Nucleic Acids Research, Vol. 19, No. 23 6533–6539

The structure of the gene coding for the mouse cell adhesion molecule uvomorulin

M.Ringwald, H.Baribault+, C.Schmidt and R.Kemler*
Max-Planck-Institut für Immunbiologie, Stübeweg 51, D-7800 Freiburg, FRG

Received August 6, 1991; Revised and Accepted October 6, 1991 EMBL accession nos X60961–X60976 (incl.)

ABSTRACT

We have recently shown that the Ca^{2+} dependent cell adhesion molecule uvomorulin is encoded by a single gene, localized on mouse chromosome 8. Here we describe the organization of the uvomorulin gene and give an initial characterization of the uvomorulin promoter. Uvomorulin is encoded by 16 exons, which are distributed over a region of more than 40 kb genomic DNA. The exon structure of the genes for uvomorulin and its chicken homologue L-CAM are nearly identical and thus highly conserved. The relationship between the exon structure and the structure of the uvomorulin protein is analysed. The initiation site of transcription of the uvomorulin gene is located 127 bp upstream of the translation start site in a GC-rich region with no TATA-box, but with a GC-box in position −48 and a CCAAT-box starting at position −65 with respect to the transcription start site. 1.6 kb of the uvomorulin promoter (−1492 to + 92) confer cell type specific promoter activity to the CAT reporter gene. Homologies to known cis acting elements of other promoters are discussed.

INTRODUCTION

Specific cell-cell adhesion controlled by cell adhesion molecules is considered to play an important role during the development of multicellular organisms (1, 2, 3). The cell adhesion molecule (CAM) uvomorulin belongs to the group of Ca^{2+} dependent C-AM's. These CAM's require calcium for their cell binding activity and constitute a gene family of structurally and functionally related transmembrane glycoproteins termed cadherins (2, 3). Uvomorulin is identical to E-cadherin (4). Chicken L-CAM (5), canine Arc-1 (6) and human cell-CAM 120/80 (7, 8) are species homologues of mouse uvomorulin. P-, N-, R- and M-cadherin (9, 10, 11, 12) are other subclasses of this family. A sixth subclass may be EP-cadherin (13, 14). Cadherins mediate cell adhesion by homophilic binding. The specific interaction of identical cadherin molecules from neighbouring cells is thought to represent the basis for cell sorting and morphogenesis (15). The developmentally regulated expression of cadherins correlates with a variety of morphogenetic events that involve cell aggregation or disaggregation (16).

Uvomorulin is already expressed during mouse preimplantation development, where it mediates the compaction process at the morula stage (17, 18). Later in development and in adult tissues, uvomorulin is exclusively expressed in epithelial cells. In polarized epithelial cells uvomorulin is transported to the basolateral compartment and is concentrated in the zonula adherens (19). The clustering of uvomorulin in the zonula adherens is thought to be mediated by catenins, which complex with the cytoplasmic domain of uvomorulin and which link uvomorulin to cytoskeletal actin (20, 21). The formation of this molecular complex is a prerequisite for the function of uvomorulin (21). The extracellular part of uvomorulin, which mediates the selective cell adhesiveness, is largely composed of three repeating domains, each with two putative calcium binding motifs (22). These motifs are well conserved and are located at analogous positions in all cadherins (2). Recently we have shown that one of the Ca^{2+}- binding motifs of uvomorulin complexes Ca^{2+}. A single amino acid substitution in this calcium binding site renders the molecule more susceptible to proteolytic degradation and abolishes completely the adhesive function of the whole molecule (23).

Uvomorulin is encoded by a single gene, localized on chromosome 8 (24). Here we report the molecular cloning of the uvomorulin gene and describe its organization. By comparison with the genomic organization of chicken L-CAM (25) we show that the exon structure of these genes is highly conserved. Since exons often are structural and functional modules in protein architecture, we analyse relationships between the exon structure and the structure of the uvomorulin protein. The differential expression of cadherins in development must be under the strict control of regulatory genes (16). Knowledge about these genes should provide deep insights into the genetic mechanism of morphogenesis. As a first step in this direction we have now isolated 1.5 kb of the uvomorulin promoter. We show that this promoter fragment confers cell type specific expression to a reporter gene and analyse its sequence for homologies to known cis acting elements of other promoters.

* To whom correspondence should be addressed

+Present address: La Jolla Cancer Research Foundation, 1091 N.Torrey Pines Rd, La Jolla, CA 92037, USA

MATERIALS AND METHODS

Isolation of uvomorulin genomic clones

Clones UG1, 2, 3 and 5 were isolated from a genomic library in EMBL3 containing DNA from t^{AE5}/129Sv mice, by using the 3′ BglI-EcoRI fragment of uvomorulin cDNA clone M2 (20) and cDNA clone F5H3 (22) as probes. This library was constructed by T. Grünfeld (Oslo) and B. Herrmann (Tübingen) and kindly provided by Hans Lehrach. The isolation of clone UG4 has already been described (24). Subclone S11 harbours an 11 kb SacI fragment of a cosmid clone isolated from a mouse BALB/c-genomic-DNA library in cosmid vector Lorist B, kindly provided by M.Steinmetz (Basel), by using the 5′ EcoRI-AvaI fragment of cDNA clone F5 (22) as probe. Clone P1 was isolated from a genomic 6 kb-EcoRI-sublibrary: EcoRI digested 129/Sv-DNA was size fractionated on an agarose gel. DNA fragments from 5.8 kb to 6.2 kb were isolated using Geneclean (Bio101 Inc., La Jolla CA), ligated into lambda-gt10 vector arms and packaged in vitro (Gigapack Gold; Stratagene, La Jolla). 1×10^5 pfu were screened with the radiolabeled 5′ EcoRI-BglI fragment of cDNA clone M2 (20). Fragments of the isolated phage and cosmid clones were subcloned in pBluescript II KS+ (Stratagene, La Jolla) in both orientations. Screening and cloning was performed using standard methods (26).

