

## Extensive Homology between the Carboxyl-terminal Peptides of Mouse $\alpha 1(\text{IV})$ and $\alpha 2(\text{IV})$ Collagen\*

(Received for publication, October 9, 1986)

Markku Kurkinen<sup>†§</sup>, Michael R. Condon<sup>§¶</sup>, Bruce Blumberg<sup>||\*\*</sup>, Denise P. Barlow<sup>††§§</sup>,  
Susan Quinones<sup>§</sup>, Juan Saus<sup>§</sup>, and Taina Pihlajaniemi<sup>†¶¶</sup>

From the Departments of <sup>†</sup>Biochemistry and <sup>§</sup>Medicine, University of Medicine and Dentistry of New Jersey-Robert Wood Johnson Medical School, Piscataway, New Jersey 08854, the <sup>||</sup>Department of Biology, University of California, Los Angeles, California 90024, and the <sup>††</sup>Imperial Cancer Research Fund, Mill Hill, London NW7 1AD, Great Britain

We have determined the complete primary structure for the carboxyl-terminal peptides of mouse  $\alpha 1(\text{IV})$  and  $\alpha 2(\text{IV})$  collagen, which have 229 and 227 amino acids, respectively. The amino acid sequences are 63% identical and conservatively substituted in 28 positions. A striking feature of these peptides is that the first half of each sequence is homologous with the second half, 37% in  $\alpha 1(\text{IV})$  and 36% in  $\alpha 2(\text{IV})$ . These results suggest that the carboxyl-terminal peptides of type IV collagen are closely related in their structure and evolution. Presumably, they were first derived by internal duplication of a common ancestral DNA sequence which later, by gene duplication, gave rise to the two different but homologous carboxyl-terminal peptides of type IV collagen.

Type IV collagen is a specific component of basement membranes and contains two genetically distinct polypeptides, the  $\alpha 1(\text{IV})$  and  $\alpha 2(\text{IV})$  chains, each of about 1700 amino acids (1). The subunit composition of type IV collagen may be heterogeneous and vary between different basement membranes (2). For example, type IV collagen from bovine kidney is a heterotrimer of two  $\alpha 1(\text{IV})$  and one  $\alpha 2(\text{IV})$  chains (3), whereas type IV collagen produced by cultured rat parietal yolk sac cells is composed of three  $\alpha 1(\text{IV})$  chains (4). Type IV collagen differs from the fibrillar collagens (types I-III) in that it contains several interruptions in the characteristic Gly-X-Y repeat sequence (5, 6), is not known to be proteolytically processed after secretion, and does not form ordered fibrillar structures. Recently, network models have been proposed for the assembly of type IV collagen. In one model, four molecules are associated at the amino-terminal domains, and the tetramers are linked end-to-end at their carboxyl-terminal domains (7). In another model, type IV collagen molecules are shown to assemble into hexagonal lattices which also

allow their lateral association to each other (8).

In order to understand the assembly and interactions of type IV collagen in more detail, several laboratories have been characterizing amino acid sequences of type IV collagen chains. The complete amino acid sequence has been determined for a 914-residue fragment from the triple-helical domain of human  $\alpha 1(\text{IV})$  collagen (5). A cysteine-rich sequence of 216 residues at the amino terminus of human  $\alpha 1(\text{IV})$  collagen was recently reported (9) as well as 511-residue sequences for the end of the triple helical domains of mouse  $\alpha 1(\text{IV})$  and  $\alpha 2(\text{IV})$  collagens (6). We previously (10) deduced from the cDNA sequence a 143-residue fragment for the helical domain of mouse  $\alpha 2(\text{IV})$  collagen, which is now known to be located between residues 514-371 from the end of the triple helical sequence (see Ref. 6).

The complete primary structure for the carboxyl-terminal peptide from mouse and human  $\alpha 1(\text{IV})$  collagen has been determined. A striking feature of this 229-residue peptide is extensive homology between the two half-sequences, suggesting that it has evolved by gene duplication (11-13). We have determined from nucleotide sequences the complete primary structure for the carboxyl-terminal peptides of mouse  $\alpha 1(\text{IV})$  and  $\alpha 2(\text{IV})$  collagen. The peptides are 63% identical in sequence. Furthermore, they contain 33 invariant amino acids, including 6 cysteine residues, that are found in homologous positions in the two halves of both peptides. These results suggest that the carboxyl-terminal peptides of type IV collagen are closely related in their structure and evolution.

