Rodent Somatic Cell Hybrids

To localize the $\alpha 5(\text{IV})$ collagen gene, a panel of rodent \times human somatic cell hybrids were analyzed by Southern blot hybridization with cloned cDNAs. Human, rodent, and hybrid genomic DNAs were digested with either *Bam*HI or *Hind*III. Two different cDNA fragments, of 0.68 and 1.2 kb, of the cDNA clone PF17 (fig. 1) were used as probes. Using the 0.68-kb probe, the unique human bands were approximately 9.7 and 5 kb with *Bam*HI, and 10 and 5.5 kb with *Hind*III. Using the 1.2-kb probe, the human bands were approximately 12 and 5.5 kb with *Bam*HI, and 7.4, 5.4, 4.2

Table 1

Segregation of the $\alpha 5$ (IV) Collagen Gene with Human Chromosomes in Human \times Rodent Hybrid Cell DNA

A.

HYBRID	HUMAN CHROMOSOME																						X
	$\alpha 5$ (IV)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
CTP34B4	+	+	+	+	-	+	+	+	+	-	-	+	-	+	+	-	+	+	+	-	-	-	+
CTP412A2	+	+	+	-	-	+	+	+	-	-	-	-	-	+	+	-	-	+	+	-	-	-	+
DI1.2.4	+	-	+	-	-	-	-	-	-	-	+	-	-	-	-	+	+	+	+	-	+	+	+
Dur4.3	+	-	+	-	-	+	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+
FG10	+	+	+	-	-	+	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	+
F4SC13C112	-	+	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	+
Horp9.5	+	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Mog13/9	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+
Mog34A4	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Sif4A31	+	-	-	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+
Sir74ii	+	+	+	+	+	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+
Twin19/D12	-	+	+	+	+	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
3W4C15	+	-	-	-	-	-	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+

B.

HYBRID TYPE (marker/chromosome)	HUMAN CHROMOSOME																						X
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	
Concordant:																							
+ / +	3	4	7	3	5	4	4	3	0	5	5	6	3	8	4	3	6	6	1	2	6	4	11
- / -	0	2	1	1	1	2	2	1	1	2	1	1	2	0	2	1	1	1	2	1	2	1	1
Discordant:																							
+ / +	8	7	4	8	6	7	7	8	11	6	6	5	8	3	7	8	5	5	10	9	5	7	0
- / -	2	0	1	1	1	0	0	1	1	0	1	1	0	2	0	1	1	1	0	1	0	1	1 ^b
Discordancy (%) ^c	77	54	38	69	54	54	54	69	92	46	46	46	62	38	54	69	46	46	77	77	38	62	8

NOTE.—In section A, + = presence of chromosome in the hybrid; - = absence of chromosome. In section B, values are numbers of hybrids.

^a Presence of X-chromosome inferred by G6PD enzyme assay only; thus, this hybrid may contain trace amounts of X or fragments of X.

^b Hybrid contains trace amount of chromosome.

^c Discordancy percentage was calculated by dividing the number of discordant hybrids by the total number of hybrids analyzed.

and 1.2 kb with *Hind*III (not shown). The panel contained human \times rat, human \times mouse, and human \times hamster hybrids. In all cases, the corresponding rodent parental line was used with each hybrid.

Table 1 shows the results of experiments designed to determine the cosegregation frequency of the $\alpha 5(\text{IV})$ collagen gene locus with various human chromosomes. Results were identical for both enzyme digests. The hybrids were scored for the presence (+) or absence (-) of the $\alpha 5(\text{IV})$ human bands. The chromosome content of each hybrid except one (see below) had been established by a combination of karyotyping as well as enzyme and DNA markers representative of each human chromosome. The $\alpha 5(\text{IV})$ collagen gene was scored as concordant or discordant with respect to the presence or absence of each chromosome. These values are presented for each chromosome at the bottom of table 1 except for hybrids containing only the Y chromosome, which were later found to be negative (data not shown). Every chromosome except X had at least five discordancies. All hybrids containing the X chromosome were positive for human bands, including the X-only hybrid MOG13/9. One hybrid which was scored negative for the $\alpha 5(\text{IV})$ bands is indicated as having the human X based upon its being positive for G6PD. However, this hybrid had not been karyotyped or tested for other X-linked markers after being regrown for the current set of experiments and may therefore contain only a fragment of the X. Otherwise, results from the hybrids show unequivocally that the $\alpha 5(\text{IV})$ collagen gene is located on the X chromosome.

