Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 4655–4659, May 1993
Neurobiology

Molecular cloning of olfactomedin, an extracellular matrix protein specific to olfactory neuroepithelium

(chemoreception/olfaction/mucus/neuronal growth)

HIROKO YOKOE AND ROBERT R. H. ANHOLT*

Department of Neurobiology, Box 3209, Duke University Medical Center, Durham, NC 27710

Communicated by Morris E. Friedkin, March 4, 1993

ABSTRACT **The extracellular mucous matrix of olfactory neuroepithelium is a highly organized structure in intimate contact with chemosensory cilia that house the olfactory transduction machinery. Here we describe the molecular cloning and primary structure of olfactomedin, which is the major component of this extracellular matrix. Olfactomedin is expressed exclusively in olfactory neuroepithelium and its amino acid sequence shows no homologies to any known protein. This olfactory tissue-specific glycoprotein contains cysteines which form disulfide-linked polymers that constitute the primary architecture of the olfactory extracellular matrix. By analogy to other extracellular matrix proteins of the nervous system, olfactomedin may influence the maintenance, growth, or differentiation of chemosensory cilia on the apical dendrites of olfactory neurons.**

Vertebrate olfaction is mediated via dendritic cilia of olfactory receptor neurons (1–4). These neurons are continuously replaced throughout the lifespan of the animal from neurogenic stem cells in the base of the epithelium (5, 6). The axons of newly generated olfactory neurons establish functional connections with postsynaptic cells in the olfactory bulb of the brain while their dendrites grow toward the nasal lumen; when they reach the chemosensory surface they form a dendritic knob that carries a group of chemosensory cilia. The signals which trigger dendritic differentiation of developing olfactory neurons are unknown but most likely reside in the mucus which lines the chemosensory surface. Although the pathways of olfactory transduction have been documented extensively (1–4, 7), little is known about the molecular organization of the extracellular mucous matrix which coats the chemosensory surface.

Histochemical studies with lectins (8) and electron microscopic studies with antibodies (9) have demonstrated that the olfactory mucus is a highly structured compartment. The major glycoprotein of this extracellular matrix, olfactomedin, has been purified from *Rana catesbeiana* (10). This 57-kDa protein is found exclusively in olfactory tissue, where it is produced by Bowman's glands and sustentacular cells and deposited into the deep mucus layer. Furthermore, previous biochemical studies have shown that olfactomedin forms disulfide-linked polymers which adhere to and copurify with olfactory cilia (10, 11).

To gain insights into the molecular organization of the extracellular mucous matrix of olfactory neuroepithelium, we investigated the primary structure of olfactomedin. Here we report the molecular cloning of this olfactory glycoprotein and its complete amino acid sequence deduced from the nucleotide sequence of a corresponding cDNA clone. This sequence† shows no homologies to that of any known protein. Moreover, Northern blot analysis confirms that olfactomedin is expressed only in olfactory tissue. Analysis of olfactomedin's sequence, together with previous biochemical observations, supports a structural model which indicates that olfactomedin represents an olfactory tissue-specific extracellular matrix protein which may influence the growth and differentiation of chemosensory cilia that house the olfactory transduction machinery.

MATERIALS AND METHODS

Construction and Screening of an Olfactory cDNA Library from *R. catesbeiana*. Bullfrogs (*R. catesbeiana*) were purchased from Lemberger (Oshkosh, WI). RNA was extracted from olfactory tissue according to an acid guanidinium thiocyanate/phenol/chloroform extraction procedure (12), and mRNA was purified by oligo(dT) chromatography using PolyA Quick columns from Stratagene. A cDNA library was constructed in λZAPII, using Stratagene's ZAP-cDNA synthesis kit according to the manufacturer's protocol. The average size of inserts in the library was ≈1 kb. Olfactomedin was purified from olfactory membranes of *R. catesbeiana*, as described previously (10), and the amino acid sequence of its N terminus was determined by conventional protein sequencing. Based on this sequence, a degenerate oligonucleotide probe of the sequence 5′-GCTAC(A/C)GG(A/C)ATCCTIGCTGG(A/C)AAGGATCACGATGT(G/T)IIIGAAGATCT-3′ (where I is deoxyinosine) was synthesized and used to screen ≈4 × 10^5 independent recombinants. Hybridization was performed on nitrocellulose membranes overnight at 42°C in 6× standard saline citrate (SSC)/20% formamide/5× Denhardt's solution/0.1% SDS containing denatured salmon sperm DNA at 100 μg/ml. The membranes were washed at 42°C with 6× SSC/0.1% SDS for 1 hr. The sizes of inserts of positive clones were determined by PCR amplification, and two clones, one of 1.9 kb and the other of 2.0 kb, were sequenced. Both strands of the DNA were sequenced by using Sequenase version 2.0 (United States Biochemical) according to the dideoxynucleotide chain-termination method (13).