Oligonucleotide synthesis and Southern blot hybridisation

Oligonucleotide synthesis and hybridisation of radiolabeled oligonucleotides to Southern blots were performed as described (24). The hybridisation was done at 10°C below the melting temperature calculated for the respective oligonucleotide.

DNA sequence analysis

DNA sequencing was performed by the dideoxy method (27). Single stranded DNA, derived from subclones in pBluescript II KS+ by infection with helper phage VCS M13 (Stratagene, La Jolla) was used as template, and oligonucleotides complementary to insert DNA were used as primers in addition to T3 and T7 primers. Computer analysis of nucleotide sequences was done using the University of Wisconsin Genetics Computer Group package of computer programs (28) and a compilation of transcription regulating proteins (29).

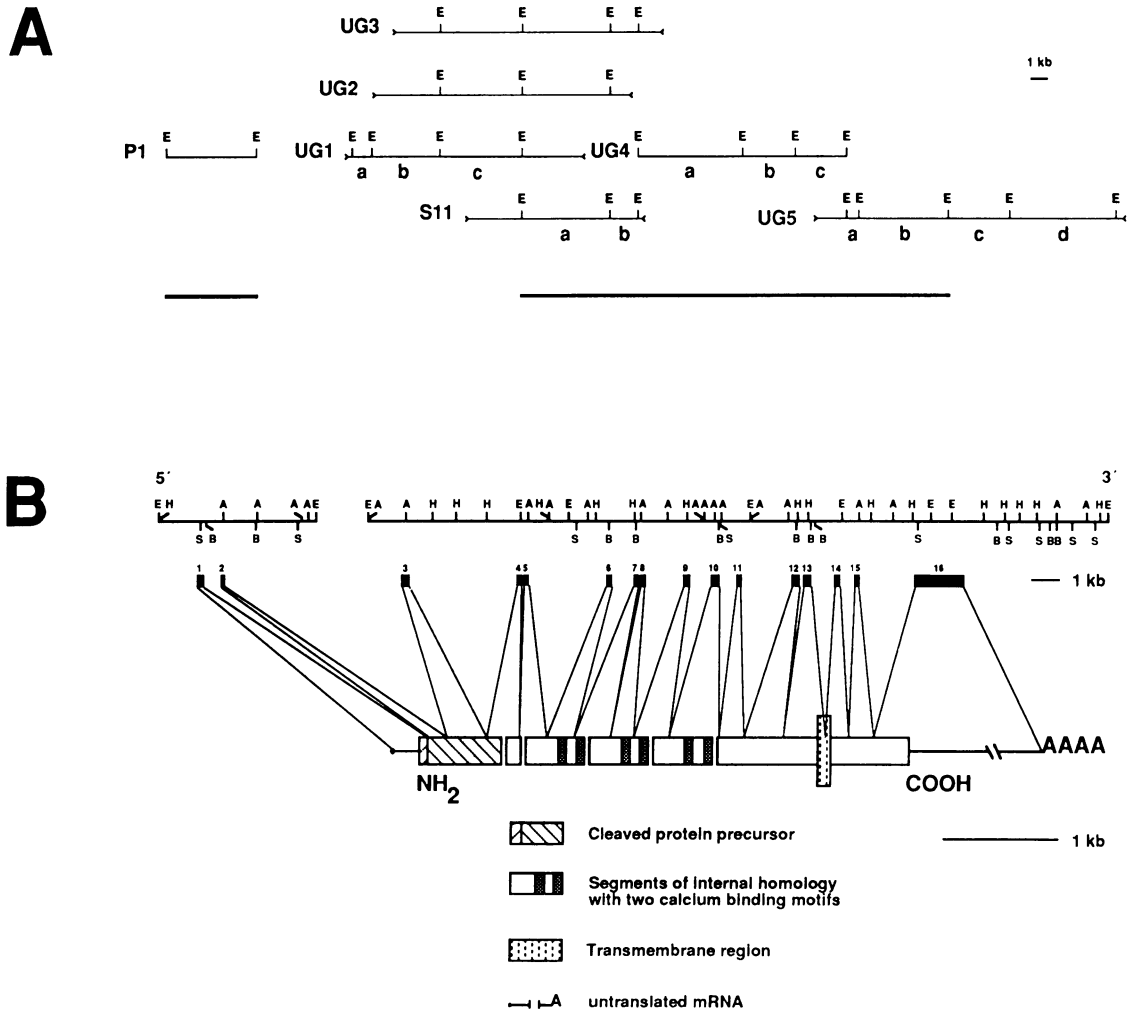

Figure 1. **(A)** Molecular cloning of the uvomorulin gene. Genomic clones isolated from phage and cosmid libraries are shown. The lane at the bottom indicates subclones carrying coding regions of uvomorulin. These subclones were further characterized to determine the exon structure of the uvomorulin gene. E = EcoRI. **(B)** Exon structure of the uvomorulin gene. Uvomorulin mRNA includes 16 exons. Upper part: Restriction fine map of the genomic clones carrying coding regions and location of the exons. A = AccI, B = BamHI, E = EcoRI, H = Hind2, S = SacI. Lower part: Schematic representation of the uvomorulin protein and the untranslated regions of the mRNA, showing the location of the exon boundaries.