Chromosomal Localization of the $\alpha 5(\text{IV})$ Collagen Gene by In Situ Hybridization

The $\alpha 5(\text{IV})$ collagen gene was independently mapped by in situ hybridization of DNA clones to human metaphase chromosomes. Initially, the 5' 0.68-kb *Eco*RI fragment of PF17 (fig. 1) was used as the probe. In this experiment, 42 metaphase spreads were examined, and 215 grains were found on chromosomes. Of these, 37 grains (17.2%) were distributed on the X chromosome with a large fraction, 26 grains (70%), at the Xq22 \rightarrow q23 region. A representative autoradiograph is shown in fig. 3.

A subsequent in situ hybridization was performed by combining the 5' probe with the adjacent 1.2-kb *Eco*RI fragment of PF17 (fig. 1): together, they code for 1.9 kb of contiguous $\alpha 5(\text{IV})$ cDNA sequences. In this experiment, 58 metaphase spreads were analyzed, and 216 grains were located on chromosomes. The cumulative distribution of grains in this experiment is il-

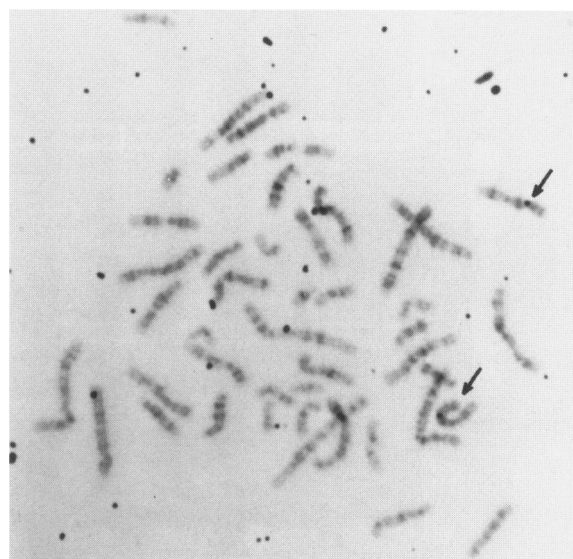


Figure 3 Autoradiograph from in situ hybridization of the $\alpha 5(\text{IV})$ cDNA clone to metaphase chromosomes. Arrows indicate grains on the long arms of both chromosome X homologs.

lustrated in figure 4. Consistent with the results of the previous experiment, 46 grains (21.3%) were scored on the X chromosome. Of these, 29 grains (63%) were found at bands Xq22 \rightarrow q23. Sublocalization of the $\alpha 5(\text{IV})$ collagen gene to this region of the X chromosome is diagrammatically shown in fig. 5.

Discussion

We have presented data showing the existence of the $\alpha 5$ chain of type IV collagen by molecular cloning and have mapped the gene to the q22 \rightarrow q23 region of the X chromosome involved in the X-linked form of Alport-type familial nephritis (Alport 1927). To put these results in perspective, it is necessary to introduce some information on the structure and function of type IV collagen and its proposed role in kidney diseases.

