Molecular Cloning of the Human Goodpasture Antigen Demonstrates It To Be the \( \alpha3 \) Chain of Type IV Collagen

Neil Turner,* Philip J. Mason,† Ron Brown,# Margaret Fox,¹ Sue Povey,¹ Andrew Rees,* and Charles D. Pusey*

*Renal Unit, Department of Medicine, and †Department of Haematology, Royal Postgraduate Medical School, London W12 0NN, United Kingdom; ‡Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2, United Kingdom; and §Medical Research Council Human Biochemical Genetics Unit, The Galton Laboratory, UCL, London NW1 2HE, United Kingdom

Abstract

To characterize the autoantigen of Goodpasture's (anti-glomerular basement membrane) disease, a molecule of 26-kD reactive with autoantibodies from patients' sera was purified from collagenase digests of sheep glomerular basement membrane. Short internal amino acid sequences were obtained after tryptic or cyanogen bromide cleavage, and used to deduce redundant oligonucleotides for use in the polymerase chain reaction on cDNA derived from sheep renal cortex. Molecules of 175 bp were amplified and found to come from two cDNA sequences. One was identical to that of a type IV collagen chain (\( \alpha5 \)) cloned from human placenta and shown to be expressed in human kidney. The other was from a type IV collagen chain with close similarities to \( \alpha1 \) and \( \alpha5 \) chains, and was used to obtain human cDNA sequences by cDNA library screening and by further polymerase chain reaction amplifications. The correspondence of the derived amino acid sequence of the new chain with published protein and cDNA sequences shows it to be the \( \alpha3 \) chain of type IV collagen. Its gene, COL4A3, maps to 2q36-2q37. The primary sequence and other characteristics of this chain confirm that it carries the Goodpasture antigen. (J. Clin. Invest. 1992. 89:592–601.) Key words: glomerular basement membrane • anti-glomerular basement membrane disease • glomerulonephritis • Alport's syndrome • chromosome 2

Introduction

Goodpasture's disease is an autoimmune condition in which rapidly progressive glomerulonephritis and lung hemorrhage are associated with antibodies to glomerular and alveolar basement membranes. The autoantibodies have been shown to be pathogenic when transferred to primates (1), and to have a highly restricted specificity (2-4). Their target, the Goodpasture antigen, is limited to certain basement membranes (2, 5), and is known to be closely associated with the major, COOH-terminal noncollagenous (NC1) domain of type IV collagen.

Address correspondence to Neil Turner, M.D., Renal Unit, Department of Medicine, Royal Postgraduate Medical School, Ducane Road, London W12 0NN United Kingdom.

Received for publication 8 March 1991 and in revised form 3 October 1991.

1. Abbreviations used in this paper: CS, collagenase-solubilized; GBM, glomerular basement membrane; hGBM, human GBM; NC1, COOH-terminal noncollagenous domain of type IV collagen; sGBM, sheep GBM; PBS/Tw, PBS containing 0.05% Tween 20; PCR, polymerase chain reaction.

Histochimical and protein sequence data (largely from bovine material) have suggested that the autoantigenic epitopes are carried on a novel chain of type IV collagen, designated \( \alpha3 \) (6-8).

The ubiquitous \( \alpha1 \) and \( \alpha2 \) chains of type IV collagen have been extensively characterized and their genes cloned (9). cDNAs for two new type IV collagen chains have been isolated recently. Hostikka et al. identified a cDNA encoding the \( \alpha5 \) chain from a human placental cDNA library, and showed it to be expressed in human kidney (10). Its gene was localized to the q22 region of the X chromosome, making it an obvious candidate for involvement in Alport's syndrome. Morrison et al. (11) have obtained a bovine cDNA encoding the \( \alpha3 \) chain by amplifying sequences from the bovine antigenic chain (6). This report describes the characterization and purification of the Goodpasture antigen from sheep kidney, and the subsequent cloning of human cDNAs encoding the molecule. Its identity is confirmed as the noncollagenous domain of the \( \alpha3 \) chain of type IV collagen.

