The Collagens of Hydra Provide Insight into the Evolution of Metazoan Extracellular Matrices

A collagen-based extracellular matrix is one defining feature of all Metazoa. The thick sheet-like extracellular matrix (mesoglea) of the diploblast, hydra, has characteristics of both a basement membrane and an interstitial matrix. Several genes associated with mesoglea have been cloned including a basement membrane and fibrillar collagen and an A and B chain of laminin. Here we report the characterization of a further three fibrillar collagen genes (Hcol2, Hcol3, and Hcol5) and the partial sequence of a collagen gene with a unique structural organization consisting of multiple von Willebrand factor A domains interspersed with interrupted collagenous triple helices (Hcol6) from Hydra vulgaris. Hcol2 and -5 have major collagenous domains of classical length (~1020 amino acid residues), whereas the equivalent domain in Hcol3 is shorter (969 residues). The N-propeptide of Hcol2 contains a whey acidic protein four-cysteine repeat (WAP) domain, and the equivalent domain of Hcol3 contains two WAP and two von Willebrand factor A domains. Phylogenetic analyses reveal that the hydra fibrillar collagen genes form a distinct clade that appears related to the protostome/deuterostome A clade of fibrillar collagens. Data base searches reveal Hcol2, -5, and -6 are highly conserved in Hydra magnipapillata, which also provided preliminary evidence for the expression of a B-clade fibrillar collagen. All four of the H. vulgaris collagens are expressed specifically by the ectoderm. The expression pattern for Hcol2 is similar to that previously reported for Hcol1 (Deutzmann, R., Fowler, S., Zhang, X., Boone, K., Dexter, S., Boot-Handford, R. P., Rachel, R., and Sarras, M. P., Jr. (2000) Development 127, 4669 – 4680) but distinct from the pattern shared by Hcol3 and Hcol5. The characterization of multiple collagen genes in relatively simple diploblastic organisms provides new insights into the molecular evolution of collagens and the origins of the collagen-based extracellular matrix found throughout the multicellular animal kingdom.

Collagen genes that are associated with basement membranes and interstitial matrices in vertebrates (type IV and fibrillar collagens) are found throughout the metazoan kingdom (1, 2). However, the extracellular matrices (ECM) with which these collagen genes are associated, namely basement membranes and interstitial matrices, only become apparent as distinct entities in triploblasts, three cell-layered organisms (3). The simplest metazoan forms of life (Parazoa such as the sponges) exhibit both fibrillar and basement membrane collagens, and some sponges exhibit thin fibrils and evidence of a basement membrane-like structure (4, 5). Diploblasts, such as the cnidarian, hydra, are intermediate between the simplest metazoans and the triploblasts (protostomes and deuterostomes) that constitute the majority of the multicellular animal kingdom. The thick sheet-like ECM of hydra (mesoglea) appears to be a “composite” matrix having characteristics of both a basement membrane and an interstitial matrix (6). The mesoglea may represent an intermediate form of matrix bridging the gap between the relatively unorganized matrices secreted by sponges and the large variety of highly structured matrices found in triploblasts, i.e. basement membranes and interstitial matrices containing large, well developed fibrils.

Previously, we have cloned and characterized several matrix genes from hydra associated with ECM including a type IV collagen, Hcol-IV, hereafter referred to as Hcol4 (7), a fibrillar collagen Hcol-I, hereafter referred to as Hcol1 (8), and laminin A and B chains (9, 10). Immunohistochemical and ultrastructural analyses have indicated that the mesoglea is organized as two sub-epithelial thin layers of basal lamina sandwiching a central interstitial zone (11). Expression studies have revealed that the collagen genes (Hcol1 and Hcol4) appear to be expressed exclusively by ectoderm (7, 8), whereas laminin genes are expressed by endoderm (9, 10). When decapitated, hydra is capable of regenerating the missing head and tentacles. The regeneration involves cell migration to cover the wound and de novo ECM synthesis at the regenerating tip (12). The ECM re-synthesis and assembly at the regenerating tip provides an attractive and simple multicellular model for studies of ECM structure and function in vivo in the laboratory setting.

