Characterization of Hydra Type IV Collagen

TYPE IV COLLAGEN IS ESSENTIAL FOR HEAD REGENERATION AND ITS EXPRESSION IS UP-REGULATED UPON EXPOSURE TO GLUCOSE*

Received for publication, July 5, 2000, and in revised form, August 14, 2000 Published, JBC Papers in Press, August 23, 2000, DOI 10.1074/jbc.M005871200

Susan J. Fowler[‡], Sheba Jose[‡], Xiaoming Zhang[§], Rainer Deutzmann^{||}, Michael P. Sarras, Jr.[§], and Raymond P. Boot-Handford^{‡**}

From the Wellcome Trust Centre for Cell-Matrix Research, School of Biological Sciences, University of Manchester, Manchester M13 9PT, United Kingdom, the §Department of Anatomy and Cell Biology, University of Kansas Medical Centre, Kansas City, Kansas 66160, the ||Institut fur Biochemie I, University of Regensburg, Regensburg, D-8400, Germany, and the ¶Department of Biomedical Sciences, Southwest Missouri State University, Springfield, Missouri 65804-0094

Hydra vulgaris mesoglea is a primitive basement membrane that also exhibits some features of an interstitial matrix. We have characterized cDNAs that encode the full-length hydra $\alpha 1(IV)$ chain. The 5169-base pair transcript encodes a protein of 1723 amino acids, including an interrupted 1455-residue collagenous domain and a 228-residue C-terminal noncollagenous domain. N-terminal sequence analyses of collagen IV peptides suggest the molecule is homotrimeric. Denatured hydra type IV collagen protein occurs as dimers and higher order aggregates held together by nonreducible cross-links. Hydra collagen IV exhibits no functional evidence for the presence of a 7 S domain. Type IV collagen is expressed by the ectoderm along the entire longitudinal axis of the animal but is most intense at the base of the tentacles at the site of battery cell transdifferentiation. Antisense studies show that inhibition of collagen IV translation causes a blockage in head regeneration, indicating its importance in normal hydra development. Exposure of adult hydra to 15 mm glucose resulted in up-regulation of type IV collagen mRNA levels within 48 h and significant thickening of the mesoglea within 14 days, suggesting that basement membrane thickening seen in diabetes may be, in evolutionary terms, an ancient glucose-mediated response.

The mesoglea is a $0.3-3.-\mu$ m-thick sheet of extracellular matrix that extends throughout the freshwater coelenterate *Hydra vulgaris*, with the exception of the mouth and the aboral pore (1). In 1961, Fawcett (2) first noticed the physical similarities between mesoglea and vertebrate basement membrane, the specialized sheets of ECM¹ composed mainly of type IV

The nucleotide sequence(s) reported in this paper has been submitted to the GenBankTM/EBI Data Bank with accession number(s) AF282902.

** To whom correspondence should be addressed: Wellcome Trust Centre for Cell-Matrix Research, 2.205 Stopford, School of Biological Sciences, University of Manchester, Manchester M13 9PT, UK. Fax: 161-275-5082; E-mail" Ray.Boot-Handford@man.ac.uk.

¹ The abbreviations used are: ECM, extracellular matrix; BM, basement membrane; HM, hydra medium; LEP, localized electroporation; UTR, untranslated region; bp, base pair(s); kb, kilobase(s); CHAPS, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonic acid; FITC, fluorescein isothio

collagen, laminin, nidogen/entactin, and heparan sulfate proteoglycan that underlie or surround cells. Although the mesoglea lies between two organized epithelial cell layers and has an obvious role in physical support of the organism, this ECM is also implicated in several cellular processes. In vitro studies showed that epithelial cells (3) and nematocysts (4) could attach to and migrate along isolated mesoglea or substrata coated with purified ECM components such as collagen IV and laminin. Zhang and Sarras (5) proposed that in addition to cell-cell interactions and chemotactic gradients, cell migration also depends on cell-matrix interactions. Using two modified grafting procedures, in which normal cell-matrix interactions were perturbed, they showed that interstitial cell migration is sensitive to alterations in collagen structure and that migration is inhibited by molecules that can compete with cell-fibronectin interactions.

The relation of mesoglea to higher metazoan connective tissue has been speculated on for at least the past century. By the early 1950s it was recognized that mesoglea was structurally and chemically similar to vertebrate connective tissues. Since then several biochemical studies have provided evidence of collagen-like molecules in isolated hydra mesoglea (6-8). More recent studies on mesoglea by Sarras et al. (9, 10) using indirect immunocytochemistry suggested the presence of molecules similar to type IV collagen, laminin, and heparan sulfate proteoglycan core protein (all characteristic of vertebrate BM) in mesoglea, as well as fibronectin, a component of interstitial ECM. In frozen sections, antibodies to ECM components appeared localized to the mesoglea layer between the epithelial bilayer. The presence of a hydra laminin in mesoglea was confirmed when cDNA clones encoding a β 1 chain of laminin were isolated (11). A fibrillar collagen has also recently been isolated from mesoglea and characterized (12), providing further evidence that mesoglea represents a primitive ECM exhibiting composite basement membrane and interstitial matrix properties.

So far six genetically different mammalian type IV α chains, $\alpha 1(IV)$ to $\alpha 6(IV)$, have been identified which interact to form various heterotrimeric triple-helical isoforms (13, 14). Type IV α chains have also been characterized in invertebrates including *Drosophila*, *Caenorhabditis elegans*, sea urchin, and sponge (15–19). The primary structure of all the $\alpha(IV)$ chains is similar, each chain having a short noncollagenous sequence at the N terminus, a central collagenous region consisting of Gly-X-Y repeats, and a highly conserved C-terminal noncollagenous

cyanate; contig, group of overlapping clones.

^{*} This work was funded by grants from the Dr. Hadwen Trust for Humane Research and the Royal Society (to R. B.-H.), National Institutes of Health Grant DK47840 (to M. P. S.), and Grant SFB 521/A4 from the Deutsche Forschungsgemeinschaft (to R. D.). 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.

(NC1) domain (20). Unlike fibrillar collagens, the long collagenous domain is interrupted at several sites by short, noncollagenous, sequences that are thought to impart flexibility on the molecule (21). Type IV collagen molecules interact to form a complex, irregular network that provides the BM with a high level of stability.

Zhang *et al.* (22) produced hydra cell aggregates and exposed the regenerating hydra to levels of glucose that mimic those in human diabetic patients (15 mM). Regeneration occurred as normal although within 72 h, glucose-treated reaggregating animals synthesized mesoglea twice the normal thickness, suggesting that the hydra may be a nonmammalian and temporally rapid model for the study of the much longer term glucoseinduced basement membrane thickening seen in diabetic microangiopathy.

In this study, using a hydra fibrillar collagen cDNA as a probe in a low stringency screen, overlapping clones were isolated that encode a complete hydra collagen IV α chain. In addition, type IV collagen protein was purified from hydra mesoglea and partially sequenced. All of the peptide sequences obtained are present within the primary sequence deduced from the cDNAs. The importance of collagen IV in hydra development is demonstrated using antisense thio-oligonucleotides, which block head regeneration. In addition, we show that the mesoglea of adult (rather than the previously demonstrated regenerating) hydra is rapidly thickened and collagen IV expression is up-regulated upon exposure to glucose.

EXPERIMENTAL PROCEDURES

Hydra Culture—H. vulgaris were maintained at 18 °C in hydra medium (HM) as described previously by Sarras *et al.* (9). Cultures were washed regularly and fed *Artemia salina* (brine shrimp) three times a week.

Isolation and Characterization of cDNA Clones-ECM-enriched segmented hydra were prepared (9) and allowed to regenerate for 24 h. Poly (A+) RNA was extracted and used as template for cDNA synthesis with either oligo(dT) or random primers. The randomly primed cDNA was cloned into Stratagenes \U03b7ZAPII vector (the HZAPII library), whereas the 3'-biased oligo(dT)-primed cDNA was directionally cloned into Stratagenes UniZAP vector (the HUZAP library). 3×10^5 recombinant clones of the HZAPII library were screened with a ³²P-labeled 2.3-kb hydra fibrillar collagen probe, HC1, which encodes the 5' half of hydra type I/II collagen (12) and washed first at low stringency (45 °C, $0.2 \times$ SSC, 0.1% SDS) then at high stringency (65 °C, 0.1 × SSC, 0.1% SDS). After each wash, filters were exposed to x-ray film for 16 h. Plaques that hybridized only at low stringency were chosen for secondary and tertiary screening. In vivo excised (ZAPPED) clones were cycle sequenced using an ABI PRISM XL 377 DNA sequencer in combination with a Big Dye terminator chemical kit (PerkinElmer Life Sciences). The DNA clones were sequenced in both directions using a primer walking strategy shown in Fig. 1. Comparisons of DNA sequences were conducted using the Blastn and Blastx programs and the GenBankTM data base (NCBI, National Institutes of Health). Protein alignments were obtained using the Blastp, GCG, and MacVector 5.0

