Kaplan, J., Ray, F. A., & Keogh, E. A. (1981) *J. Biol. Chem. 256*, 7705–7707.
Kurecki, T., Kress, L. F., & Laskowski, M. (1979) *Anal. Biochem. 99*, 415–420.
Lottenberg, R., & Jackson, C. M. (1983) *Biochim. Biophys. Acta 742*, 558–564.
Pizzo, S. V., Schwartz, M. L., Hill, R. L., & McKee, P. A. (1972) *J. Biol. Chem. 247*, 636–645.
Pochon, F., & Bieth, J. (1982) *J. Biol. Chem. 257*, 6683–6685.
Pochon, F., Amand, B., Lavalette, D., & Bieth, J. (1978) *J. Biol. Chem. 253*, 7496–7499.
Pochon, F., Favaudon, V., Tourbez-Perrin, M., & Bieth, J. (1981) *J. Biol. Chem. 256*, 547–550.
Rinderknecht, H., Fleming, R. M., & Geokas, M. C. (1975) *Biochim. Biophys. Acta 377*, 158–165.
Robbins, K. C., Summaria, L., Hsieh, B., & Shah, R. J. (1967) *J. Biol. Chem. 242*, 2333–2342.
Salvesen, G. S., & Barrett, A. J. (1980) *Biochem. J. 187*, 695–701.
Salvesen, G. S., Sayers, C. A., & Barrett, A. J. (1981) *Biochem. J. 195*, 453–461.
Sottrup-Jensen, L., Peterson, T. E., & Magnusson, S. (1980) *FEBS Lett. 121*, 275–279.
Sottrup-Jensen, L., Peterson, T. E., & Magnusson, S. (1981) *FEBS Lett. 128*, 127–132.
Swenson, R. P., & Howard, J. B. (1979a) *J. Biol. Chem. 254*, 4452–4456.
Swenson, R. P., & Howard, J. B. (1979b) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4313–4316.
Van Leuven, F., Cassiman, J.-J., & Van Den Berghe, H. (1979) *J. Biol. Chem. 254*, 5155–5160.
Wyckoff, M., Rodbard, D., & Chramback, A. (1977) *Anal. Biochem. 78*, 459–482.

Isolation and Characterization of Pepsin-Solubilized Human Basement Membrane (Type IV) Collagen Peptides†

Robert S. MacWright,* Virginia A. Benson, Katherine T. Lovello, Michel van der Rest,‡ and Peter P. Fietzek§

ABSTRACT: Native type IV collagen was isolated from human placental tissue by pepsin digestion, fractional salt precipitation, reduction and alkylation, a second pepsin digestion, and chromatography on diethylaminoethyl- and carboxymethylcellulose. After denaturation, 10 distinct peptides were isolated from this material by molecular sieve, ion-exchange, and high-performance liquid chromatography. All of the peptides were found to have amino acid compositions characteristic of type IV collagen. Analysis of the eight major peptides by amino-terminal amino acid sequencing and by cyanogen bromide and tryptic peptide mapping has revealed the manner in which they are derived from type IV collagen. Pepsin liberates two large peptides by attacking non-triple-helical regions, one derived from the α1(IV) chain (F2, M_r 90 000) and one derived from the α2(IV) chain (F3, M_r 75 000). The α1(IV)-derived F2 peptide is also represented in the pepsin digest by amino-terminal and carboxy-terminal subfragments [F4c (M_r 41 000) and F4a (M_r 60 000)], as is the α2(IV)-derived F3 peptide [F5 (M_r 28 000) and F4b (M_r 50 000), respectively]. These findings indicate that the molecular regions from which the larger peptides are derived in themselves contain pepsin-sensitive (non-triple-helical) domains. In addition, several of the peptides examined were found to be present in two slightly different forms, suggesting that closely adjacent pepsin-sensitive sites often exist within the type IV collagen molecules. The methods outlined here provide a reliable means by which identifiable type IV collagen peptides can be isolated. Furthermore, the above conclusions and the data from which they have been drawn provide a basis from which the previously described type IV collagen peptides can be more clearly identified and related to a common structure of origin.

Basement membranes are sheetlike extracellular matrices which support and compartmentalize soft tissue structures (Vracko, 1974) and also serve as selective barriers to permeability (Caufield & Farquhar, 1978). Chemical and immunological studies of basement membranes have shown the presence of the glycoproteins laminin (Timpl et al., 1979a) and fibronectin (Stenman & Vaheri, 1978), a heparin sulfate containing proteoglycan (Hassell et al., 1980), and a major collagenous component known as type IV collagen. Type IV collagen is chemically and structurally distinct from the interstitial collagens (Kefalides, 1973) and is believed to be the primary structural element of basement membranes (Bornstein & Sage, 1980).

Recent studies on the type IV collagen secreted by cell and tissue cultures have shown that two distinct polypeptides are produced, pro-α1(IV) and pro-α2(IV), with molecular weights of 185 000 and 175 000, respectively (Crouch & Bornstein, 1979; Alitalo et al., 1980; Crouch et al., 1980; Fessler & Fessler, 1980; Tryggvason et al., 1980). Although processing of these biosynthetic products to smaller chains has not been observed under culture conditions, type IV collagen polypeptides with molecular weights of 160 000 and 140 000 have been extracted from bovine lens capsules (Gay & Miller, 1979) and from the matrix of the EHS murine tumor when grown

† From the Department of Biochemistry, University of Medicine and Dentistry of New Jersey—Rutgers Medical School, Piscataway, New Jersey 08854. *Received March 8, 1983.* This work was supported by National Institutes of Health Grant AM22051. Studies reported were submitted by R.S.M. in partial fulfillment of the requirements for the Ph.D. Degree of Rutgers—The State University of New Jersey and the University of Medicine and Dentistry of New Jersey—Rutgers Medical School.

* Address correspondence to this author at the Department of Microbiology, University of Medicine and Dentistry of New Jersey—Rutgers Medical School.
‡ Present address: Genetics Unit, Shriner's Hospital, Montreal, Canada H3G 1A6.
§ Present address: Carl Freudenberg/Collan, D-6940, Weinheim, West Germany.

0006-2960/83/0422-4940$01.50/0

in lathyritic animals (Orkin et al., 1977; Timpl et al., 1978).

The small absolute mass of basement membranes in a given tissue and the low solubility of type IV collagen have hampered attempts to examine its chemical and structural features. In order to obtain sufficient quantities of material, it has generally been necessary to solubilize the collagen in basement membrane rich tissues by limited pepsin digestion, liberating a mixture of type IV collagen peptides with molecular weights ranging from 15 000 to 200 000 (Bailey et al., 1979; Daniels & Chu, 1975; Kefalides, 1971; Dehm & Kefalides, 1978; Dixit, 1978, 1979, 1980; Dixit & Kang, 1979, 1980; Glanville et al., 1979; Glanville & Rauter, 1981; Kresina & Miller, 1979; Mayne & Zettergren, 1980; Sage et al., 1979; Timpl et al., 1979b; Tryggvason & Kivirikko, 1978). This heterogeneity is caused by the pepsin cleavage of globular interruptions in the type IV collagen triple helix (Glanville et al., 1979; Timpl et al., 1979b). The major type IV collagen pepsin digestion products obtained by different investigators have varied considerably, apparently due to the kinetics involved in the cleavage of these globular domains, which appear to be greatly complicated by the presence of large numbers of inter- and intrachain disulfide bonds (Timpl et al., 1979b).