Basement membranes separate epithelial, endothelial, and parenchymal cells from the interstitial connective tissue, and are thought to have major roles in several biological processes, such as tissue morphogenesis, cell migration, and filtration of macromolecules (reviewed by Martinez-Hernandez and Amenta [1983]). Type IV collagen is exclusively located in basement membranes where it interacts with noncollagenous matrix glycoproteins to form a supramolecular aggregate needed to maintain the architecture of the thin, sheetlike structures (reviewed by Timpl [1989]). Until lately, the mo-

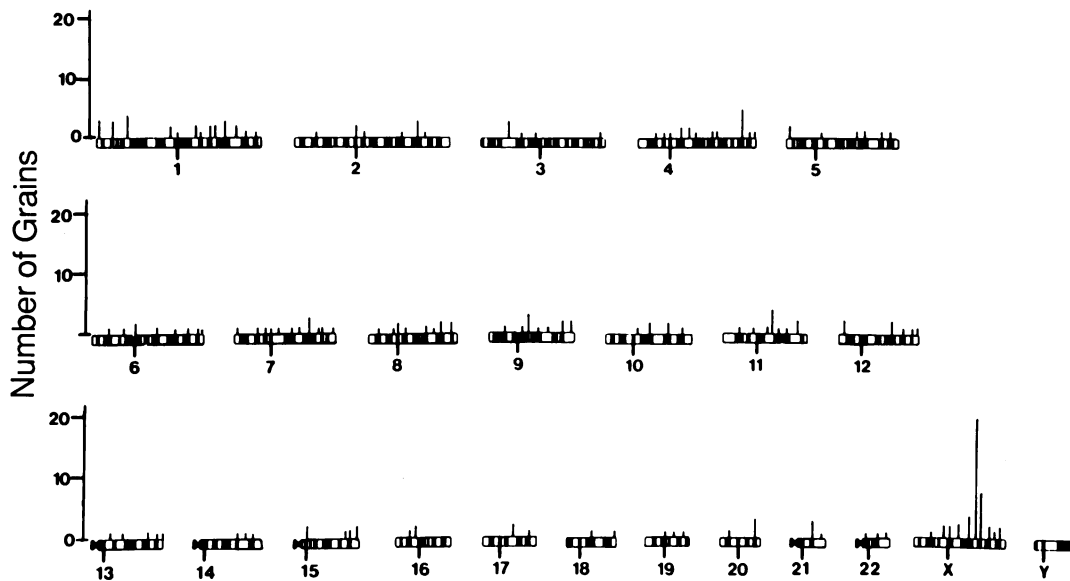


Figure 4 Histogram of grains distributed on metaphase chromosomes from hybridization of PF17 5' 0.68-kb and adjacent 3' 1.2-kb fragments. The abscissa represents the chromosomes in their relative size proportions; the ordinate shows the number of silver grains.

lecular composition of type IV collagen was believed to be $\alpha1(IV)_2\alpha2(IV)$. Each polypeptide, which consists of about 1,700 residues, is comprised of three domains: the NH₂-terminal intermolecular cross-link region, the central helical region containing Gly-X-Y triplet inter-

ruptions, and the COOH-terminal globular region. Physical and chemical properties of the latter domain provided the criteria for the recent identification of two new type IV collagen chains named $\alpha3$ and $\alpha4$. Very little sequence information is currently available on these two chains, and what there is is confined to the last few Gly-X-Y triplets and the adjacent part of the COOH-terminal NC1 domain (Butkowski et al. 1987; Saus et al. 1988). The $\alpha3$ chain has generated particular interest with the knowledge that its NC1 domain contains the antigenic determinant of Goodpasture syndrome, an autoimmune disorder characterized by glomerulonephritis and lung hemorrhage (reviewed by Hudson et al. [1989]). Results of several immunohistochemical studies suggest a relationship between Goodpasture syndrome and Alport-type familial nephritis, a genetically heterogeneous renal disease frequently associated with sensorineural deafness (Feingold et al. 1985; Hasstedt et al. 1986). An absence of the $\alpha3(IV)$ NC1 component from renal basement membrane of Alport patients has been reported (Kleppel et al. 1987), whereas in other studies a partial rather than complete loss of the Goodpasture antigen has been found (Kleppel et al. 1989; Savage et al. 1989). The actual connection between the $\alpha3(IV)$ chain and Alport syndrome, therefore, remains to be established and may reflect the added complexity of the existence of type IV collagen molecules containing the $\alpha5(IV)$ chain.