Isolation and Analyses of Type IV Collagen from Hydra Mesoglea-For isolation of collagen, stocks of H. vulgaris were used that had been stored at -78 °C after shock freezing at -196 °C in 10 mM Tris buffer, pH 7.5, containing 1% Triton X-100 and 10 μ g/ml of the protease inhibitors phenylmethylsulfonyl fluoride and *p*-chloromercuribenzoate. For preparative extractions typically polyps from 20 plastic dishes $(20 \times 20 \text{ cm})$ grown to a density of about 6000 hydranths/tray were used. All preparations were done at 4 °C. After thawing, the hydras were washed with the same buffer as used for freezing until the supernatant was colorless. This procedure removed most of the cellular proteins. To remove cell surface proteins still bound to the mesoglea, the material was washed once more with 50 mM Tris buffer, pH 7.5, containing 1 M NaCl, 0.5% Triton X-100, and the same protease inhibitors as above. Subsequently, the mesoglea collagens were extracted with 10 ml of 50 mM Tris buffer containing 1 M NaCl, 10 mM EDTA, and 10 μ g/ml phenylmethylsulfonyl fluoride for 5 h at 4 °C. This fraction contained the majority of the type IV collagen and in addition a fibrillar collagen (12). To isolate denatured collagen chains for sequence analysis, the EDTA extract was made 1% in mercaptoethanol and separated

by reversed phase chromatography on a Vydac C-18 column (Separations Group, 250×2.2 mm) at a flow rate of 0.2 ml/min using a gradient of 5-70% acetonitril in 50 min. Two major peaks at 21 and 35% of the organic modifier were obtained, the latter of which contained the type IV collagen. After lyophilization, the isolated chains were dissolved in 70% formic acid or 0.2 M ammonium bicarbonate for digestion with cyanogen bromide and trypsin, respectively. The peptides were separated by reversed phase chromatography and sequenced on a Procise 492A sequencer (PerkinElmer Life Sciences) with on-line detection of the PTH amino acids according to the manufacturer's instructions. To isolate the type IV collagen under nondenaturing conditions, the EDTA extract, supplemented with mercaptoethanol, was centrifuged at 45000 rpm in a Beckman 50 TI rotor. The supernatant was dialyzed against 20 mM Tris buffer, pH 7.4, and separated on a HiTrap Q column (1 ml) using a gradient of 0-1 M NaCl in 40 min. Hydra collagen IV eluted as a broad peak in the range of 0.2-0.4 M NaCl. Samples were analyzed on 4-10% SDS-polyacrylamide gels. For electron microscopy, samples (about 50 μ g/ml) were mixed with an equal volume of glycerol and, after spraying onto mica discs, rotary shadowed with carbon/platinum (23). The replicas were viewed with a Philips CM 12 electron microscope.

In Situ Hybridization-Whole mount in situ localization of mRNA was performed using a digoxygenin-labeled RNA probe generated from clone 194.1 encoding the most 3' 2505 bp presented in Fig. 2. The probe is made by EcoRV linearization and T3 promoter transcription of the entire insert including 2078 bp of the open reading frame and 427 bp of 3'-untranslated sequence. Fixation, processing, hybridization, and visualization of the riboprobe in whole mount preparation was performed as described previously by Grens et al. (24, 25). Briefly, hydra were fixed with 4% paraformaldehyde after relaxation of the polyps with 2% urethane. Specimens were subsequently treated with ethanol and proteinase K to facilitate diffusion of the probes into the epithelial bilayer. To stabilize digested tissues, specimens were refixed with 4% paraformaldehyde and then prehybridized in hybridization solution (50% formamide, $5 \times$ SSC, $1 \times$ Denhardt's, 200 mg/ml tRNA, 0.1% Tween 20, 0.1% CHAPS, 100 mg/ml heparin) to block nonspecific hybridization sites. This was followed by a 48 h hybridization with the digoxygenin-labeled RNA probe and a subsequent wash in hybridization solution and $1 \times$ SSC. Specimens were next washed in MAB (100 mM maleic acid, 150 mM NaCl, pH 7.5) and preblocked (2-6 h) in MAB with 20% sheep serum and 1% bovine serum albumin. This was followed by the 16-h incubation at 4 °C in the same solution with anti-digoxygenin antibody (1:2000). Animals were washed eight times with MAB and then briefly in alkaline phosphatase buffer (100 mM Tris-HCl, pH 9.5, 50 mM MgCl₂, 100 mM NaCl, 0.1% Tween 20). Specimens were then stained with BM purple alkaline phosphatase substrate (Roche Molecular Biochemicals), dehydrated with ethanol, and mounted in euparal (Asco Laboratories).

Functional Analysis of Hydra Collagen IV Using Localized Electroporation and Thio-oligonucleotide Antisense Constructs-Because transfection approaches have not been successfully applied to Cnidarians, we developed a procedure to specifically test the effect of antisense constructs on head or foot regeneration in hydra. This approach utilizes a localized electroporation technique (LEP) to introduce antisense oligonucleotides into the apical or basal pole of hydra. This procedure has been applied to the functional analysis of a number of hydra genes (26-28). Applying the LEP procedure, we tested the hypothesis that de novo biosynthesis of hydra collagen IV is required for normal head morphogenesis following decapitation. Based on the work of Dr. Richard W. Wagner (29-31), a series of 20-mer oligonucleotides with phosphorothioate linkages were designed. Six oligonucleotides were synthesized to include four antisense sequences to portions of the 5'-UTR, initiation site, coding sequence, and 3'-UTR; a sense strand of the 5'-UTR; and a mismatch construct (the sequence of these molecules is given in the legend of Fig. 9). The oligonucleotides were introduced into hydra cells using electroporation (Bio-Rad Gene Pulser) with a micropipette drawn on a pipette puller. The ends of the micropipettes brought in contact with the hydra were polished with a micro-forge. Because hydra collagen IV is expressed in the ectoderm layer of cells, LEP was performed on the surface of the organism at the apical pole. Based on preliminary experiments, we found that decapitation had to be performed 2-4 h after LEP. This was necessary because if decapitation was performed prior to electroporation, an extensive loss of electroporated cells would occur at the cut edge of the apical pole. To maximally retain the oligonucleotides within the area where the electroporation was performed and also to reduce movement of the animal during LEP, hydra were chilled in a 10% heptanol solution (heptanol in hydra medium) for a maximum of 1 h. A 100 µM stock solution of 20-mer thiolated oligonucleotide (Genset Corp.) was mixed with FITC-dextran with a M_r 10,000 (Molecular Probes, Eugene, OR) in a ratio of 3:1



FIG. 1. **Illustration of overlapping cDNA clones and scheme of the primary structure of the** α (**IV**) **collagen chain.** Top panel, white boxes indicate the five overlapping cDNA clones obtained from regenerating hydra cDNA libraries, HZAPII and HUZAP. Clone G4.1.2 was obtained following screening with a hydra fibrillar collagen probe, HC1, at low stringency. A small 3' restriction fragment from each purified clone was used as a probe in subsequent rounds of library screening. *Eco*RI restriction sites are shown by *E*. All clones were sequenced using the primers indicated. *Middle panel*, the complete mRNA contig contains a 159-bp 5'-UTR, a 5196-bp open reading frame, and a 521-bp 3'-UTR. *Bottom panel*, the primary structure of the entire α (IV) chain, deduced from the cDNA sequence. The 5' noncollagenous domain, including the signal peptide, is indicated by the *first shaded box*, the large collagenous domain is indicated by the *long box* interrupted by 25 short noncollagenous sequences, and the 3' NC1 domain is indicated by the *final shaded box*.

(typically 6 μ l of DNA + 2 μ l of FITC-dextran). After fitting the micropipette over the microelectrode (World Precision Instruments, Inc.), the oligonucleotide/FITC-dextran mixture was loaded into the micropipette. The Bio-Rad gene pulser was set at 100 ohms, 25 µFD, and 50 V, and the average time for the pulse was 3.6 ms. Hydra were placed on a Nytex net attached to a plastic Petri dish with soft wax. The dish was placed on an incline to facilitate positioning of the micropipette. The micropipette was placed in contact with the ectoderm at the apical pole. After charging the gene pulser to 50 V, the pulse was initiated for a 3.6-ms period. An additional pulse could be applied if the width of the ectoderm area at the pole was greater than the micropipette diameter. To ensure that the DNA/FITC mixture was retained in cells at the apical pole after decapitation, electroporation was performed with two pulses on the ectoderm just inferior to the mouth region to make sure that a more extensive area of the apical pole received the construct. In this way, cells with the DNA/FITC mixture composed the apical pole following decapitation. The hydra were then placed in hydra medium, and at 24 h all animals were screened on a Leitz fluorescent microscope to ensure that each specimen had retained the DNA/FITC-dextran mixture at its regenerating apical pole. Electroporated hydra were observed every 24 h, and the degree of regeneration was compared with mock electroporated controls. In the current study animals were cut in the neck region just inferior to the mouth and tentacle ring. The degree of head regeneration was monitored by (i) observing the morphology of the head process under a dissection microscope and determining the degree of tentacle eruption and hypostome formation and (ii) analyzing the cellular morphology of cells of the hypostome and tentacles using Normarski optics. In controls, head regeneration is normally completed within 72 h, and, therefore, experimental groups in which blockage was observed were monitored for an additional 5 days to determine whether recovery from blockage occurred. Control animals were treated with mismatched thio-oligonucleotide 20-mer constructs (randomized sequence) or sense constructs if a particular antisense construct was found to block morphogenesis. Control and experimental groups were statistically compared using a Chi squared test and an analysis of variance test.