In this study, type IV collagen was isolated from human placental tissue by a three-step procedure (pepsin digestion, reduction/alkylation, and pepsin digestion) which is similar to but more severe than that described previously (Glanville et al., 1979). Ten distinct peptides were then purified from the resulting mixture and were characterized by a variety of techniques. This analysis has allowed us to reveal the relationships between the various peptides and to establish the characteristics by which each can be clearly identified.

Materials and Methods

Isolation of Human Type IV Collagen. Normal human placentas which had been frozen immediately following delivery were thawed and washed extensively in cold tap water to remove blood. After the umbilical cords and fetal membranes had been removed by dissection, the villi and attached chorionic plate were ground with crushed ice in a meat grinder. The ground material was suspended in 0.4 M sodium acetate, pH 4.8, and solid material was pelleted by centrifugation. This washing was repeated until all residual blood had been removed. The ground villi were then washed once with 0.5 M acetic acid, and the solid material was suspended in 0.5 M acetic acid to give a final concentration of about 0.1 g of tissue/mL. Pepsin (Boehringer Mannheim, 2500 Anson units/mg) was slowly added to this slurry with constant agitation. About 1 g of pepsin was used for each 1000 g of tissue. The slurry was maintained at 20 °C, and digestion was allowed to continue for 20 h. All insoluble material was then removed by centrifugation. The salt concentration of the supernatant was raised to 0.9 M by addition of solid NaCl, and after the mixture was stirred overnight, the precipitate was collected and was dissolved in 0.2 M NaCl/0.05 M tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl), pH 7.5. The pH was adjusted to 7.5 by the addition of 1 N NaOH, and the solution was dialyzed against the same buffer. The NaCl concentration was then adjusted to 1.9 M, and the precipitate was collected by centrifugation. This precipitate was then dissolved in 0.2 M NaCl/0.05 M Tris-HCl, pH 7.5, and the 1.9 M NaCl precipitation was repeated.

The 1.9 M precipitate was then dissolved in 0.5 M Tris-HCl, pH 8.0, for reduction and alkylation in the native state (Kühn et al., 1966). β-Mercaptoethanol was added to a final concentration of 0.01 M, and after incubation at 20 °C for 20 h, iodoacetate was added to a final concentration of 0.009 M. The sample was then incubated at 4 °C for 1 h and was dialyzed extensively against 0.5 M acetic acid.

Following reduction and alkylation, 1 mg of pepsin/g of original 1.9 M precipitate was added, and the sample was stirred at 20 °C for 20 h. The collagen was then precipitated by adjusting the NaCl concentration to 2 M. This 2 M NaCl precipitate was then dissolved in 0.2 M NaCl/0.05 M Tris-HCl, pH 7.4, which contained 2 M deionized urea. The sample was extensively dialyzed against this same buffer to adjust the ionic strength and pH and to inactivate the pepsin. The sample was then passed through a 2.5 × 15 cm column of DEAE-cellulose (Whatman DE52) which had been equilibrated at 15 °C in the same buffer. Most of the collagen did not bind to the resin, and the material passing through the DEAE column was dialyzed against 0.04 M sodium acetate, pH 4.8, containing 2 M deionized urea. The solution was then applied to a 5 × 40 cm column which was packed with carboxymethylcellulose (Whatman CM52) equilibrated at 15 °C in the same buffer. Bound protein was eluted with a linear 0–0.2 M NaCl gradient (3000-mL total volume). Peak fractions were pooled, dialyzed extensively against 0.1 N acetic acid at 4 °C, and lyophilized.

Purification of Type IV Collagen Peptides. Lyophilized samples of human type IV collagen were dissolved in 0.05 M Tris-HCl, pH 7.5, containing 1 M calcium chloride. After heat denaturation at 42 °C for 30 min, these samples were applied to 2 × 120 cm columns of agarose A5M or A1.5M (200–400 mesh; Bio-Rad) and were eluted in the same buffer. Peak fractions were pooled, desalted on a 2 × 50 cm column of Bio-Gel P2 (100–200 mesh; Bio-Rad) containing 0.1 N acetic acid, and lyophilized.

Peak fractions from molecular sieve chromatography were then chromatographed on a 2.5 × 10 cm column of carboxymethylcellulose (Whatman CM52) which had been equilibrated at 42 °C in 0.04 M sodium acetate, pH 4.8, containing 2 M urea. Peptides were eluted with a linear gradient of 0–0.16 M NaCl which had been prepared by using the same buffer (total volume 1000 mL). Peak fractions were pooled, dialyzed against 0.1 N acetic acid, and lyophilized. Alternatively, some samples were desalted and lyophilized as described above.

If necessary, peak fractions from CM-cellulose chromatography were dialyzed against an appropriate starting buffer and were then rechromatographed on CM-cellulose, agarose A5M, or agarose A1.5M. Alternatively, final purification of individual peptides was achieved by reverse-phase high-performance liquid chromatography (HPLC) as described by van der Rest & Fietzek (1982). Lyophilized peptides were dissolved in aqueous 0.01 M heptafluorobutyric acid (sequencer grade, Beckman) to a final concentration of 1–6 mg/mL and were applied to a Vydac 201 TP 0.46 × 25 cm C18 reverse-phase column (Separation Group) which had been equilibrated with aqueous 0.01 M heptafluorobutyric acid containing 30% acetonitrile (HPLC grade, Burdick and Jackson). Peptides were eluted with a linear gradient of 30–60% acetonitrile over a period of 60 min at room temperature; a constant flow rate of 1 mL/min was maintained by using a two-pump system (Waters Associates). Peak fractions were pooled and were dried in a Speedvac concentrator (Savant).

Sodium Dodecyl Sulfate (SDS)–Polyacrylamide Gel Electrophoresis. Electrophoresis was carried out as described by Laemmli (1970), using 8%, 10%, or 15% separating gels which were 1.5 mm thick. From 20 to 150 μg of protein was applied to each well. Gels were electrophoresed at 100 V (constant voltage) until the tracking dye had penetrated 15

cm into the separating gel (8–9 h). Gels were stained for 60 min with 20% trichloroacetic acid containing 0.25% Coomassie Brilliant Blue R250 (Bio-Rad) and were then destained in 10% methanol/7% acetic acid.

Amino Acid Analysis. Samples were hydrolyzed for 24 h at 108 °C in 6 N HCl under nitrogen. Hydrolysates were dried down in a Speedvac concentrator (Savant) and were then analyzed on a Beckman Model 121 MB amino acid analyzer. Separation was achieved on a single column by an elution program which used four sodium citrate buffers: pH 2.99, $[Na^+]$ 0.2 M; pH 3.28, $[Na^+]$ 0.2 M; pH 3.7, $[Na^+]$ 0.35 M; and pH 4.9, $[Na^+]$ 1.4 M. Elution was carried out over 2 h at temperatures of 39.3 and 59 °C.