* The work was supported by National Institutes of Health Grant DK065123 (to M. P. S. and X. Z.). The Royal Society, and Biotechnology and Biological Sciences Research Council, Swindon, United Kingdom. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The on-line version of this article (available at http://www.jbc.org) contains supplemental Fig. 15.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) DQ679807, DQ679808, DQ679809, and DQ679810.

1 These are co-first authors.

2 To whom correspondence should be addressed. Tel.: 44-161-275-5097; Fax: 44-161-275-5082; E-mail: ray.boot-handford@manchester.ac.uk.

3 The abbreviations used are: ECM, extracellular matrix; WAP, whey acidic protein; vWF, von Willebrand factor; Chaps, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonic acid.
Here we report the characterization of a further three hydra fibrillar collagen genes (Hcol2, Hcol3, and Hcol5). In addition, we report the partial characterization of a novel hydra collagen gene (Hcol6) with a unique organization consisting of vWF A domains interspersed with interrupted triple-helical domains. We demonstrate by phylogenetic analyses that the hydra fibrillar collagen genes form a distinct clade that appears related to the protostome/deuterostome A clade of fibrillar collagens. We demonstrate by searching EST databases that most of these genes are highly conserved in the related *Hydra magnipapillata* and provide the first evidence for the existence of B-clade fibrillar collagens in diploblasts. Each collagen gene is expressed in highly specific developmentally regulated patterns solely within the ectoderm. These four novel *Hydra vulgaris* collagen genes provide new insights into the molecular evolution of collagens and the organization and structure of metazoan extracellular matrices.

**EXPERIMENTAL PROCEDURES**

Animal Maintenance—*H. vulgaris*, L2 stain was used for in situ hybridization experiments. Animals were cultured in M solution (13) and fed with freshly hatched *Artemia* nauplii three times a week. All experiments were carried out 24–48 h after feeding.

**cDNA Cloning and Sequencing**—The isolation of the original cDNA clones for the different collagen genes by differential screening of regenerating hydra cDNA libraries and the subsequent cloning and sequencing of overlapping cDNAs for each gene were performed as described previously (7, 8).

**Phylogenetic Analysis**—The novel fibrillar collagen Hcol2, -3, and -5 genes identified in *H. vulgaris* were aligned with representative metazoan fibrillar collagens and the previously identified Hydra Hcol1 gene (8) using ClustalX (14). Gap-containing sites were removed from each alignment, and a Neighbor Joining tree was inferred. The topology of this tree was tested using four independent methods. Neighbor Joining and Maximum Parsimony bootstrap replicates were obtained using PROTDIST, NEIGHBOR, and PROTPARS from the PHYLIP package (15). Maximum Likelihood trees were inferred using PROML from the PHYLIP package (15). The JTT model of amino acid substitutions was used with global rearrangements and correction for rate heterogeneity (α value was obtained from TREEPUZZLE (16)). Bayesian tree inference values were produced from the MrBayes program (17).

Identification of vWF A Domains and WAP Domains—vWF A and WAP domains were originally identified using the rps-blast and cdart protocols (www.ncbi.nlm.nih.gov/BLAST). Alignments were performed using Clustal programs as indicated in the legends to the appropriate figures. vWF A domain MIDAS motifs, α helices, and β strands were predicted based upon Jackson et al. (18).

**In Situ Hybridization**—Digoxigenin-labeled riboprobes were obtained through the following method. PCR primers were syn-