Effects of Glucose on Mesoglea Thickness—Hydra were cultured in normal HM or in HM supplemented with either 15 mM glucose or 15 mM sorbitol for 14 days prior to being anesthetized with 10 mg/ml sodium pentobarbital to prevent body column contraction. Anesthetized organisms were fixed in 2.5% gluteraldehyde, post-fixed in 2% osmium tetroxide, stained with 2% uranyl acetate and 0.3% osmium tetroxide, and embedded in Spurr's resin. Animals were sectioned exactly perpendicular to the axis of the body column to produce circular rather than oval cross-sections. The body columns were sectioned near the head, at the middle, and near the base. Eight randomly selected regions from the circumference of each section were photographed, and the data from these micrographs were used to calculate the mean mesoglea thickness at each of the three locations in the body column.

RNA Extraction—RNA was isolated from adult hydra that had been incubated for 48 h in HM or HM supplemented with 15 mM glucose. 2–10 hydra were added to an Eppendorf tube and 400 μ l solution D (4 M guanidine thiocyanate; 0.5% sarkosyl; 25 mM sodium citrate, pH 7.0, and 0.1 M β -mercaptoethanol added). The tube was vortexed for 30 s. 40 μ l of 2 M sodium acetate, pH 4.0, 400 μ l of phenol, and 40 μ l of chloroform were added, and the tube was briefly vortexed before incubation on ice for 15 min and centrifugation at 6,000 \times g at 4 °C for 30 min. The upper layer was transferred to a fresh tube and an equal

volume of ice-cold isopropanol (propan-2-ol) was added. The tube was vortexed, incubated at -70 °C for 1 h, and spun for 30 min at 4 °C at 10,000 \times g. The pellet was resuspended in 100 μ l of solution D, and 100 μ l ice-cold isopropanol was added. The tube was incubated for a further hour at -70 °C and centrifuged (10,000 - g, 5 min), and the pellet was washed with 70% ethanol to remove any residual salt. After air drying, the pellet was resuspended in 10–20 μ l of diethylpyrocarbonate-treated water. The concentration and purity of the RNA was determined by measuring $A_{260 \text{ nm}}$ and $A_{280 \text{ nm}}$. Northern blotting and probing were performed as described previously (32).

RESULTS

Characterization of Type IV Collagen-The previously described hydra fibrillar type I/II collagen cDNA clone HC1 (12) was used to screen a regenerating hydra cDNA library at low and high stringency. A clone, G4.1.2, was identified that hybridized to the fibrillar collagen probe only under conditions of low stringency. The clone was purified and found to contain 1.7 kb of interrupted collagenous sequence. Using an EcoRI restriction fragment of G4.1.2 to rescreen the library, several overlapping clones were isolated and purified that together cover the entire coding sequence of a H. vulgaris type IV collagen chain. The relationship of these clones to each other is shown in Fig. 1. The complete nucleotide sequence of hydra type IV collagen chain together with its deduced amino acid sequence is shown in Fig. 2. There is a 5'-untranslated region of 159 bp, an open reading frame of 5,169 bp, and a 504 bp 3'-untranslated region. The deduced translation product of 1723 residues with a predicted molecular weight of approximately 160,000 comprises a 24-residue putative signal peptide, a short 5' noncollagenous domain of 16 residues, a 1455-residue collagenous domain, and a 228-residue 3' NC1 domain. The triple helical domain (residues 41-1495) contains 450 Gly-X-Y repeats interspersed with 24 noncollagenous interruptions of 1-15 residues. The primary sequence and domain sizes (Table I) are similar to other reported type IV collagen chains. The locations of threequarters of the interruptions in the hydra collagenous domain are conserved across species (data not shown). As in other species, twice as many long imperfections (5 or more residues) occur in the N-terminal half of the hydra type IV collagenous domain. The hydra type IV collagen chain also has the single N-linked glycosylation site (Fig. 2, amino acids 122–124) conserved in other species and type IV chain types. Cysteine residues are particularly highly conserved. The 4 cysteines in the 16-residue 5' noncollagenous domain (Fig. 2, amino acids 25-40) are also found in all other type IV chains and are required for N-terminal association of collagen IV molecules into tetramers to form the '7S domain'. There are five cysteine residues in the triple helical domain, mostly located in imperfections, with four being in the N-terminal half of the collagenous domain. This is comparable with other vertebrate and invertebrate

Hydra Type IV Collagen

	gaactcactattattettetaacaatagogt taatagacagtgcacacettgatattetaacaaacgtagtttttttecccaattttgttagttgetttegtatatggeegtttaategaggataaaaggagaaaatategeaatataaaa	a
1	atgaatcacacgaccaaatggactatttggattgtgttgattttcacgttacattcgtgtttagttttactcagttgtgtggtggtggtggtggtgtaatcaatgcagtaatgtttgtgt MNHTTKWTIWIVLIFTLHSCLVFTQLCGCGGGGGONQCSNVCV	t 120
41	ggtcaaaaaggagacaggggttctgttggtctgcctggatttaaaggacctattggtgagcctggatttcctggaggagaaggtcctgctggaaggcctggagagaaggagatggtgg . G Q K G D R G S V G L P G F K G P I G E P G F P G G E G P A G R P G E K G D G G	t 240
81	gcccctggtttacaaggagaagttggtgctagaggtaaagtaggaccacctggaacacctggaatcccaggttctcatggtcgacctggagatgaaggaag	a 360
121	ccagggtgcaacggaaccaagggagatattggtcctcctggtcctcgtggtaaagatggagttcaaggaccatctggattacccggaccaattggtcctccaggagaaccaggcgattc . P G 8 N G 7 K G D I G P P G P R G K D G V Q G P S G L P G P I G P P G E P G D S	£ 480
161	actgcaacaaagttaaaaggtcaaaaaggagaaccaggaccaaaagtcaagacggaccagctggtgacccaggaacaaaaggcgataaaggagaatctggaaaagcaggaagaatggg . <mark>T A T K L K</mark> G Q K G E P G P K <mark>S Q D</mark> G P A G D P G T K G D K G E S G K A G R M G	t 600
201	ccacaaggtttgcgtggcgagaaaggggctagaggtgattctaatattaccatatttggagaacgtggtgataaagggggacattggcttaccaggaccaccaggcagagactgtaatgg . P Q G L R G E K G A <u>R G D</u> S <mark>N I T I P</mark> G E <u>R G D</u> K G D I G L P G P P G R D <mark>C N G</mark>	a 720
241	agttcagttactggagaacttataacaaatattcaaggcccaaaaggtgagcaaggtaaaaaaggagatcaaggacaaaaaggagaaccaggacaacctggtcaagctggagaagcggg 8 S V T G E L I T N I Q G P K G E Q G K K G D Q G Q K G E P G Q P G Q A G E A G	t 840
281	caggatggtcaaaaaggagaaaaaggtgataaaggagagattggtagtggtg	a 960
321	ccagggtctgatggctctctaggttcaccaggccagaaaggtgatcatggaaccaaaggttctgaaggtcctgttggtgatcgaggtaaacaaggagaaagcataactggaccacctgg . P G S D G S L G S P G Q K G D H G T K G S E G P V G D R G K Q G E S L G G P P G	g 1080
361	caaataggtgaaaagggacagcgtggaccagaaggaaaaaaggagaacctggtccagaagggcetectggacctaatggtgaagtaggagatactggacetecaggtettaaaggtt Q_I_G_E_K_G_Q_R_G_P_E_G_K_K_G_E_P_G_P_E_G_P_P_G_P_N_G_E_V_G_D_T_G_P_P_G_L_K_G_L	a 1200
401	aaaggtgagataggtatgactggtccaagtggtgaacctggaaaaactggagcagaaggaatgaaaggaccaattggacctgctggtcagcgaggagaaacaggatcaaaaggagagtc K G E I G M T G P S G E P G K T G A E G M <u>K G P I G P A G Q R G E T G S K G E S</u>	t 1320
441	ggaagaccaggtcaaagtgtacaaggtagtcctggtatggatgg	a 1440
481	gtigtaaatgaictaggtgaaacagtattaccaggccctccaggtgatcaaggtcctcaaggtattcaaggtgttgcaggaccacaaggaagtccaggtattagaggagataaaggaga V <mark>V N D L G E T V L P</mark> G P P G D Q G P Q G I Q G V A G P Q G S P G I <u>R G D</u> K G D	t 1560
521	gcttgcaaatcatgcccaagtggcccgcaaggagaaaagggtgtagcaggacaagatggtcttccgggtgctaatggagaaaaaggtgataaaggagatgaaggtaaacctggtcttga L A <mark>C K S C P S</mark> G P Q G E K G V A G Q D G L P G A N G E K G D K G D E G K P G L E	a 1680
561	ggaagtgttggtgatccaggtgaaagtggactaaccggaccacaaggacctattggtgaaaaaggtgacacaggtgctaaaggaaataaaggggagcgtggatctgataggattgttca L G S V G D P G E S G L T G P Q G P I G E K G D T G A K G N K G E R G S D <mark>R I V Q</mark>	a 1800
601	ggagaacgtggtgaaaaaggtactgatggtatgccaggatctcctggttccataggatttaaaggtgaacagggggttcctggtgatccaggtactgatggaccacaaggacctcaagg G E R G E K G T D G <u>M P G S P G S I G F K G E Q G V P G D P G T D G P Q G P Q G</u>	a 1920
641	cttcaaggtcctcaaggtgatcgtggtgctcctggtaaatcaggtataccaggcaataacggagagaaaggtgaaaagggaaataaaggtgatggcttacgtggatttaaaggtgagg L L Q G P Q G D R G A P G K S G I P G N N G E K G E K G N K G D G L R G F K G E R	t 2040
681	ggtcgtgatggagaaccaggaccaaaaggatctcaaggaccaatgggtaaaccaggacctaaaggagaaccattttctacagctttaaaaggacaacaaggtgacaaaggagatgttgg L G R D G E P G P K G S Q G P M G K P G P K G E P <mark>F S T A L K</mark> G Q Q G D K G D V G	t 2160
721	tcagaaggtccacctggacctcaaggagctgaagggcaagtaggaaaccagggatttaaaggtgagccaggactcagaggtgcaccaggaagtggagaaaaaggagagaagggtgatgc L S E G P P G P Q G A E G Q V G N Q G F K G E P G L R G A P C S G E K G E K G D A	t 2280
761	ggcacaccaggaatagcaggqaaaaatggagaaaatggtaaggatggtcaagatggaccgaaaggtgacaaaggtgaaaaaggcctgcaaggaccaattggtatttctggtccaattaa L G T P G I A G K N G E N G K D G Q D G P K G D K G E K G L Q G P I G I S G P I K	a 2400
801	ggtgaacetggtttaaaaggtgatecaggtaaacetgggettgatggegeteetggaaaatetgtaceaaaaggtgaacetggtagagatggaaaagatggagateeaggagetaaagg L G E P G L K G D P G K P G L D G À P G K S <mark>V P K</mark> G E P G R D G K D G D P G A K G	t 2520
841	gaacoaggcactcctggtgaaagcggtcgtgatggtggaaaaaggtgaaattggacctcaaggccctcagggtgataaaggagatcaagggcaatcagtgataggtccaaagggtgaaac	t 2640