Molecular Weight Determination. The molecular weights of type IV collagen peptides were determined by chromatography on columns of agarose A5M and A1.5M (Bio-Rad) which had been calibrated by using a variety of collagenous peptides whose sizes and amino acid sequences were known. Molecular weights were also determined from the migration of peptides on SDS–polyacrylamide gels which were run using appropriate standards.

Amino-Terminal Sequence Analysis. Amino acid sequence analysis was done by automated Edman degradation using a Beckman Model 890C sequenator. Peptides were dissolved in 300 μL of 0.1 N acetic acid, dried in the cup under vacuum, and sequentially degraded by using the 1 or 0.1 M Quadrol program (supplied by Beckman). The resulting anilinothiazolinones were manually converted to the phenylthiohydantoin derivatives and were identified by high-pressure liquid chromatography (HPLC) (Bhown et al., 1978), thin-layer chromatography (TLC), or back-hydrolysis and subsequent amino acid analysis (Van Orden & Carpenter, 1964). Repetitive yields were determined by comparing HPLC peak heights with those of standards whose concentrations were known.

Cyanogen Bromide Peptide Pattern Determination. Peptides were dissolved in 70% formic acid containing 12 mg of cyanogen bromide/mL. After being flushed with nitrogen, samples were incubated for 4 h at 30 °C. Samples were diluted with 10 volumes of distilled water and were dried down in a Speedvac concentrator (Savant). The dried samples were then dissolved in gel sample buffer and were run on 15% SDS–polyacrylamide gels, as described earlier.

Tryptic Peptide Mapping. For digestion with trypsin, samples were first reacted with phenyl isothiocyanate, as described by Dewes et al. (1979), in order to prevent cleavage at lysine residues. Samples were then dissolved in 0.2 M NH_4HCO_3, and trypsin was added to give a ratio of 1:50 (trypsin:peptide). Samples were incubated at 37 °C for 4 h. The samples were then lyophilized, and the tryptic peptides were resolved by HPLC by using the procedure outlined above.

Results

Preparation of Type IV Collagen. Throughout the preparation of type IV collagen, the presence of the various collagen types in the supernatants and precipitates was monitored by polyacrylamide gel electrophoresis and by amino acid analysis (data not shown). The initial pepsin digestion of the placental villi solubilized nearly all of the collagenous material, as indicated by a low content of hydroxylysine and hydroxyproline in the insoluble pellet. The supernatant of the pepsin digest contained type I, type III, type IV, and type V collagen. This material was then subjected to a series of fractional salt precipitation steps to remove as much as possible of the collagen other than type IV. The first precipitation at 0.9 M NaCl removed type III and type IV collagen, while type I and type V collagen remained in solution. When the pellet was then dissolved and equilibrated in neutral buffer, a small amount of hydroxyproline-containing material remained insoluble and was removed by centrifugation. This material was not further characterized. A second precipitation was then carried out by raising the NaCl concentration to 1.9 M. The supernatant contained type I and type V collagen, and the precipitate contained mainly type III and IV collagen. In order to remove residual type I and type V collagen from the precipitate, it was necessary to repeat the 1.9 M NaCl precipitation. The type IV collagen in the final 1.9 M NaCl precipitate still contained small amounts of type III collagen, and the type IV collagen itself was in a high molecular weight form. When applied to a 6% polyacrylamide gel, only a small portion of the material entered the gel (data not shown).

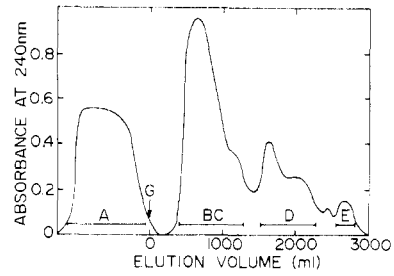

FIGURE 1: Final purification of pepsin-solubilized human type IV collagen by chromatography on carboxymethylcellulose. Bars indicate fractions pooled. An arrow denotes the beginning of the gradient.

The partially purified type IV collagen was reduced and alkylated under nondenaturing conditions. This material was then dialyzed against 0.5 M acetic acid and was digested a second time with pepsin. The collagenous material was recovered from this digest by NaCl precipitation and was passed through a DEAE-cellulose column in order to remove various acidic contaminants. The collagenous material, which did not bind to this column, was then chromatographed on CM-cellulose (Figure 1). The material eluted with the starting buffer (peak A) and the material in the first peak eluted by the NaCl gradient (peak BC) both contained type IV collagen which was essentially free of contamination by collagen type I, type III, and type V. The total yield of type IV collagen (peaks A + BC) from 1000 g of placenta was about 1 g. The material eluting at higher ionic strength contained some type IV collagen in addition to type III collagen (peak D) and type V collagen (peak E).

Characterization of Pepsin-Solubilized Type IV Collagen. The amino acid compositions of the collagen eluted from the CM-cellulose column (Figure 1) are presented in Table I. The glycine content in similar to that found for the interstitial collagens, whereas the hydroxylysine and hydroxyproline contents are high in comparison with those of type I collagen. The alanine and arginine contents are also comparatively low. These features are characteristic of type IV collagen. No major differences were found between the amino acid compositions of the type IV collagen which bound to the CM-cellulose column (peak BC) and that which did not bind (peak A).

The molecular composition of the pepsin-solubilized type IV collagen was analyzed by SDS–polyacrylamide gel electrophoresis; a typical analysis of the material in peak BC is shown in Figure 2. Peaks A and BC from the CM-cellulose purification both contained an identical array of major peptides, ranging in molecular weight from 12000 to 90000, along with a large amount of high molecular weight material. Peak A contained somewhat more high molecular weight material than peak BC (data not shown). Reduction with mercapto-

Table I: Amino Acid Compositions of Pepsin-Solubilized Human Type IV Collagen

amino acid	sample A	sample BC
Hyp	102	92
Asp	54	53
Thr	18	16
Ser	13	12
Glu	102	99
Pro	86	89
Gly	335	343
Ala	32	35
Val	29	27
Met[a]	12	12
Ile	29	32
Leu	52	53
Tyr	9.7	10
Phe	30	30
Hyl	50	53
Lys	6.8	6.7
His	7.6	7.5
Arg	32	29

[a] Determined as methionine + methionine sulfoxide.

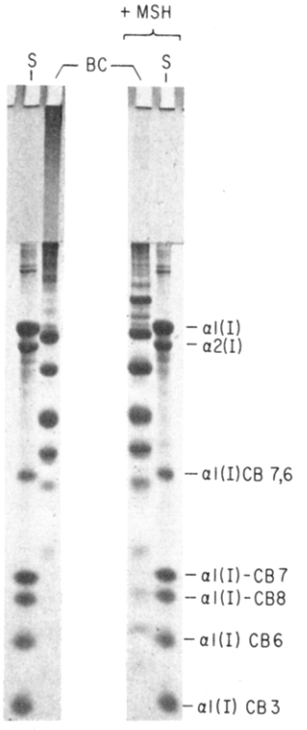

FIGURE 2: SDS–polyacrylamide gel electrophoresis pattern of pepsin-solubilized human type IV collagen (10% gel). BC refers to that peak eluted from CM-cellulose chromatography (see Figure 1). Type I collagen and selected cyanogen bromide peptides were used as standards. Samples applied to lanes on the right were reduced with β-mercaptoethanol prior to electrophoresis.

ethanol before electrophoresis caused a decrease in the amount of high molecular weight material and caused an increase in the number of bands below 100 000 in molecular weight. Bands of 140 000 and 120 000 molecular weight were also noted following reduction.

