Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 7617–7621, October 1988
Developmental Biology

Structure of the gene for the liver cell adhesion molecule, L-CAM

(calcium-dependent adhesion/precursor processing/intron/exon junctions)

BARBARA C. SORKIN, JOHN J. HEMPERLY*, GERALD M. EDELMAN, AND BRUCE A. CUNNINGHAM

Department of Molecular and Developmental Biology, The Rockefeller University, 1230 York Avenue, New York, NY 10021

Contributed by Gerald M. Edelman, July 5, 1988

ABSTRACT **The liver cell adhesion molecule, L-CAM, mediates calcium-dependent cell–cell adhesion in early embryos and in nonneural epithelia in adult tissues. Earlier studies of cDNAs for chicken L-CAM established the amino acid sequence of the mature protein. The sequence has now been extended in the 5′ direction through the precursor and signal sequences and past a consensus translation initiation site. The combined cDNAs were used to isolate genomic clones covering the entire L-CAM coding sequence. The structural gene for chicken L-CAM contains 16 exons ranging in size from 115 to over 1045 base pairs with an average size of 222 base pairs. Single exons do not correspond to known structural elements such as the signal sequence, precursor segment, internal repeats, or membrane-spanning region of L-CAM. Hybridization of restriction digests of chicken genomic DNA with cDNA and genomic probes indicated that there is a single L-CAM gene in the chicken. In contrast to genes for other cell–cell or cell–substrate adhesion molecules, there is no evidence for alternative splicing of exons in this gene.**

Cell adhesion is of fundamental importance in establishing and maintaining tissue form and function (1). A number of molecules involved in direct cell–cell adhesion have now been isolated. Among the best characterized of these are the neural cell adhesion molecule, N-CAM, and the liver cell adhesion molecule, L-CAM. Both are primary CAMs that appear in the earliest embryonic cells and act by homophilic mechanisms (a CAM on one cell binds to the same CAM on another). N-CAM is distinguished by its content of polysialic acid and its resemblance to members of the immunoglobulin superfamily (2), whereas L-CAM is the exemplar of a different group of closely related glycoproteins that mediate calcium-dependent cell–cell adhesion. Structural (2, 3) and functional (4–6) studies indicate that the two kinds of CAMs are evolutionarily unrelated.

L-CAM was initially purified from embryonic chicken liver (7) but is found in most nonneural epithelial cells (8). This molecule is expressed on the cell surface as a glycoprotein of M_r 124,000 (3, 9). Proteins with tissue distributions, biochemical properties, and amino acid sequences similar to those of L-CAM have been detected in other species, including mouse [uvomorulin (10), E-cadherin (11)]; dog [Arc-1 (12); rrl (13)], human (14), and *Xenopus laevis* (15). Uvomorulin (16) and E-cadherin (17) are identical to each other in amino acid sequence; they are 65% identical to L-CAM overall and close to 90% identical in the cytoplasmic domain, supporting the notion that they are murine equivalents of L-CAM.

Transfection of LMTK$^-$ cells with L-CAM cDNA clones (6, 17) leads to expression of L-CAM at the cell surface and confers on the cells the property of calcium-dependent aggregation, supporting the notion that L-CAM acts by a homophilic mechanism. Perturbation experiments using specific antibodies have illustrated the biological importance of L-CAM-mediated adhesion. Addition of antibodies to L-CAM to cultures of chicken skin alters the response to inductive signals that is necessary for feather development (5). In addition, antibodies to uvomorulin prevent the compaction of mouse embryos (18), and antibodies to rrl block the formation of tight junctions in MDCK cells (13). Moreover, the coordinate appearance of L-CAM and N-CAM in adjacent populations of cells at sites of embryonic induction as well as in later inductive events (4, 8, 19) suggests that the regulated expression of both CAMs plays an important role in these processes (1).