Analysis of the Sequence of Olfactomedin. The sequence of olfactomedin was analyzed by the GCG version 7 sequence-analysis package (Genetics Computer Group, Madison, WI). Homology searches through the GenBank, EMBL, and Swiss-Prot databases were performed according to the algorithm of Pearson and Lipman (14) using 50% amino acid identity in a window of 20 amino acids as a minimal criterion for homology. The Swiss-Prot database contained 80,958 nonredundant sequences when we analyzed olfactomedin's sequence.

Genomic Analysis. Genomic DNA was extracted from frog liver and 10-μg samples were digested with *Eco*RI or *Xho* I. The endonuclease-treated DNA was subjected to agarose gel

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

*To whom reprint requests should be addressed.

†The sequence reported in this paper has been deposited in the GenBank data base (accession no. L13595).

electrophoresis and blotted onto a nitrocellulose membrane. A 1.2-kb fragment corresponding to the coding region of olfactomedin was generated by PCR amplification, radiolabeled by random-primer synthesis, and used to identify hybridizing DNA fragments. Hybridization was performed overnight at 42°C in 6× SSC/50% formamide/5× Denhardt's solution/0.1% SDS containing denatured salmon sperm DNA at 100 μg/ml. The membrane was washed with 0.1× SSC/0.1% SDS for 1 hr at 60°C and then subjected to autoradiography.

Analysis of mRNA. RNA was extracted from frog tissues by an acid guanidinium thiocyanate/phenol/chloroform extraction procedure (12) and its quality was verified by assessing the integrity of the 2.0- and 4.5-kb bands of rRNA. RNA extracts were quantitated by measuring their absorbances at 260 nm and 10-μg samples were subjected to denaturing agarose gel electrophoresis. The RNA was immobilized onto a Hybond nylon membrane (Amersham) and hybridized to the same radiolabeled olfactomedin-encoding probe under the same conditions as described under *Genomic Analysis*.

RESULTS

Primary Structure of Olfactomedin. We extracted mRNA from olfactory epithelium of *R. catesbeiana* and constructed a cDNA library in λZAPII. Two clones were isolated and sequenced. Fig. 1 shows the nucleotide sequence and the deduced amino acid sequence of a 2.0-kb clone. It contains 93 bases of 5′ untranslated region, 1392 bases of coding region, and 517 bases of 3′ untranslated region, including a poly(A) tail (Fig. 1). The 464-amino acid open reading frame includes a 16-amino acid leader peptide with the initiating methionine. The mature protein, which consists of 448 amino acids, has a calculated molecular weight of 50,873, which is in close agreement with the experimentally observed size of deglycosylated olfactomedin (M_r 51,000) (15). Olfactomedin contains six potential N-linked glycosylation sites, evenly distributed throughout its sequence, and seven cysteines which can form disulfide bonds. Comparison of olfactomedin's amino acid sequence with all sequences stored in the GenBank, EMBL, and Swiss-Prot databases as of January 1993 did not reveal any significant homologies.

Expression of Olfactomedin. In agreement with the previously observed abundance of olfactomedin in olfactory tissue (10), we found 0.6% positive clones when we screened the olfactory cDNA library with a random-primer-labeled probe corresponding to olfactomedin's open reading frame. To investigate the number of genes which encode olfactomedin, we hybridized olfactomedin-encoding cDNA to *Eco*RI- or *Xho* I-digested genomic DNA from *R. catesbeiana*. The result reveals single hybridizing bands (Fig. 2). This, together with the observation that two independent clones of olfactomedin yielded the same amino acid sequence of the mature protein, indicates that olfactomedin is likely to be encoded by a single gene.