Purification of Type IV Collagen Peptides. Pepsin-solubilized type IV collagen was heat denatured, and the peptides present in this material were separated according to molecular weight by chromatography on agarose A5M (Figure 3). Peptides eluting in each peak are primarily those indicated on the polyacrylamide gel electrophoresis pattern of the starting material (see inset, Figure 3). The first peak, eluting

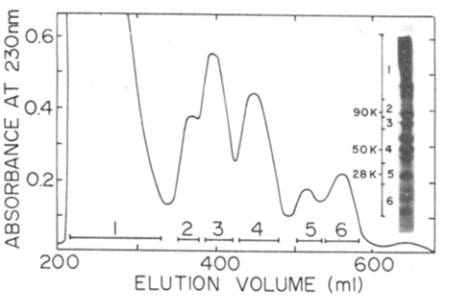

FIGURE 3: Chromatography of pepsin-solubilized human type IV collagen on agarose 5M. Fractions pooled for further purification are indicated by bars. Peptides contained in each peak are primarily those indicated in the gel electrophoresis pattern of the starting material (see inset, 10% gel).

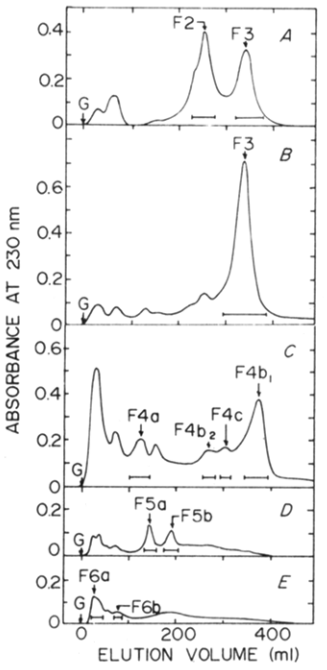

FIGURE 4: CM-cellulose chromatography of peaks 2–6 from agarose 5M chromatography of pepsin-solubilized human type IV collagen (see Figure 3). Arrows indicate the elution volumes of the various peptides. Fractions pooled for further purification are indicated by bars. (A) CM-cellulose chromatography of agarose 5M peak 2; (B) CM-cellulose chromatography of agarose 5M peak 3; (C) CM-cellulose chromatography of agarose 5M peak 4; (D) CM-cellulose chromatography of agarose 5M peak 5; (E) CM-cellulose chromatography of agarose 5M peak 6. "G" denotes the beginning of the gradient.

in the void volume of the column, represents about 60% of the type IV collagen applied. This peak contained the high molecular weight material which was present in the starting mixture. This material, designated F1, was not further studied. The second and third peaks which eluted each contained primarily a single peptide. These peptides were designated F2 and F3, respectively. The fourth peak which eluted contained three major peptides, designated F4a, F4b, and F4c, as well as some more minor components. Peak five contained primarily a single peptide, designated F5, and the sixth peak contained at least two major peptides (F6a and F6b) as well as several minor ones.

In order to further resolve the various components, each of the peaks 2–6 was chromatographed on carboxymethylcellulose (Figure 4). The elution positions of the major peptides were determined by polyacrylamide gel electrophoresis; those peaks not identified were not found to contain protein (data not shown). The peptides F4b and F5 were found upon chro-

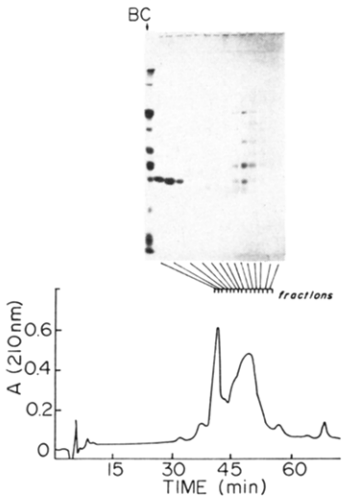

FIGURE 5: SDS–polyacrylamide gel electrophoresis of fractions obtained by HPLC purification of the human type IV collagen peptide F4c. Lines show the correspondence between the HPLC elution profile and the samples applied to the gel. Human type IV collagen from CM-cellulose fraction BC (Figure 1) was applied to the gel as a standard (8% gel).

Table II: HPLC Elution Times for Type IV Collagen Peptides

peptide	elution time (min)
F2	44.3
F3	48
F4a	34
$F4b_1$	45.6
F4c	40
F5a	42.8
F5b	41.6
F6a	28
F6b	33.6

matography to be heterogeneous, yielding peptides $F4b_1$, $F4b_2$ and F5a, F5b, respectively. Peptides from agarose peak F6 were generally recovered in poor yield.

In order to obtain improved resolution of the various peptides, each major fraction from CM-cellulose chromatography was rechromatographed on agarose A5M or A1.5M, or on CM-cellulose (data not shown). Alternatively, final purification of each peptide was achieved by the use of high-performance liquid chromatography. The major peptides eluted upon purification by HPLC were identified by polyacrylamide gel electrophoresis; an example of such an analysis is shown in Figure 5. The elution times for all of the major peptides are listed in Table II.

Characterization of the Purified Type IV Collagen Peptides. The high purity of the peptides which were obtained by the above procedures is demonstrated by the polyacrylamide gel electrophoresis pattern presented in Figure 6. The molecular weights of the individual peptides, as determined by SDS–polyacrylamide gel electrophoresis and by molecular sieve chromatography, are listed in Table III. These purified peptides were further characterized as described below. Unfortunately, the peptides F6a and F6b could not be obtained in sufficient quantity to allow for their complete characterization.

The amino acid compositions of the various peptides are presented in Table IV. All are quite similar and exhibit the features characteristic for type IV collagen. There are, however, several differences between the compositions of the individual peptides. The number and type of hydrophobic residues present show the greatest variation. The number of

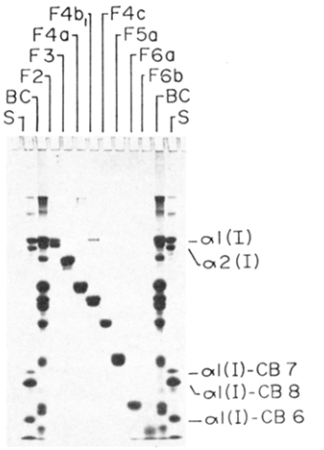

FIGURE 6: SDS–polyacrylamide gel electrophoresis of purified type IV collagen peptides (10% gel). Human type IV collagen from CM-cellulose fraction BC (Figure 1), type I collagen, and selected cyanogen bromide peptides were used as standards.