Methods

Preparation of collagenase-solubilized glomerular basement membrane (GBM). The technique described is based on the method of Spiro (12). Kidneys were frozen at \(-20°C\) within 24 h of death. Partially thawed kidneys were decapsulated, sliced, and the medulla removed and discarded. The cortex was minced, pushed through a 150-\( \mu \)m stainless steel sieve, and repeatedly washed with cold PBS pH 7.4. Separated glomeruli were passed through a 250-\( \mu \)m sieve to remove large fragments and collected on a 63-\( \mu \)m sieve. Isolated glomeruli were then washed three times in cold PBS and examined by light microscopy to ensure that tubular contamination was \(<5\%\). The glomeruli were sonicated at 18-\( \mu \)m amplitude in 30-s bursts on ice until disrupted, and the mixture spun at 700 g for 15 min. The supernatant containing cellular debris was discarded, and the pellet containing basement membrane fragments was washed three times with ice cold distilled water and lyophilized. Freeze-dried material was resuspended at 10 mg/ml in 0.1 M Tris/0.005 M calcium acetate buffer pH 7.4, and digested with 0.7% by weight of type I collagenase (Sigma Chemical Co., St. Louis, MO) which had been purified as described by Seifer and Gallop (13). 5 mM N-ethylmaleimide, 1 mM PMSF, and 25 mM amino-n-caproic acid were added as protease inhibitors to preparations of sheep glomerular basement membrane (sGBM) but not to human glomerular basement membrane (hGBM). Digestion was carried out by stirring at 37°C for 1 h, after which collagenase was inactivated by heating at 60°C for 10 min. Insoluble material was removed by centrifugation at 700 g for 15 min and the protein concentration of the collagenase-solubilized glomerular basement membrane (CS-GBM) estimated.

Chromatographic purification of antigenic components. Collagenase-solubilized sheep GBM (CS-GBM) was lyophilized and resuspended in 6 M guanidine hydrochloride in Tris-HCl pH 7.0 at \(-25\) mg/ml. 20–30 mg at a time was heated at 100°C for 15 min before applying to a Sephacryl 200 HR (Pharmacia LKB Biotechnology Inc., Piscataway, NJ) 17 mm × 90 cm gel-filtration column equilibrated in the same buffer. The column was eluted at 0.5 ml/min and 4-min fractions collected while monitoring the optical density of the eluate at 280 nm. 0.1 ml of each fraction was ethanol precipitated, resuspended

592 N. Turner, P. J. Mason, R. Brown, M. Fox, S. Povey, A. Rees, and C. D. Pusey
in 10 µl of SDS-PAGE loading buffer, and analyzed on miniature 12.5% polyacrylamide gels (PhastSystem; Pharmacia). Duplicate gels were silver stained and immunoblotted by diffusion to nitrocellulose overnight under a weighted stack of paper towels. Nitrocellulose sheets were blocked in 0.5% Triton X-100, 0.3% Tween 20 for 30 min, and incubated with serum from a patient with anti-GBM antibodies for a dilution of 1 in 20 in PBS containing 0.05% Tween 20 (PBS/Tw) for 1 h. Bound antibody was recognized by alkaline phosphatase–coupled secondary antibodies (Sigma) using the method described by Blake et al. (14). Fractions found to contain “monomer” components (24–30 kD) alone were pooled, acidified by adding trifluoroacetic acid to 0.5%, and injected onto a Dynamax 10 × 100 mm C18 reverse-phase column (Rainin Instrument Co. Inc., Woburn, MA) equilibrated in 20% acetonitrile/0.1% trifluoroacetic acid. Elution was by a gradient of acetonitrile 20–40% over 35 min (HPLC system; Gilson Co., Inc., Worthington, OH). 1-min fractions were collected, and aliquots evaporated to dryness in a SpeedVac (Savant Instruments, Inc., Hicksville, NY), resuspended in water, and analyzed on a mini-SDS-PAGE system as above. Fractions containing a single band on silver staining that were also antigenic on immunoblotting were pooled and lyophilized before further analysis.