To examine expression of olfactomedin, we extracted total RNA from various tissues and performed Northern blot analysis. mRNA for olfactomedin is expressed in olfactory tissue, but not in lung, liver, heart, or brain (Fig. 3*A*), whereas mRNA for actin is readily detected in all of these tissues (Fig. 3*B*). This result agrees with the tissue specificity observed previously with antibodies (10, 11). The size of the mRNA which encodes olfactomedin is ≈2 kb, similar to that of the full-length cDNA.

Model for the Structure of Olfactomedin. Chou–Fasman analysis (16) predicts a predominantly α-helical structure of the N-terminal 150 amino acids of olfactomedin, mostly β-sheet configuration near its C-terminal third, and a region characterized by several turns in the center. This analysis, together with previous biochemical observations (10, 11, 15), supports a structural model in which cysteines at positions 160, 177, 283, and 290 form intramolecular disulfide bonds (Fig. 4 *Upper*). This generates a large central "head" domain consisting of a loop of 106 amino acids, a "neck" region consisting of a loop of 26 amino acids, and two "legs" of equal length. Olfactomedin forms disulfide-linked polymers which can be reduced to monomers (10, 11). However, even under reducing conditions, a relatively stable residual 120-kDa dimer is observed (10, 11), which may result from the formation of adjacent intermolecular disulfide bridges via cysteines at positions 16 and 18. This is in line with the observation that a site-directed antibody against a synthetic peptide corresponding to amino acids 7–16 can bind only to the unobstructed monomer (15). The remaining cysteine at position 377 on the C-terminal leg is proposed to form an intermolecular disulfide bond with its counterpart on a neighboring molecule to give rise to disulfide-linked polymers, as observed experimentally (Fig. 4 *Lower*) (10, 11).

DISCUSSION

Previously, we discovered in frog olfactory neuroepithelium an olfactory glycoprotein which we named "olfactomedin" (10). This 57-kDa glycoprotein was found to be unique to olfactory tissue, where it is produced in massive quantities by Bowman's glands and sustentacular cells and exported into the extracellular mucus which lines the chemosensory surface (10). Immunohistochemical studies at the electron microscopic level have shown that olfactomedin remains confined to the deep layer of the extracellular mucous matrix (9, 10), and biochemical studies have demonstrated that it copurifies with chemosensory cilia detached from the olfactory neuroepithelium (10, 11). Olfactomedin forms polymers which can be converted to monomers by treatment with reducing agents. In addition, a relatively stable 120-kDa dimeric form of olfactomedin is often observed (10, 11). Previous studies have estimated that olfactomedin represents up to 5% of total protein of frog olfactory tissue (10). The abundance of this tissue-specific protein, along with its distinct localization, which is confined to the immediate vicinity of the chemosensory surface (9, 10), suggests that disulfide-linked polymers of olfactomedin constitute the primary structural building blocks of the extracellular mucous matrix of olfactory neuroepithelium.

To isolate olfactomedin-encoding cDNAs we screened our library with a degenerate oligonucleotide probe which corresponded to the N-terminal amino acid sequence of electrophoretically pure olfactomedin (10) as determined by conventional protein sequencing. The notion that the clone which was identified and sequenced indeed encodes olfactomedin is consolidated by the following observations. (*i*) Immediately following the leader peptide the open reading frame of the nucleotide sequence of the cDNA contains a segment which encodes the amino acid sequence of the N terminus of olfactomedin previously identified by conventional protein sequencing. (*ii*) An antibody raised against a synthetic peptide which corresponds to this N-terminal sequence recognizes olfactomedin on Western blots and in tissue sections of frog olfactory epithelium (15). (*iii*) The calculated molecular weight of the protein encoded by our cDNA clone corresponds precisely to the molecular weight of the enzymatically deglycosylated protein identified by this antibody (15). (*iv*) The relative abundance of mRNAs that encode olfactomedin, as reflected by the large percentage of clones in the library that can hybridize to olfactomedin-encoding cDNA, is in line with the previously observed abundance of this protein in olfactory tissue (10). (*v*) The tissue-specific expression of mRNA identified with our cDNA probe correlates with previous immunochemical studies which revealed olfactomedin immunoreactivity only in olfactory tissue (10, 11). (*vi*) The structural model which we