F2 Tyr-Phe-Asp-Leu-Gly-Leu-Gly-Gly-Asp-Gly-Gly-Asp-Hyp-Gly-Phe-Hyp
Leu-Gly-Leu-Gly-Gly-Asp-Gly-Gly-Asp-Hyp-Gly-Phe-Hyp

F4c Tyr-Phe-Asp-Leu-Gly-Leu-Gly-Gly-Asp-Gly-Gly-Asp-Hyp-Gly-Phe-Hyp
Leu-Gly-Leu-Gly-Gly-Asp-Gly-Gly-Asp-Hyp-Gly-Phe-Hyp

F4a Glu-Ala-Ile-XXX-Gly-Leu-Hyp

F3 Glu-Ala-Ile-Gly-Pro-Gly-Gly-Ile-Gly-Pro-Hyp

F5a Glu-Ala-Ile-Gly-Pro-Gly-Gly-Ile-Gly-Pro-Hyp

$F4b_1$ Asp-Met-Gly-Val-Met-XXX-Gly-Gln-Glu-Gly-Asp-Gln-Gly
Gly-Val-Met-XXX-Gly-Gln-Glu-Gly-Asp-Gln-Gly

$F4b_2$ Asp-Lys-Met-Asp-Met-Gly-Val-Met-XXX-Gly-Gln-Glu-Gly-Asp-Gln-Gly

FIGURE 7: Amino-terminal amino acid sequences of the peptides purified from pepsin-solubilized human type IV collagen. Gly-X-Y repeating sequences are underlined.

Table III: Molecular Weights of Peptides Isolated from Pepsin-Solubilized Human Type IV Collagen

peptide	M_r by gel electrophoresis	M_r by agarose 5M chromatography
F2	90 000	100 000
F3	75 000	80 000
F4a	60 000	55 000
$F4b_1$, $F4b_2$	50 000	49 000
F4c	41 000	44 000
F5a, F5b	28 000	30 000
F6a	19 000	22 000
F6b	12 000	

alanine residues is also variable, and F4a has a remarkably low content of arginine.

The amino-terminal amino acid sequences of the major peptides are presented in Figure 7. All of the sequences consist of short non-Gly-X-Y regions followed by regions composed of a repeating Gly-X-Y structure. The non-Gly-X-Y regions apparently are the remnants of larger non-triple-helical regions in the type IV collagen chains which are responsible for the pepsin sensitivity of the type IV collagen molecules.

Edman degradation of peptide F2 yielded two major phenylthiohydantoin derivatives in most cycles. Examination of these results clearly indicated that two identical sequences were present, one being longer than the other by three residues at the amino-terminal end. The existence of two forms of F2, as indicated by these results, may explain why the purified F2

Table IV: Amino Acid Compositions of Peptides Purified from Pepsin-Solubilized Human Type IV Collagen

amino acid	peptide								
	F2	F3	F4a	$F4b_1$	F4c	F5a	F5b	F6a	F6b
Hyp	118	116	105	113	142	121	122	89	103
Asp	35	37	32	36	34	33	34	34	52
Thr	20	29	17	30	19	11	12	32	33
Ser	50	29	43	33	50	38	34	60	55
Glu	70	66	86	72	65	65	62	115	71
Pro	87	62	82	71	79	82	72	82	62
Gly	338	344	349	350	337	345	354	303	318
Ala	30	52	35	66	31	48	49	42	37
Val	30	24	26	25	31	13	14	23	22
Met[a]	16	17	19	12	7.8	16	13	6.3	13
Ile	31	48	42	37	24	29	32	21	32
Leu	55	69	56	46	60	83	86	41	60
Tyr	3.2	4.1	2.7	3.4	3.1	3.2	2.6	10	7.7
Phe	24	26	19	29	28	38	38	29	29
Hyl	59	47	73	36	51	44	45	48	49
Lys	7.1	7.3	8.4	15	7.8	3.6	3.4	14	9.3
His	6.8	8.1	4.4	8.1	9.9	8.5	8.2	8.9	9.8
Arg	18	14	3.0	16	22	19	20	42	36

[a] Determined as methionine + methionine sulfoxide.

peptide often appears as two closely migrating bands upon gel electrophoresis (see Figure 6). Degradation of peptide F4c gave the same results; two sequences identical with those of peptide F2 were obtained, one being three residues longer than the other. This indicates that peptide F4c represents the amino-terminal portion of peptide F2.

Two overlapping sequences were also found for peptide $F4b_1$, one being longer than the other by two residues. These sequences were distinct from those of F2 and F4c. The peptide $F4b_2$ gave a single sequence which is the same as the longest form of $F4b_1$, but several residues longer at the amino-terminal end. This finding indicates that the peptides $F4b_1$ and $F4b_2$ have a common origin. The slight difference between the amino-terminal sequences of $F4b_1$ and $F4b_2$ may account for our ability to resolve them by ion-exchange chromatography.

Degradation of each of the other peptides yielded only a single sequence. Peptides F3 and F5a were found to have identical amino-terminal sequences, indicating that F5a represents the amino-terminal end of the peptide F3. The peptide F4a yielded a sequence which was unique from the sequences of all other peptides.

In order to reveal other possible common origins, each of the major peptides was cleaved with cyanogen bromide. The peptides generated by this digestion were analyzed by SDS–polyacrylamide gel electrophoresis, as shown in Figure 8. Even when the peptides were reduced before cleavage (Adelstein & Kuehl, 1970), the cyanogen bromide peptide patterns of peptides F3 and $F4b_1$ both contained large amounts of partially cleaved material in the high molecular weight region of the gel (data not shown).

Analysis of the peptide patterns obtained confirms the conclusions drawn from the sequence data alone. Peptide patterns generated by cleavage of the largest fragments, F2 and F3, exhibit no similarities, and there are no peptides in common, indicating that they are distinctly different. The peptides generated by cleavage of F4c are, except for a single peptide, completely represented in the pattern obtained for F2, indicating that F4c represents a portion of F2. Likewise, the peptides generated by cleavage of F5a are largely found in the pattern obtained for F3, indicating that F5a represents a portion of F3. Furthermore, the peptides generated by cleavage of $F4b_1$ and $F4b_2$ were found to be identical, confirming the conclusion that they share a common origin (data not shown).

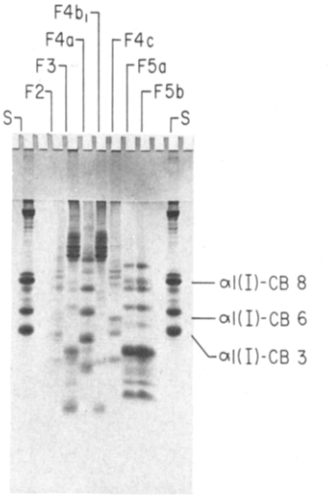

FIGURE 8: SDS–polyacrylamide gel electrophoresis of peptides generated by cyanogen bromide digestion of the purified type IV collagen peptides (15% gel). Type I collagen and selected cyanogen bromide peptides were used as standards.