RNA isolation, RNAase protection assays and primer extension analysis

Total RNA was prepared according to the guanidinium thiocyanate method (26).

RNAase protection analysis was performed as described (30). Radiolabeled riboprobes were generated by T3 polymerase from the NotI-PstI and NotI-SacI fragments indicated in Figure 3. Then 60 μg of total RNA or tRNA and 3×10^5 cpm of riboprobe were hybridized at 55°C overnight.

Primer extension analysis was performed as described (31). The synthetic oligonucleotide PRE, indicated in Figure 5, was end labeled with [γ-^{32}P]ATP using T4 polynucleotide kinase and then purified by chromatography on a BioGel A-0.5m column (BioRad). 1×10^5 cpm of radiolabeled oligonucleotide were hybridized to 60 μg of total RNA.

Table 1. Sequences around the exon/intron junctions of uvomorulin.

Exon number	Exon size (bp)	Sequence at exon-intron junction: 5′ Splice donor	3′ Splice acceptor
1	175	CTG CTG CAG 116 gtatgt	tacctccatcccag 117 GTC TCC TCA
2	121	GGC AGA G 237 gtaagg	tctccatcctgtag 238 TG AGA TTT
3	224	CAC CAC CGC 461 gtacgt	ttcttttctcttag 462 GAC CCT GCC
4	144	CTG GTT CAG 605 gtagga	tgttttgtgttcag 606 ATC AAA TCC
5	156	AAG TAC ATC 761 gtgagt	tgactctcttgcag 762 CTC TAT TCT
6	145	GTT CCA G 906 gtaaat	ctctcctttgacag 907 GA ACC TCC
7	176	GAC CGA GAG 1082 gtcagg	tttggtgtctacag 1083 AGT TAC CCT
8	129	CCA AGC ACG 1211 gtaact	cctttcctcttcag 1212 TAT CAG GGT
9	183	ACA GCC AAG 1394 gtttgt	ctcctgttctctag 1395 GGC TTG GAT
10	245	AAG ATC AC 1639 gtgagt	gcttgtgccctcag 1640 G TAT CGG
11	146	GAT GAT G 1785 gtatgg	atgttttctcatag 1786 GT TCA CCC
12	225	GAC GCA G 2010 gtgtga	tttcctttcccctag 2011 CT CAA GAA
13	228	CTG CTG A 2238 gtaagt	ttccttgtctccag 2239 TT CTG ATC
14	131	GAA GAC CAG 2369 gtgggt	gctcttctctctag 2370 GAC TTT GAT
15	144	ATC GAT GAA 2513 gtaagt	gctcttcctttag 2514 AAC CTG AAG
16	1841	TGTTAAACTGT 4354 tgtgag	

Exon sequences are in capital letters and intron sequences in lower case letters. Residues matching the consensus splice sequence (62) are underlined. The number below the DNA sequence denotes the nucleotide position at which the intron interrupts the uvomorulin mRNA. Nucleotide positions refer to Figure 1 in reference 4.

Cells

The following cell lines were used: mouse embryonal carcinoma (EC) cells F9 (32) and LT-1 (33), the mouse colon carcinoma cell line CMT (34), NIH-3T3 fibroblasts and L tk$^-$ cells. The cells were grown in Dulbecco's modified Eagle's medium (DMEM), supplemented with 15% fetal calf serum (FCS), in a 10% CO_2 atmosphere at 37°C.

CAT assays

CAT constructs: The 5′ EcoRI-SacI fragment of clone P1 was inserted via BamHI linkers into the BamHI site of pBLCAT3 (35) in both orientations to generate the test plasmid and a negative control. As positive control, pHDCAT (36) was used. 20 μg of each construct respectively were used for transfection.

CAT assays were performed by the calcium phosphate method as described (37). The day before transfection cells were replated at 5% (F9, CMT) and 10% (LT-1, NIH-3T3) confluency on a 10 cm dish. Cells were harvested 48 h after removal of the precipitate. 20 μl cell extract of total 100 μl were used for the assay (30 min at 37°C). Autoradiography was performed for 72 h.

RESULTS

Genomic clones encoding 40 kb of the uvomorulin gene were isolated by screening mouse genomic libraries with uvomorulin cDNA probes (see Figure 1A and Materials and Methods). To determine the sequence of all exons and neighbouring intronic regions, subclones P1, S11a, b, UG4a, b, c and UG5a, b were partially sequenced, using oligonucleotides complementary to uvomorulin cDNA as primers. Subsequently the subclones were fine mapped by restriction digests and Southern blot hybridization to localize the exons. The results of this analysis are shown in Figure 1B. Uvomorulin is encoded by 16 exons, which are distributed over a region of more than 40 kb of genomic DNA. Between exons 2 and 3 there is a large intron of at least 12 kb. Because this intron has not yet been completely cloned, the actual size of the entire uvomorulin gene remains to be determined. The precursor part of the uvomorulin protein is encoded by four exons. Exon 1 encodes the complete 5′ untranslated region of the mRNA and almost all of the signal peptide (see also Figure 3). The three internal repeats in the extracellular part of uvomorulin (22) are encoded by 6 exons. Separate exons encode