**FIGURE 1. Domain structure of the hydra collagens.** A schematic drawing of the hydra Hcol2 (accession number DQ679807), Hcol3 (accession number DQ679808), and Hcol5 (accession number DQ679809) domain structures, shown with the previously identified Hcol1 sequence (8) and representative human fibrillar collagen genes from the A, B, and C clades (collagens Iα1, Vα1, and XXVIIα1). B, schematic drawing of the Hcol6 (accession number DQ679810) domain structure. The asterisk indicates the sequence is incomplete.
thesized according to the sequences of hydra collagens. The primer sequences are: Hcol2, accacctggtc-aaagag and aagccccaattcttccttgtg; Hcol3, tctgcctctactggtagag and taccctggcccttgttc; Hcol5, accacctggtc-aaagag and aagccccaattcttccttgntt; Hcol6, accacctggtc-aaagag and aagccccaattcttccttccttggt. PCR fragments were cloned into pCR II vector (Invitrogen) and linearized with appropriate restriction enzymes. In vitro transcription was carried out using "Maxiscript" kit (Ambion) and Dig labeling mix (Roche Applied Science) according to manufacturer protocols. Both wild type adult polyps and polyps that were decapitated and followed during head regeneration for different times were used for in situ. The whole-mount in situ hybridization reaction was carried out the same way as described previously (7–10). Briefly, hydra polyps were “relaxed” in 2% urethane and subsequently fixed with freshly prepared 4% paraformaldehyde. Samples were treated with ethanol and proteinase K, re-fixed with 4% paraformaldehyde, and hybridized with probes in hybridization solution (50% formamide, 5× SSC (0.15 M NaCl and 0.015 M sodium citrate), 1× Denhardt’s solution, 200 mg/ml tRNA, 100 mg/ml heparin, 0.1% Chaps, and 0.1% Tween 20). Hybridization was carried out for 65 h followed by washing steps using SSC. Anti-digoxigenin antibody (AP-conjugated, 1:2000 dilution, Roche Applied Science) was used to detect probe labeling, and BM Purple (Roche Applied Science) was used as an alkaline phosphatase substrate for color reaction. Some samples were embedded in JB4 after the color reaction and sectioned (20-μm thickness) with a glass knife for microscopic examination.

RESULTS

Four Novel Collagen Genes Identified in H. vulgaris—Differential screening of a cDNA library using a combination of non-stringent and stringent conditions resulted in the identification of four novel collagen genes in addition to the previously described fibrillar Hcol1 (8) and basement membrane Hcol4 genes (5).

Characterization of the Three Novel Fibrillar Collagen Genes—Three of the novel genes (Hcol2, Hcol3, and Hcol5) are members of the fibrillar collagen family based on their conserved C-terminal non-collagenous and uninterrupted collagenous domain structures (Fig. 1 and 2; Table 1). Fibrillar collagen NC1 domains have 8 cysteine residues (19), and 6 of these (numbers 1 and 4–8) are conserved in all of the hydra fibrillar collagen genes. In the hydra fibrillar collagens, both cysteines 2 and 3 were partially conserved whereas cysteines 1 and 4–8 are completely conserved. Levels of sequence conservation are boxed. Levels of sequence conservation are indicated (red, >50% identical; blue, conservative substitutions).
lar collagens (Fig. 2). Cysteine 2 is present in Hcol2, and cysteine 3 is in Hcol1 (Fig. 2). All of the hydra fibrillar collagens have the short chain selectivity sequence (Fig. 2) exhibited by invertebrate fibrillar collagen chains (3).

The major helical domains of Hcol2 and Hcol5 are 1020–1023 amino acid residues in length, which is very similar to most other fibrillar collagens, whereas the same domain in Hcol3 is significantly shorter (969 residues) (Table 1, Fig. 2). The three novel hydra fibrillar collagen genes contain uninterrupted N-terminal minor triple helical domains separated from their respective major collagenous domains by telopeptide-like sequences.

By amino acid composition, all of the hydra fibrillar collagenous domains have ~33% glycine content (Table 2). However, the hydra chains (which appear related to the A-clade; see below) contain lower levels of alanine and higher levels of lysine than most A-clade fibrillar collagen domains (Table 2). Proline levels in the hydra chains are considerably lower than those in the classical vertebrate collagens (types I–III, V, and XI) but are typical for invertebrate fibrillar collagens (Table 2).

N-terminal Non-collagenous Domains of the Novel Fibrillar Collagens—The sequence for Hcol2 is full-length, and the N-terminal non-collagenous domain consists of a signal peptide followed by a whey acidic protein 4 disulfide core (WAP) domain (Fig. 3). The sequence of Hcol3 N-terminal to the minor collagenous domain consists of a tandem repeat of a vWF A and WAP domain (Fig. 1A, Table 1); this sequence is not full-length as we have yet to identify a start codon followed by a signal peptide. The Hcol2 and Hcol3_1 WAP domains have complete 8-cysteine complements with the position of cysteine 2 being least constrained as seen in other WAP domains (Fig. 3). The more C-terminal Hcol3_2 WAP domain is missing the cysteine at position 7 of the 8-cysteine motif (Fig. 3). The vWF A domains of Hcol3 are discussed in detail below. The N-terminal non-collagenous region of Hcol5 has yet to be cloned.