The cyanogen bromide peptide patterns also reveal information about the origins of several other peptides. First, the patterns generated by cleavage of the peptides F5a and F5b are virtually identical, indicating that they share a common origin. Second, the peptides generated by cleavage of $F4b_1$ are largely found in the pattern obtained for F3. Since the amino-terminal sequences of F3 and $F4b_1$ differ, and since the cyanogen bromide peptide patterns for $F4b_1$ and F5a are quite different, it appears that the F4b peptides represent the carboxyl-terminal end of F3, which is not represented by the F5 peptides. And third, the peptides produced from F4a are represented by faint bands in the pattern obtained from peptide F2. Since the amino-terminal amino acid sequences of F4a and F2 differ, and since the cyanogen bromide peptide patterns for F4a and F4c are quite different, it appears that F4a represents the carboxy-terminal portion of F2, which is not represented by F4c.

Each of the major pepsin digestion peptides was also digested with trypsin, and the tryptic peptides were separated by HPLC. A comparison of the various elution profiles obtained in this manner (Figure 9) largely confirms the above conclusions. Peptides F2 and F3 are distinctly different; $F4b_1$

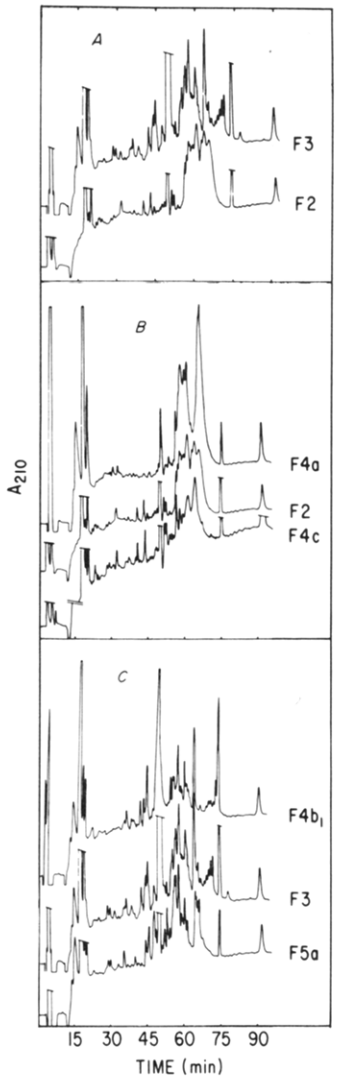

FIGURE 9: HPLC elution profiles of peptides generated by trypsin digestion of the various type IV collagen peptides. The scales for absorbance at 210 nm were omitted to allow the profiles to be moved closer together for easier comparison. (A) Profiles for tryptic peptides of F2 and F3. (B) Profiles for tryptic peptides of F4a and F4c; the profile for tryptic peptides of F2 is repeated for comparison. (C) Profiles for tryptic peptides of $F4b_1$ and F5a; the profile for tryptic peptides of F3 is repeated for comparison.

and F5a are derived from F3, and F4c and F4a are derived from F2.

Discussion

The procedure used here for the isolation of type IV collagen is an adaption of that outlined by Dehm & Kefalides (1978) and later modified by Glanville et al. (1979). This procedure has proven to be a reproducible means of obtaining highly purified native type IV collagen from the highly complex tissue of the human placenta. The ability of this procedure to remove other contaminating collagen types should make it applicable to the study of type IV collagen from other tissues, even those which contain relatively large amounts of interstitial collagen.

In earlier studies which used a similar procedure for the isolation of type IV collagen (Dehm & Kefalides, 1978; Glanville et al., 1979), those type IV collagen peptides found by gel electrophoresis to have molecular weights of less than 75 000 were only minor components of the final purified product. It appears that the differences between those results and ours are due to differences in the temperature and duration of the pepsin digestion steps. This is supported by our finding

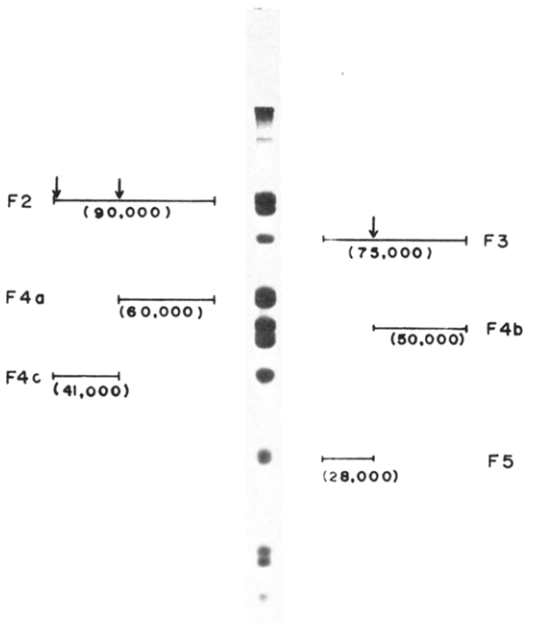

FIGURE 10: Relationship between the various pepsin-derived fragments of human type IV collagen. Lines represent the approximate molecular weights of the peptides. Arrows represent probable pepsin cleavage sites. The polyacrylamide gel electrophoresis pattern of pepsin-solubilized type IV collagen is provided for comparison, and lines representing each of the major peptides lie parallel to the corresponding bands.

that the relative amounts of the smaller fragments obtained were increased when the temperatures and/or durations of the pepsin digestion steps were further increased (data not shown).

In this research, we have established reproducible procedures for the purification of several distinct peptides present in pepsin-solubilized type IV collagen. These procedures largely incorporate classical methods for the purification of collagen peptides. However, in our most recent experiments, final purification of each peptide was achieved by reversed-phase high-performance liquid chromatography (HPLC), and this technique was found to offer several distinct advantages over other chromatographic systems. First, HPLC provided a resolution of peptides from contaminating species which was far superior to that available by using other existing techniques (see Figure 5). Second, recoveries from HPLC columns were generally over 90%, far better than the recoveries generally obtained with other chromatographic systems. And third, the buffer system used was completely volatile, thus allowing us to completely avoid the losses which occur when samples must be dialyzed to remove buffer ions. The only limitation of the system which was found was that sample size was generally limited to 1 mg of protein. However, since a chromatographic run could be completed in little more than an hour, and since column reequilibration could be done in a matter of minutes, this limitation proved to be of little practical consequence.

Our studies on the peptides isolated from pepsin-solubilized type IV collagen have allowed us to draw detailed conclusions concerning their origin. These conclusions are shown diagrammatically in Figure 10 and can be summarized as follows. Upon pepsin digestion, peptide F3 is liberated from type IV collagen by uniform cleavage of a non-Gly-X-Y domain. This peptide itself contains an internal non-Gly-X-Y domain, the cleavage of which by pepsin results in the liberation of the amino-terminal peptide F5 and the carboxy-terminal peptide F4b. There is more than one site in this domain at which pepsin can cleave, giving rise to slightly different forms of each

of the peptides F5 and F4b. In a similar fashion, peptide F2 is liberated from type IV collagen by a nonuniform pepsin cleavage. Peptide F2 contains an internal non-Gly-X-Y domain which is uniformly cleaved by pepsin, liberating the amino-terminal peptide F4c and the carboxy-terminal peptide F4a. It must be noted that the Gly-X-Y regions of the various peptides are parts of triple-helical structures during this process and are therefore not degraded by pepsin.