FIG. 2. Nucleotide sequence of full-length cDNA, as shown in Fig. 1, and deduced amino acid sequence of *H. vulgaris* α 1(IV) polypeptide.² The overlapping cDNAs correspond to a 5855-nucleotide mRNA, including 159-nucleotide 5'.UTR and 521-nucleotide 3'.UTR. Numbering begins at the start codon. The signal peptide is *yellow*, and the 5' noncollagenous domain is *underlined*. Cysteine residues are *red*, RGD sequences are *double underlined*, and the potential *N*-linked glycosylation site is in *light blue*. Imperfections in the triple helical domain are in green, and the 3' NC1 domain is *blue*. The termination codon (taa) is in *bold type*. Peptide sequences obtained from purified type IV collagen are *underlined* with a *dotted line*.

chains (Table I). The C-terminal NC1 domain of hydra is the most conserved region of the molecule (Table I). The number and relative position of the cysteine residues in the NC1 domain is identical to all other type IV chains (Fig. 3) regardless of species.

The hydra sequence is slightly more closely related to the $\alpha 1$ family ($\alpha 1$, $\alpha 3$, and $\alpha 5$) than to the $\alpha 2$ ($\alpha 2$, $\alpha 4$, and $\alpha 6$) family of collagen IV chains based on: (i) the levels of sequence identity across species in the NC1 domain (Table I and Fig. 3); and (ii) the presence of the highly conserved $\alpha 1$ -like sequence GC-

NGTK (Fig. 2, residues 122–127) in the 7 S region, whereas in α 2-like chains, the Lys is replaced by Arg. Accordingly, we designate the reported sequence hydra α 1(IV) collagen.

Northern analyses revealed that hydra RNA contains two $\alpha 1(IV)$ mRNA species, the more abundant mRNA being approximately 7 kb and the lesser, 6 kb (see Fig. 11). The most 3' clone we sequenced contained a poly(A⁺) addition signal approximately 10 bases upstream of a poly(A⁺) tail and completes a 5.9-kb contig. It is probable that a second poly(A⁺) addition signal lies 3' and accounts for the longer mRNA transcript as is