Phylogenetic Analysis—The major triple helical and C-terminal non-collagenous domain sequences of the four known hydra fibrillar collagen sequences (Hcol1–3 and 5) were aligned with representative protostome and deuterosome orthologues as listed in Table 3. A condensed version of this alignment containing hydra and representative human fibrillar collagen sequences is presented in Fig. 2. These alignments were then used for phylogenetic analyses to determine the evolutionary relationships of the hydra genes. The results indicate that the four hydra fibrillar collagen sequences cluster weakly with the A rather than B or C clade genes in all four of the different phylogenetic methods used in this analysis (Fig. 4). The clustering of the hydra genes in the deepest region of the A clade indicates that these genes have evolved within the hydra...
lineage and are not directly related to individual fibrillar collagen genes in the vertebrate lineage. The analyses provide very strong support for the contention that Hcol3 and -5 and Hcol1 and -2 are paralogous pairs that have evolved after the divergence of diploblasts such as hydra and triploblasts such as arthropods (apis), echinoderms (urchin), or chordates (ciona and vertebrates) (Fig. 4).

Characterization of a Fourth Novel Collagen Gene—The fourth novel hydra collagen gene, Hcol6, exhibits a unique domain structure (Fig. 1). To date we have characterized over 7 kilobases of open reading frame encoding 2439 amino acid residues (Fig. 5) but have not yet located either the start or termination codons. The gene has a unique structure encoding five vWF A domains interspersed with interrupted collagenous domains (Fig. 1).

Multiple vWF A Domains Are Found in Hcol3 and Hcol6—Approximate A domain boundaries in the N-propeptide of hydra collagen gene Hcol3 and Hcol6 are indicated (Fig. 3).

### TABLE 1

| N-terminal non-collagenous domain | N-terminal minor helical domain and telopeptide | Major helical domain | C-terminal non-collagenous domain |
|-----------------------------------|-----------------------------------------------|---------------------|----------------------------------|
| Hcol2 (1476 aa residues)          | 23 aa SP plus 63 aa sequence including 49 aa WAP domain | 54 aa Gly-...-repeat plus 42 aa telopeptide-like region | 1020 Gly-X-...-repeat 274 aa residues |
| Hcol3 (1883 aa residues)          | 539 aa residues* including 2 WAP (50 and 45 aa) and 2 vWF A (151 and 179 aa) | 54 aa Gly-X-...-repeat plus 36 aa telopeptide-like region | 969 Gly-X-...-repeat 285 aa residues |
| Hcol5 (1371 aa residues)*         | Not cloned                                      | 33 aa Gly-X-...-repeat plus 36 aa telopeptide-like region | 1023 Gly-X-...-repeat 279 aa residues |
| Hcol1* (1412 aa residues)         | 15 aa SP plus 12 aa                             | 57 aa Gly-X-...-repeat plus 8 aa telopeptide-like region | 1020 Gly-X-...-repeat 300 aa residues |

* Sequence not complete at the N terminus.

### TABLE 2

Amino acid composition of the collagenous domain of specified α chains

Chains included in B* are α1(V), α3(V), α1(XI), α2(XI), and α2(V), which is an α-V clade chain. Chains included in A* are α1(I), α2(I), α1(II), and α1(III). Sequences analyzed were the collagenous domains from the N-terminal end of the major helix to the C-terminal end of the major helix inclusive. Lys residues in the Y position are lysine residues located at the X Y Gly position in the collagenous repeat that are the substrate for lysyl hydroxylase.

### TABLE 3

Alignment of hydra Hcol2 and Hcol3 WAP domains with human WAP domain-containing proteins. Alignment was performed using ClustalX. Levels of sequence conservation are indicated (red, >70% identical; blue, conservative substitutions.), NCBI accession numbers for human sequences: WDFC1 (WAP four-disulfide core domain 1, NP067020), WDFC2 (WAP four-disulfide core domain 2, NP060994), WDFC3 (WAP four-disulfide core domain 3, NP542181), KAL (Kallmann syndrome 1 protein NP000207), P13 (peptide inhibitor 3, elafin preprotein, NP002629), SLPI (secretory leukocyte peptide inhibitor precursor, NP003055), SPNLW1 (serine peptide inhibitor-like with Kunitz and WAP domains 1, NO065131).