L-CAM and its mammalian homologues appear to be part of a larger family of calcium-dependent adhesion molecules. Several similar calcium-dependent CAMs have been described recently: P-cadherin was first detected in mouse placenta (20), N-cadherin was initially found in chicken brain (21), and A-CAM was isolated from adherens junctions of chicken cardiac cells (22). A-CAM closely resembles N-cadherin, and all three molecules have striking similarities to L-CAM and its homologues. L-CAM and the cadherins are initially synthesized as larger precursors that are processed to mature proteins of 120–130 kDa. Each molecule contains three to five repeats (110–115 amino acids each) of similar amino acid sequence (3, 16, 17, 23, 24). Overall, the amino acid sequences of mouse L-CAM and P-cadherin are 58% identical, and chicken L-CAM and N-cadherin are 50% identical, establishing the relationship among these proteins. Despite their similarities, L-CAM, N-cadherin, and P-cadherin are not immunologically cross-reactive and have distinct adhesive specificities and tissue distributions.

Although L-CAM mRNA and protein are detected as species of single sizes in all tissues in which they are expressed, Southern blots with a small cDNA probe (25) raised the possibility that in chickens there may be more than one L-CAM gene. Other studies have suggested that alternative RNA splicing is important in regulating the functions of some CAMs. For example, the expression of N-CAM involves alternative splicing of RNAs transcribed from a single gene (2, 26) to give at least five different polypeptides, and other molecules with cell adhesion functions such as cytotactin (27) and fibronectin (28) are present in variant forms produced by alternative splicing. To assess the number of L-CAM genes, to look for alternative L-CAM exons, and to provide a foundation for studying the regulation of L-CAM expression, we have examined the structure of the L-CAM gene in detail. Overlapping genomic clones that encompass the entire L-CAM structural gene were isolated. This gene contains 16 exons, all of which are present in all known chicken L-CAM cDNAs. The structure of the gene† and Southern blot analyses explain the results obtained in the

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.

Abbreviations: L-CAM, liver cell adhesion molecule; N-CAM, neural cell adhesion molecule.

*Present address: Becton Dickinson and Company, Research Center, P.O. Box 12016, Research Triangle Park, NC 27709-2016.

†The sequence reported in this paper is being deposited in the EMBL/GenBank data base (IntelliGenetics, Mountain View, CA, and Eur. Mol. Biol. Lab., Heidelberg) (accession no. J04074).

earlier studies (25) and indicate that there is a single chicken L-CAM gene.

MATERIALS AND METHODS

Chicken genomic DNA was purchased from Clontech Laboratories (Palo Alto, CA) or prepared as described (29). cDNA was synthesized from poly(A)$^+$ RNA from 9-day chicken embryo livers (29).

Screening of libraries, digestion of DNA, gel electrophoresis of DNA or RNA, transfer of gels to nitrocellulose, and hybridization were performed as described (29, 30). For use as probes, DNA fragments or plasmids were radiolabeled (31). DNA was subcloned into pBluescript KS vectors, deletions were performed as described (32), and sequencing was performed with either reverse transcriptase or Sequenase (United States Biochemical, Cleveland, OH) (33).

RESULTS AND DISCUSSION

Processing of Chicken L-CAM. In accord with earlier results (3, 9, 16, 17, 34), pulse–chase experiments showed that chicken L-CAM is synthesized as a larger precursor that was processed to the mature form with a half-time of 20–30 min (Fig. 1*a*). Processing persisted in the presence of tunicamycin (Fig. 1*b*), indicating that it is not dependent on the presence of N-linked carbohydrate. It does involve carbohydrate processing: all of the N-linked carbohydrate in the precursor is removed by endoglycosidase H (Fig. 1*c*; ref. 37), whereas only one of the four N-linked carbohydrates in the mature L-CAM is of the high mannose type (9). Maturation of L-CAM also involves an increase in phosphorylation. Gel slices containing the precursor or processed forms were obtained from a hepatocyte culture double-labeled with $^{32}PO_4$ and [^{3}H]leucine. The ratio of ^{32}P to ^{3}H in the precursor was half that in the mature form (1000 cpm of ^{32}P/4000 cpm of ^{3}H vs. 25,000 cpm of ^{32}P/50,000 cpm of ^{3}H), indicating that the precursor is phosphorylated but that additional sites are phosphorylated in the mature molecule, which is known to contain phosphothreonine and phosphoserine (9).