We have found that the peptide F2 elutes from carboxymethylcellulose at a lower ionic strength than the peptide F3. We have also found that peptides F2, F4c, and F4a have a lower alanine content than the peptides F3, F5a, F5b, $F4b_1$, and $F4b_2$. Furthermore, we have found that the peptide F2 and its proteolytically derived peptides (F4c and F4a) have cyanogen bromide and tryptic peptide patterns which are quite distinct from those of peptide F3 and its derived peptides (F5 and F4b). These results clearly separate the peptides studied here into two discrete sets, a conclusion which is consistent with the recent proposal that the pepsin digestion products of type IV collagen arise from two distinct polypeptide chains (Crouch & Bornstein, 1979; Alitalo et al., 1980; Crouch et al., 1980; Fessler & Fessler, 1980; Tryggvason et al., 1980; Gay & Miller, 1979; Timpl et al., 1978; Bailey et al., 1979; Dixit, 1980; Glanville et al., 1979; Kresina & Miller, 1979; Mayne & Zettergren, 1980). Furthermore, based upon the above characteristics and the criteria presented by Crouch et al. (1980), we conclude that the peptide F2 and its proteolytic products F4c and F4a are derived from the pro-α1(IV) chain and that the peptide F3 and its proteolytic products F5a, F5b, $F4b_1$, and $F4b_2$ are derived from the pro-α2(IV) chain. We have not found any evidence that a third type of type IV collagen chain exists, as has been suggested by several investigators (Timpl et al., 1979b; Hung et al., 1980).

The results of our analysis of human type IV collagen peptides allow us to better understand the research on pepsin-solubilized type IV collagen which has already appeared in the literature. The results of a comparison between the peptides studied here and those studied by other investigators are shown in Table V. Comparisons between these various peptides have been made in the following manner. First, the peptides were assigned as belonging to either the α1(IV) or the α2(IV) chain, largely on the basis of their relative alanine and arginine contents (Crouch et al., 1980). Their molecular weights, relative chromatographic behavior, cyanogen bromide peptide patterns, and amino-terminal sequences, if available, were then compared with those of the type IV collagen peptides examined in this report. From these comparisons, it appears that the pepsin digestion products of type IV collagen from human placenta, human glomerulus, bovine lens capsule, bovine kidney, and porcine kidney are remarkedly similar.

A number of investigators have described an apparent precursor to the peptide F3, with a molecular weight of 95 000. Reduction of this precursor liberates an F3-like fragment and a 15 000-dalton fragment which migrates as two bands upon gel electrophoresis (Dixit & Kang, 1980). From the sizes of these small peptides and an examination of their amino acid compositions, it seems likely that they are identical with the F6a and F6b peptides that we have described. However, a more extensive characterization of the F6 peptides will be needed before this relationship can be confirmed.

A comparison of the human type IV collagen peptides studied here with those obtained by pepsin digestion of the EHS tumor (Timpl et al., 1979b) indicates that the pepsin digestion products of the EHS tumor type IV collagen are quite different from those of human type IV collagen. However,

Table V: Comparison of type IV collagen peptides described by other investigators with those in this report[a]

Study (tissue used)	Peptide Name	Mol. wt. M.S.	Mol. wt. PAGE	of Origin	Same as or Similar to	Basis for Comparison MW	CB	CMC	Seq.
Glanville et al., 1979 (placenta)	α	95		α1(IV)	F2	✓	✓	✓	
	α'	95		α1(IV)	F2	✓	✓	✓	
	α''	70		α2(IV)	F3	✓	✓	✓	
Glanville and Rauter, 1981 (placenta)	α1(IV)95,000	95		α1(IV)	F2	✓			✓
	B	36	39	α1(IV)b	F4c	✓			✓
	D		19	α1(IV)b	F4a	X			✓
	α2(IV)70,000	70		α2(IV)	F3	✓			✓
	A	45	47	?	?				
	C	27	28	?	?				
	E		11	?	?				
	F		+8	?	?				
	G		+4	?	?				
Bailey et al., 1979 (placenta)	170K		170	α1(IV)	F2	X	✓	✓	
	100K		100	α1(IV)	F2	✓	✓	✓	
	70K		70	α2(IV)	F3	✓	✓	✓	
Dixit, 1979 (glomerulus)	C	95	95	α1(IV)	F2	✓			
	D	95	95	α2(IV)	F3	✓			
Kresina and Miller, 1979 (placenta)	C'	120	140	α1(IV)	F2	X	✓	✓	
	C	95		α1(IV)	F2	✓	✓	✓	
	$50K_1$	50	60	α1(IV)	F4a	✓		✓	
	D	95		α2(IV)	F3	X	✓	✓	
	80K	80		α2(IV)	F3	✓	✓	✓	
	$50K_2$	40		α2(IV)	F4b	✓		✓	
	15K[c]	15		α2(IV)c	F6a + F6b	✓		✓	
Sage et al., 1979 (placenta)	140K	140		α1(IV)	F2	X	✓	✓	
	100K	100		α1(IV)	F2	✓	✓	✓	
	70K	70		α1(IV)	?2	X	✓	✓	
	$70K_2$	70		α2(IV)	F3	✓	✓	✓	
Daniels and Chu, 1975 (bovine gomeruli)	C	140		α1(IV)	F2	X		✓	
	A	93		α1(IV)	F2	✓		✓	
	D	140		α2(IV)	F3	X		✓	
	B	93		α2(IV)	F3	X		✓	
Dixit, 1979 (porcine kidney)	C	95	95	α1(IV)	F3	✓	✓	✓	
	D	95	75	α2(IV)	F3	✓	✓	✓	
Dixit, 1980 (bovine kidney)	C-1	140	110	α1(IV)	F2	X	✓	✓	
	C	95		α1(IV)	F2	✓	✓	✓	
	$50K_1$	50		α1(IV)	F4a	✓		✓	
	D-1	120		α2(IV)	F3	X	✓	✓	
	D	90		α2(IV)	F3	X		✓	
	75K	75		α2(IV)	F3	✓	✓	✓	
	$50K_2$	50		α2(IV)	F4b	✓		✓	
	15K[c]	15		α2(IV)c	F6a + F6b	✓		✓	
Dixit and Kang, 1979 (bovine lens)	C-1	110		α1(IV)	F2	X	✓		
	C	95		α1(IV)	F2	✓	✓	✓	
	D	95	75	α2(IV)	F3	✓	✓	✓	
Dixit and Kang, 1980 (porcine kidney)	D	90		α2(IV)	F3	X			
	75K	75		α2(IV)	F3	✓			
	15K	15		α2(IV)c	F6a + F6b	✓			
Gay and Miller, 1979 (bovine lens)	C	~100		α1(IV)	F2	✓	✓	✓	
	$50K_1$	50		α1(IV)	F4a	✓		✓	
	D	~100		α2(IV)	F3	X	✓	✓	
	80K	80		α2(IV)	F3	✓			
	50K	50		α2(IV)	F4b	✓		✓	
	$15K^2$	15		α2(IV)c	F6a + F6b	✓			

[a] Peptide molecular weights (in thousands) obtained by other investigators using molecule sieve chromatography (M.S.) and polyacrylamide gel electrophoresis (PAGE) are listed separately. Those parameters used as a basis for the relationships drawn are molecular weight (M.W.), cyanogen bromide peptide patterns (C.B.), relative elution upon carboxymethyl cellulose chromatography (CMC), and amino terminal amino acid sequences (Seq.).