Inhibition radioimmunoassay for detection of antigenic material. The RIA for detecting anti-GBM antibodies has been described previously (15). Collagenase-solubilized human GBM (CS-hGBM) was coated to flexible polyvinylchloride microtitre plates (Dynatech Laboratories, Inc., Alexandria, VA), and test or control serum at 1 in 8 in PBS/Tw for 1 h, added to the coated wells in triplicate, incubated at 37°C for 1 h, washed with PBS/Tw, and drained. Bound IgG was detected by incubating wells with 125I-labelled anti–human IgG (2 × 10^6 cpm) in PBS/Tw for 1 h. Plates were washed with PBS/Tw and dried for 10 min before counting the wells in a gamma counter. This assay was modified to detect soluble antigen by the incorporation of an inhibition step. The primary (anti-GBM) antibody was preincubated in solution with test or control material in PBS/Tw at 37°C for 1 h. CS-hGBM and hemoglobin were used as positive and negative controls. The reduction in counts per well (in triplicate) induced by preincubation with the putative antigen was measured. By using the test serum at a dilution that gives 50% of maximum counts (estimated previously), this system can detect antigen contained in 0.1–1.0 µg of CS-hGBM per well. Positive sera were obtained from patients diagnosed as having Goodpasture’s disease by clinical criteria, linear deposition of IgG on the GBM by direct immunofluorescence of renal biopsy specimens, and positive results in the RIA for circulating antibodies.

Peptide cleavage and amino acid sequencing. Purified antigenic monomer from sGBM was pooled, lyophilized, and subjected to NH2-terminal sequence analysis on a pulsed-limited automated amino acid sequencer (477A; Applied Biosystems Inc., Foster City, CA) or cleaved and the separated peptides analyzed in the same way. Digestion with trypsin (Boehringer Mannheim Corp., Indianapolis, IN) was performed in 100 mM Tris-HCl pH 8.5 at 37°C for 4 h. For cyanojen bromide cleavage, the lyophilized protein was taken up in 70% formic acid, a crystal of cyanojen bromide added, and the tube was left at room temperature in the dark overnight. The products of both methods of cleavage were separated by reverse-phase HPLC using a 2.1 × 30-mm column with a flow rate of 0.3 ml/min (Aquapore RP300; Brownlee Labs, Santa Clara, CA) eluting with a gradient of 1–90% acetonitrile in 0.1% trifluoroacetic acid over 90 min.

Amplification of sequences from cDNA. The overall cloning strategy is shown in Fig. 1. Methods involving recombinant DNA were performed using standard techniques (16). Lamb kidneys were obtained fresh from an abattoir. Human fetal kidney was obtained from the Medical Research Council Tissue Bank (Dr. L. Wong). RNA was extracted by the guanidine thiocyanate lysis/caesium chloride gradient method as originally described by MacDonald et al. (17). Cortex was dissected from the kidneys and snap-frozen in liquid nitrogen, stored at −70°C, and reduced to a powder under liquid nitrogen before homogenizing with guanidine thiocyanate solution. Oligo-dT was used to prime first strand cDNA synthesis of 100 mg of total RNA (18) (AMV reverse transcriptase XL; Life Sciences Inc., St. Petersburg, FL). One-fiftieth of this reaction was used in each polymerase chain reaction (PCR) experiment, which was performed under standard conditions with the addition of 0.5% Tween and 0.5% NP-40. For the reactions on lamb kidney cDNA the forward primer was GC(GATCC)(GATCT)(GATCTT)(TCTAT)(TCA)(GAG)(ATG) (redundancies in brackets), and the reverse primer TTCAT(GA)(CA)(GATC)(TCT)(CTG)(GA)(CA). Because of the high level of redundancy these primers were used at a concentration of 50 µg/ml, 10 times usual. Magnesium ion concentrations of 1–4 mmol/liter were compared. 30–40 cycles of 1 min each at 94°C, 55°C, and 72°C were used, followed by 10 min at 72°C, in automated thermal cycling devices manufactured by Perkin-Elmer Cetus Instruments, Norwalk, CT, or Techne Inc., Princeton, NJ. Tag polymerase was purchased from Perkin-Elmer Cetus. Reactions were analyzed on a 1.8% agarose gel and product bands purified by electrophoresis from polyacrylamide gels. The eluted DNA was phosphorylated and blunt-end ligated into Smal-cut M13mp8. Single-stranded DNA was prepared from clones hybridizing with the oligo-labelled PCR product and sequenced. Amplification of sequences from human fetal kidney cDNA was performed in the same way. The forward primer (HGRGTC) was CTGAA(TTCCACGGAAGGAGGAACTG), and the reverse primer (ENIIRS) CTTCTGCAAGGACTTTATGTTGACGAGGAGGAGGAACTG.
T(TC)/TC. Purified products were digested with EcoRI and PstI to cut at restriction sites within primers and cloned into M13 vectors.