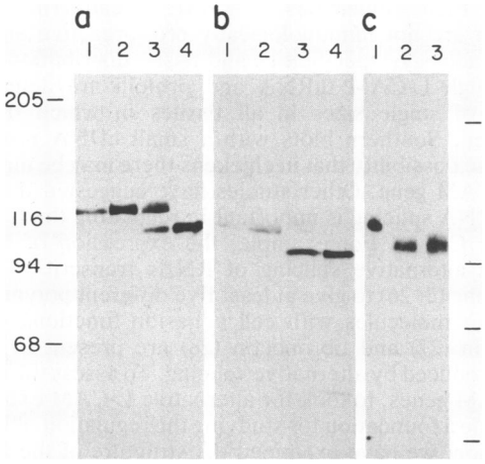

FIG. 1. Processing of L-CAM. (*a* and *b*) Hepatocytes from 10-day chicken embryos were pulse-labeled with [^{35}S]methionine for 10 min, chased, and extracted, and extract components were immunoprecipitated (9); components were resolved on 6% or 7% acrylamide gels (35) and detected by fluorography. Migration of the L-CAM bands varies with the acrylamide concentration in the gel. Cultures were labeled in the absence (*a*) or presence (*b*) of tunicamycin. Lanes: 1, no chase; 2, 10-min chase; 3, 30-min chase; 4, 90-min chase. (*c*) Cultures were labeled for 30 min and treated with; no enzyme (lane 1), 90 milliunits of endoglycosidase F (36) for 14 hr at 37°C (lane 2), or 0.25 milliunit of endoglycosidase H for 14 hr (9) (lane 3). Migrations of standard proteins (M_r/1000) are indicated on the left. Marks on the right are the same standards for *c*.

Isolation of cDNA Clones Encoding the Signal and Precursor Sequences of L-CAM. The previously reported L-CAM cDNA clone pEC320 (3) extended through the amino and carboxyl termini of the mature L-CAM polypeptide but did not encode sufficient polypeptide to account for all of the precursor and lacked a translational start site and a signal sequence. To extend the L-CAM cDNAs in the 5′ direction, a cDNA library was prepared by using an oligonucleotide primer complementary to nucleotides 100–118 of pEC320. From this library two new clones, pEC330 and pEC331 (Fig. 2), were isolated. These clones overlapped the 5′ end of pEC320, continued upstream to a termination codon (nucleotides 45–47), and included a sequence (nucleotides 49–54) that matches the consensus sequence for initiation of translation in eukaryotic mRNAs (38). As in most vertebrate mRNAs (38), there were no additional ATG sequences in the 5′ untranslated region of the cDNA. The 75 base pairs (bp) following the translational start site encode a typical signal sequence (39), which should be cleaved (40) after residue −135 (glycine). The 402-bp segment that follows is sufficient to account for the difference in molecular mass of the mature polypeptide and the precursor (Figs. 1 and 2).

The presumptive signal and precursor polypeptide sequences are 42% identical to those of murine L-CAM (16, 17). P-cadherin (23) and N-cadherin (24) are also synthesized as precursor polypeptides; the precursor segments of L-CAM