881	ggtgc G A	tat	agg G	K	aaaa K	agga G	E	T	ggt G	Q	D	ggc G	L	aaa K	ggc G	aga R	gat D	gga G	I	G	aaa K	Q.	gga G	A	aac N	ggt G	gaa E	aag K	ggt G	gaa E	aaa K	gga G	N	stca	gta V	v.	ggt	ttgc L	ca P	2760
921	ggtga <u>G</u> D	tco	tgg	aga E	aaaa K	aggt G	tca	P	igga G	gaa E	iaaa K	ggc G	aaa K	gat D	gga G	att I	tct S	gga G	aaa K	Q	igga G	aaa K	P	gga G	cca P	cct P	gga G	aaa K	gat D	ggt G	gtt V	gat D	gga G	ract T	cct P	ggt G	aaa K	igatg D	gt G	2880
961	ttgcc L P	tgg	aaa N	P	aggo G	P	E	G	P	aaa K	G	egaa E	aaa K	gga G	gtt V	act T	gga G	cca P	act T	ggt G	tta L	gat D	ggg G	igaa E	aaa K	gga G	gag E	cct P	gga G	tac Y	agt S	gga G	Q	iaaa K	iggt G	gcc A	aag K	iggag G	aa E	3000
1001	actgg T G	tgt V	tcg R	G	aaaa K	P	gga G	igco A	aca T	s	aat N	gta V	M	gga G	cca P	aaa K	gga G	aat N	cgt R	ggg G	igag E	s	ggt G	gct	aaa K	ggc G	gaa E	cca P	gga G	aaa K	caa Q	gga G	P	P	ggq G	igct A	gag E	iggac G	ect P	3120
1041	caagg Q G	tga E	aaa K	agg G	agaa E	aga R	ggt G	att I	ggc G	aca	gat D	aaa K	ggt G	act T	aaa K	ggg G	gat D	aca T	ggt G	gat D	ata I	gga G	L	Q	gga G	P	aaa K	gga G	gaa E	aga R	gga G	gaa E	R R	ggt G	gaa E	icaa Q	ggt G	cgaa R	aa K	3240
1081	ggaga G D	tca	agg G	P	tata I	gga G	P	iaaa K	igga G	D D	gtt V	gga G	gaa E	cct P	ggt G	cca P	gct A	ggg G	tca S	ato I	gga G	aac N	aca T	ggt G	att I	tct S	ggc G	aat N	gat D	gga G	cct P	aaa K	iggt G	gaa E	S	igga G	P	ccag P	gt G	3360
1121	attaa I K	agg	tga	P	tggt G	stca S	V	gga G	agt S	I	gga G	P	aaa K	gga G	agc S	aaa K	gga G	gag E	cag Q	ggt G	gaa E	aaa K	gga G	gag E	aga R	gga G	agc S	aca T	ctg L	aaa K	gga G	aat N	cct P	ggt G	P	igaa E	gga G	P	gc R	3480
1161	ggccc G P	cac	agg G	P	P	cgga G	P	igct A	gga G	iaaa K	D.	tat X	ata I	ctt L	gat D	ctt L	gaa E	gtc V	ggt G	cct	ttg	ggg G	gaa E	aaa K	cca P	acc T	aaa K	ggt G	gag E	cca P	gga G	gaa E	gat D	gga G	P	iaag K	gga G	igata D	aa K	3600
1201	ggtga G D	tat	tgg G	iga E	aact T	ggo G	stca	Q	igga G	act T	aaa K	iggt G	gaa E	aaa K	ggt G	gtt V	ccg P	ggt G	tct	aaa K	igga G	gaa E	aaa K	gga G	gag E	caa Q	ggt G	tta L	gaa E	gga G	cca P	aaa K	ggc G	ttg L	igaa E	igga G	aga R	gttg V	gt G	3720
1241	ttagc L A	tgg G	atca S	aaa K	aggo G	ggac D	caaa K	ggt G	ata I	iaaa K	gga G	gtt V	att I	gga G	aag K	cct P	ggt G	cca P	act T	gga G	ngat D	aaa K	iggt G	gaa E	aaa K	ggt G	tta L	caa Q	ggt G	ttt F	aaa K	gga G	gat D	aca T	igga G	P	caa Q	ggaa G	tc I	3840
1281	caagg Q G	gaa K	aga E	agg G	agaq E	gcca P	igga G	aga R	gaa E	igga G	gca A	iaaa K	gga G	ctt L	aaa K	gga G	gaa E	gta V	gga G	gct A	gaa E	gga G	aaa K	atg M	gga G	cca P	cct P	tgt	gat D	gcc A	gga G	gct A	aca T	I	att	.ggt G	gaa E	aaag K	gt G	3960
1321	gatag D R	agg G	aga E	aaa K	aggt G	igaa E	A	gga G	Q	R	ggt G	tac Y	aaa K	gga G	gaa E	cca P	ggt G	cca P	tct S	agt S	caa Q	cta L	T	att	aat N	cca P	ggc G	gct A	aaa K	ggt G	gaa E	caa Q	gga G	gat D	aga R	igga G	att I	.aaag K	ga G	4080
1361	gagaa E K	ggg G	gaat N	at	tgga G	P	iaaa K	iggt G	ata I	aat N	ggt G	cca P	ttt F	ggt G	aaa K	gat D	ggt G	tca S	att I	gga G	P	gag E	igga G	cca P	aaa K	gga G	gat D	tct S	ggt G	caa Q	aaa K	gga G	gaa E	act T	ggt G	.tta L	gat D	.ggtc G	ca P	4200
1401	aaagg K G	aga E	aaa K	agg G	gcaç Q	jcca P	igca A	ata	igga G	aca	ggt G	eca P	aaa K	ggt G	gaa E	aaa K	ggt G	gat D	caa Q	ggg G	P	cct P	ggt G	cca P	caa Q	ggc G	ccc P	cag Q	ggt G	gat D	att I	gga G	gaa E	iaag K	iğgt G	.gat D	att I	ggac G	cg P	4320
1441	ccagg P G	aga E	aact T	gg G	cgat D	att	gga G	P	Q	igga G	ata I	iaaa K	gga G	gaa E	ttc F	ggt G	ata I	gct A	ggt G	att	aaa K	ggt G	gaa E	act T	ggt G	cct P	cct P	ggt G	cta L	caa Q	ggt G	cct P	att	ggt G	ttg L	aag K	gga G	gacc D	aa Q	4440
1481	ggacc G P	gcc F	tgg G	ga D	tgca A	iggt G	aaa K	gta V	ggt G	cca P	cct P	ggt G	cgt R	agt S	ggc G	s	ggt G	gat D	ttt F	ggg G	F	tat ¥	ctt L	gtc V	aag K	cac H	agt S	caa Q	tct S	atc I	aag K	gtt V	cca P	s	tgt	.ccg P	gca A	ggaa G	tg M	4560
1521	caaac Q T	tat	gtge	gga E	aggt G	tat Y	agt S	F	tta L	tat Y	gca A	Q	ggt G	aat N	gaa E	cgt R	gcg A	ttt F	ggt G	caa Q	ngat D	tta L	ggt G	caa Q	cct P	ggc G	tct S	tgt	ctt L	aaa K	cga R	ttt F	ago S	aca T	atg M	P	ttc F	ttat L	tt F	4680
1561	tgcga	tat	tca:	iaa N	taaq K	gtgc	gtt V	gtt V	gca A	s	aga R	iaac N	gac D	tat Y	tct S	ttc F	tgg W	ctt L	tct S	act T	gca A	gaa E	aaa K	cca P	aag K	gaa E	gca A	cca P	tcg S	agt S	gga G	gct A	gat D	ctt	gag E	aac N	tac Y	attt	ca S	4800
1601	cgttg R	tat	tgta V	ntg	cgao E	igca A	P	s	cac	gtt V	cta	igct	gtt V	cac H	agt S	caa Q	agt S	gaa E	cta L	gat D	cca P	aaa K	tgt	cca P	gat D	gga G	tgg W	gaa E	aat N	cta L	tgg W	act T	gga G	F	agt S	ttc F	cta L	atgt M	ac Y	4920
1641	aacag N S	tgc A	cggt G	gc A	caa Q	igga G	s	ggt G	caa Q	ctg	tta L	tct S	tca S	tct S	ggc G	tct S	tgt	ttg L	gaa E	gat D	F	cgt R	gtt V	aat N	cca P	tat ¥	att I	gaa E	tgc	cac H	ggt G	cgg R	gga G	aca T	tgt	tgg W	tac Y	tacg Y	gc G	5040
1681	cctac P T	tct	tagi S	F	ttgg	tta L	s	act	att	ggc G	gaa E	agt S	aac N	atg M	F	caa Q	gtt V	cca P	aaa K	ttt F	gaa E	att	ctt L	gaa E	aga R	aat N	ctt L	aaa K	gca A	cga R	gtc V	agt S	cgc R	tgt	gct A	gtt V	tgc	atga M	aa K	5160
1721	tetgt s v	tcc	tta	aga	eteç	iddo	gaa	cta	caa	caa	aga	acc	tta	gtg	ttg	caa	aag	agc	ctt	aat	att	ata	caa	gga	tca	ctg	tgt	gat	ccg	tat	aca	att	tat	ato	aac	tga	cat	tttt	t 5	280
aaaa acag ttgg	tcaaa ttttt ttatt	att ggt ttg	cati taa tgto	cat cat	ttt	aaa cat ttc	caci	tat att act	tta tag ttt	tat ctg taa taa	tta tcc tta	igat att icat	atc att gtt	gca tta aac	tga aaa	tct ttt	aal	ttt. tac	aga cta att	atg gaa att 696	icaa iaac taa	ttt cta gtt	tta aat gtt	aaq caa ctt	ggga igta :ttt	att ata aaa	taca atga agai	atg aac ttt	att caa ata	tga tca gga	tta aac aac	cgt caa ata	aaa atgi atti	aaaa ttco ttta	ata jaa act	atti ttt: ttt:	tgt ccc tta	tacg: acgt tata	aa tt aa	5400 5520 5640

FIG. 2—continued

the case in several other collagen IV genes (20, 33).

Molecular and Supramolecular Structure—Extraction of hydra mesoglea with EDTA, 1 M NaCl solubilized two major collagens with molecular masses of 155 and 290 kDa (Fig. 4). The smaller component has recently been shown to represent a fibrillar collagen, hydra collagen I/II, that forms a large network of thin fibrils (12). Here we show that the larger component is type IV collagen, because all the peptide sequences obtained from the 290-kDa chain (Fig. 2) are found within the hydra $\alpha 1$ (IV) collagen-deduced sequence. The peptide data demonstrate that the polypeptide chain is not processed by removal of the NC1 domain, as expected, and that the hydra collagen IV frequently contains hydroxyproline in the Y position of the triplet (data not shown). In many positions where

lysine should be found according to the cDNA sequence, Edman degradation failed to detect any amino acid residue. This is most probably because of the presence of glycosylated lysine (or hydroxylysine) residues that are not seen in normal protein sequencing protocols.

As shown in Fig. 4, the hydra type IV collagen was not able to enter an SDS-polyacrylamide gel in the unreduced state, indicating extensive polymerization via disulfide bonds. However, even after reduction, only bands down to the size of a dimer (290 kDa) were found. Monomers were never detected, indicating that two polypeptide chains are held together via nonreducible cross-links. Most probably intermolecular bonds between two adjacent molecules are formed rather than intramolecular bonds within the same triple helical molecule

Hydra Type IV Collagen

TABLE I

Comparison of the hydra with other vertebrate and invertebrate type IV collagen α chains

Domain sizes are given as numbers of amino acid residues.

	Hydra $\alpha 1$	$\operatorname{Human}_{\alpha 1^a}$	$\operatorname{Human}_{\alpha 2^b}$	$\frac{\text{Drosophila}}{\alpha 1^c}$	$ \underset{\alpha 2^d}{\operatorname{Drosophila}} $	C. elegans $\alpha 1^e$	C. elegans α2 ^f	$\operatorname{Urchin}_{\alpha 1^g}$	$\operatorname{Urchin}_{\alpha 2^h}$	A. suum $\alpha 2^i$
Translation product	1723	1669	1712	1775	1761	1758	1758	1752	1752	1763
Signal peptide	24	27	36	23	24	28	26	28	27	26
5' noncollagenous	16	15	32	70	16	14	15	14	15	16
Collagenous domain	1455	1413	1449	1456	1461	1487	1486	1484	1483	1487
NCl domain	228	229	227	231	259	229	231	226	227	234
Interuptions in collagenous domain	24	21	23	21	24	17	15	23	19	17
Cys in 5' noncollagenous domain	4	4	4	3	3	4	4	4	4	4
Cys in collagenous domain	5	4	5	6	3	9	3	4	3	3
Cys in NCl domain	12	12	12	12	13	12	12	12	12	12
Percentage of identity in NCl domain compared to hydra α chain	100	58.1	50.4	45.1	44.6	55.3	53.5	58.4	50.5	54.1

1.0

Т

20

. .

1 00

^a Ref. 38.

^b Ref. 39.

^c Ref. 15.

^d Accession number AAB6408.

^e Ref. 16.

^f Ref. 40.

^g Ref. 17.

^h Ref. 18.

^{*i*} Ref. 41.

1101. 41.

FIG. 3. Amino acid sequence comparison of the NC1 domain of hydra type IV collagen (Hy) with all six human type IV collagen chains (A1-A6). Dashes represent identities with the amino acid sequence of the hydra NC1 domain. Gaps (\cdot) are introduced to achieve best alignment. Conserved cysteine residues are indicated by arrows.