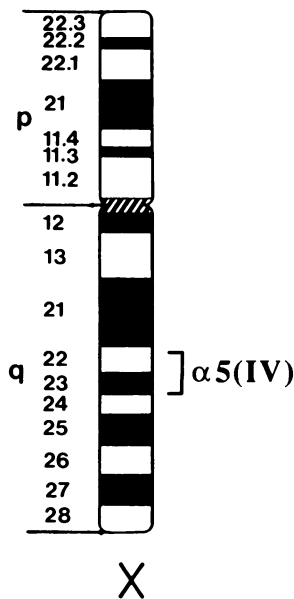


Figure 5 Idiogram of the X chromosome. The bracket shows the region of the X chromosome, q22→q23, where the $\alpha5(IV)$ collagen gene is located.

Whether the few cases of autosomal dominant and

rare instances of autosomal recessive modes of inheritance actually belong to the Alport syndrome classification remains controversial. The X-linked form of transmission explains at least the vast majority of pedigrees (Kashtan and Michael 1989). One of the main difficulties in genetic analysis of Alport families stems from the lack of or severely reduced number of offspring of affected young males, who often die early of terminal renal failure. Regardless of the form of inheritance, the disease is manifested by ultrastructural glomerular abnormalities including splitting of the basement membrane's lamina densa (Rumpelt 1980; Yoshikawa et al. 1981). Generally, affected males display progressive pathological changes, while mild expression of the disease is seen in females. Recently, results of linkage studies employing anonymous DNA markers confirmed X-linked inheritance in several types of Alport syndrome and assigned the gene to the middle of the long arm of the X chromosome. In two independent RFLP analyses, the authors suggested that the Alport locus is distal to a DNA marker located at Xq21.3→q22 (Atkin et al. 1988; Flinter et al. 1989), and, in another study using the same approach, the Alport gene was mapped to Xq21.2→q22.2 (Brunner et al. 1988).

The clinical, biochemical, and immunological findings are all compatible with Alport syndrome's being due to a mutation in type IV collagen (Kleppel et al. 1987, 1989; Savage et al. 1989). However, involvement of the genes encoding the $\alpha 1(\text{IV})$ and $\alpha 2(\text{IV})$ chains seemed unlikely because of the restricted location of the basement membrane lesions in the affected individuals. The possibility of altered $\alpha 1(\text{IV})$ and $\alpha 2(\text{IV})$ chains was completely dismissed once the autosomal locus of these genes at 13q34 was learned (Emanuel et al. 1986; Griffin et al. 1987; Killen et al. 1987; Boyd et al. 1988). The $\alpha 5(\text{IV})$ chain is a prime candidate for mutations causing the Alport syndrome, especially considering the chromosomal assignment of the corresponding gene to Xq22→q23. A clearer picture awaits identification of RFLPs associated with the $\alpha 5(\text{IV})$ locus and screening of Alport kindreds. It is also important to establish whether either of the genes coding for the $\alpha 3(\text{IV})$ or $\alpha 4(\text{IV})$ chains is located on the X chromosome and whether they are arranged in tandem with the $\alpha 5(\text{IV})$ gene similar to the organization of the $\alpha 1(\text{IV})$ and $\alpha 2(\text{IV})$ genes on chromosome 13.

Acknowledgments

This work was supported by grants AM20553 and HL41882 from the National Institutes of Health and by grants

from the Medical Research Council of the Academy of Finland. We are grateful to Aila Jokinen and Liisa Äijälä for expert technical assistance and to John Hoyer, Kari Kivirikko, and Billy Hudson for constructive discussions.