Screening of human cDNA libraries. The short product from the amplification of a new type IV collagen mRNA sequence from human kidney was gel purified and labelled with $[^{32}P]dCTP$ using the Klenow fragment of DNA polymerase (Amersham Corp., Arlington Heights, IL) and the primers used in the amplification reaction. Filters were hybridized with $[^{32}P]dCTP$-labelled probes in 5X standard saline/citrate (5X Denhardt’s/0.1% SDS and washed in 0.2X standard saline citrate (SSC) at 65°C. Positive clones were obtained from a cDNA library in Agt10 made from adult human renal cortex (Clontech Laboratories, Inc., Palo Alto, CA). Inserts from positive clones were subcloned into M13 and pUC vectors and analyzed by restriction endonuclease digestion and sequencing.

Extension of cDNA sequence. Because of difficulties in obtaining clones that extended beyond the midpoint of the NC1 domain (Fig. 1), a PCR-based strategy was developed based on the observation that the motif Gly-Pro-Pro-Gly-Pro (GPPGP) occurs frequently in type IV collagens. A primer based on this sequence, ACGTCGACGG(TA)CC-(TAC)CC(TA)GG(TAC)CC, was used together with a reverse primer, ACATTCTTTCTGGTITAAAGTA, corresponding to a sequence (SLNPERM) from the most variable region of the known chains. First strand cDNA was made from 1 µg of poly(A)+ RNA (isolated by oligo-dT–cellulose column chromatography) from human fetal kidney, by specific priming with the reverse PCR primer. 30 cycles of 45°C 1 min, 72°C 2 min, 94°C 1 min, were used. PCR products were digested with Sall (site in forward primer) and EcoRI (site present in the target molecule) and cloned into M13mp18 and mp19. Plaques were screened with the small EcoRI fragment of KcD5 described below.

Nucleotide sequencing. Single-stranded DNA from M13 plaques or double-stranded DNA in pUC vectors was sequenced by the dideoxy chain termination method using T7 DNA polymerase (Sequenase; United States Biochemical Corp., Cleveland, OH). All PCR-derived sequences were verified in more than two clones, except nucleotides 1–280, which were sequenced in both strands of one clone.

Northern analysis. Total RNA prepared as described above was separated in 1% agarose/formaldehyde gels in 3-(4-Morpholino)propane sulfonic acid buffer (16) and blotted to Genescreen Plus (Dupont Co., Wilmington, DE) membranes as recommended by the manufacturer. Filters were hybridized with the large EcoRI fragment of the cDNA clone KcD5 described below (and Fig. 1), or with a probe encoding the NC1 domain of α1(IV) collagen derived from the plasmid pH-21 (19).

Chromosomal localization. The human–rodent hybrid cell lines used (Table I) have been described previously (20, 21). The subclones of FG10E8 listed in Table I have been characterized by further karyotyping since their original description. Genomic DNA was isolated from hybrids and rodent cells as described by Edwards et al. (22). Oligonucleotides were designed for amplification of part of the 3’ noncoding region of the α3 cDNA sequence, as this region is likely to be the most divergent region between different species. The sequences, ACTGCTCATACGTTGATTGATGAA and TCAGGGAATCCCCCTATT-GCGTGA, from nucleotides 1294–1318 and 1600–1624 were designed to amplify a fragment of 330 bp from genomic DNA. 25 PCR cycles were performed as described above, with an annealing temperature of 50°C and 1-min extension times at 72°C. Products were analyzed by agarose gel electrophoresis and Southern blotting.

Figure 2. Purification of the antigenic component from sheep GBM by reverse-phase HPLC. Upper panel shows chromatogram. Middle and lower panels show SDS-PAGE of alternate fractions analyzed by silver staining (middle panel) and immunoblotting with Goodpasture serum (lower panel). The antigenic band elutes later than most of the monomer bands, and is barely visible on the silver-stained gels.
Tryptic fragments

1 AHPFIEC (170–176)
2 CQVCMK (223–228)

Cyanogen bromide fragments

3 XXLFNCINDVCNXA (61–74)
4 FTSAGXE (146–152)

Figure 3. Amino acid sequences obtained from peptides derived from antigenic sheep monomer, with their corresponding residues in the α1 chain (numbered as in Fig. 9) shown in brackets.