### EXPERIMENTAL PROCEDURES

**Materials**—Restriction endonucleases, DNA polymerase I (Klenow fragment), T4 DNA ligase, and the 17-mer oligonucleotide sequencing primer were from New England Biolabs and International Biotechnologies, Inc. (New Haven, CT). Reverse transcriptase was obtained from Seikagaku America Inc. (St. Petersburg, FL). The M13 vectors were from Amersham Corp. (Arlington Heights, IL) and sequencing nucleotides from Pharmacia P-L Biochemicals. Isopropyl thiogalactoside, Bluo-gal, and 1-kb ladder were from Bethesda Research Laboratories.  $\alpha$ -<sup>32</sup>P-labeled dCTP (3000 Ci/mmol) and  $\alpha$ -<sup>35</sup>S-labeled dATP (1100 Ci/mmol) were from New England Nuclear. Nitrocellulose filters were from Schleicher and Schuell.

**cDNA Cloning**—A 340-base pair cDNA clone, pPE1180, specific for one of the mouse type IV collagen chains, was isolated from a mouse parietal endoderm (PE)<sup>1</sup> cDNA library (14). This clone was used to isolate 55 additional clones from the PE cDNA library. To identify clones containing the longest cDNA inserts, plasmid DNA was prepared as described (15) and analyzed on agarose gels. Partial nucleotide sequence analysis showed that clones pPE41 (1.4-kb insert), pPE131 (1.1 kb), pPE90 (1.4 kb), pPE132 (1.5 kb), and pPE123 (1.8 kb) had the same 5' sequence. PE123 cDNA was sequenced completely. The PE cDNA library was screened with the 3' end

\* This work was supported in part by National Institutes of Health Grant GM 34090. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) J02765.

<sup>†</sup> Present address: Laboratory of Cancer Genetics and Cytogenetics, Sloan-Kettering Institute, New York, NY 10021.

<sup>\*\*</sup> Recipient of National Research Service Award CA-09056.

<sup>§§</sup> Present address: European Molecular Biology Laboratory, 6900 Heidelberg, Federal Republic of Germany.

<sup>¶¶</sup> Present address: Dept. of Medical Biochemistry, University of Oulu, 90220 Oulu 22, Finland.

<sup>1</sup> The abbreviations used are: PE, parietal endoderm; kb, kilobase.

*EcoRI* fragment (3.3 kb) from mouse  $\alpha 2(IV)$  collagen genomic clone HA3 (10) to isolate cDNA clones for  $\alpha 2(IV)$  collagen. By Northern analysis, this genomic fragment was known to contain coding sequences. From this screening, we isolated 7 clones, and the longest cDNA clone, pPE10A (1.9 kb), was sequenced completely.  $^{32}P$ -labeled DNA probes were prepared by nick-translation or oligo-labeling (16), and duplicate filter copies of the PE cDNA library were hybridized overnight at +68 in  $5 \times$  SSC,  $5 \times$  Denhardt's solution, 0.1% SDS, and 0.1% denatured salmon sperm DNA. Filters were washed at +68 in  $0.1 \times$  SSC, 0.1% sodium dodecyl sulfate and exposed overnight to Cronex film with intensifying screens.

**DNA Sequence Analysis**—The PE123 cDNA sequence was determined by the Maxam-Gilbert technique (17). The PE10A cDNA sequence was obtained by M13 cloning (18) and dideoxy-sequencing (19) of overlapping restriction fragments using  $^{35}S$ -labeled dATP (20) and reverse transcriptase (21). For some experiments, the M13-PE10A cDNA was progressively shortened from its 3' end (22) before sequencing. Both cDNA strands were sequenced at least twice.

**Computer Analysis**—DNA and amino acid sequences were analyzed using the programs of the University of Wisconsin Genetics Computer Group (23) and the Los Alamos portable sequence homology package (24, 25) running on a Digital Equipment Corp. VAX11/780 computer.