References

- Alport AC (1927) Hereditary familial congenital haemorrhagic nephritis. *Br Med J [Clin Res]* 1:504–506
- Atkin CL, Hasstedt SJ, Menlove L, Cannon L, Kirschner N, Schwartz C, Nguyen K, et al (1988) Mapping of Alport syndrome to the long arm of the X chromosome. *Am J Hum Genet* 42:249–255
- Boyd CD, Toth-Fejel S, Gadi IK, Litt M, Condon MR, Kolbe M, Hagen IK, et al (1988) The genes coding for human pro $\alpha 1(\text{IV})$ collagen and pro $\alpha 2(\text{IV})$ collagen are both located at the end of the long arm of chromosome 13. *Am J Hum Genet* 42:309–314
- Brinker JM, Gudas LJ, Loidl H, Wang SY, Rosenbloom J, Kefalides NA, Myers JC (1985) Restricted homology between human $\alpha 1$ type IV and other procollagen chains. *Proc Natl Acad Sci USA* 82:3649–3653
- Brunner H, Schröder C, van Bennekom C, Lambermon E, Tuerlings J, Menzel D, Olbing H, et al (1988) Localization of the gene for X-linked Alport's syndrome. *Kidney Int* 34:507–510
- Burgeson RE (1988) New collagens, new concepts. *Annu Rev Cell Biol* 4:551–577
- Butkowski RJ, Langeveld JPM, Wieslander J, Hamilton J, Hudson BG (1987) Localization of the Goodpasture epitope to a novel chain of basement membrane collagen. *J Biol Chem* 262:7874–7877
- Edwards YH, Parker M, Povey S, West LF, Farrington FM, Solomon E (1985) Human myosin heavy chain genes assigned to chromosome 17 using a human cDNA as probe. *Ann Hum Genet* 44:101–109
- Emanuel BS, Cannizzaro LA, Seyer JM, Myers JC (1985) Human $\alpha 1(\text{III})$ and $\alpha 2(\text{V})$ procollagen genes are located on the long arm of chromosome 2. *Proc Natl Acad Sci USA* 82:3385–3389
- Emanuel BS, Sellinger BT, Gudas LJ, Myers JC (1986) Localization of the human procollagen $\alpha 1(\text{IV})$ gene to chromosome 13q34 by in situ hybridization. *Am J Hum Genet* 38:38–44
- Feinberg AP, Vogelstein B (1983) A technique for radiolabeling DNA fragments to high specific activity. *Anal Biochem* 132:6–13 and addendum (1984) 137:117–126
- Feingold J, Bois E, Chompret A, Broyer M, Gubler M-L, Grunfeld GP (1985) Genetic heterogeneity of Alport syndrome. *Kidney Int* 27:672–677
- Flinter FA, Abbs S, Bobrow M (1989) Localization of the gene for classic Alport syndrome. *Genomics* 4:335–338
- Griffin CA, Emanuel BS, Hansen JR, Cavenee WK, Myers JC (1987) Human collagen genes encoding basement mem-