PCR product from human kidney. This clone contained a 16-kb segment of human DNA including the DNA encoding the COOH-terminal 65 amino acids and the 3′ noncoding region of COL4A3. The whole genomic clone was biotinylated and used as a probe for in situ hybridization according to procedures of Lichter et al. (24) with certain modifications. Human cot-1-DNA (Bethesda Research Laboratories, Gaithersburg, MD) was used to suppress hybridization of repeat DNA sequences and added to the hybridization mixture at 50× excess of probe DNA concentration, where the final concentration of probe DNA in the hybridization cocktail was 10 µg/ml. 20 µl was used per slide under 22-mm diameter cover slips. After denaturation of the probe mixture, preannealing of repeat sequences was allowed for 30 min at 37°C before application to chromosome preparations from normal cultured lymphocytes, separately denatured. After overnight incubation and posthybridization washing, the signal was detected using FITC-conjugated avidin (Vector Laboratories, Inc., Burlingame, CA) and amplified by alternate applications of biotinylated antivitin and FITC avidin. Procedures took place at room temperature using 5% nonfat dried milk as a blocking agent. Preparations were mounted in antifade solution p-phenylenediamine dihydrochloride (Sigma) to which diamidino-phenylindole and propidium iodide had been added for counterstaining and banding. Preparations were evaluated using epifluorescence microscopy. Images were collected electronically separately for each fluorochrome using a confocal laser microscope (MRC 600; Bio-Rad Laboratories, Richmond, CA) and narrow band pass filters of 550 nm for FITC and 610 nm for propidium iodide, respectively. The images were then merged.

(a) GCCCCATCCATTATTTGAGTTGTCATGGCGGGGACCTGCAACTATTTACGCGCACTCTCTACAGCTTCTTGCGGACCTAGCGAGCTGCA
AHPFIECHGRGTNCNYYANSSYSFWLASRTDVVS
(b) GCCCCATCCATTATTTGAGTTGTCATGGCGGGGACCTGCAACTATTTACGCGCACTCTCTACAGCTTCTTGCGGACCTAGCGAGCTGCA
AHPFIECHGRGTNCNYYANSSYSFWLASRTDVVS
(c) CTCAGATTCCCAGGGAAGGAGACGCTGCACTACATTTCAATTCTACAGTTTCTGCGTCCTAGTAAACCCAGAA
HGRFTCNYSNSYSYSFWLASLNPEDAK

(a) GACATGTTCGCAAACCCTCTAATGGAACCTGGAAGGCACTAGCGCACTCTCTACAGCTTCTTGCGGACCTAGCGAGCTGCA
DMFSPKQSETLKDRLTRISRQCQVMK
(b) AGAAATTTGACGAATACCCATCATCTACACTGGAAGCCTGGGAGGAGATGAGAAACATATTAGTCCGTGCAAGGCTACAGCGAGCTGCA
RMFRKPISTVKAGELENISRCQVCVMK
(c) AGAAATTTGACGAATACCCATCATCTACACTGGAAGCCTGGGAGGAGATGAGAAACATATTAGTCCGTGCAAGGCTACAGCGAGCTGCA
RMFRKPISTVKAGELENISRCQVCVMK

Figure 4. cDNA and derived amino acid sequences from products of amplification reactions. Oligonucleotide sequences are underlined. (a) and (b) Products from amplifications of lamb renal cortex cDNA between oligonucleotides designed according to peptides 1 and 2 in Fig. 3. The derived amino acid sequence of (a) is identical to that of the human type IV collagen α5 chain (10) in this region. Sequence (b) has distinct differences from α5 and α1 chains. (c) The product of amplification of human fetal kidney cDNA between oligonucleotides made to sequence (b).

Results

Chromatographic purification and amino acid sequencing. NC1 domains of the various α-chains are released as hexamers by collagenase digestion of GBM. Under dissociating conditions these can be separated into individual NC1 domain “monomers” of 24–30 kD (7). Fig. 2 shows the reverse-phase HPLC separation of isolated monomers from CS-hGBM, with analysis of the eluate by SDS-PAGE and immunoblotting. Eluting later than other monomer bands were two distinct 26-kD components, one of which gave a dense band in the silver stain, although it was barely antigenic on the Western blot, while the other gave rise to a weakly silver-staining band in later fractions that was strongly antigenic on Western blotting. Fractions 22–24 inhibited the binding of patients’ autoantibodies to CS-hGBM by 80% in the RIA; other fractions gave < 20% inhibition. The intact purified monomer gave no signal when subjected to Edman degradation, and was assumed to have become NH₂-terminal blocked during the purification procedure. Short protein sequences were obtained by microsequencing of fragments obtained by cleaving this antigenic material. Many of these were double or triple and difficult to analyze, but distinct sequences were obtained from four peptides (Fig. 3). Two tryptic fragments of six and seven amino acids were chosen as suitable for deduction of redundant oligonucleotides. The amino acid sequence used for the α₅ oligonucleotide differed from α₁ and α₂ chain sequences in the first two of its seven residues, but the sequence used for the α₅ oligonucleotide is conserved in other type IV collagen chains. 32-fold and 384-fold redundant synthetic oligonucleotides were generated for use in the next set of experiments.