			1 15	↓ J0	40	v 60
1	Hydra		.GSGDFGFYLVKHSQ	SIKVPSCPAGMQTMW	EGYSFLYAQGNERAF	GQDLGQPGSCLKRFS
2	Human	A1	V-HLVTR	T-DD-QS-TKILY	H L V H	RK
3	Human	A3	ATWTTR VFTR	TTAI E - TVPLY	S-FFVQH	QT
4	Human	A5	VAHLITR	${\tt TTDA-QQ-TLQVY}$	F-LKH	GTAR
5	Human	A2	VSI-YL	TDQE-MVNKL-	S L FE - Q - K - H	NGLAA
6	Human	Α4	YLGLL	TDQE-TLPRL-	T L L E - Q - K - H	NGLAPV
7	Human	A6	MRV-YT	-EQPISQL-	VL-FVE-Q-K-H	NGFAP
			↓ ↓75	90	105	↓ ↓ 120
1	(Hy)		TMPFLFCDIONKCVV	ASRNDYSFWLSTAE.	.KPKEAPSSGADLEN	YISRCIVCEAPSHVL
2	(A1)		N-N-V-NF	YP-P	MPMSMIT-ENIRP	FAEAM-M
3	(A3)		NVNDV-NF	PAL	MPMNMIT-RAP	TEG-AIAI
4	(A5)		MN-N-V-NF	P-P	MPMSMO-LK-OSIOP	FAEAV-I
5	(A2)		Y-NPGDV-YY	K-YTAP	LPMM-VAEDEIKP	SEAIAI
6	(A4)		-LAY-NIHQV-HY	-QR-YSAAP	LPMM-LSEEAIRP	-VAEAOAV
7	(A6)		HY-NINEV-HY	-RK-YTAP	IPMM-VSQTQIPQ	SEQAI
						_
			↓135	150	165	↓ ↓180
1	(Hy)		AVHSQSELDPKCPDG	WENLWTGFSFLMYNS	AGAOGSGOLLSSSGS	CLEDFRVNPYIECHG
2	(A1)		TIQI-PS-	-SSI-YV-HT-	EA-A-P	ESA-F
3	(A3)		H-	-ISKI-FT-	SE-TA-A-P	EAS-FL
4	(A5)		TIQI-HQ-	-DSI-YM-HT-	EA-A-P	ESA-F
5	(A2)		DVSI-HA-	-RSI-YHTA	-GDE-GS-V-P	N-
6	(A4)		DQSI-PQT	-RSI-YHTG	-GDGA-M-P	AA-FLO-
7	(A6)		DITI-QL-	-RSI-YHTA	E-GS-V-P	
			↓ 195	210	225	↓ ↓ 235
1	(Hy)		↓ 195 .RGTCWYYGPTLSFW	210 LSTIGESNMFO.VPK	225 SEILER.NLK.ARVS	$\downarrow \downarrow 235$ RCAVCMKSVP
1 2	(Hy) (A1)		↓ 195 .RGTCWYYGPTLSFW NANAY	210 LSTIGESNMFQ.VPK -AER-EK.K-T	225 SEILER.NLK.ARVS PST-KAGE-R.TH	↓ ↓ 235 RCAVCMKSVP 0RRT.
1 2 3	(Hy) (A1) (A3)		↓ 195 .RGTCWYYGPTLSFW NANAY NSNSY	210 LSTIGESNMFQ.VPK -AER-EK.K-T -ASLNPERR.K-I	225 SEILER.NLK.ARVS PST-KAGE-R.TH PSTVKAGE-E.KII-	↓ ↓ 235 RCAVCMKSVP QRRT. OKRH
1 2 3 4	(Hy) (A1) (A3) (A5)		↓ 195 .RGTCWYYGPTLSFW NANAY NSNSY NANSY	210 LSTIGESNMFQ.VPK -AER-EK.K-T -ASLNPERR.K-I -A-VDV-DS.K-Q	225 SEILER.NLK.ARVS PST-KAGE-R.TH- PSTVKAGE-E.KII T-KAGD-R.T-I-	↓ ↓ 235 RCAVCMKSVP QRRT. QKRH ORT.
1 2 3 4 5	(Hy) (A1) (A3) (A5) (A2)		↓ 195 .RGTCWYYGPTLSFW N-ANAY N-SNSY GH-ANKY	210 LSTIGESNMFQ.VPK -AER-EK.K-T -ASLNPERR.K-I -A-VDV-DS.K-Q -TP.QSGS-S	225 SEILER.NLK.ARVS PST-KAGE-R.TH PSTVKAGE-E.KII- -T-KAGD-R.T-I- ADT-KAGLIR.THI-	↓ ↓ 235 RCAVCMKSVP QRT. QKRH QRT. QNL.
123456	(HY) (A1) (A3) (A5) (A2) (A4)		↓ 195 .RGTCWYYGPTLSFW NANAY NSNSY NANSY GHFFANKY RQHFFANKY	210 LSTIGESNMFQ.VPK -AER-EK.K-T -ASUNPERR.K-I -A-VDV-DS.K-Q -TPQSGS-S -T-VKADLQ-SSA-A	225 SEILER.NLK.ARVS PST-KAGE-R.TH PSTVKAGE-E.KII- T-KAGLE.THI- ADT-KAGLIR.THI- PDT-KESQAQROKI-	↓ ↓ 235 RCAVCMKSVP QRT. QRT. QRT. QNL. QYS.
1234567	(HY) (A1) (A3) (A5) (A2) (A4) (A6)		↓ 195 .RGTCWYYGPTLSFW NANAY NANSY GHANSY RQHFFANKY AH-FANKY	210 LSTIGESNMFQ.VPK -AER-EK.K-T -ASLNPERR.K-I -A-VDV-DS.K-Q -TP.QSGS-S -T-VKADLQ-SSA-A -T-VE-RQQ-GEL-V	225 SEILER.NLK.ARVS PST-KAGE-R.TH PSTVKAGE-E.KII- T-KAGD-R.T-I- ADT-KAGLIR.THI PDT-KESQAQRQKI- T-KAGQ-H.T	↓ ↓ 235 RCAVCMKSVP QRRT. QRT. QRT. QNL. QV-YS. QL.

that should result in a mixture of dimers and monomers.

To get information about the structure of the dimer we isolated type IV collagen by ultracentrifugation and ion exchange chromatography in the presence of mercaptoethanol but in otherwise nondenaturing conditions. Electron microscopy after rotary shadowing of the protein showed dimers connected via the C-terminal globular domains (Fig. 5). Surprisingly, in marked contrast to the vertebrate type IV collagens, the triple helix exhibits many bends and kinks so that the predicted length of almost 400 nm is never observed. As seen in one molecule in Fig. 5, the collagenous domain is even able to fold back on itself. The supramolecular structure of the type IV collagen in the unreduced state was visualized by electron microscopy of the unfractionated EDTA extract. Although this fraction contains a mixture of the hydra fibrillar collagen I/II and the type IV collagen, assignment of the type IV structures was very easy because collagen I/II forms a clearly distinguishable fibrillar network lacking globular domains (12). Some selected electron microscopy pictures are shown in Fig. 6. Irregular networks are seen typically, but in addition, smaller oligomeric structures formed by lateral association of a few molecules can also be detected.

In Situ Hybridization—In situ hybridization analysis in adult hydra revealed that collagen IV expression is observed along the entire longitudinal axis of the animal but is most intense at the base of the tentacles, at the site of battery cell



FIG. 4. Isolation of collagen IV from hydra mesoglea. By extraction with EDTA/NaCl, a mixture of collagen I/II (155 kDa) and collagen IV (290 kDa) was solubilized and analyzed by electrophoresis on a 4-10% SDS gel prior to (*lane 1*) and after (*lane 2*) reduction with mercaptoethanol. Collagen IV was purified from the reduced extract by ion exchange chromatography using nondenaturing buffers (*lane 3*). The molecular weights of the chains have been determined with appropriate protein standards.



FIG. 5. Panels of electron micrographs of rotary shadowed dimeric type IV collagen molecules. Note the extremely high flexibility of the triple helical domain. The *bar* denotes 100 nm (applies to all images).

transdifferentiation (Fig. 7, A and B). Interestingly, collagen IV expression is restricted to the ectoderm (Fig. 7, C and D), whereas laminin β 1, the other basement membrane component characterized in hydra, is expressed solely by the endoderm (see "Discussion"). During head regeneration type IV collagen levels are increased basal to the decapitation site. Initially, an apical cap of type IV-positive cells form (Fig. 8, A–C), and then, by 72 h, additional patches of stronger expression appear in the tentacle region (Fig. 8D).