RESULTS

**Isolation and Characterization of cDNA Clones**—We have previously described a type IV collagen-specific mouse cDNA clone, pPE1180 (14). With this cDNA probe we isolated a clone, pPE123, from a mouse parietal endoderm cDNA library. It contains 612 nucleotides encoding 204 amino acids, followed by 1164 nucleotides of 3' untranslated sequence. Within the translated sequence, an exact match was found to a 14-residue sequence (PFLFXNINNVXNFA, residues 61–74 in Fig. 1) available from the carboxyl-terminal peptide of human  $\alpha 1(IV)$  collagen.<sup>2</sup> From this result, we conclude that the mouse PE 123 cDNA, reported here, is specific for the  $\alpha 1(IV)$  collagen chain. Subsequent to these studies, the sequence of a 229-residue carboxyl-terminal peptide of human and mouse  $\alpha 1(IV)$  collagen was reported (11–13). We obtained the sequence for the first 25 residues, missing from PE 123 cDNA, from the genomic sequence. The amino acid sequence presented here is identical with that published previously (13).

In order to isolate 3' cDNA clones for the  $\alpha 2(IV)$  collagen, the PE cDNA library was screened with a 3.3-kb *EcoRI* fragment of the  $\alpha 2(IV)$  collagen genomic clone, HA3 (10). From this screening, one cDNA clone, pPE10A, was isolated and subsequently sequenced. This clone contains 290 nucleotides encoding a sequence identical to the end of the triple helical sequence of  $\alpha 2(IV)$  collagen (6), 681 nucleotides encoding a 227-residue carboxyl-terminal peptide, and about 950 nucleotides of 3' untranslated sequence. From these results we conclude that pPE10A is specific for the  $\alpha 2(IV)$  collagen.

**Nucleotide and Amino Acid Sequences: Sequence Homology**—Fig. 1 shows the comparison of amino acid and nucleotide sequences for the carboxyl-terminal peptides of mouse  $\alpha 1(IV)$  and  $\alpha 2(IV)$  collagen of 229 and 227 amino acids, respectively. The sequences have been aligned for maximum amino acid homology. As shown, in total, 142 amino acids are found in homologous positions throughout the sequences, including all the 12 cysteine residues for each peptide. Furthermore, most of the proline (81%), glycine (80%), tryptophan (80%), and leucine (74%) residues are also conserved. Overall, the amino acid sequences are 63% identical. In addition, in 28 positions amino acids substitutions are by chemically similar residues.

The nucleotide sequences encoding the two carboxyl-terminal peptides of type IV collagen are 62% identical. It is of