Enzymatic amplification and sequencing of cDNA. An amplification product of the predicted size of 173-bp was shown to contain two different cDNA sequences that were represented approximately equally in 13 clones examined (Fig. 4, a and b). When translated into amino acid sequence, both were found to be closely related to human type IV collagen α₁ and α₂ chains. One encoded an amino acid sequence identical to that of the human cDNA for the α₅ chain of type IV collagen (10). The
Fig. 5. Nucleotide and derived amino acid sequence of human type IV collagen α3 chain cDNA. Cysteine residues in the NC1 domain are ringed and the beginning of the NC1 domain is indicated. These data are available from EMBL/GenBank/DDBJ under accession number M81379.
other came from a similar chain with identical derived amino acid sequence to that of the subsequently described bovine \( \alpha_3 \) chain cDNA (11). Using primers made from this sequence, a cDNA encoding an almost identical derived amino acid sequence was amplified from RNA from human fetal kidney (Fig. 4).

**Human cDNA cloning.** The short PCR-derived human cDNA was used as a probe to isolate a 1.05-kb cDNA (KcD5) from a human kidney library in \( \lambda gt10 \). This extended from approximately the middle of the NC1 domain of the new chain to the COOH terminus and included 670 bp of 3' noncoding region, but did not contain a consensus polyadenylation signal (25). Other cDNA clones proved to be shorter than KcD5 and not to extend the sequence in either direction. A PCR strategy based on the recurrent collag enous motif GPPGP was used to extend the sequence into the collagenous region, and gave rise to a product of 940 bp which hybridized with the small EcoRI fragment of KcD5. The overlapping sequences obtained from KcD5 and these PCR-derived clones span 1,691 bp, including 108 residues of collagenous sequence (Gly-Xaa-Yaa triplets), the entire 232 amino acids of the NC1 domain, and 670 bp of the 3' noncoding region. Nucleotide and derived amino acid sequences are shown in Fig. 5. A GPPGP motif at residues 46-50 was not recognised by the 5' oligonucleotide, as the

don for the second glycine, GGG, was not included in the redundant primer mix.

**Northern analysis.** Hybridization of a probe from the 3' end of KcD5 was seen to a mRNA species present in human kidney of very similar size to that hybridizing with an \( \alpha_1(IV) \) probe (Fig. 6). As expected from what is known of the relative amounts of \( \alpha_1 \) and \( \alpha_3 \) protein in the kidney, the signal from \( \alpha_3 \) was weaker. Expression of the \( \alpha_3 \) chain is much more restricted than that of the \( \alpha_1 \) chain. In particular, there was insignificant or absent expression in cells with very high levels of \( \alpha_1 \) chain expression: glomerular mesangial cells and HT1080 cells.

**Chromosomal mapping.** Two oligonucleotides were synthesized from the sequence of the 3' noncoding region. When used in a PCR with human genomic DNA a single DNA fragment of 330 bp was produced; this band was not present using rat, mouse, or hamster DNA as the template. A series of DNAs from human–rodent hybrid cells were tested in this PCR and the presence of this band appeared to correlate with the presence of human chromosome 2. A series of DNAs specifically designed to test the hypothesis that the COL4A3 gene is on chromosome 2 were then investigated. The results (Fig. 7 and Table I) localize the COL4A3 gene to chromosome 2.

At least 12 metaphases were examined after fluorescent in situ hybridization with the probe XCOL4A3-1. In all cells a signal was seen on at least one of the homologues of chromosome 2, in a terminal position 2q36-2q37 (Fig. 8). The small amount of background fluorescence observed was randomly distributed.