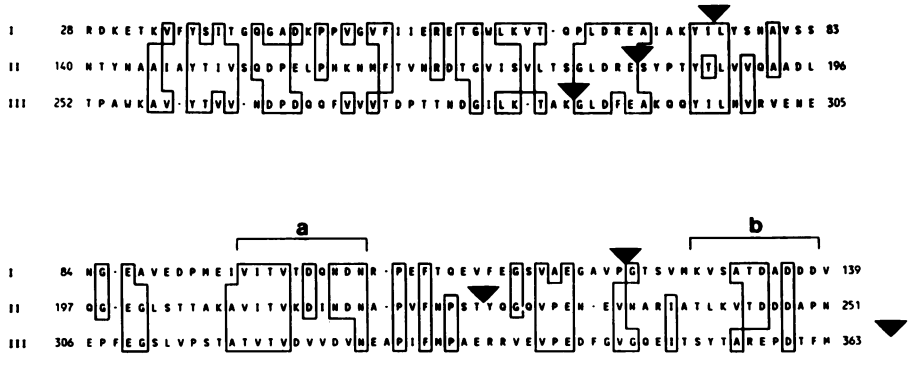

Figure 2. Location of the exon boundaries in the region encoding the three homologous protein domains. The repeating domains of uvomorulin are aligned. Positions identical in at least two domains are enclosed in boxes. The putative Ca^{2+}-binding sites a and b are indicated. Triangles denote the exon boundaries. Amino acid positions refer to Figure 1 in reference 22. The exon boundaries do not coincide with the homology of the protein domains.

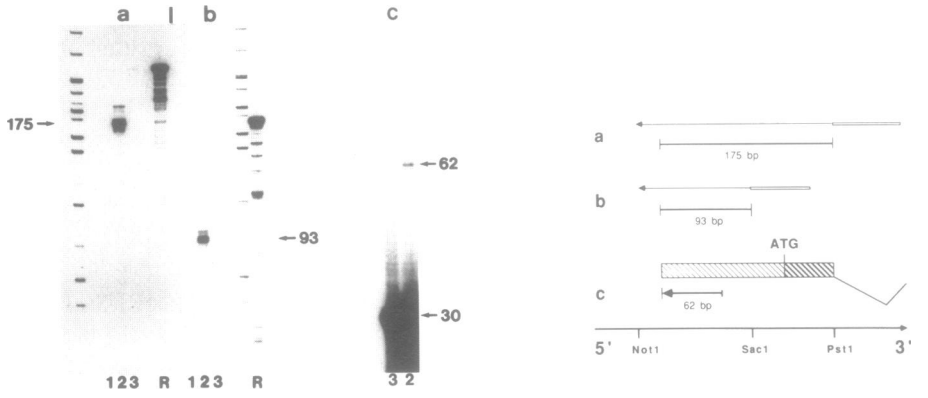

Figure 3. Initiation site of transcription of the uvomorulin gene. Left side: a and b: RNAase protection experiments. Two riboprobes, starting from different positions in the first exon of the uvomorulin gene, were hybridized to tRNA (lane 1), RNA from F9 cells (lane 2) and RNA from L-cells (lane 3). Lane R: undigested riboprobe. After RNAase digestion 175 nucleotides of riboprobe a and 93 nucleotides of riboprobe b were specifically protected by RNA from F9 cells. MspI digested pBR322-DNA was used as molecular weight standard. c: Primer extension analysis. Oligonucleotide PRE (see Figure 5) was hybridized to RNA from L-cells (lane 3) and F9-cells (lane 2). Only on RNA from F9-cells, the oligonucleotide was extended for 32 nucleotides. Right side: a and b: Map of the 5′ region of the uvomorulin gene, indicating the position of riboprobes and respective protected fragments. Regions of the riboprobes corresponding to plasmid sequences are indicated by the open boxes. c: Map of the reaction product obtained by primer extension of oligonucleotide PRE. Hatched box represents the first exon of the uvomorulin gene. ATG: Translation start codon.

the four putative calcium binding motifs from the first two repeats. The cytoplasmic domain is encoded by 3 exons. Exon 16 encodes the c-terminal end of uvomorulin, which is necessary for the association with catenins (21), as well as the complete 3′ untranslated region of the uvomorulin mRNA.

The sequences of the exon/intron boundaries of uvomorulin and the length of the exons are shown in more detail in Table 1. All 15 introns have consensus splice sequences. The complete sequence data are available from the EMBL data library under the accession Nos. X60961 to X60976. Nine nucleotide differences between the genomic sequences reported here and the 4.3 kb cDNA (4, 22) were detected. Only at two positions would these differences result in amino acid changes of the predicted protein sequence (for details see EMBL data library). Release 28.0 from the PIR and release 27.0 from the EMBL data library were searched for homologies to each exon sequence. Besides the members of the cadherin gene family, no significant homologies were found.

Comparison of the uvomorulin gene to chicken L-CAM (25) reveals a high conservation of the exon structure, although the two genes are quite different in size — the exons of L-CAM are distributed over less than 10 kb of genomic DNA. Both genes have not only the same number of exons, but in an end to end alignment of the protein sequences all but one of the exon boundaries are located at identical positions (not shown). Moreover, all introns in uvomorulin are integrated in the same phase of the reading frame as the respective introns of L-CAM.

A surprising feature of the exon structure of uvomorulin (and L-CAM) is seen in an alignment of the three internally repeated domains of uvomorulin, as shown in Figure 2. Despite the clear cut homology of the repeats, the exon boundaries do not coincide, but show a 'permutation'. This raises interesting questions about the generation of the repeated domains during evolution. They cannot have been generated simply by duplication of (present day) exons but rather by a more complex process, in which gene recombination and/or integration of introns could also have been involved. All these events must have already occurred before the divergence of the mouse and chicken species.