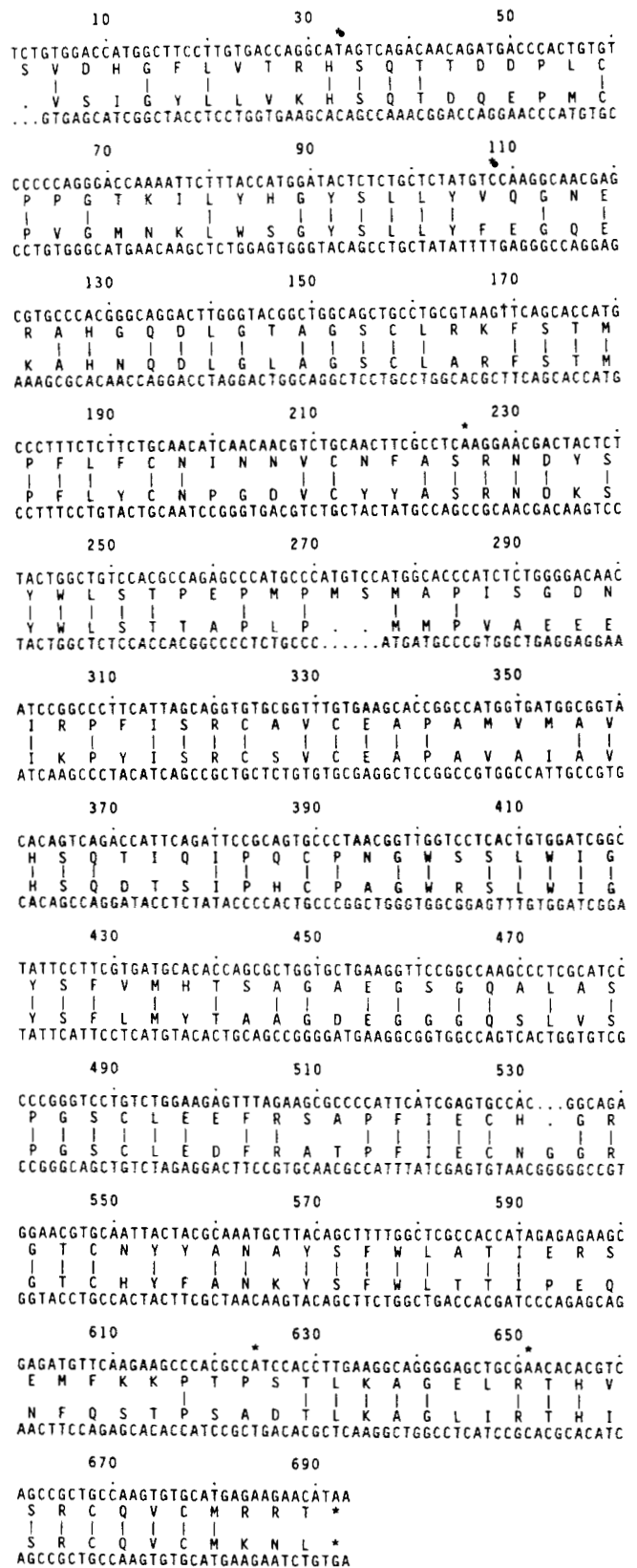


FIG. 1. DNA and amino acid sequences of the carboxyl-terminal peptides of mouse  $\alpha 1(IV)$  and  $\alpha 2(IV)$  collagen. The sequences have been aligned for maximum homology between the  $\alpha 1(IV)$  and  $\alpha 2(IV)$  peptides (top and bottom lines), respectively. The first 75 nucleotides in the  $\alpha 1(IV)$  sequence were obtained from exon sequence. In the  $\alpha 1(IV)$  sequence, five nucleotides (marked with asterisk) differ from those reported earlier (13), presumably due to nucleotide sequence polymorphism.

<sup>2</sup> T. Voss, R. Glanville, and K. Kuhn, personal communication.

**FIG. 2. Sequence duplication in the two halves of the carboxyl-terminal peptides of  $\alpha 1(IV)$  and  $\alpha 2(IV)$ .** The two halves of  $\alpha 1(IV)$  and  $\alpha 2(IV)$  were aligned for maximum amino acid homology. Identical residues found in both halves of both peptides appear in the consensus line. Less conserved residues are boxed.

$\alpha 1$ Half 1	1	SVDHGFLVTR	HSQ	TDDPLC	PPG	KILYHG	YS	LLYVQGN	ERA	HD	LD	GT	AGS	CL	RKF	ST	60
$\alpha 1$ Half 2		..APAMVMAV	HSQ	TQIPQC	PNG	WSSLWIG	YS	FVMHTSAG	AEG	SG	Q	ALAS	PGS	CLE	EFF	RS	
$\alpha 2$ Half 1		.VSIGYLLVK	HSQ	TQDEPMC	PVG	MNKLWSG	YS	LLYFEGQ	EKA	HN	QDLGL	AGS	CL	A	R	F	ST
$\alpha 2$ Half 2		..APAVAIIV	HSQ	TSIPHC	PAG	WRSLWIG	YS	FLMTAAG	DEG	GG	QSLVS	PGS	CLE	D	F	R	A
Consensus		-----	HSQ	-----	P-C	P-G	---	L--G	YS	-----	-----	Q-L	--	-GSCL	--	F--	
$\alpha 1$ Half 1	61	MPFLFCN.IN	NVC	NFASRND	YSY	WLSTPEP	MP	M	S	MAP...	ISGD	NI	IPF	ISRC	AV	CE..	121
$\alpha 1$ Half 2		APFIEC.HGR	GTC	NYFA.NA	YSF	WLATIEE	SEM	F	K	KPTPS	TLK	AG	L	R	T	H	VSRC
$\alpha 2$ Half 1		MPFLYCNPG.	DVC	YASRND	KSY	WLSTTAP	LP	.	M	MP...	VAEE	E	I	K	P	Y	ISRC
$\alpha 2$ Half 2		TPFIEC.HGR	GTC	YFA.NK	YSF	WLTTIPE	QNF	Q	S	TSPAD	TLK	AG	L	R	T	H	ISRC
Consensus		-PF--C----	--C	-----	N-	-S-WL-T	---	-----	P--	-----	-----	-SRC	-VC	---			

**TABLE I**  
Sequence homologies of the carboxyl-terminal peptides of mouse  $\alpha 1(IV)$  and  $\alpha 2(IV)$  collagen

	Amino acids		Amino acid similarity	Nucleotide similarity
	Identical	Conservative <sup>a</sup>		
$\alpha 1/\alpha 2$	142	28	63	62
$\alpha 1/\alpha 1^b$	40	20	37	46
$\alpha 2/\alpha 2^b$	40	15	36	49

<sup>a</sup> Chemically similar residues.