**Discussion**

Species other than man have frequently been used in studies of the Goodpasture antigen and other basement membrane components because of the difficulties in obtaining adequate quantities of well-preserved human material. We chose to work with sheep kidney because of the evidence that an autoimmune response to the Goodpasture antigen is mounted in Steblay nephritis (2, 26–29), which is induced in sheep by immunization with human GBM in Freund's adjuvant (30). In this study we show that the biochemical and chromatographic properties
of the antigenic component of GBM isolated from sheep kidney (Fig. 2) are very similar to those of antigen purified from bovine (6, 31) and human kidney (8, 32). In particular, it copurifies with larger quantities of similar, but nonantigenic components of collagenase digests, presumed to be NC1 domains of α1 and α2 chains. These are resolved by reverse-phase HPLC, in which the antigenic component elutes later than similar nonantigenic molecules (Fig. 2). The conservation of the antigenicity of this molecule between species suggests that it has a functional role in the basement membrane. The evidence from Alport's syndrome, mentioned below, and from the distribution of the antigen in specialized basement membranes, supports this view. However, it is not clear why this particular component should give rise to a severe autoimmune response in certain circumstances and in the right genetic setting (33).

Until recently the α3 and α4 chains of type IV collagen have only been known from limited amino acid sequence data from the triple helical/NC1 domain junction, almost all obtained from bovine material. In those studies, the component that reacted with Goodpasture autoantibodies was the 28-kD NC1 monomer 'M2' (6, 34), the equivalent of a more cationic protein, 'M28+++', in human GBM (7, 8). This component was labelled the α3 chain of type IV collagen on the basis of its NH2-terminal sequence. Recently Morrison et al. succeeded in using this sequence data to obtain a bovine cDNA clone for the NC1 domain of the α3 chain (11). Fig. 9 shows alignment of the bovine α3 sequence with the sequence obtained in our study, and with human α1, α5, and α2 chains. Our derived amino acid sequence agrees very closely with the NH2-terminal sequence of the bovine α3 NC1 domain apart from changes at positions 1–3, and is 90% homologous with the bovine sequence across the NC1 domain. In comparison it is 70% homologous with the human α1 chain and 57% with α2. We conclude that the cDNA sequence presented here encodes the COOH-terminal 340 amino acids of the human type IV collagen α3 chain.

The evidence that the α3 chain of type IV collagen carries the Goodpasture antigen is now compelling. Hudson and co-workers have published NH2-terminal sequence from purified components of bovine basement membranes that are recognized by autoantibodies in Goodpasture's disease (6, 7). These sequences gave the α3 chain its name, and were used by Morrison et al. to clone the bovine α3 cDNA (11). Antibodies (including monoclonals) raised to the bovine α3 chain replicate the binding pattern of human autoantibodies in histochemical studies and in Western blotting studies (5, 8). A monoclonal antibody raised to human GBM in our laboratory also replicates the binding of human autoantibodies to tissue sections and on Western blots (2, 3). On two-dimensional Western blotting, sera from patients and the monoclonals bind to the strongly cationic 28-kD molecule of human GBM that has been proposed to be the α3 chain (35, 36). The purified human antigen reacts with antibodies made to the bovine α3 chain, whereas other components of human GBM do not (37). The purified bovine molecule can substitute for whole collagenase-solubilized human GBM in a solid phase assay for anti-GBM antibodies in Goodpasture's disease (38); we have found that the purified sheep molecule is equally effective (unpublished observations). Segelmark et al. (39) have recently described studies of 37 positive sera in ELISAs using purified human α1, α2, α3, and α4 chains (the latter two characterized with antisera and monoclonal antibodies to the purified bovine molecules). 31 had antibody responses that were almost completely limited to the α3 chain; levels of antibodies in the other patients tended to be low and less specific for any one type IV collagen chain.

The findings in this paper are independent of these previous results, and they confirm that the Goodpasture antigen purified from GBM contains sequences from the α3 chain of type IV collagen. The sequences of the cyanogen bromide peptides, numbers 3 and 4 in Fig. 3, are present in the derived bovine α3 chain sequence. In particular, peptide 3 is identical to the bovine sequence at positions 61–70 in Fig. 9, and differs from all other known chains. Further support for the identity of the antigen as the NC1 domain of α3(IV) comes from observations on the derived amino acid sequence. From the usual site of
cleavage by collagenase (6, 8, 34), residue 97 in Fig. 5 and -15 in Fig. 9, to the COOH terminus, the molecule has 244 residues with a calculated mol wt of 26,849. It contains 14 arginine and 8 lysine residues, 9.0% of the total. Equivalent proportions for other chains are: α1, 7.4%; α5, 7.0%; and bovine α3 (known to be less cationic than its human equivalent [8, 34, 40]), 8.2%. In contrast the content of acidic residues is very similar in all known chains. This high content of basic amino acids accounts