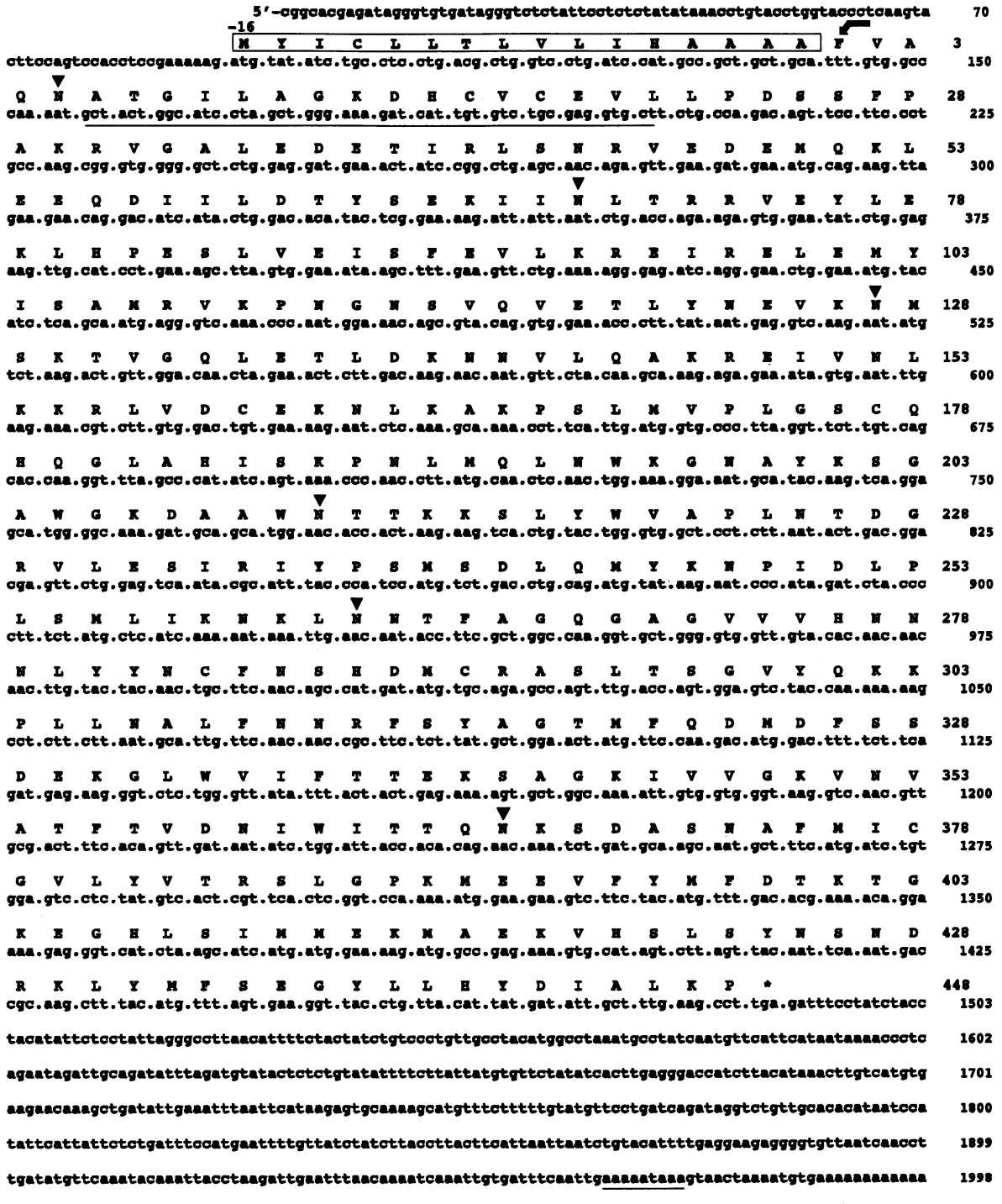

FIG. 1. Amino acid sequence of olfactomedin deduced from the nucleotide sequence of a cDNA clone. The N terminus of the mature protein was identified by comparing the translated open reading frame with 21 amino acids of the N-terminal sequence of purified olfactomedin obtained by conventional protein sequencing. The N-terminal amino acid of mature olfactomedin is the phenylalanine indicated by the arrow. It is preceded by a 16-amino acid leader peptide, indicated by the boxed area, which contains the initiating methionine. Olfactomedin contains six potential N-linked glycosylation sites, indicated by arrowheads. The nucleotide sequence which corresponds to the degenerate oligonucleotide probe initially used to screen the cDNA library is underlined. Both cysteine at position 16 and valine at position 20 were initially misidentified as aspartic acid by conventional protein sequencing, whereas position 18 could not be determined. A characteristic polyadenylylation signal near the 3' end is also underlined. Amino acids are indicated by single-letter code. The star indicates the termination codon.

propose for olfactomedin based on the nucleotide sequence of our cDNA clone can account for the previously documented glycosylation, polymerization, and characteristic immunoreactivity patterns of olfactomedin (10, 11, 15).

We deduced the complete amino acid sequence of olfactomedin from the nucleotide sequence of its corresponding cDNA. Olfactomedin is synthesized with a characteristic hydrophobic leader peptide and contains six consensus N-linked glycosylation sites. Identification of partially glycosylated and nonglycosylated precursors of olfactomedin with a site-directed antibody and enzymatic deglycosylation studies have indicated that most, if not all, of these potential

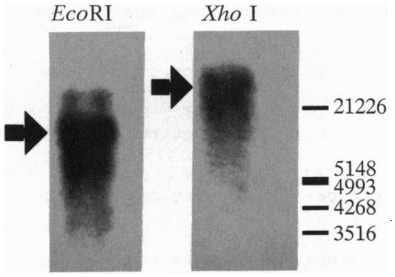

FIG. 2. Hybridization of olfactomedin-encoding cDNA to *Eco*RI- or *Xho* I-digested frog genomic DNA. Fragments of *Eco*RI/*Hin*dIII-digested phage λ DNA were used as standards (sizes in base pairs at right).