<sup>b</sup> Comparison between the two half-sequences.

interest that at the end of the sequences there is a stretch of 24 nucleotides which is identical in both. This sequence encodes the amino acids SRCQVC, and they represent the most conserved peptide segment of the two peptides (see below). Since each of these residues is specified by multiple codons, in principle, 1152 ( $6 \times 6 \times 2 \times 2 \times 4 \times 2$ ) different 18-nucleotide sequences could encode that sequence. However, taking into account the actual codon usage in the carboxyl-terminal peptides of type IV collagen (data not shown), this figure is reduced to 21 different sequences. It is not clear why this part of the nucleotide sequence has been conserved. Elsewhere, in only three other regions, sequences of 26–31 nucleotides show 84–86% identity.

Previously, chemical analysis of the carboxyl-terminal domain of type IV collagen revealed about 3–6% of the total mass as glucosamine and galactosamine, suggesting the presence of *N*- and *O*-linked sugar chains (26). The amino acid sequences of the two carboxyl-terminal peptides (Fig. 1), however, do not contain any potential sites (NXS or NXT) for *N*-linked sugars. Since in metabolic labeling experiments mannose was incorporated into type IV collagen (27), it is likely that *N*-linked sugars are present in the triple helical or amino-terminal domains. We have also reported previously (28) that the type IV collagen chains synthesized in the presence of tunicamycin migrate slightly faster in sodium dodecyl sulfate-polyacrylamide gel electrophoresis than the chains of control samples, an indication for *N*-linked sugar chains.

**Sequence Duplication**—As has been reported earlier (11–13), a striking feature of the carboxyl-terminal peptide of  $\alpha 1(IV)$  collagen is that the first half of the sequence is homologous with the second half. Similar intrasequence homology is also found for the carboxyl-terminal peptide of  $\alpha 2(IV)$  collagen as shown in Fig. 2. It should be noted that positions for the gaps in the two half-sequences were chosen to maximize the homologies while introducing a minimum number

and length of gaps. The alignment for  $\alpha 1(IV)$  is slightly different from those shown in three earlier papers (11–13), which are not identical either; however, two of those alignments (11, 12) were made by eye with no rigorous consideration for gap number and length. The two  $\alpha 1(IV)$  half-peptide sequences contain 40 identical and 20 chemically similar residues; the two  $\alpha 2(IV)$  half-peptides contain 40 identical and 15 chemically similar residues. Nucleotide sequences for the two half-peptides are 46% identical in  $\alpha 1(IV)$  and 49% identical in  $\alpha 2(IV)$ .

In summary, the nucleotide and amino acid sequence comparisons for the carboxyl-terminal peptides of type IV collagen are presented in Table I.

## DISCUSSION

In this paper we have presented the nucleotide and amino acid sequences for the carboxyl-terminal peptides of mouse  $\alpha 1(IV)$  and  $\alpha 2(IV)$  collagen, 229 and 227, respectively. The peptides are highly conserved in sequence, and both display a similar, internal sequence duplication. Overall, identical or chemically similar residues represent 74% of these sequences. Most of the proline, glycine, tryptophan, and leucine, and all 12 cysteine residues are conserved. A comparison of the mouse and human (11, 12) sequences revealed only 7 amino acid differences in the carboxyl-terminal peptides of  $\alpha 1(IV)$  collagen. Such a high degree of conservation between mouse and human proteins is unusual and suggests that the precise amino acid sequence is important in the structure and function of the peptide. It also reflects the known multiple interactions of the peptides during the biosynthesis and assembly of type IV collagen molecules. By analogy with the carboxyl-terminal propeptides of fibrillar collagens, some of these interactions could be in chain selection and initiation of the correct alignment for the chains during formation of the triple helix (29).

In the carboxyl-terminal domain of type IV collagen, a number of interactions between the two peptides are known. They all are associated noncovalently with each other, but disulfide and nonreducible bonds are found only between peptides of separate type IV collagen molecules which are linked end-to-end at their carboxyl-terminal domains (1). Subunit structure for the domain of mouse type IV collagen produced by the Engelbreth-Holm-Swarm tumor was recently reported by Weber *et al.* (26). As shown by partial sequence analysis of purified monomeric and dimeric peptides, disulfide-bonded peptides were derived from the  $\alpha 1(IV)$  chains, whereas most of the monomeric peptides were of  $\alpha 2(IV)$  chain origin. Dimeric peptides containing disulfide and nonreducible bonds revealed sequences from both chains. In addition,

it is of interest that the domain was found to contain both peptides in about equal amounts, which points to a possible heterogeneity in the type IV collagen structure (26).