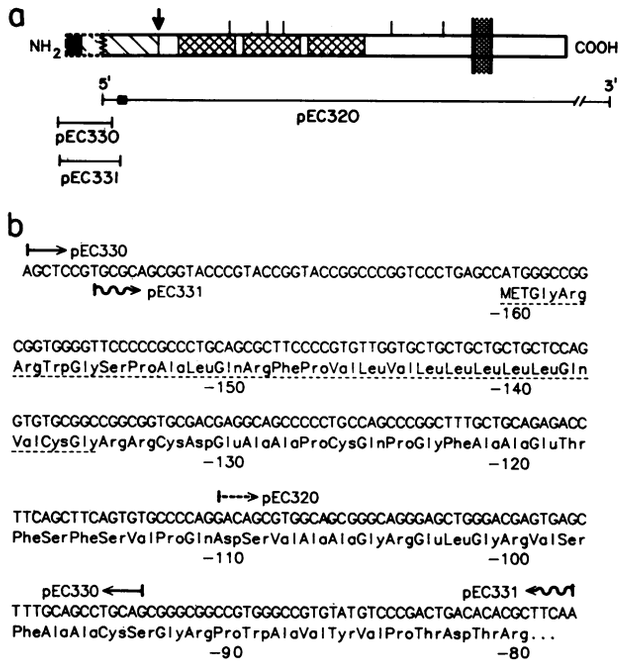

FIG. 2. Extended sequence of the L-CAM cDNA. (*a*) Synthesis of pEC330 and pEC331 was primed with an oligonucleotide (small black box) complementary to L-CAM cDNA clone pEC320. The new sequences (dashed outline) encode the signal peptide (black box) and part of the precursor (hatched box). pEC320 encoded the repeating units of similar amino acid sequence (crosshatched boxes), transmembrane segment (stippled area), and the 3′ untranslated region of the mRNA. The vertical arrow indicates the amino terminus of the mature form of L-CAM. Vertical lines above the protein are potential glycosylation sites. (*b*) Sequences of pEC330 and pEC331. The ends of pEC330 are indicated by straight arrows; the ends of pEC331, by wavy arrows; and the 5′ end of pEC320, by a dashed arrow. Amino acid numbering is based on the aspartic acid at the amino terminus of the mature L-CAM polypeptide as 1. The presumed translational start site is at −160, and the presumptive signal peptide is indicated by the dashed underline.

and P- and N-cadherin all end in the sequence Leu-(Arg/Lys)-Arg-(Gln/Arg)-Lys-Arg, and the mature proteins all have the amino-terminal sequence (Asp/Glu)-Trp-Val-(Ile/Met)-Pro-Pro-Ile, suggesting that proteases of similar specificities are involved in their processing.

Isolation and Analysis of L-CAM Genomic Clones. A chicken genomic library (41) was screened by hybridization with cDNA probes derived from pEC320 (3). Three overlapping clones were independently isolated (Fig. 3): λcL1 contains most of the L-CAM gene, λcL2 includes the region encoding the amino terminus of the mature L-CAM, and λcL3 contains 3′ untranslated sequences of pEC320. Large parts of λcL1, λcL2, and λcL3 were sequenced, providing the complete sequences of all exons and of all but the largest intron. The sequence begins at the 5′ end of the λcL1 insert and extends 3′ through the regions encoding the L-CAM protein and the 3′ end of pEC320. λcL3 has a polyadenylylation consensus sequence (AATAAA) in exon 16, 13 bp upstream from the region where pEC320 contains a poly(A) tract, suggesting that this is the site recognized in the initial RNA transcript.

To confirm that the map derived from the sequences of λcL1, λcL2, and λcL3 is consistent with the fragments detected by hybridization of L-CAM probes with genomic DNA, digests of chicken DNA were hybridized with the cDNA probes pEC331, pEC301, or pEC325 or with genomic probe A (Figs. 3 and 4). As expected from the map, pEC331, pEC301, and probe A all hybridize with the same *Bam*HI fragment, whereas pEC325 hybridizes with a smaller *Bam*HI fragment. In the *Eco*RI digests, pEC331 and probe A detect a fragment of approximately 12 kilobases (kb), while pEC301 and pEC325 hybridize with a slightly smaller fragment; pEC331, which also overlaps this fragment by 64 bp, hybridizes with it weakly. In the *Pst* I digests, each probe hybridized with a different small (<3 kb) fragment or fragments as expected from the large number of *Pst* I sites in the gene. In the *Sst* I digests, pEC331 and probe A both