glycosylation sites indeed carry carbohydrate moieties consisting of 8–10 sugar residues characterized by *N*-acetylglucosamine and terminal galactose (15). Comparison of olfactomedin's sequence with sequences stored in the GenBank, EMBL, and Swiss-Prot databases revealed no homologies with any known protein. In addition, Northern blot analysis confirmed the tissue-specific expression of olfactomedin previously observed in immunochemical studies (10, 11). The tissue-specific production of abundant quantities of this glycoprotein with a striking localization in the immediate vicinity of the surface of the chemosensory neuroepithelium implies an important function for olfactomedin in olfaction.

Analysis of olfactomedin's sequence according to Chou–Fasman predictions of secondary structure (16), together with evidence from previous biochemical studies, has led us to propose the structural model presented in Fig. 4. This model is consistent with all currently available biochemical evidence: (*i*) the formation of disulfide-linked polymers and the occurrence of a relatively stable dimer (10, 11); (*ii*) the observation that a site-directed antibody against a synthetic peptide corresponding to amino acids 7–16 only has access to the monomeric form of olfactomedin (15); and (*iii*) the identification of a monoclonal antibody that recognizes a carbohydrate determinant and reacts with dimers and polymers, but not with monomers (10); this supports the symmetry of apposed carbohydrate attachment sites which is a characteristic feature of our proposed structural model (Fig. 4) and which allows the formation of antigenic determinants

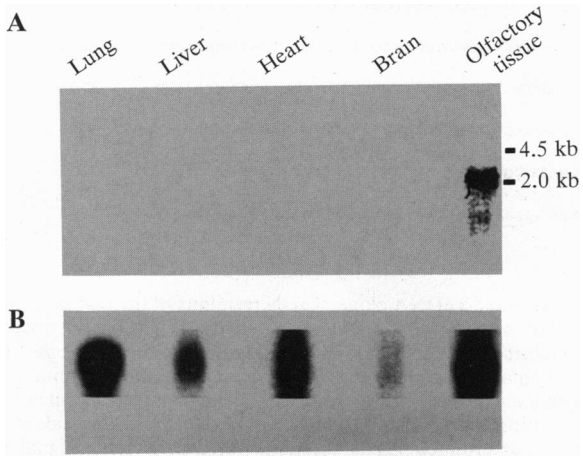

FIG. 3. Identification of mRNA encoding olfactomedin. (*A*) Tissue-specific expression of olfactomedin. mRNA slightly larger than 2.0 kb which hybridizes to an olfactomedin probe is observed only in olfactory tissue. The positions of the 2.0- and 4.5-kb bands of rRNA are indicated. (*B*) Presence of mRNA for actin in all five tissues. Note that under our experimental conditions, mRNA for actin in adult frog brain appears to be less abundant than in the other tissues.