A striking feature of the carboxyl-terminal peptides of type IV collagen is that the first half of the sequence is homologous with the second half. In the  $\alpha 1(\text{IV})$  chain, the two half-sequences are 37% identical, and in the  $\alpha 2(\text{IV})$  chain they are 36% identical. Moreover, a consensus sequence containing 33 invariant amino acids, including 6 cysteine residues, is present in both halves of the two peptides (Fig. 2). What is the significance of this sequence duplication? It is tempting to speculate that this duplication is important and has evolved for the end-to-end assembly of type IV collagen molecules and thus should be found in all type IV collagens across species. Recent data show that the carboxyl-terminal peptide of *Drosophila* type IV collagen (30) is duplicated in sequence and contains 25 of the 33 consensus amino acids referred to above.

The nucleotide sequence homologies presented here (Table I) suggest a model for the evolution of the carboxyl-terminal peptides of type IV collagen. Presumably, duplication of an ancestral DNA sequence first established a primordial peptide containing two domains and later, by gene duplication, the two different but very homologous carboxyl-terminal peptides of type IV collagen evolved. The internal nucleotide sequence homologies for the two peptides are 46 and 49%, respectively, whereas it is 62% between them. This result is consistent with the proposed model where the peptides diverged after gene duplication. If these peptides had diverged before the internal duplication, one would expect to see less nucleotide sequence similarity between the peptides than between the two domains of a given peptide.

An interesting question is how the  $\alpha 1(\text{IV})$  and  $\alpha 2(\text{IV})$  collagen genes have evolved. Exon structures in the triple helix coding domain of these genes are clearly distinct from those of fibrillar collagen genes, suggesting a different type of evolution (10, 31, 32). This notion is further supported by the results reported here and elsewhere (11–13) indicating only little similarity between the carboxyl-terminal peptides of fibrillar and type IV collagens, whereas between the fibrillar collagens, the carboxyl-terminal peptides are 52–69% identical in sequence (33). Furthermore, size and distribution of the four exons encoding them are conserved among species and different collagen types (34). Since it is believed that the fibrillar collagen genes have evolved from a common origin by repeated gene duplications (35), it is likely that the ancestral sequence encoding the carboxyl-terminal peptide was already present in the primordial gene. In this context, the related evolution of the carboxyl-terminal peptides of type IV collagen, suggested by our data, could imply that the  $\alpha 1(\text{IV})$  and  $\alpha 2(\text{IV})$  collagen genes have evolved similarly. However, this may not be the case, because so far we have found no indication for similar exon sizes in the helix coding domain

of the two genes.<sup>3</sup> Furthermore, Schwarz *et al.* (6) recently reported 511-residue sequences for the ends of triple helical domains of mouse  $\alpha 1(\text{IV})$  and  $\alpha 2(\text{IV})$  collagens, indicating a remote evolutionary relationship between the two chains.

*Acknowledgments*—We thank Tilmann Voss, Robert Glanville, and Klaus Kuhn for amino acid sequence data, John Fessler for computing facilities, and Francine Mittleman for preparing this manuscript.