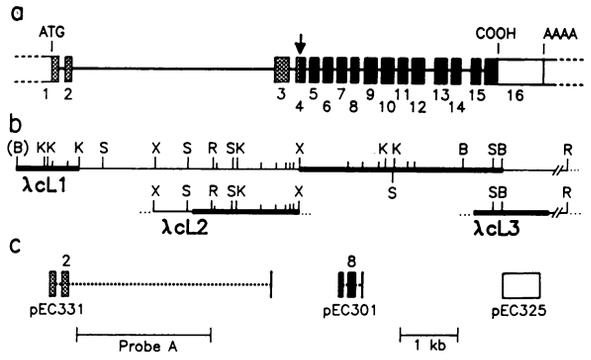

FIG. 3. Structure of the L-CAM gene. (*a*) Diagrammatic representation of the intron/exon structure of the L-CAM gene. Exons are represented by vertical bars and introns by connecting lines. The presumptive translational start site (ATG), the amino terminus of the mature L-CAM polypeptide (arrow), the carboxyl terminus (COOH), and the site of poly(A) addition (AAAA) are indicated. The region indicated by stippling, including all of exons 2 and 3 and portions of exons 1 and 4, encodes peptide sequences present only in the precursor form of L-CAM. (*b*) Partial restriction maps of genomic clones λcL1, λcL2, and λcL3. All three phage inserts extend beyond the regions shown. Thick lines represent segments that were sequenced. Sites for cleavage by *Bam*HI (B), *Eco*RI (R), *Kpn* I (K), *Xho* I (X), *Sst* I (S), and *Pst* I (short vertical bars) are indicated. That portion of the large intron that was not sequenced was not analyzed for *Pst* I sites. The *Bam*HI site at the 5′ end of λcL1 may result from the fusion of an *Mbo* I site in the gene with the *Bam*HI site of the vector. (*c*) Probes used for genomic DNA blots: cDNA probes are shown as boxes (exons) connected by dotted lines. The scale for the entire figure is indicated at lower right. kb, kilobases.

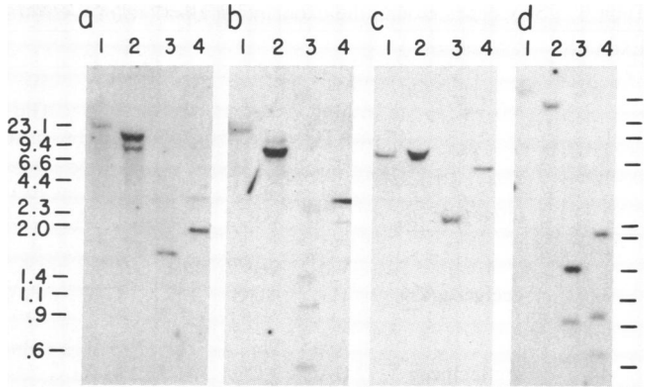

FIG. 4. Hybridization analysis of chicken genomic DNA. Chicken DNA (7.5 μg per lane) was digested with *Bam*HI (lane 1), *Eco*RI (lane 2), *Pst* I (lane 3), or *Sst* I (lane 4); resolved on 0.7% agarose gels; transferred to nitrocellulose; and hybridized with one of the ^{32}P-labeled probes shown in Fig. 3. (*a*) pEC331. (*b*) pEC301. (*c*) pEC325. (*d*) genomic probe A. Migrations of *Hin*dIII λ and *Hae* III φX174 fragments of indicated size (kb) are shown; marks on the right are the same standards for *d*.

hybridized with a 1.7-kb fragment that presumably extends through exons 1 and 2 to the first *Sst* I site in the large intron. Genomic probe A also hybridized to 0.9- and 0.7-kb fragments that are completely within the intron and do not hybridize to the cDNA probes. pEC331 and pEC301 both detect a *Sst* I fragment of slightly less than 3 kb, although pEC331, which overlaps this fragment by only 64 nucleotides, hybridizes weakly; pEC325 detects a single fragment that should extend 3′ from the 3′-most *Sst* I site mapped.

In more than 3 kb of cDNA sequence, nine nucleotide differences were detected between earlier results (3) and those reported here. Only one of these results in a change in the amino acid sequence; nucleotide 278 (cytidine) in pEC320 is a thymidine in the genomic sequence, substituting methionine for threonine. The other differences are: guanosine at position 21 of pEC320 is an adenosine in pEC330, pEC331, and λcL1; nucleotides 282, 468, and 477 (all guanosine residues) in pEC320 are adenosine residues in the genomic clones; the cytidine at position 1605 of pEC320 is a thymidine in λcL1, and there are two single nucleotide replacements and one extra base pair in the 3′ untranslated region of pEC320. These differences may be transcriptional errors arising during construction of the cDNAs or natural variant L-CAM sequences present in the population.