Hydra Extracellular Matrix Collagen and Evolution

Hydra Extracellular Matrix Collagen and Evolution
Hcol3 and in Hcol6 were determined by sequence alignment with vWF A domains from zebrafish matrilin-1 and human matrilin-3. The sequence alignment (Fig. 6) was produced using ClustalW with manual adjustment and is more reliable in the N-terminal portion of the domain because of the relative lack of sequence conservation toward the C terminus. The hydra sequences all show the characteristic pattern of alternating B strands and α helices predicted for vWF A domains (Fig. 6). The alignment reveals that the complete set of residues required for MIDAS cation-coordination are conserved in the vWF A domain repeats Hcol6-1, Hcol6-4, and Hcol6-5 (Fig. 6).

Localization of Fibrillar Collagen Gene Expression in Hydra during Development and during Tissue Regeneration—The expression patterns of the novel fibrillar collagens, Hcol2, -3, and -5, were examined in adult and budding hydra as well as during tissue regeneration after decapitation (Figs. 7–9, respectively).

Relatively high expression of Hcol2 was apparent throughout the adult polyp (Fig. 6, A and B). In buds Hcol2 expression was marginally increased compared with adults and also widely distributed throughout the developing organism (Fig. 7C). Cross-sections of the body tube reveal the expression of Hcol2 to be concentrated in the ectoderm (Fig. 7D). Hcol2 expression is strongly up-regulated during tissue regeneration within 6 h of decapitation and remains high throughout the regeneration process (Fig. 7, E–H).

The expression of Hcol3 was focused at the base of the tentacles and the peduncle region in adult polyps and in buds (Fig. 8A). A much weaker signal was apparent throughout the body column. Nevertheless, the expression of Hcol3 appears concentrated in the ectoderm (Fig. 8B). Hcol3 appears only weakly expressed during tissue regeneration, reaching detectable levels 12 h after decapitation (Fig. 8, C and D).

The expression pattern of Hcol5 was similar to that of Hcol3 being focused upon tentacles and the peduncle region in adult and through developing buds (Fig. 9, A–D). Expression of Hcol5 is up-regulated during tissue regeneration within 6 h of decapitation (Fig. 9, E and F).
Expression of the Novel vWF A Domain-containing Hc06—The expression pattern of Hc06 is limited to the base of the tentacles (Figs. 8, 9). In Hydra, the expression level for Hc06 appears lower than that detected for any of the incomplete terminal sequences for H. vulgaris collagen genes described above. However, BLAST searching with the known hydra fibrillar collagen sequences yielded two overlapping clones from the H. magnipapillata database (Hc07). Hc07 consists of an N-terminal 39-amino acid collagenous domain (Gly-X-Y) repeat followed by 128 amino acids of a fibrillar collagen NC1 domain. This sequence appears most similar to B clade collagen chains (supplemental Fig. S1). Although too short for any meaningful phylogenetic analysis, Hc07 provides preliminary evidence for the presence of a B clade fibrillar collagen in diploblasts.

Amino Acid Composition of the Collagenous Domain and Fibril Diameter—One possible explanation for the thin fibrils formed in the hydra mesoglea is that the N-propeptide of Hc07 is not cleaved during assembly into the matrix (8). It is unknown whether the bulkier N-propeptides of the newly described hydra collagens (see below) are processed. Nevertheless, a second possible factor that may influence fibril diameter is amino acid composition of the collagenous domain. The collagens are known to form thin, 10-nm non-striated fibrils (8, 11). Similarly, we have recently immunolocalized collagen XXVII to 10-nm non-striated fibrils,4 and we assume this to be a characteristic of other C-clade collagens. The vertebrate minor fibrillar collagens (types V and XI) are predominantly composed of B-clade α chains and assemble to form significantly thinner fibrils than their A-clade collagen counterparts.