Functional Analysis of the Role of Hydra Collagen IV in Head Morphogenesis Using Antisense Thio-oligonucleotides—As shown in Fig. 9., head regeneration was blocked with antisense thio-oligonucleotides designed to a region of the 5'-UTR and coding region of hydra collagen IV. Statistically significant ($p \le$ 0.001) blockages of 84 and 45% were observed for these antisense oligonucleotides, respectively, when compared with mismatch and sense controls. No blockage was observed with antisense thio-oligonucleotides designed to the initiation site or to



FIG. 6. Electron micrographs of rotary shadowed type IV collagen oligomers and polymers seen in the unfractionated mesoglea extract. The *arrows* denote regions of parallel association of two type IV collagen molecules which can be recognized because of the short distances between their globular NC1 domains. The *bar* is 100 nm (applies to all images).

a region of the 3'-UTR. Inhibited animals (5'-UTR and coding groups) recovered from blockage by 5-7 days after LEP.

Effect of Glucose on Mesoglea and Type IV Collagen—Adult hydra were incubated for 14 days in HM containing 15 mM glucose. Control groups were incubated either in HM alone or in HM containing 15 mM sorbitol (control for osmolarity because the polyol sorbitol is not readily transported across cell membranes). TEM micrographs of sections through the three different regions of the body columns revealed that the mesoglea in the two control groups had the same thickness irrespective of the level of sectioning, whereas the mesoglea in the glucose-treated animals was significantly thickened (p < 0.001; Fig. 10 and Table II).

To investigate the effect of glucose on type IV collagen expression, hydra were incubated for 48 h in hydra medium without glucose or supplemented with 15 mM glucose. A Northern blot of RNA extracted from each experimental group was probed with the hydra collagen IV cDNA (Fig. 11). Compared with hydra in normal medium, there is 2–3-fold increase of type IV collagen mRNA levels in animals exposed to 15 mM glucose.

DISCUSSION

Although the presence of type IV collagen in hydra had previously been suggested by a study using antibodies raised against vertebrate BM components (9, 10), the elucidation of the primary structure of collagen IV (Figs. 1 and 2) provides the first molecular identification of this protein in hydra. Sequence



FIG. 7. Expression pattern of type IV collagen in adult hydra. Whole mount *in situ* hybridization using an RNA probe generated from clone 194.1 is shown. Localization of type IV collagen is observed along the entire longitudinal axis of the polyp (A) but is most intense at the base of the tentacles (B). The localization is restricted to the ectoderm (A and B). This is best seen in thin sections of whole mounts embedded in JB4 compound as shown in C and D. A bright field image is shown in C with the ectoderm containing reaction product for type IV collagen mRNA indicated by an *arrowhead*. The endoderm is negative and is indicated by an *asterisk* (C). The intervening ECM (*arrow*) is shown (*asterisk*) are also shown in D. Bar magnifications: A, 500 μ m; B, 100 μ m; C and D, 20 μ m.

analysis clearly shows the molecule is similar to all known $\alpha(IV)$ chains but most closely resembles vertebrate and invertebrate α 1(IV) collagen chains (Table I and Fig. 3).

Our data suggest that the hydra type IV collagen molecule is homotrimeric because all the peptide sequence data obtained from purified protein (Fig. 2) originates from the $\alpha 1(IV)$ chain. Solubilization of type IV collagen from vertebrate basement membranes usually requires pepsin digestion, and it was therefore very surprising that the hydra type IV collagen could be solubilized in the form of dimeric molecules under reducing but otherwise nondenaturing conditions (Figs. 4 and 5). Sequence analysis of peptides obtained by harsh protease treatment of the insoluble residue remaining after extraction revealed that all the type IV collagen had been solubilized (data not shown), thus excluding the possibility that solubilized material is only a minor fraction. Electron microscopical investigation (Fig. 6) showed that the dimers assemble into a tight network stabilized by disulfide bonds. For vertebrate type IV collagen a model has been proposed in which four molecules aggregate via their N-terminal domains forming spider like structures. The interactions are stabilized via disulfide bonds and lysine derived cross-links, resulting in a highly protease-resistant 7 S domain that can easily be isolated after pepsin digestion. In addition, the C-terminal globular domain, NC1, binds to itself, mainly via disulfide bridges, to form a linear dimer (14). Both interactions at the N-terminal and C-terminal ends lead to the proposal of an open network structure that can further polymerize via lateral aggregation of the triple helical domains (34). Our data certainly support the idea of lateral aggregation of type IV molecules, which is very difficult to observe experimentally in vertebrate basement membranes, although there are also marked differences in the molecular architecture. Firstly, two molecules of hydra collagen type IV are held together by nonreducible cross-links between the globular NC1 domains rather than by disulfide bonds (Fig. 5). Secondly, we do not have any structural evidence for the existence of a 7 S domain in hydra collagen IV despite the presence of the 4 conserved cysteine residues in this region. Pepsin digestion of the hydra type IV collagen did not yield the 7 S domain but instead resulted in a complete destruction of the protein; a protease-resistant domain could not be detected by SDS gel electrophoresis or by electron microscopy (data not shown). It is interesting to note that lack of a 7 S domain is not unique to hydra because similar results have previously been obtained for the collagen type IV molecule of the helminth Ascaris suum (35). Thus, formation of compact, protease-resistant 7 S domain might be an important step in the polymerization of vertebrate collagen type IV but appears not be essential for invertebrates. For hydra, a less compacted and highly flexible organization of the collagen IV network is presumably advantageous given the organisms physiological requirements for an extremely flexibile mesoglea.

The hydra is a member of the phylum Cnidaria, one of the oldest metazoan phyla and the highly conserved nature of type IV collagen indicates its critical role in hydra ECM formation. The presence of both collagen IV and laminin in hydra and their co-localization in the mesoglea (11) provide compelling evidence that the mesoglea is essentially composed of the same molecular components found in vertebrate BMs, and it is likely that cell-substrate interactions involve the same ligands and receptors. However, it is noteworthy that in addition to classical BM components, the hydra mesoglea also contains a centrally located core of another collagen, which, based on sequence comparisons, is a classical fibrillar collagen (12). It is therefore possible that the first functional extracellular matrices evolved with composite properties of what are now considered classical basement membrane and interstitial matrices and that later in evolution, the uses of components in these "primitive basement membranes" were refined to produce the wide array of connective tissues seen in higher orders.

The restriction of hydra type IV expression to the ectoderm

Hydra Type IV Collagen

FIG. 8. Expression of hydra type IV collagen during head regeneration. Following decapitation just basal to the neck region, hydra were fixed and processed for whole mount *in situ* hybridization using a digoxigenin-labeled RNA probe generated from clone 194.1. The *arrow* indicates localization of reaction product for type IV collagen in the ectoderm layer beginning at 4 h and through 72 h when higher levels are associated with the base of the tentacles. The *bar* denotes 20 μ m (applies to all images).



(Fig. 7) contrasts with laminin, which is expressed solely in the endoderm but is localized to the subepithelial zones adjacent to both the ectoderm and endoderm layers (11). In an immunocytochemical study of hydra ECM, polyclonal antibodies generated against mammalian type IV collagen were localized throughout the entire mesoglea (9, 10), suggesting that, like laminin, type IV collagen can assemble into basement membranes on cells that have not produced it. The sources and final locations of type IV collagen have been determined in some vertebrate and invertebrate systems. In co-cultures of fetal intestinal chick mesenchyme with rat endoderm, type IV collagen in the resulting basement membrane was derived only from the mesenchyme (36), whereas in *C. elegans* type IV has been detected at sites distant from its site of synthesis (37). The assembly of type IV collagen in hydra, away from the ectoderm cells that express it, suggests there is a mechanism regulating its assembly that is directed by interaction with other (cell surface-associated) molecules. Functional antisense studies show inhibition of type IV collagen translation causes a subsequent blockage in head regeneration (Fig. 9), re-emphasizing the importance of this molecule and the mesoglea in regeneration and development.

In mammals, thickening of the basement membranes in blood capillaries is the hallmark of diabetic microangiopathy, a



FIG. 9. Effect of hydra collagen IV antisense thio-oligonucleotides on head regeneration. The percentage of blockage of head regeneration is shown for each experimental and control group tested. The *asterisk* indicates groups that were significantly different from controls at a $p \leq 0.0001$. N denotes the number of animals per group. Sequence of thio-oligonucleotides are as follows: 5'-UTR, 5'-AAT CGG CCA TAT ATC GAA AG-3'; Initiation (*Init*), 5'-CAT TTG GTC GTG GTG TGA TTC AT-3'; Coding, 5'-AAT AGG TCC TTG TGG TCC GG-3'; 3'-UTR, 5'-AAA ACT GTT TCG TAA CAA AT-3'; Mismatch, 5'-GCT AAT ATG CAT GTA TCC GA-3'; and Sense 5'-UTR, 5'-CTT TCG ATA TAT GGC CGT TT-3'.



FIG. 10. Ultrastructural analyses of mesoglea formation under normal hydra medium and elevated glucose conditions. Representative TEM micrographs of sections through the middle region of the body columns of three separate adult hydra (a-c) incubated for 14 days in control hydra medium (Control) and medium containing either 15 mm sorbitol (control for osmolarity) or 15 mm glucose. The mesoglea (me) and ectoderm (ec) are labeled (×25,000). TEM, transmission electron microscope.