As expected, the structural genes for L-CAM and N-CAM are radically different. The chicken N-CAM gene extends over more than 50 kb (26) and includes sites for alternative RNA splicing (2). The L-CAM gene extends less than 10 kb. It includes 16 exons (Fig. 3; Table 1) that range in size from 115 to more than 1000 bp and encode 23–81 amino acids. All 15 introns have consensus splice sequences (Table 1; ref. 42). The introns range in size from 65 bp to approximately 3.5 kb, but 14 have fewer than 210 bp. Fragment A of λcL1 (*Kpn* I–*Eco*RI), which includes most of the 3.5-kb intron, did not hybridize with any chicken liver or brain poly(A)$^+$ RNA species (data not shown), and no consensus junction sequences were detected within the introns sequenced. These results and the detection of a single-size mRNA in all L-CAM-expressing tissues by all L-CAM cDNA probes suggest that all exons are included in our cDNA sequences and that additional L-CAM mRNA sequences will not be found.

In initial studies (25), pEC301 reacted with multiple fragments in genomic blots, suggesting that there might be more than one L-CAM gene. The genomic map indicates that this probe spans portions of three exons, and additional studies

Table 1. Sequences around the exon/intron junctions of L-CAM

Exon no.	Acceptor	Exon	Donor	Exon size, bp	Intron size, bp
		-160 -138			
		...Met Gly Arg ... Leu Gln			
1		...ATG GGC CGG... CTC CAG	gtgcggggctgc	>119 (23)	
		-137 -100			75
		Val Cys Gly ... Gly Arg			
2	tctctcgggcag	GTG TGC GGC... GGA CGA G	gtaggacccgtg	115 (38)	
		-99 -23			≈3500
		Val Ser Phe ... Leu Gln			
3	ccctgcccgcag	TG AGC TTT... CTG CAG	gtgaagtcccca	230 (77)	
		-22 23			111
		Asp Thr Thr ... Val Gln			
4	ctgttcctgcag	GAC ACG ACA... GTG CAG	gtacggggctga	135 (45)	
		24 75			105
		Ile Lys Ser ... Tyr Thr			
5	gcatcgccccag	ATC AAA TCC... TAC ACC	gtgagtgggggc	156 (52)	
		76 123			65
		Leu Leu Ser ... Lys Pro			
6	tgccccccgcag	CTC TTA TCC... AAG CCA G	gtgcggggxtgg	145 (48)	
		124 182			105
		Gly Thr Ser ... Arg Glu			
7	cgctctgcccag	GC ACC TCC... CGG GAG	gtgggcgctctc	176 (59)	
		183 225			110
		Thr Thr Pro ... Thr Met			
8	ctcatcctgcag	ACX ACT CCC... ACC ATG	gtatgtgttccc	129 (43)	
		226 287			99
		Tyr Glu Gly ... Ala Lys			
9	ccctctctgcag	TAC GAA GGT... GCC AAG	gtgggtcatgtc	186 (62)	
		288 368			89
		Gly Leu Asp ... Lys Ile			
10	cccaccctgcag	GGC CTG GAT... AAG ATC AC	gtgagtgggggc	245 (81)	
		369 416			77
		Thr Tyr Arg ... Asp Asn			
11	ccctccgcccag	G TAC CGC... GAC AAT G	gtgaggctgtgc	143 (48)	
		417 491			80
		Gly Ile Pro ... Gly Gln			
12	ctggctttgtag	GG ATA CCG... GGC CAA G	gtaggtgctgtg	225 (75)	
		492 565			203
		Asp Glu Leu ... Leu Leu			
13	cccttccacaag	AT GAG CTG... CTG CTG A	gtgagtacggtg	222 (74)	
		566 609			77
		Ile Leu Leu ... Asp Gln			
14	cctctgctgcag	TC CTG CTG... GAC CAG	gtgggagcgggg	131 (44)	
		610 657			207
		Asp Tyr Asp ... Asp Glu			
15	atccccacgcag	GAC TAC GAC... GAC GAG	gtgagccccaac	144 (48)	
		658			111
		Asn Leu Lys			
16	ccccccccccag	AAC CTG AAG ...TCAAAAGAATAAATTGT-		>1000 (69)	
		AAGAAAACCAGCAGCCCTG			
		↑			

Exon sequences are in uppercase letters and introns in lowercase letters. Exon and intron sizes are given in base pairs (bp) and the number of amino acids encoded by each exon is in parentheses. The consensus sequences AG and GT are underlined as is the polyadenylylation signal. The arrow in exon 16 indicates the site of poly(A) addition.

indicate that some of the fragments detected previously resulted from partial digestion. Repeated probings of digests of the same DNA used in those experiments and of commercially available chicken genomic DNA yielded the patterns expected from our analyses of L-CAM genomic clones with no additional fragments (Fig. 4), indicating that there is only one L-CAM gene in chicken.