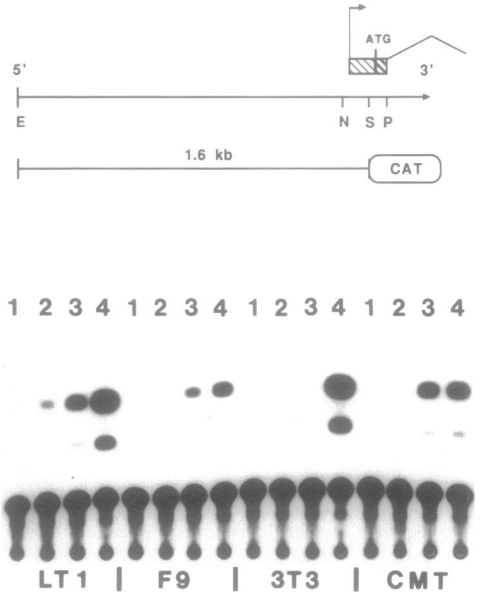

Figure 4. Analysis of the uvomorulin promoter in transient CAT-assays. Upper part: The 1.6 kb EcoRI/SacI fragment of genomic clone P1 (see also Figure 1B) was used to regulate the CAT reporter gene. The initiation site of transcription is indicated by an arrow. E = EcoRI, N = NotI, P = PstI, S = SacI. Lower part: CAT-assays: Transfected cell lines: LT-1, F9, CMT, NIH-3T3. Transfected vectors: 1: untransfected cells, 2: negative control: CAT gene, driven by the 1.6 kb fragment in the reverse orientation, 3: construct, described in the upper part, 4: positive control: pHD-CAT; CAT gene driven by SV40 enhancer/tk promoter.

To study the developmental and tissue specific regulation of uvomorulin gene expression, we have started to analyse 5′ genomic region of uvomorulin. Figure 3 shows the determination of the initiation site of transcription by primer extension and RNAase protection experiments. Using primer extension analysis a specific band with a length of 62 nucleotides was detected in uvomorulin positive F9 cells, but not in uvomorulin negative L tk$^-$ cells (Figure 3c). By RNAase protection experiments using

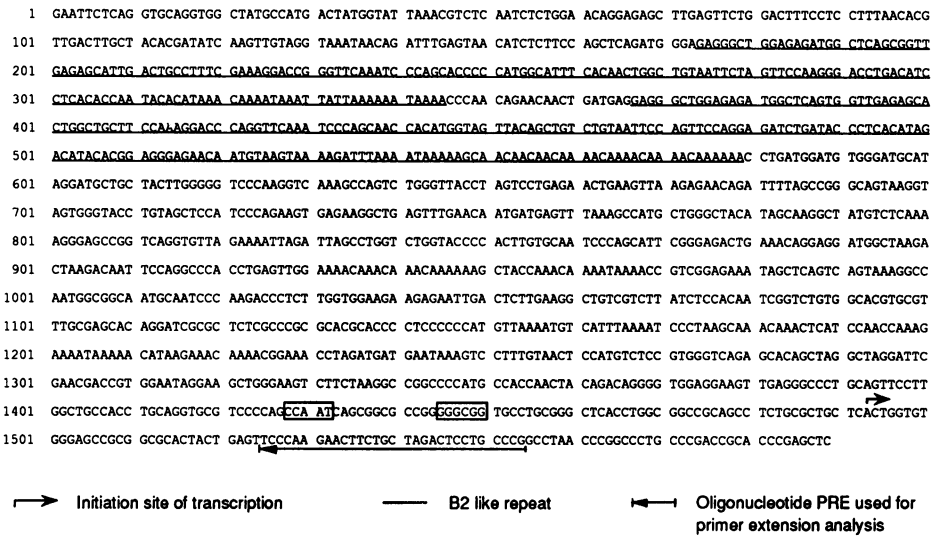

Figure 5. Sequence of the uvomorulin promoter (EcoRI/SacI fragment shown in Figure 4). The initiation site of transcription is located in a GC-rich region with no TATA-box, but with a CCAAT-box starting at position −65 and a GC-box at position −48 with respect to the transcription start site. The 5′ part of the sequence harbours two direct repeats homologous to the B2 gene family. Oligonucleotide PRE was used for the primer extension experiments shown in Figure 3c.

two different riboprobes specific protected fragments of 175 bp and 93 bp, respectively, were detected in uvomorulin positive F9 cells (Figure 3a and b). Both primer extension and RNAase protection experiments demonstrate that the initiation site of transcription is located 127 bp upstream of the translation start site. In addition, RNAase protection assays show that exon 1 encodes the 5′ untranslated and a short translated region of the uvomorulin mRNA (Figure 3, right side).

The 5′ 1.6 kb EcoRI-SacI fragment of uvomorulin genomic clone P1 (Figure 4 upper part and Figure 1) was analysed for promoter activity in transient CAT assays. Three different uvomorulin positive mouse cell lines, the embryonal carcinoma cell lines LT-1 and F9 and the colon carcinoma cell line CMT, and one uvomorulin negative cell line, NIH 3T3, were transfected with the CAT construct indicated in the upper part of Figure 4 or with control constructs. The results of the CAT assays are shown in the lower part of Figure 4. In all three uvomorulin positive cell lines CAT activity is observed when the promoter fragment is cloned in the right orientation (as can be seen by comparison of lanes 3 and 2); we cannot explain the weak CAT activity observed with the negative control in LT-1 cells (lane 2). No CAT activity is found in uvomorulin negative NIH 3T3 cells. From these results we conclude that the uvomorulin promoter fragment tested contains at least part of the cis acting elements necessary for tissue specific uvomorulin gene expression.