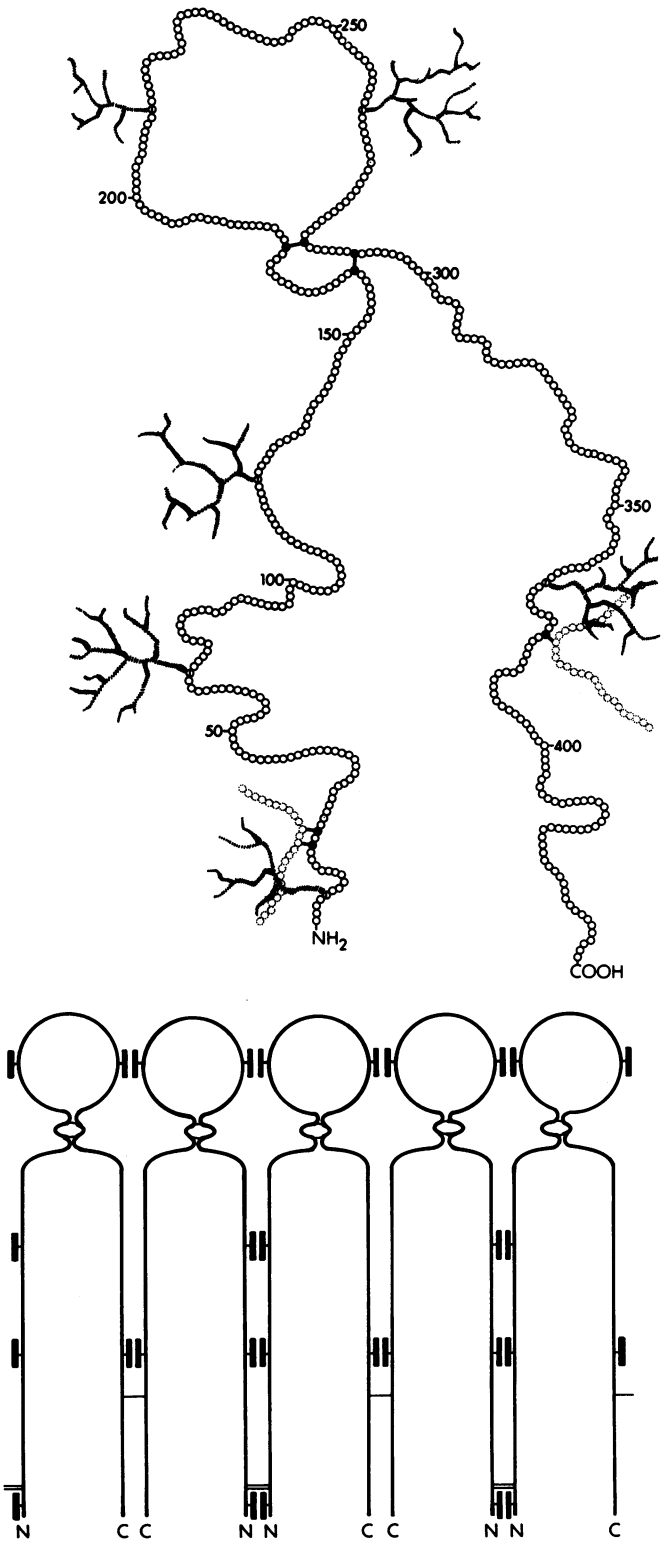

FIG. 4. Diagrammatic representation of olfactomedin and its polymers. (*Upper*) Model for the structure of olfactomedin. Each amino acid is represented by a circle. Cysteines are represented by solid circles and attachment sites for carbohydrates are indicated at positions 5, 69, 137, 212, 263, and 367. Olfactomedin contains truncated complex carbohydrates with terminal galactose residues devoid of sialic acid (11, 15). Parts of neighboring olfactomedin molecules are depicted by dotted lines. (*Lower*) How olfactomedin may generate long disulfide-linked polymers which represent the building blocks for the olfactory extracellular matrix. Filled rectangles indicate carbohydrate moieties.

through interactions between carbohydrate groups present on adjacent polypeptide chains.