Although the probes used here do not cross-react at high stringency with other genes, we occasionally detected clones at lower hybridization stringency that may correspond to genes for other calcium-dependent CAMs. Our results are similar to those obtained in mouse (43), where a single L-CAM (uvomorulin) gene was mapped to chromosome 8 and a cDNA probe used in the *in situ* hybridization experiments had weak cross-reactivity with chromosome 10.

In accord with recent analyses of the genes for other proteins (44), the known structural elements of chicken L-CAM are not specified by single exons (Fig. 5). The signal and precursor sequences, membrane-spanning region and the three internal repeats of similar amino acid sequences (3) are each specified by at least two exons. Moreover, the intron/exon boundaries do not apparently coincide with the borders of any of these structural elements.

The expression of similar calcium-dependent adhesion molecules in disparate tissues, often at times correlated with inductive events (4, 8, 19, 23), and the dramatic results of

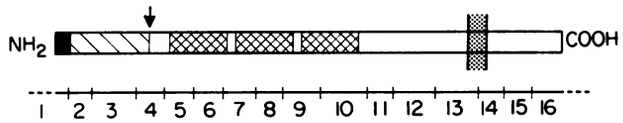

FIG. 5. Schematic representation of the L-CAM protein, showing the locations of exon boundaries superimposed on known features of the polypeptide. Junctions are indicated by vertical lines, and the exons are numbered. Cleavage to generate the mature L-CAM polypeptide occurs at the site marked with the arrow. The presumptive signal peptide (black box), the precursor polypeptide (hatched box), the repeating units of sequence similarity (cross-hatched boxes), and the transmembrane segment (stippled area) are all encoded by multiple exons. Exon 2 encodes three amino acids of the signal peptide.

antibody perturbation experiments (5, 13, 18, 45) suggest that the calcium-dependent CAMs have essential but distinct functions in forming different epithelia. This analysis of the structural gene for L-CAM lays the groundwork for studies of the genetic regulatory elements required for the temporal and tissue-specific transcriptional regulation of L-CAM expression during embryogenesis.

We thank Dr. Hidesaburo Hanafusa for the use of the chicken genomic library and Vivian Sone and Vivian Yeh for excellent technical assistance. This work was supported by Public Health Service Grants HD-16550, HD-09635, and DK-04256, by Biomedical Research Grant S15 CA 46157, and by Senator Jacob Javits Center for Excellence in Neuroscience Grant NS-22789.