- brane $\alpha 1(IV)$ and $\alpha 2(IV)$ chains map to the distal long arm of chromosome 13. *Proc Natl Acad Sci USA* 84:512–516
- Hanson IM, Gorman P, Liu VCH, Cheah KSE, Solomon E, Trowsdale J (1989) The human $\alpha 2(XI)$ collagen gene (COL11A2) maps to the centromeric border of the major histocompatibility complex on chromosome 6. *Genomics* 5:925–931
- Harper ME, Saunders GF (1981) Location of single copy DNA sequences of G-banded human chromosomes by *in situ* hybridization. *Chromosoma* 83:431–439
- Hasstedt SJ, Atkin CL, San Juan AC (1986) Genetic heterogeneity among kindreds with Alport syndrome. *Am J Hum Genet* 38:940–953
- Henry I, Bernhein A, Bernard M, van der Rest M, Kimura T, Jeanpierre C, Barichard F, et al (1988) Mapping of a human fibrillar collagen gene, pro $\alpha 1(XI)$ (COL11A1), to the p21 region of chromosome 1. *Genomics* 3:81–90
- Hostikka SL, Tryggvason K (1988) The complete primary structure of the $\alpha 2$ chain of human type IV collagen and comparison with the $\alpha 1(IV)$ chain. *J Biol Chem* 263:19488–19493
- Hudson BG, Wieslander J, Wisdom BJ, Noelken ME (1989) Biology of disease Goodpasture syndrome: molecular architecture and function of basement membrane antigen. *Lab Invest* 61:256–269
- Jones EA, Goodfellow PN, Kennet RH, Bodmer WF (1976) The independent expression of HLA and B2 microglobulin on human-mouse hybrids. *Somat Cell Mol Genet* 2:483–496
- Kashtan CE, Michael AF (1989) Hereditary nephritis. *Semin Nephrol* 9:135–146
- Kielty CM, Povey S, Hopkinson DA (1982) Regulation of expression of liver specific enzymes. III. Further analysis of a series of rat hepatoma and human somatic cell hybrids. *Ann Hum Genet* 46:307–327
- Killen PD, Francomano CA, Yamada Y, Modi WS, O'Brien SJ (1987) Partial structure of the human $\alpha 2(IV)$ collagen chain and chromosomal localization of the gene (COL4A2). *Hum Genet* 77:318–324
- Kimura T, Cheah KSE, Chan SDH, Lui VCH, Mattei M-G, van der Rest M, Ono K, et al (1989a) The human $\alpha 2(XI)$ collagen (COL11A2) chain. *J Biol Chem* 264:13910–13916
- Kimura T, Mattei M-G, Stevens JW, Goldring MB, Ninomiya Y, Olsen BR (1989b) Molecular cloning of rat and human type IX collagen cDNA and localization of the $\alpha 1(IX)$ gene on the human chromosome 6. *Eur J Biochem* 179:71–78
- Kleppel MM, Kashtan CE, Butkowski RJ, Fish AJ, Michael AF (1987) Alport familial nephritis: absence of 28 kilodalton non-collagenous monomer of type IV collagen in glomerular basement membrane. *J Clin Invest* 80:263–266
- Kleppel MM, Kashtan C, Santi PA, Wieslander J, Michael AF (1989) Distribution of familial nephritis antigen in normal tissue and renal basement membranes of patients with homozygous and heterozygous Alport familial nephritis. *Lab Invest* 61:278–289
- Martinez-Hernandez A, Amenta PS (1983) The basement membrane in pathology. *Lab Invest* 48:656–677
- Miller EJ, Gay S (1987) The collagens: an overview and update. *Methods Enzymol* 144:3–41
- Myers JC, Emanuel BS (1987) Chromosomal localization of human collagen genes. *Coll Relat Res* 7:149–159
- Myers JC, Howard PS, Jelen AM, Dion AS, Macarack EJ (1987) Duplication of type IV collagen COOH-terminal repeats and species-specific expression of $\alpha 1(IV)$ and $\alpha 2(IV)$ collagen genes. *J Biol Chem* 262:9231–9238
- Pajunen L, Tamminen M, Solomon E, Pihlajaniemi T (1989) Assignment of the gene coding for the $\alpha 1$ chain of collagen type XIII to human chromosome region 10q11–qter. *Cytogenet Cell Genet* 52:190–193
- Philips JR, Shephard EA, Povey S, Davis MB, Kelsey G, Monteiro M, West LF, et al (1985) A cytochrome P-450 gene family mapped to human chromosome 19. *Ann Hum Genet* 49:267–274
- Pihlajaniemi T, Tryggvason K, Myers JC, Kurkinen M, Lebo R, Cheung MC, Prockop DJ, Boyd CD (1985) cDNA clones coding for the pro $\alpha 1(IV)$ chain of human type IV procollagen reveal an unusual homology of amino acid sequences in two halves of the carboxyl-terminal domain. *J Biol Chem* 260:7681–7687
- Pöschl E, Pollner R, Kühn K (1988) The genes for the $\alpha 1(IV)$ and $\alpha 2(IV)$ chains of human basement membrane collagen type IV are arranged head-to-head and separated by a bidirectional promoter of unique structure. *EMBO J* 7:2687–2695
- Povey S, Jeramiah SJ, Barker RF, Hopkinson DA, Robson EB, Cook PJJ, Solomon E, et al (1980) Assignment of human locus determined phosphoglycolate phosphatase (PGP) to chromosome 16. *Ann Hum Genet* 43:241–248
- Reed KC, Mann DA (1985) Rapid transfer of DNA from agarose gels to nylon membranes. *Nucleic Acids Res* 13:7207–7221
- Rumpelt HJ (1980) Hereditary nephropathy (Alport syndrome): correlation of clinical data with glomerular basement membrane alterations. *Clin Nephrol* 13:203–207
- Saus J, Wieslander J, Langeveld JPM, Quinones S, Hudson BG (1988) Identification of the Goodpasture antigen as the $\alpha 3(IV)$ chain of collagen IV. *J Biol Chem* 263:13374–13380
- Savage COS, Noel L-H, Crutcher E, Prieck SRG, Grunfeld GP, Lockwood RM (1989) Hereditary nephritis: immunoblotting studies of the glomerular basement membrane. *Lab Invest* 60:613–618
- Shows TB, Tikka L, Byers MG, Eddy RL, Haley LL, Henry WM, Prockop DJ, et al (1989) Assignment of the human collagen $\alpha 1(XIII)$ chain gene (COL13A1) to the q22 region of chromosome 10. *Genomics* 5:128–133
- Soininen R, Huotari M, Hostikka SL, Prockop DJ, Tryggvason K (1988) The structural genes for $\alpha 1$ and $\alpha 2$ chains of human type IV collagen are divergently encoded on opposite DNA strands and have an overlapping promoter region. *J Biol Chem* 263:17217–17220