Elucidation of the primary structure of olfactomedin establishes this protein as a member of a diverse group of extracellular matrix proteins which includes the neural cell adhesion molecule (CAM), laminin, fibronectin, and proteoglycans (17, 18). The monomer of olfactomedin is smaller than any of these proteins, but olfactomedin nevertheless can generate an extracellular matrix through the formation of polymers. The size of its proposed intramolecular loop and the conformational importance of its disulfide bonds are reminiscent of the immunoglobulin superfamily, which includes cell adhesion molecules such as N-CAM (19). The tissue-specific expression of olfactomedin contrasts with the widespread occurrence of other extracellular matrix proteins. N-CAM, laminin, fibronectin, and proteoglycans all have been implicated in adhesion, maintenance, migration, growth, or differentiation of neurons (17, 18). Olfactomedin makes close contact with the highly specialized chemosensory surface of the olfactory neuroepithelium and, by analogy, may contribute to the maintenance, growth, or differentiation of olfactory cilia.

We thank Dr. James R. Burke and Dr. Gerald D. Fasman for helpful advice and valuable discussions, Alison Steube and Cristina Davis for technical assistance, and Drs. William Matthew and Trudy Mackay for critical reading of the manuscript. This work was supported by grants from the National Institutes of Health (DC00394) and the U.S. Army Research Office (DAAL03-89-K-0178 and DAAL03-92-G-0390).

1. Anholt, R. R. H. (1991) *Chem. Senses* **16**, 421–427.
2. Bakalyar, H. A. & Reed, R. R. (1992) *Curr. Opin. Neurobiol.* **1**, 204–208.
3. Breer, H. & Boekhoff, I. (1992) *Curr. Opin. Neurobiol.* **2**, 439–443.
4. Firestein, S. (1992) *Curr. Opin. Neurobiol.* **2**, 444–448.
5. Graziadei, P. P. C. & Monti-Graziadei, G. A. (1978) in *Handbook of Sensory Physiology: Development of Sensory Systems*, ed. Jacobson, M. (Springer, Berlin), pp. 55–83.
6. Morrison, E. E. & Costanzo, R. N. (1992) in *Science of Olfaction*, eds. Serby, M. & Chobor, K. L. (Springer, New York), pp. 31–50.
7. Buck, L. & Axel, R. (1991) *Cell* **65**, 175–187.
8. Foster, J. D., Getchell, M. L. & Getchell, T. V. (1992) *Cell Tissue Res.* **267**, 113–124.
9. Menco, B. P. M. & Farbman, A. I. (1992) *Cell Tissue Res.* **270**, 47–56.
10. Snyder, D. A., Rivers, A. M., Yokoe, H., Menco, B. P. M. & Anholt, R. R. H. (1991) *Biochemistry* **30**, 9143–9153.
11. Anholt, R. R. H., Petro, A. E. & Rivers, A. M. (1990) *Biochemistry* **29**, 3366–3373.
12. Chomczynski, P. & Sacchi, N. (1987) *Anal. Biochem.* **162**, 156–159.
13. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
14. Pearson, W. R. & Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2444–2448.
15. Bal, R. S. & Anholt, R. R. H. (1993) *Biochemistry* **32**, 1047–1053.
16. Chou, P. Y. & Fasman, G. D. (1978) *Adv. Enzymol.* **47**, 45–148.
17. Sanes, J. R. (1989) *Annu. Rev. Neurosci.* **12**, 491–516.
18. Reichardt, L. F. & Tomaselli, K. J. (1991) *Annu. Rev. Neurosci.* **14**, 531–571.
19. Cunningham, B. A., Hemperly, J. J., Murray, B. A., Prediger, E. A., Brackenburg, R. & Edelman, G. M. (1987) *Science* **236**, 799–806.