Therefore, the sequence of this promoter fragment was determined and analysed for homologies to regulatory elements of other promoters. The sequence is shown in Figure 5. The initiation site of transcription, determined in the experiments shown in Figure 3, is localized in a G+C-rich region with no TATA-box, but with a CCAAT-box in position −65 and a GC-box starting at position −48 with respect to the transcription start site. The 5′ part of the promoter fragment harbours two direct repeats homologous to the B2 gene family (38). Several motifs with sequences identical to known cis acting elements are found. Those which may be of interest for the regulation of uvomorulin gene expression, because the transacting factors specific for these sequences are active in uvomorulin expressing tissues, are listed in Table 2.

Table 2. Sequence motifs of the uvomorulin promoter found to be homologous to cis acting elements of other promoters. The transcription factors listed were shown to be specific for the respective sequence motifs in the references cited.

Transcription factor	Motif	Ref.
AP2	1422 - CCCCAGCC	39
CP1	1419 - CGTCCCCAGCCAATCAGCG	40
CTF/ NF1	997 - GGCCAATG 1426 - AGCCAATC	41
EF - C	r-1183 - TGTTTGCT	42
GCF 2	1338 - GGCCGGCCCC 1433 - CAGCGGCGCC r-1452 - CACCGCCCCC 1506 - CCGCGGCGCA r-1513 - CGCCGCGGCT 1560 - ACCCGGCCCT	43
Glucocorticoid-Receptor	r- 687 - GCTAAAATCTGTTCT	44
LVα	59 - GAACAG 674 - GAACAG	45
NFκ E2 E12/E47	14 - CAGGTGGC r-1412 - CAGGTGGC	46 47
Progesterone-Receptor	r- 152 - TGTTACT r- 520 - TGTTCTCC 1247 - AGTCCTTT	48
Sp1	r-1146 - GGGGAGGGGT 1445 - GGGCGG	49 50

DISCUSSION

Here we report the molecular cloning of the uvomorulin gene, describe its organization and give an initial characterization of the uvomorulin promoter.

The exon structure of uvomorulin was determined by

sequencing all genomic regions complementary to uvomorulin cDNA, evaluating the exon boundaries by the breakpoint of homology and the consensus splice sequence. The 5′ end of the uvomorulin mRNA was determined by primer extension and RNase protection experiments, using RNA from F9 embryonal carcinoma cells which was also the source for the isolation of uvomorulin cDNA (22). By this approach only exons used in the uvomorulin mRNA of F9 cells are detected. But in all uvomorulin expressing cells so far tested, only a single size mRNA of 4.3 kb and a single protein species with an apparent molecular weight of 120 kDa is found. Moreover, no additional exons are found in the gene of the chicken homologue of uvomorulin L-CAM, which has been almost completely sequenced (25). Therefore there is no evidence for alternative splicing, and we most likely have characterized the exon structure of the uvomorulin gene completely.

Uvomorulin mRNA is made up of 16 exons, which are distributed over a region of more than 40 kb genomic DNA. Exons often are structural and functional modules in protein architecture (52). Upon cursory examination this correlation is hard to see for uvomorulin. For example according to Van Heijne (53) the signal peptide of the uvomorulin precursor should be cleaved off after S in amino acid position 19 (for numbering of amino acid sequence see reference 4). Exon 1 encodes almost all of this putative signal peptide, but the last three amino acids are encoded by exon 2. Each of the three repeated domains of uvomorulin are encoded by at least two exons. This would be in line with the subhomologies found and the functional subdivision postulated for each domain (22). However, the exon boundaries do not coincide with the boundaries of homology as one would expect. Finally, the membrane spanning region is specified by two exons.

Comparison of the uvomorulin gene to its chicken homologue L-CAM reveals a high conservation of the exon structure, although the two genes are quite different in size. The respective exon boundaries are located at identical positions in an alignment of the two proteins. This might argue for a functional significance of the exon structure with respect to the protein architecture. In line with this we found that exon 16 encodes the domain of uvomorulin which is a prerequisite for the association of catenins (21). Also it is interesting to note that exons 13 and 14 encode almost exactly that part of uvomorulin around the membrane spanning region which is not homologous to desmoglein, a more distantly related member of the cadherin family (54). Information about the organization of other genes of the cadherin family might be required to better understand the molecular mechanisms which generated specific calcium dependent CAM's during evolution.

Uvomorulin can already be detected on the cell surface of unfertilized and fertilized eggs. It is not synthesized in these cells, and thus is presumably maternally derived, persisting from oogenesis (55). Uvomorulin transcription starts at the late two cell stage, when stored maternal RNA is broken down and the embryonic genome is activated. Until implantation uvomorulin is expressed in all cell types, the first disappearance of the molecule is observed during the formation of mesoderm at the primitive streak stage. Later in development and in adult tissues uvomorulin is exclusively expressed in epithelial cells, independent of their germ layer origin. To study the developmental and tissue specific expression of uvomorulin we have started to analyse the promoter of the uvomorulin gene.

The initiation site of transcription was mapped 127 bp upstream of the translation start site by RNAase protection and primer extension experiments. Based on this analysis, we infer that we have isolated 1.5 kb of the uvomorulin promoter. This promoter fragment (exactly the genomic region from −1492 to +92 with respect to the transcription start site) exhibits cell type specific promoter activity, as we have shown in transient CAT assays: Promoter activity is found in three uvomorulin expressing mouse cell lines, the embryonal carcinoma cell lines LT-1 and F9 and the colon carcinoma cell line CMT, and no activity can be detected in NIH-3T3 fibroblasts. Therefore, at least part of the cis acting elements necessary for tissue specific uvomorulin gene expression should be localized on this DNA fragment. Its regulating capacity must be tested further in a broader range of cell types and in transgenic mice.