## REFERENCES

1. Timpl, R., and Martin, G. R. (1982) *Immunocytochemistry of the Extracellular Matrix*, Vol. 2, pp. 119–150, CRC Press, Boca Raton, FL
2. Kefalides, N. A., Howard, P., and Ohno, N. (1985) in *Basement Membranes* (Shibata, S., ed) pp. 73–87, Elsevier Scientific Publishing Co., Amsterdam
3. Trüeb, B., Gröbli, B., Spiess, M., Odermatt, B. F., and Winterhalter, K. H. (1982) *J. Biol. Chem.* **257**, 5239–5245
4. Haralson, M. A., Federspiel, S. J., Martinez-Hernandez, A., Rhodes, R. K., and Miller, E. J. (1985) *Biochemistry* **24**, 5792–5797
5. Babel, W., and Glanville, R. W. (1984) *Eur. J. Biochem.* **143**, 545–556
6. Schwarz, U., Schuppan, D., Oberbaumer, I., Glanville, R. W., Deutzmann, R., Timpl, R., and Kühn, K. (1986) *Eur. J. Biochem.* **157**, 49–56
7. Timpl, R., Wiedemann, H., Van Delden, V., Furthmayr, H., and Kühn, K. (1981) *Eur. J. Biochem.* **120**, 203–211
8. Yurchenco, P. D., and Furthmayr, H. (1984) *Biochemistry* **23**, 1839–1850
9. Glanville, R. W., Qian, R., Siebold, B., Risteli, J., and Kühn, K. (1985) *Eur. J. Biochem.* **152**, 213–219
10. Kurkinen, M., Bernard, M. P., Barlow, D. P., and Chow, L. T. (1985) *Nature* **317**, 177–179
11. Pihlajaniemi, T., Tryggvason, K., Myers, J. C., Kurkinen, M., Lebo, R., Cheung, M.-C., Prockop, D. J., and Boyd, C. D. (1985) *J. Biol. Chem.* **260**, 7681–7687
12. Brinker, J. M., Gudas, L. J., Loidl, H. R., Wang, S.-Y., Rosenbloom, J., Kefalides, N. A., and Myers, J. C. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 3649–3653
13. Oberbaumer, I., Laurent, M., Schwarz, U., Sakurai, Y., Yamada, Y., Vogeli, G., Voss, T., Siebold, B., Glanville, R. W., and Kühn, K. (1985) *Eur. J. Biochem.* **147**, 217–224
14. Kurkinen, M., Barlow, D. P., Helfman, D. M., Williams, J. G., and Hogan, B. L. M. (1983) *Nucleic Acids Res.* **11**, 6199–6202
15. Holmes, D. S., and Quigley, M. (1981) *Anal. Biochem.* **114**, 193–197
16. Feinberg, A. P., and Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13
17. Maxam, A. J., and Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–500
18. Messing, J. (1983) *Methods Enzymol.* **101**, 20–78
19. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467
20. Biggin, M. D., Gibson, T. J., and Hong, G. R. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80**, 3963–3965
21. Karanthanasis, S. (1982) *Focus* **4**, 6–7
22. Dale, R. M., McClure, B. A., and Houchins, J. P. (1985) *Plasmid* **13**, 31–41
23. Devereaux, J., Haeblerli, P., and Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395
24. Kanehisa, M. I. (1982) *Nucleic Acids Res.* **10**, 183–196
25. Goad, W. B., and Kanehisa, M. I. (1982) *Nucleic Acids Res.* **10**, 247–263
26. Weber, S., Engel, J., Wiedemann, H., Glanville, R. W., and Timpl, R. (1984) *Eur. J. Biochem.* **139**, 401–410
27. Clark, C. C., and Kefalides, N. A. (1982) in *New Trends in Basement Membrane Research* (Kuhn, K., Timpl, R., and Schone, H. H., eds) pp. 155–162, Raven Press, New York
28. Kurkinen, M., Foster, L., Barlow, D. P., and Hogan, B. L. M. (1982) *J. Biol. Chem.* **257**, 15151–15155
29. Prockop, D. J., and Kivirikko, K. I. (1984) *N. Engl. J. Med.* **311**, 376–386
30. Blumberg, B., MacKrell, A. J., Olson, P. F., Kurkinen, M., Monson, J. M., Natzle, J. E., and Fessler, J. H. (1987) *J. Biol. Chem.*, in press
31. Soinenen, R., Tikka, L., Chow, L., Pihlajaniemi, T., Kurkinen, M., Prockop, D. J., Boyd, C. D., and Tryggvason, K. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 1568–1572
32. Sakurai, Y., Sullivan, M., and Yamada, Y. (1986) *J. Biol. Chem.* **261**, 6654–6657
33. Sandell, L. J., Prentice, H. L., Kravis, D., and Upholdt, W. B. (1984) *J. Biol. Chem.* **259**, 7826–7834
34. Ramirez, F., Bernard, M., Chu, M.-L., Dickson, L., Sangiorgi, F., Weil, D., de Wet, W., Junien, C., and Sobel, M. (1985) *Ann. N. Y. Acad. Sci.* **460**, 117–129
35. Yamada, Y., Liau, G., Mudryj, M., Obici, S., and de Crombrughe, G. (1984) *Nature* **310**, 333–337

<sup>3</sup> M. Kurkinen, *et al.*, unpublished data.