4 D. A. Plumb, V. Dhir, A. Mironov, R. Poulos, K. E. Kadler, D. J. Thornton, M. D. Briggs, and R. P. Boot-Handford, manuscript in preparation.
such as type I or II collagen (21). When arranged in this order, all of the chains to the left in Table 2 have reduced levels of alanine (the smallest amino acid normally occupying the X or Y position of the Gly-X-Y repeat) in comparison with the chains to the right (with the exception of the Apis A clade chain). Likewise, all of the chains on the left have much higher lysine contents (in particular, lysine located at the Y position of the Gly-X-Y repeat that are the substrate for forming hydroxyllysine; Table 2) than those on the right (again with the exception of the Apis A clade chain). Fibril thickness has previously been inversely related to the abundance of bulky hydrophobic amino acids and glycosylated hydroxylysine (21). Assuming that the level of lysine in the Y position of the collagen repeat is related to the final level of lysine hydroxylation and glycosylation, it would appear that the relatively high levels of alanine and low levels of lysine may be a predictor of the capacity to form thick fibrils and vice versa for thin fibrils. Furthermore, it is hypothesized that, based on their alanine and lysine contents, the fibrils formed by both A and B clade collagens in the bee (Apis) are likely to be more akin to the thin fibrils formed by hydra collagens than the well developed cross-striated fibrils of vertebrates.

**Novel N-propeptide Domains in Hydra Fibrillar Collagens—**As reported previously, Hcol1 has little sequence N-terminal to its minor collagenous helix other than the signal peptide and start codon (8), and in this respect Hcol1 is reminiscent of the vertebrate proα2(I) collagen chain and the Strongylocentrotus purpuratus α1 chain (22). However, it is of interest to note that the N-terminal propeptides of Hcol2 and Hcol3 have completely novel arrangements of previously characterized protein domains. Both share WAP domains, which have not previously been reported associated with collagen genes (Fig. 1). WAP domains consist of about 50 amino acid motifs containing eight highly conserved cysteine residues that form four disulfide bridges (Fig. 3 and Ref. 23). WAP domains in some proteins have protease inhibitory activity, and others have antibacterial activities contributing to innate immunity (24), particularly important in invertebrates, which have no acquired immunity. It is an intriguing possibility that the N-propeptides of Hcol2 and Hcol3 function as part of the hydra innate immune system providing antibacterial activities upon and if released during collagen assembly. In addition, Hcol3 has vWF A domains in its N terminus (Figs. 1 and 6), and these domains could serve to concentrate the released N-propeptides to the mesoglea by way of interaction with the collagenous helices (see Ref. 29).

**How Do Hydra Fibrillar Collagens Relate to the Fibrillar Collagen Family—**Detailed phylogenetic analyses (Fig. 4) indicated that the four H. vulgaris genes described in the results are weakly related to each other in two paralogous pairs (Hcol3 and -5 and Hcol1 and -2) and that these pairs cluster with the A rather than B or C clade. The phylogenetic analyses coupled with the fact that Hcol2 and 3 share, uniquely to date compared with all other known fibrillar collagens, a WAP domain in their N-propeptides raise the possibility that these hydra fibrillar collagens arose by duplication from a common progenitor after the divergence of the diploblast and triploblast lineages. The N-propeptide organizations of these hydra fibrillar collagens are in contrast to those of the protostome/deuterostome A-, B-, and C-clade N termini that are characterized by either vWF C domains (A clade) or a thrombospondin N-terminal and variable domain (B and C clade) (3, 26–28). The lack of homology in the N-propeptide of hydra fibrillar collagens in compar-
ison with homologous genes from higher order Metazoa raises the possibility that diploblasts diverged before the establishment of the A and B clades, which are found in both protostome and deuterostome lineages (28). However, the discovery of a short fibrillar collagen sequence in \textit{H. magnipapillata} that appears most similar to B-clade \(H9251\) chains (\(Hcol7\); see above and supplemental Fig. S1) raises the alternative possibility that the A and B clades of fibrillar collagens were already established in diploblasts and subsequently rearranged their N-propeptide domains as lineages diverged.