Agreement between these parameters and those reported for a particular peptide studied in this report is denoted by a check (✓). Disagreement is denoted by an X. Where no mark appears the necessary data was unavailable.

[b] The chain of origin for B and D was determined by comparison of sequence data. Amino acid compositions for fragments A through G were too similar to allow their chain of origin to be proposed.

[c] The 15K fragment is liberated by reduction of the D fragment, and therefore must be derived from α2(IV).

[+] Molecular weights determined by amino acid analysis.

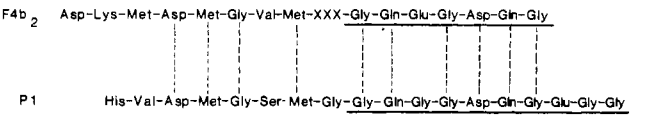

FIGURE 11: Apparent sequence homologies between the human placental type IV collagen peptide $F4b_2$ and the type IV collagen peptide P1 from the EHS mouse tumor. Underlined sequences have a Gly-X-Y repeating structure. Dotted lines between sequences indicate that the same residue appears in that location in both chains. [Sequence of P1 is from Timpl et al. (1979b).]

if we compare the amino-terminal amino acid sequence of the mouse tumor peptide P1 (Timpl et al., 1979b) with that of our human peptide $F4b_2$, we find that peptide P1 has a sequence which agrees with that of peptide $F4b_2$ in 10 positions and disagrees only in 4 positions (Figure 11). The greatest variation between the two sequences occurs in the non-Gly-X-Y regions. Using the amino acid composition of $F4b_2$ to calculate relative frequencies for each amino acid, and giving glycine in Gly-X-Y a probability of 1, we calculate that the chance occurrence of this homology has a probability of 2×10^{-5}. It remains unclear whether the peptides P1 and $F4b_2$ represent analogous regions in one of the type IV collagen chains or whether the similarities between their amino-terminal

sequences reflect the existence of similar sequences in different regions, within one chain or in separate chains, that have arisen in response to similar functional requirements.

Acknowledgments

The technical assistance of Richard Woller and Hamiyet King is gratefully acknowledged.

Registry No. Pepsin, 9001-75-6.

References

Adelstein, R. S., & Kuehl, W. M. (1970) *Biochemistry 9*, 1355.
Alitalo, K., Vaheri, A., Krieg, T., & Timpl, R. (1980) *Eur. J. Biochem. 109*, 247.
Bailey, A. J., Sims, T. J., Duance, V. C., & Light, N. D. (1979) *FEBS Lett. 99*, 361.
Bhown, A. S., Mole, E. J., Weissinger, A., & Bennett, J. C. (1978) *J. Chromatogr. 148*, 582.
Bornstein, P., & Sage, H. (1980) *Annu. Rev. Biochem. 49*, 957.
Caufield, J. P., & Farquhar, M. G. (1978) *Lab. Invest. 39*, 505.
Crouch, E., & Bornstein, P. (1979) *J. Biol. Chem. 254*, 4197.
Crouch, E., Sage, H., & Bornstein, P. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 745.
Daniels, J. R., & Chu, G. H. (1975) *J. Biol. Chem. 250*, 3531.
Dehm, P., & Kefalides, N. A. (1978) *J. Biol. Chem. 253*, 6680.
Dewes, H., Fietzek, P. P., & Kühn, K. (1979) *Hoppe-Seyler's Z. Physiol. Chem. 360*, 821.
Dixit, S. N. (1978) *FEBS Lett. 85*, 153.
Dixit, S. (1979) *FEBS Lett. 106*, 379.
Dixit, S. (1980) *Eur. J. Biochem. 106*, 563.
Dixit, S., & Kang, A. H. (1979) *Biochemistry 18*, 5686.
Dixit, S., & Kang, A. H. (1980) *Biochemistry 19*, 2692.
Fessler, L. I., & Fessler, J. H. (1980) *J. Supramol. Struct. 4*, 178.
Gay, S., & Miller, E. J. (1979) *Arch. Biochem. Biophys. 198*, 370.
Glanville, R. W., & Rauter, A. (1981) *Hoppe-Seyler's Z. Physiol. Chem. 362*, 943.
Glanville, R. W., Rauter, A., & Fietzek, P. P. (1979) *Eur. J. Biochem. 95*, 383.
Hassell, J. R., Gehron Robey, P., Barrach, H. J., Wilczek, J., Rennard, S. I., & Martin, G. R. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 4494.
Hung, C., Butowski, R. J., & Hudson, B. G. (1980) *J. Biol. Chem. 255*, 4964.
Kefalides, N. A. (1971) *Biochem. Biophys. Res. Commun. 45*, 226.
Kefalides, N. A. (1973) *Int. Rev. Connect. Tissue Res. 6*, 63.
Kresina, T. F., & Miller, E. J. (1979) *Biochemistry 18*, 3089.
Kühn, K., Fietzek, P. P., & Kühn, J. (1966) *Biochem. Z. 344*, 418.
Laemmli, U. (1970) *Nature (London) 227*, 680.
Mayne, R., & Zettergren, J. G. (1980) *Biochemistry 19*, 4065.
Orkin, R. W., Gehron, P., McGoodwin, E. B., Martin, G. R., Valentine, T., & Swarm, R. (1977) *J. Exp. Med. 145*, 204.
Sage, H., Woodbury, R. G., & Bornstein, P. (1979) *J. Biol. Chem. 254*, 9893.
Stenman, S., & Vaheri, A. (1978) *J. Exp. Med. 147*, 1054.
Timpl, R., Martin, G. R., Bruckner, P., Wick, G., & Wiedmann, H. (1978) *Eur. J. Biochem. 84*, 43.
Timpl, R., Rhode, H., Ott-Ulbricht, V., Risteli, L., & Bachinger, H. P. (1979a) in *Glycoconjugates* (Schauer, R., Boer, P., Buddecke, E., Kramer, M. F., Villiegenthart, J. F. G., & Wiegandt, H., Eds.) p 145, George Thieme, Stuttgart, West Germany.
Timpl, R., Bruckner, P., & Fietzek, P. P. (1979b) *Eur. J. Biochem. 95*, 255.
Tryggvason, K., & Kivirikko, K. I. (1978) *Nephron 21*, 230.
Tryggvason, K., Gehron Robey, P., & Martin, G. R. (1980) *Biochemistry 19*, 1284.
van der Rest, M., & Fietzek, P. (1982) *Eur. J. Biochem. 125*, 491.
Van Orden, H. O., & Carpenter, F. H. (1964) *Biochem. Biophys. Res. Commun. 14*, 399.
Vracko, R. (1974) *Am. J. Pathol. 77*, 314.