The activity of a promoter is determined by its architecture, consisting of the type of control elements contained in the promoter and the context in which they appear (40). From this point of view we have analysed the sequence of the isolated uvomorulin promoter region. No TATA-box, but a CCAAT-box and a GC-box are found upstream of the initiation site of transcription at distances similar to those found in many other promoters. The transcription start site is located in a G+C rich region, which resembles a CpG rich island. CpG rich islands are found around the transcription start site of many housekeeping genes and are thought to identify sequences that are to be constantly available in the nucleus (56). Uvomorulin expression starts with the activation of the embryonic genome, as does the expression of housekeeping genes. The putative CpG rich island in the uvomorulin promoter may therefore be important for the expression of the gene in early mouse development. Between positions −1298 to −913 with respect to the transcription start site two direct repeats are found which are homologous to the B2 gene family. B2 genes are rodent specific elements of no known function, transcribed by RNA polymerase III. The gene family most likely arose by retroposition of processed tRNA (57). Retroposons can provide new regulatory elements for an existing gene (58), but this is unlikely to be the case for the B2 repeats in the uvomorulin promoter since B2 genes are rodent specific, while the distribution of uvomorulin expression is very similar in chicken, mouse and human. Furthermore, deletion of the B2 repeats does not affect the activity of the promoter fragment in transient CAT assays (B.Kinzel, unpublished observations).

Several putative binding sites for trans acting regulatory proteins are found in the uvomorulin promoter (Table 2). The listed transcription factors are all functional in uvomorulin expressing tissues and could therefore be involved in regulation of uvomorulin gene expression. For example it will be of interest, whether AP2 could interact with the uvomorulin promoter. AP2 transmits the effect of phorbolester tumor promoters (51), which could represent one mechanism of down-regulation of uvomorulin during carcinogenesis (3, 59). The occurrence of putative glucocorticoid receptor and progesterone receptor binding sites suggests a hormone dependent regulation of uvomorulin gene expression. Interestingly the three putative progesterone receptor binding sites are identical to functional cis acting elements in the promoter of the uteroglobin gene (48). Uteroglobin is regulated differentially by various steroid hormones in tissues known to express uvomorulin (60). It has been reported that cadherins can be stimulated by steroid hormones, but in these cases the molecular mechanism of this stimulation is not known (14, 61). The sequence of the uvomorulin promoter, reported here, provides a first hint for a direct transcriptional regulation by steroid hormones.

ACKNOWLEDGEMENTS

We would like to thank Johann Wirsching and Sabine Mayer for excellent technical assistance, Veronika Person for preparing oligonucleotides and Rosemary Schneider for typing the manuscript. Plasmids pBLCAT3 and pHDCAT were a kind gift from Dr. Günther Schütz (Heidelberg), genomic libraries were kindly provided by Drs Thomas Grünfeld (Oslo) and Michael Steinmetz (Basel). We are grateful to Dr Erwin Wingender (Braunschweig) for providing an unpublished update of his compilation of transcription regulating proteins, to Christoph Gartmann for help with computing and to Dr Klaus Schughart for helpful discussions. This work was supported by the Deutsche Forschungsgemeinschaft.