\textbf{vWF A Domains Have Combined with Collagenous Domains on Several Different Occasions}—The hydra \(Hcol6\) has an unusual domain organization consisting of a string of vWF A domains separated by interrupted Gly-X-Y repeats (Fig. 1). This combination of multiple vWF A domains and collage-
Hydra Extracellular Matrix Collagen and Evolution

Hydra Collagen Genes Are Expressed in Distinct Developmentally Controlled Patterns—Based on the model proposed by Sarras and Deutzmann (20), hydra mesoglea is organized into two sub-epithelial basal lamina zones and a central (acellular) interstitial matrix. This interstitial component occupies 80–90% thickness of the mesoglea and is composed of fibrous molecules (31). The three new hydra fibrillar collagens described here are candidates to participate in the formation of the interstitial component of the mesoglea. Using a monoclonal antibody, we have previously shown that Hcol1 is localized to 10-nm microfibrils in the interstitial component of the mesoglea (8), and more recent work shows it forms a grid-like fibrous network.5 The expression pattern of Hcol2 (Fig. 7) is very similar to that previously reported for Hcol1. The possibility that Hcol1 and Hcol2 could co-assemble in a heterotypic trimer or co-assemble into heterotypic fibrils must await further biochemical and immunocytochemical characterization of the mesoglea. It is of interest to note, however, that all of the collagenous peptide sequences generated from a biochemical characterization of collagen in the mesoglea of adult polyps originated from either the Hcol1 or Hcol4 genes (8). It is, therefore, possible that Hcol2 may have a lower relative abundance in the mesoglea of adults than Hcol1 despite its widespread and apparently strong expression throughout the polyp (Fig. 7).

The other two fibrillar collagens, Hcol3 and Hcol5, present different expression patterns from Hcol1 and Hcol2 but similar expression patterns to each other (Figs. 7–9). Both Hcol3 and Hcol5 are expressed more intensely on tentacles and buds, where mesoglea is thinner, and tissue extension/cell migration is undertaken and at much lower levels in the body column (Figs. 8 and 9). The restricted distribution of these two collagens perhaps explains why no evidence for their existence was revealed by the previous biochemical characterization of mesoglea collagens (8). It seems unlikely that Hcol3 and Hcol5 would co-trimerize since the collagenous domain of Hcol3 is significantly shorter than that of Hcol5 (Table 1).

During head regeneration, mesoglea is synthesized at the regenerating tip de novo (12). Hcol2 shows stronger expression than Hcol3 and Hcol5, indicating Hcol2 (together with Hcol1 (6)) is more likely used for the initial assembly of the mesoglea, with Hcol3 and Hcol5 participating later, perhaps as tentacle regeneration is initiated. The hydra fibrillar collagen genes characterized to date therefore seem to fall into two groups. One group, represented by Hcol1 and -2, is expressed throughout the polyp and is probably responsible for forming the main fibrous framework of the mesoglea. The other group, represented by Hcol3 and Hcol5, is probably less involved in the general process of mesoglea synthesis but more involved in modifying the characteristics of the mesoglea where it needs to be thickened such as at the base of the tentacles.

Hydra Collagens Are Expressed Exclusively by the Ectoderm—A remarkable feature of collagen synthesis in hydra is that the mRNAs for all six of the collagen genes characterized to date are expressed exclusively by the ectoderm (Figs. 7–10; Refs 7 and 8). This is in contrast to the expression of, for instance, laminin mRNAs, which are localized to the endoderm (9, 10). Although much has been learnt recently about the role of epithelial cell plasma membrane receptors and laminin in controlling the assembly of basement membranes (32), it will be of interest to examine how the assembly of the mesoglea is controlled given that the interstitial component is sandwiched between two basement membrane-like layers of matrix and the collagenous and non-collagenous components of the mesoglea are secreted from opposite sides.

There can be little doubt, given the level of sequence conservation between ECM genes in hydra and the rest of the animal kingdom, that common principles govern the assembly, structure, and function of ECMs in the metazoan kingdom. Invertebrate matrices offer alternative model systems that can help highlight and unravel these general principles. Examples such as the independent generation of collagens with multiple vWF A domains on different metazoan lineages serve to highlight general principles, such as the requirements for complex multidomain proteins capable of multiple interactions that under-

5 H. Shimizu, R. Aufschnaiter, D. B. Borza, M. P. Sarras, Jr., L. Li, D. R. Abrahamson, Y. Sado, and X. Zhang, manuscript in preparation.
lie the evolution, structure, and function of collagen-based ECMs.

Acknowledgments—We thank Eileen Roach for assistance in preparing the figures and J. Dry, P. Walpole, E. Hancox, H. Swainson, and C. Mahoney for generating significant sequence data as part of the Life Sciences Honours Degree Projects at the University of Manchester.