- Solomon E, Hiorns LR, Spurr N, Kurkinen M, Barlow D, Hogan BLM, Dalgleish R (1985) Chromosomal assignment of the genes coding for human types II, III and IV collagen: A dispersed gene family. *Proc Natl Acad Sci USA* 82: 3330–3334
- Timpl R (1989) Structure and biological activity of basement membrane proteins. *Eur J Biochem* 180:487–502
- Tsipouras P, Schwartz RC, Liddell AC, Salkeld CS, Weil D, Ramirez F (1988) Genetic distance of two fibrillar collagen loci, COL3A1 and COL5A2, located on the long arm of human chromosome 2. *Genomics* 3:275–277
- VanHeyningen V, Bobrow M, Bodmer WF, Gardiner SE, Povey S, Hopkinson DA (1975) Chromosome assignment of some human enzyme loci: mitochondrial malate dehydrogenase to 7, mannose phosphate isomerase and pyruvate kinase to 15 and probably, esterase D to 13. *Ann Hum Genet* 38: 295–303
- Watt JL, Stephens GS (1986) Lymphocyte culture for chromosome analysis. In: Rooney DE, Czepulowski BH (eds) *Human cytogenetics: a practical approach*. IRL, Oxford, pp 39–55
- Weil D, Mattei M-G, Passage E, Van Cong N'G, Pribula-Conway D, Mann K, Deutzmann R, et al (1988) Cloning and chromosomal localization of human genes encoding the three chains of type VI collagen. *Am J Hum Genet* 42:435–445
- Wolff SS, Perry P (1974) Differential Giemsa staining of sister chromatids and the study of chromatid exchanges without autoradiography. *Chromosoma* 48:341–353
- Wong Z, Wilson V, Patel I, Povey S, Jeffreys AG (1987) Characterization of a panel of highly variable minisatellites cloned from human DNA. *Ann Hum Genet* 51:269–288
- Yoshikawa N, Cameron AH, White RHR (1981) The glomerular basal lamina in hereditary nephritis. *J Pathol* 135:199–209
- Zabel BU, Naylor SL, Sakaguchi AY, Bell GI, Shows TB (1983) High-resolution chromosomal localization of human genes for amylase proopiomelanocortin, somatostatin and a DNA fragment (D3S1) by *in situ* hybridization. *Proc Natl Acad Sci USA* 80:6932–6936