REFERENCES

1. Edelman,G.M. (1985) *Ann. Rev. Biochem.* **54**, 135–169.
2. Kemler,R., Ozawa,M. and Ringwald,M. (1989) *Curr. Opin. Cell Biol.* **1**, 892–897.
3. Takeichi,M. (1991) *Science* **251**, 1451–1455.
4. Nagafuchi,A., Shirayoshi,Y., Okazaki,K., Yasuda,K. and Takeichi,M. (1987) *Nature* **329**, 341–343.
5. Gallin,W.J., Sorkin,B.C., Edelman,G.M. and Cunningham,B.A. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 2808–2812.
6. Behrens,J., Birchmeier,W., Goodman,S.L. and Imhof,B.A. (1985) *J. Cell Biol.* **101**, 1307–1315.
7. Damsky,C.H., Richa,J., Solter,D., Knudsen,K. and Buck,C.A. (1983) *Cell* **34**, 455–466.
8. Mansouri,A., Spurr,N., Goodfellow,P.N. and Kemler,R. (1988) *Differentiation* **38**, 67–71.
9. Nose,A., Nagafuchi,A. and Takeichi,M. (1987) *EMBO J.* **6**, 3655–3661.
10. Hatta,K., Nose,A., Nagafuchi,A. and Takeichi,M. (1988) *J. Cell Biol.* **106**, 873–883.
11. Takeichi,M., Inuzaka,H., Shimamura,K., Fujimori,T. and Nagafuchi,A. *Cold Spring Harbor Symp. Quant. Biol.*, in press.
12. Donalies,M., Cramer,M., Ringwald,M. and Starzinski-Powitz,A. (1991) *Proc. Natl. Acad. Sci. USA*, in press.
13. Ginsberg,D., DeSimone,D. and Geiger,B. (1991) *Development* **111**, 315–325.
14. Choi,Y.S., Sehgal,R., McCrea,P. and Gumbiner,B. (1990) *J. Cell Biol.* **110**, 1575–1582.
15. Nose,A., Nagafuchi,A. and Takeichi,M. (1988) *Cell* **54**, 993–1001.
16. Takeichi,M. (1988) *Development* **102**, 639–655.
17. Kemler,R., Babinet,C., Eisen,H. and Jacob,F. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 4449–4452.
18. Hyafil,F., Morrelo,D., Babinet,C. and Jacob,F. (1980) *Cell* **21**, 927–934.
19. Boller,K., Vestweber,D. and Kemler,R. (1985) *J. Cell Biol.* **100**, 327–332.
20. Ozawa,M., Baribault,H. and Kemler,R. (1989) *EMBO J.* **8**, 1711–1717.
21. Ozawa,M., Ringwald,M. and Kemler,R. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 4246–4250.
22. Ringwald,M., Schuh,R., Vestweber,D., Eistetter,H., Lottspeich,F., Engel,J., Dölz,R., Jähnig,F., Epplen,J., Mayer,S., Müller,C. and Kemler,R. (1987) *EMBO J.* **6**, 3647–3653.
23. Ozawa,M., Engel,J. and Kemler,R. (1990) *Cell* **63**, 1033–1038.
24. Eistetter,H., Adolph,S., Ringwald,M., Simon-Chazottes,D., Schuh,R., Guenet,J.-L. and Kemler,R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 3489–3493.
25. Sorkin,B.C., Hemperly,J.J., Edelman,G.M. and Cunningham,B.A. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 7617–7621.
26. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual* (2nd edition). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
27. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
28. Devereux,J., Haeberli,P. and Smithies,O. (1984) *Nucl. Acids Res.* **12**, 387–395.
29. Wingender,E. (1988) *Nucl. Acids Res.* **16**, 1879–1902.
30. Gilman,M. (1989) In: *Current Protocols in Molecular Biology* (ed. Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A., Struhl,K.) pp. 471–478, New York, Wiley.
31. Nielsen,P.J. and Trachsel,H. (1988) *EMBO J.* **7**, 2097–2105.
32. Nicolas,J.F., Avner,P., Gaillard,J., Guenet,J.L., Jakob,H. and Jacob,F. (1976) *Cancer Res.* **36**, 4224–4231.
33. Nicolas,J.F., Jakob,H. and Jacob,F. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 3292–3296.
34. Franks,L.M. and Hemmings,V.J. (1978) *J. Pathol.* **124**, 35–38.
35. Luckow,B. and Schütz,G. (1987) *Nucl. Acids Res.* **15**, 5490.
36. Müller,G., Ruppert,S., Schmid,E. and Schütz,G. (1988) *EMBO J.* **7**, 2723–2730.
37. Gorman,C.M., Moffat,L.F. and Howard,B.H. (1982) *Mol. Cell Biol.* **2**, 1044–1051.
38. Rogers,J.H. (1985) *Int. Rev. Cytol.* **93**, 149–185.
39. Mitchel,P.J., Wang,C. and Tjian,R. (1987) *Cell* **50**, 847–861.
40. Chodosh,L.A., Baldwin,A.S., Carthew,R.W. and Sharp,P.A. (1988) *Cell* **53**, 11–24.
41. Jones,K.A., Kadonaga,J.T., Rosenfeld,P.J., Kelly,T.J. and Tjian,R. (1987) *Cell* **48**, 79–89.
42. Shaul,Y. and Ben-Levy,R. (1987) *EMBO J.* **6**, 1913–1920.
43. Kageyama,R. and Pastan,I. (1989) *Cell* **59**, 815–825.
44. Jantzen,H.M., Strähle,U., Gloss,B., Stewart,F., Schmid,W., Boshart,M., Miksicek,R. and Schütz,G. (1987) *Cell* **49**, 29–38.
45. Speck,N.A. and Baltimore,D. (1987) *Mol. Cell Biol.* **7**, 1101–1110.
46. Gimble,J.M., Flanagan,J.R., Recker,D. and Max,E.E. (1988) *Nucl. Acids Res.* **16**, 4967–4988.
47. Murre,C., Schonleber McCaw,P. and Baltimore,D. (1989) *Cell* **56**, 777–783.
48. Bailly,A., LePage,C., Rauch,M. and Milgram,E. (1986) *EMBO J.* **5**, 3235–3241.
49. Gustafson,T.A. and Kedes,L. (1989) *Mol. Cell Biol.* **9**, 3269–3283.
50. Dynan,W.S. and Tjian,R. (1985) *Nature* **316**, 774–778.
51. Imagawa,M., Chiu,R. and Karin,M. (1987) *Cell* **51**, 251–260.
52. Blake,C.C.F. (1985) *Int. Rev. Cytol.* **93**, 149–185.
53. Von Heijne,G. (1986) *Nucl. Acids Res.* **14**, 4683–4690.
54. Koch,P.J., Walsh,M.J., Schmelz,M., Goldschmidt,M.D., Zimbelmann,R. and Franke,W.W. (1990) *Eur. J. Cell Biol.* **53**, 1–12.
55. Vestweber,D., Gossler,A., Boller,K. and Kemler,R. (1987) *Devl. Biol.* **124**, 451–456.
56. Bird,A.P. (1986) *Nature* **321**, 209–213.
57. Weiner,A.M., Deininger,P.L. and Efstratiadis,A. (1986) *Ann. Rev. Biochem.* **55**, 631–661.
58. Brosius,J. (1991) *Science* **251**, 753.
59. Behrens,J., Mareel,M.M., VanRoy,F.M. and Birchmeier,W. (1989) *J. Cell Biol.* **108**, 2435–2447.
60. Beato,M., Arnemann,J., Menne,C., Müller,H., Suske,G. and Wenz,M. (1983) In McKerns,K.W. (ed.) *Regulation of gene expression by hormones*, Plenum Press, New York, pp. 151–175.
61. Blaschuk,O.W. and Farookhi,R. (1989) *Dev. Biol.* **136**, 564–567.
62. Padgett,R.A., Grabowski,P.J., Konarska,M.M., Seiler,S. and Sharp,P.A. (1986) *Ann. Rev. Biochem.* **55**, 1119–1150.