REFERENCES

1. Kielty, C. M., and Grant, M. E. (2002) in Connective Tissue and Its Heritable Disorders (Royce, P. M., Steinmann, B., eds) pp. 159–221, Wiley-Liss, Inc., New York

2. Prockop, D. J, and Kivirikko, K. I. (1995) Annu. Rev. Biochem. 64, 403–434

3. Boot-Handford, R. P., and Tuckwell, D. S. (2003) BioEssays 25, 142–151

4. Aouacheria, A., Geourjon, C., Aghajari, N., Navratil, V., Deleage, G., Lethias, C., and Exposito, J. Y. (2006) Mol. Biol. Evol. 23, 2288–2302

5. Nichols, S. A., Dirks, W., Pease, J. S., and King, N. (2006) Proc. Natl. Acad. Sci. U. S. A. 103, 12451–12456

6. Sarras, M. P., Jr., Madden, M. E., Zhang, X., Gunwar, S., Huff, J. K., and Hudson, B. G. (1993) Dev. Biol. 157, 383–398

7. Fowler, S. J., Jose, S., Zhang, X., Deutzmann, R., Sarras, M. P., Jr., and Boot-Handford, R. P. (2000) J. Biol. Chem. 275, 39589–39599

8. Deutzmann, R., Fowler, S., Zhang, X., Boone, K., Dexter, S., Boot-Handford, R. P., Rachel, R., and Sarras, M. P., Jr. (2000) Development 127, 4669–4680

9. Zhang, X., Fei, K., Agbas, A., Yan, L., Zhang, J., O’Reilly, B., Deutzmann R., and Sarras M. P., Jr. (2002) Dev. Genes Evol. 212, 159–172

10. 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

11. Sarras, M. P., Jr., Zhang, X., Huff, J. K., St. John, P. L., and Abrahamson, D. R. (1993) Dev. Biol. 157, 383–398

12. Shimizu, H., Zhang, X., Zhang, J., Leontovich, A., Fei, K., Yan, L., and Sarras, M. P., Jr. (2002) Development 129, 1521–1532

13. Lenhoff, H. M. (1983) in Hydra: Research Methods (Lenhoff, H. M., ed) pp. 29–34, Plenum Press, New York

14. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997) Nucleic Acids Res. 24, 4877–4882

15. Felsenstein, J. P. (1993) Phylogenetic Analysis Using Parsimony, Department of Genetics, University of Washington, Seattle

16. Strimmer, K., and von Haesler, A. (1996) Proc. Natl. Acad. Sci. U. S. A. 94, 6815–6819

17. Huelsenbeck, J. P. (2000) MrBayes: Bayesian Inference of Phylogeny, Department of Biology, University of Rochester, New York

18. Jackson, G. C., Barker, F. S., Jakkula, E., Czarny-Ratajcak, M., Makitei, O., Cole, W. G., Wright, M. J., Smithson, S. F., Suri, M., Rogala, P., Mortier, G. R., Baldock, C., Wallace, A., Elles, R., Ala-Kokko, L., and Briggs, M. D. (2004) J. Med. Genet. 41, 52–59

19. Lees, J. F., and Bullied, N. J. (1994) J. Biol. Chem. 269, 24354–24360

20. Sarras, M. P., Jr., and Deutzmann, R. (2001) BioEssays 23, 716–724

21. Mizuno, K., Adachi, E., Imamura, Y., Katsumata, O., and Hayashi, T. (2001) Micros 32, 317–323

22. Exposito, J. Y., D’Alessio, M., Solursh, M., and Ramirez, F. (1992) J. Biol. Chem. 267, 15559–15562

23. Hennighausen, L. G., and Sippel, A. E. (1982) Nucleic Acids Res. 10, 2677–2684

24. Vargas-Albores, F., Yepiz-Plascencia, G., Jimenez-Vega, F., and Avila-Villa, A. (2004) Comp. Biochem. Physiol. B 138, 415–422

25. Wood, R. L. (1961) The Biology of Hydra and Some Other Coelenterates (Lenhoff, H. M., and Loomis, W. F., eds) pp. 51–67, University of Miami Press, Coral Gables, FL

26. Sasaki, T., Fassler, R., and Hohenester, E. (2005) J. Cell Biol. 164, 959–963