TABLE II Effects of glucose on mesoglea thickness

Mean (μ m; \pm S.E.) mesoglea thickness at the top, middle, and bottom of the body column of three adult hydra incubated for 14 days in hydra medium alone (control) and medium containing either 15 mM sorbitol (control for osmolarity) or 15 mM glucose. The data are pooled to produce an overall measure of mean mesoglea thickness (body mean). For control versus sorbitol, there were no significant differences. For control or sorbitol versus glucose, p < 0.001, n = 3 for all comparisons.

	Control	Sorbitol	Glucose
	μm	μm	μm
Top	0.311 ± 0.006	0.312 ± 0.002	0.500 ± 0.01
Middle	0.317 ± 0.01	0.321 ± 0.001	0.516 ± 0.002
Bottom	0.328 ± 0.001	0.328 ± 0.001	0.520 ± 0.004
Body mean	0.319 ± 0.004	0.320 ± 0.002	0.512 ± 0.004

severe long term complication of diabetes mellitus that is the leading cause of blindness and renal failure in the developed world. The molecular mechanism(s) underlying this thickening are still undetermined, although the process is thought to be glucose-dependent. Zhang et al. (22) showed that reaggregating hydra exposed to 15 mm glucose developed a newly synthesized mesoglea twice the thickness of controls and suggested hydra as a simple and rapid model system for studying glucose induced basement membrane thickening. We now show that thickening of mesoglea upon exposure to glucose is a generalized phenomenon in hydra, occurring rapidly in adults (Fig. 10 and Table II) as well as developing (reaggregating) animals. This effect is directly caused by glucose because, unlike mammals, the hydra does not possess the endocrine system that is clearly perturbed in diabetes coincident with the onset of hyperglycaemia. Furthermore, we demonstrate that the thickening of mesoglea is preceded by an up-regulation of collagen IV expression within 48 h of glucose exposure (Fig. 11). The expression of laminin (β 1 chain mRNA) has also been found to be increased upon exposure to glucose.² These effects of glucose

² M. Sarras, Jr., unpublished observations.



Actin Collagen IV

FIG. 11. Effects of glucose on type IV collagen expression. Two groups of adult hydra were incubated for 48 h prior to RNA extraction in either hydra medium alone (tracks 1 and 2) or hydra medium supplemented with 15 mM glucose (tracks 1G and 2G). RNA was Northern blotted and probed firstly with the hydra collagen IV cDNA and secondly with a hydra actin sequence as a loading and blotting control.

are not due to a generalized increased message levels because the actin mRNA levels presented in Fig. 11 were not altered. These findings raise the possibility that the thickening of BM seen in diabetes is, in evolutionary terms, an ancient response to elevated glucose levels retained by cells that are freely permeable to glucose. Given the striking conservation of hydra and mammalian BM genes, together with the rapidity with which glucose-responses are elicited, the hydra represents an experimentally tractable, simple, in vivo model system in which to further investigate the mechanism underlying glucose-induced basement membrane thickening.

In summary, we have characterized hydra type IV collagen and shown it to be highly conserved with respect to its mammalian counterparts, although its organization and assembly at the supramolecular level shows some unique features. Type IV collagen expression is essential for hydra development and responsive to elevated levels of glucose.

Acknowledgments-Sharon Dexter, Euard Hochmuth, and Reinhard Rachel are thanked for technical support. Ian Miller is thanked for help in preparing for electronic submission of the manuscript.

REFERENCES

- Burnett, A. L., and Hausmann, R. E. (1969) J. Exp. Zool. 171, 15–24
 Fawcett, D. W. (1961) in The Biology of Hydra and of Some Other Coelenterates
- (Lenhoff, H. M., and Loomis, W. F., eds) p. 65, University of Miami Press, Miami, FL
- 3. Day, R. M., and Lenhoff, H. M. (1981) Science 211, 291-294
- Agosti, C. G., and Stidwill, R. P. (1991) Cell Motil. Cryoscel. 20, 215–227
 Zhang, X., and Sarras, M. P., Jr. (1994) Development 120, 425–432
- Shostak, S., Patel, N. G., and Burnett, A. L. (1965) Dev. Biol. 12, 434-450 6.
- Hausman, R. E., and Burnett, A. L. (1971) J. Exp. Zool. 177, 435–446
- Barzansky, B., and Lenhoff, H. M. (1974) Amer. Zool. 14, 575-581
- Sarras, M. P., Jr., Madden, E. M., Zhang, X., Gunwar, S., Huff, J. K., and Hudson, B. G. (1991) Dev. Biol. 148, 481-494
- 10. Sarras, M. P., Jr., Meador, D., and Zhang, X. (1991) Dev. Biol. 148, 495-500 11. Sarras, M. P., Jr., Yan, L., Grens, A., Zhang, X., Agbas, A., Huff, J. K., St. John,
- P. L., and Abrahamson, D. R. (1994) *Dev. Biol.* **164**, 312–324 Deutzmann, R., Fowler, S., Zhang, X., Boone, K., Dexter, S., Boot-Handford, 12. R. P., Rachel, R., and Sarras, M. P., Jr. (2000) Development, 127, 4669-4680
- 13. Johansson, C., Butkowski, R., and Wieslander, J. (1992) J. Biol. Chem. 267, 24533-24537
- 14. Kühn, K. (1994) Matrix Biol. 14, 439-445 15. Blumberg, B., MacKrell, A., and Fessler, J. H. (1988) J. Biol. Chem. 263, 18328 - 18337
- 16. Guo, X. D., Johnson, J. J., and Kramer, J. M. (1991) Nature 349, 707-709
- 17. Exposito, J-Y., D'Alessio, M., Di Liberto, M., and Ramirez, F. (1993) J. Biol.
- Chem. 268, 5249-5254 18. Exposito, J-Y., Suzuki, H., Geourjon, C., Garrone, R., Solursh, M., and
- Ramirez, F. (1994) J. Biol. Chem. 269, 13167-13171 19. Boute, N., Exposito, J. Y., Boury-Esnault, N., Vacelet, J., Noro, N., Miyazaki,
- K., Yoshizato, K., and Garrone, R. (1996) Biol. Cell 88, 37-44

- Leinonen, A., Mariyama, M., Mochizuki, T., Tryggvason, K., and Reeders, S. T. (1994) J. Biol. Chem. 269, 26172–26177
- 21. Prockop, D. J., and Kivirikko, K. I. (1995) Annu. Rev. Biochem. 64, 403-434
- 22. Zhang, X., Huff, J. K., Hudson, B. G., and Sarras, M. P., Jr. (1990) Diabetologia 33, 704-707
- 23. Engel, J., Odermatt, E., Engel, A., Madri, J. A., Furthmayr, H., Rohde, H., and Timpl, R. (1981) J. Mol. Biol. 150, 97–120
- 24. Grens, A., Mason, E., Marsh, J. L., and Bode, H. R. (1995) Development 121, 4027-4035 25. Grens, A., Shimizu, H., Hoffmeister, S. A. H., Bode, H. R., and Fujisawa, T.
- (1999) Development 126, 517–524 26. Yan, L., Fei, K. Y., Zhang, J. S., Dexter, S., and Sarras, M. P., Jr. (2000)
- Development 127, 129-141 27. Yan, L., Leontovich, A., Fei, K. Y., and Sarras, M. P., Jr. (2000) Dev. Biol. 219,
- 115 128
- Leontovich, A. A., Zhang, J. S., Shimokawa, K., Nagase, H., and Sarras, M. P., Jr. (2000) *Development* **127**, 907–920
 Wagner, R. W. (1994) *Nature* **372**, 333–335
 Wagner, R. W. (1995) *Nature Med.* **1**, 1116–1118

- 31. Flanagen, R. W., Kothavale, A., and Wagner, R. W. (1996) Nucleic Acids Res.

24, 2936–2941

- 32. Boot-Handford, R. P., Kurkinen, M., and Prockop, D. J. (1987) J. Biol. Chem. 262, 12475-12478
- 33. Oberbäumer, 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, 217-24
- 34. Yurchenco, P. D. (1990) Ann. N. Y. Acad. Sci. 580, 195-213
- 35. Noelken, M. E., Wisdom, B. J., Jr., Dean, D. C., Hung, C.-H., and Hudson, B. G. (1986) J. Biol. Chem. 261, 4706-4714
- 36. Simon-Assmann, P., Kedinger, P. M., DeArcangelis, A., Rousseau, V., and Simo, P. (1995) Experientia 51, 883-890
- 37. Graham, P. L., Johnson, J. J., Wang, S., Sibley, M. H., Gupta, M. C., and Kramer, J. M. (1997) J. Cell Biol. 137, 1171–1183
- 38. Brazel, D., Oberbaumer, I., Dieringer, H., Babel, W., Glanville, R. W, Deutzmann, R., and Kühn, K. (1987) Eur. J. Biochem. 168, 529-536
- 39. Hostikka, S. L., and Tryggvason, K. (1988) J. Biol. Chem. 263, 19488-19493 40. Sibley, M. H., Johnson, J. J., Mello, C. C., and Kramer, J. M. (1993) J. Cell Biol. **123,** 255–64
- 41. Pettitt, J., and Kingston, I. B. (1991) J. Biol. Chem. 266, 16149-16156