![Alignment of derived amino acid sequences of NC1 domains of human type IV collagen α1, α5, and α2 (9, 10) and bovine α3 (11) chains with the human α3 sequence obtained in this study. Residues that are the same as those of α1 are shown as a dash (-), and cysteine residues in the NC1 domain are ringed. Numbering is from the first residue in the NC1 domain of the α1 chain, although the noncollagenous region of human and bovine α3 chains begins 3 residues earlier.](image)
for the strongly cationic nature of the human antigenic molecule on two-dimensional gel electrophoresis, and has been predicted from studies of the human protein (8).

Although the evidence identifying the α3(IV) NC1 domain as the Goodpasture antibody is very strong, it is not quite conclusive. Definitive proof requires expression of cDNAs encoding the NC1 domains of the various chains in an in vitro system, and the demonstration that only the NC1 domain from the α3 chain binds autoantibodies and inhibits their binding to human GBM. However, these experiments, while important, are unlikely to be straightforward. All 12 cysteine residues in the NC1 domain of the α1 chain are involved in intramolecular disulfide bridges (41). The conserved position of these cysteine residues in all type IV collagen NC1 domains makes it likely that this structure also applies to the other type IV collagen chains.

Goodpasture autoantibody binding is dependent on conformation and abolished by reduction of disulfide bonds. Re-creating the correct disulfide bonding and tertiary structure in bacterial expression systems is likely to be difficult. Recombinant eukaryotic proteins are often produced in Escherichia coli in an inactive and frequently insoluble form, and this seems to be particularly true where there are intrachain disulfide bonds (42). For this reason it will probably be necessary to use a eukaryotic expression system, where correct folding of the molecule is more likely to be achieved. Corresponding domains from the other α chains should be simultaneously expressed to establish the specificity of the results obtained.

As predicted from protein studies, levels of mRNA for α3(IV) are lower than those for α1 in Northern blots of fetal or adult renal cortex. The difference is perhaps less marked than would have been anticipated from protein studies. Most tissues and many cell lines produce some α1(IV) mRNA, but some cell lines (e.g., HT1080 cells) are known to produce large amounts of matrix components. HT1080 cells were used in the cloning of the human α1 chain (19). The Northern analysis shown in Fig. 6 indicates that α3 chain transcription is not correlated with α1 chain transcription. Of the tissues tested, only kidney and primary cultures of glomerular cells show an α3 signal.

It is known that antibodies recognizing the Goodpasture antigen do not bind to the GBM of most patients with Alport’s syndrome, a hereditary nephritis that is usually transmitted in an X-linked manner (43, 44). The existence of the α5 chain of type IV collagen had not been suspected biochemically or immunologically, but its gene is located in the appropriate region of the X chromosome (10), and abnormalities of the COL4A5 gene have since been identified in 3 of 18 Alport kindreds (45). The possibility that mutations in another type IV collagen chain mapping to the same locus could produce the disease is suggested by the extremely close, head-to-head arrangement of the α1 and α2 chain genes on chromosome 13 (46, 47). However, a direct role for the α3 chain in most cases of Alport’s syndrome is ruled out by our observation that its gene maps to chromosome 2. It is possible that the α3 and other novel type IV collagen chains, including α5, may form a second network (4, 48) in some specialized basement membranes, adjacent to or interwoven with that formed by α1- and α2-chains. Mutations of one of these chains could then affect the stability of the whole network and prevent the display of the antigen carried on the α3 chain. Knowledge of the primary sequence of these constituents of basement membranes should advance understanding of Alport’s syndrome and of the immunopathogenesis of Goodpasture’s disease.

Acknowledgments

We are grateful to Kate Perks and David Stevens for excellent technical assistance, and to Professor Lucio Luzzatto for help and advice.

N. Turner was a Medical Research Council Training Fellow and is now a National Kidney Research Fund Research Fellow. C. D. Pusey was a Wellcome Senior Research Fellow in Clinical Science. The work was supported by grants from the National Kidney Research Fund and the Wellcome Trust. The MRC 600 Confocal laser microscope was provided by the Medical Research Council Human Genome Project.

References