1. Edelman, G. M. (1988) *Biochemistry* **27**, 3533–3543.
2. Cunningham, B. A., Hemperly, J. J., Murray, B. A., Prediger, E. A., Brackenbury, R. & Edelman, G. M. (1987) *Science* **236**, 799–806.
3. Gallin, W. J., Sorkin, B. C., Edelman, G. M. & Cunningham, B. A. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 2808–2812.
4. Crossin, K. L., Chuong, C.-M. & Edelman, G. M. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 6942–6946.
5. Gallin, W. J., Chuong, C.-M., Finkel, L. H. & Edelman, G. M. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 8235–8239.
6. Edelman, G. M., Murray, B. A., Mege, R.-M., Cunningham, B. A. & Gallin, W. J. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 8502–8506.
7. Gallin, W. J., Edelman, G. M. & Cunningham, B. A. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 1038–1042.
8. Thiery, J.-P., Delouvée, A., Gallin, W. J., Cunningham, B. A. & Edelman, G. M. (1984) *Dev. Biol.* **102**, 61–78.
9. Cunningham, B. A., Leutzinger, Y., Gallin, W. J., Sorkin, B. C. & Edelman, G. M. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 5787–5791.
10. Hyafil, F., Morello, D., Babinet, C. & Jacob, F. (1980) *Cell* **21**, 927–934.
11. Yoshida-Noro, C., Suzuki, N. & Takeichi, M. (1984) *Dev. Biol.* **101**, 19–27.
12. Imhof, B. A., Vollmers, H. P., Goodman, S. L. & Birchmeier, W. (1983) *Cell* **35**, 667–675.
13. Gumbiner, B. & Simons, K. (1986) *J. Cell Biol.* **102**, 457–468.
14. Damsky, C. H., Richa, J., Solter, D., Knudsen, K. & Buck, C. A. (1983) *Cell* **34**, 455–466.
15. Levi, G., Crossin, K. L. & Edelman, G. M. (1987) *J. Cell Biol.* **105**, 2359–2372.
16. Ringwald, M., Schuh, R., Vestweber, D., Eistetter, H., Lottspeich, F., Engel, J., Dolz, R., Jahnig, F., Epplen, J., Mayer, S., Muller, C. & Kemler, R. (1987) *EMBO J.* **6**, 3647–3653.
17. Nagafuchi, A., Shirayoshi, Y., Okazaki, K., Yasuda, K. & Takeichi, M. (1987) *Nature (London)* **329**, 341–343.
18. Kemler, R., Babinet, C., Eisen, H. & Jacob, F. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 4449–4452.
19. Edelman, G. M., Gallin, W. J., Delouvée, A., Cunningham, B. A. & Thiery, J.-P. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 4384–4388.
20. Nose, A. & Takeichi, M. (1986) *J. Cell Biol.* **103**, 2649–2658.
21. Hatta, K., Okada, T. S. & Takeichi, M. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 2789–2793.
22. Volk, T. & Gieger, B. (1984) *EMBO J.* **3**, 2249–2260.
23. Nose, A., Nagafuchi, A. & Takeichi, M. (1987) *EMBO J.* **6**, 3655–3661.
24. Hatta, K., Nose, A., Nagafuchi, A. & Takeichi, M. (1988) *J. Cell Biol.* **106**, 873–881.
25. Gallin, W. J., Prediger, E. A., Edelman, G. M. & Cunningham, B. A. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 2809–2813.
26. Owens, G. C., Edelman, G. M. & Cunningham, B. A. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 294–298.
27. Jones, F. S., Burgoon, M. P., Hoffman, S., Crossin, K. L., Cunningham, B. A. & Edelman, G. M. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2186–2190.
28. Kornblihtt, A. R., Vibe-Pederson, K. & Baralle, F. E. (1984) *Nucleic Acids Res.* **12**, 5853–5868.
29. Murray, B. A., Hemperly, J. J., Gallin, W. J., MacGregor, J. S., Edelman, G. M. & Cunningham, B. A. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 5584–5588.
30. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
31. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13.
32. Stratagene Cloning Systems (1986) *Bluescript Exo/Mung Instruction Manual* (Stratagene Cloning Systems, San Diego, CA).
33. Tabor, S. & Richardson, C. C. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 4767–4771.
34. Peyrieras, N., Hyafil, F., Louvard, D., Ploegh, H. & Jacob, F. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 6274–6277.
35. Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
36. Elder, J. H. & Alexander, S. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 4540–4544.
37. Plummer, T. H. & Tarentino, A. L. (1981) *J. Biol. Chem.* **256**, 10243–10246.
38. Kozak, M. (1984) *Nucleic Acids Res.* **12**, 857–887.
39. Kreil, G. (1981) *Annu. Rev. Biochem.* **50**, 317–348.
40. Von Heijne, G. (1986) *Nucleic Acids Res.* **14**, 4683–4690.
41. Vennstrom, B. & Bishop, J. M. (1982) *Cell* **28**, 135–143.
42. Mount, S. M. (1982) *Nucleic Acids Res.* **10**, 459–472.
43. Eistetter, H. R., Adolph, S., Ringwald, M., Simon-Chazottes, D., Schuh, R., Guenet, J.-L. & Kemler, R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 3489–3493.
44. Traut, T. W. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2944–2948.
45. Mege, R.-M., Matsuzaki, F., Gallin, W. J., Goldberg, J. I., Cunningham, B. A. & Edelman, G. M. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 7274–7278.