Hensin, the polarity reversal protein, is encoded by DMBT1, a gene frequently deleted in malignant gliomas

JIRO TAKITO, LUNBIAO YAN, JIAN MA, CHINAMI HIKITA, S. VIJAYAKUMAR, D. WARBURTON, AND QAIS AL-AWQATI

Hensin, the polarity reversal protein, is encoded by DMBT1, a gene frequently deleted in malignant gliomas. Am. J. Physiol. 277 (Renal Physiol. 46): F277–F289, 1999.—The band 3 anion exchanger is located in the apical membrane of a β-intercalated clonal cell line, whereas the vacuolar H⁺-ATPase is present in the basolateral membrane. When these cells were seeded at confluent density, they converted to an α-phenotype, localizing each of these proteins to the opposite cell membrane domain. The reversal of polarity is induced by hensin, a 230-kDa extracellular matrix protein. Rabbit kidney hensin is a multidomain protein composed of eight SRCR ("scavenger receptor, cysteine rich"); two CUB ("C1r/C1s Uegf Bmp1"); and one ZP ("zona pellucida") domain. Other proteins known to have these domains include CRP-ductin, a cDNA expressed at high levels in mouse intestine (8 SRCR, 5 CUB, 1 ZP), ebnerin, a protein cloned from a rat taste bud library (4 SRCR, 3 CUB, 1 ZP), and DMBT1, a sequence in human chromosome 10q25–26 frequently deleted in malignant gliomas (9 SRCR, 2 CUB, 1 ZP). Rabbit and mouse hensin genomic clones contained a new SRCR that was not found in hensin cDNA but was homologous to the first SRCR domain in DMBT1. Furthermore, the mRNA untranslated regions and the signal peptide of hensin were homologous to those of DMBT1. Mouse genomic hensin was localized to chromosome 7 band F4, which is syntenic to human 10q25–26. These data suggest that hensin and DMBT1 are alternatively spliced forms of the same gene. The analysis of mouse hensin bacterial artificial chromosome (BAC) genomic clone by sequencing and Southern hybridization revealed that the gene also likely encodes CRP-ductin. A new antibody against the mouse SRCR1 domain recognized a protein in the mouse and rabbit brain but not in the immortalized cell line or kidney, whereas an antibody to SRCR6 and SRCR7 domains which are present in all the transcripts, recognized proteins in intestine, kidney, and brain from several species. The most likely interpretation of these data is that one gene produces at least three transcripts, namely, hensin, DMBT1, and CRP-ductin. Hensin may participate in determining the polarized phenotype of other epithelia and brain cells.

In the collecting tubule of the kidney, the intercalated epithelial cells exist in a spectrum of types. One extreme form, the α-phenotype, has apical H⁺-ATPase and basolateral Cl/HCO₃ exchanger and hence secretes acid into the lumen. The anion exchanger is an alternately spliced form of band 3, the product of the AE1 gene, which we term kAE1 (28). The other subtype, at least in the rabbit, secretes HCO₃ by an apical anion exchanger and a basolateral H⁺-ATPase. When rabbits were fed an acid load, the HCO₃-secreting phenotype changed to that of an acid secretor (25). We termed this phenomenon “plasticity” in the polarity of these cells. We generated a conditionally immortalized clonal cell line of the HCO₃-secreting intercalated cell and found that cells plated at subconfluent density had apical kAE1, whereas the same clonal cell plated at confluent density relocalized the protein to the basolateral membrane. In addition, the polarity reversal was associated with the appearance of other characteristics of acid-secreting phenotype such as vigorous apical endocytosis (10, 29). This change was also accompanied by reorganization of the apical actin cytoskeleton, the induction of villin and cytokeratin 19 and the appearance of apical microplicae, all of which are markers of terminal differentiation in epithelia (30). A glycosylated 230-kDa extracellular matrix (ECM) protein, termed hensin, was purified using the induction of apical endocytosis as a functional assay. The partially purified protein induced the relocalization of kAE1, and a polyclonal antibody against recombinant hensin blocked the appearance of apical endocytosis and reorganization of cytoskeleton (26, 30).
Here we present the cDNA sequence of rabbit kidney hensin. It is a modular protein containing a signal peptide, SRCR (“scavenger receptor, cysteine rich”) domains, SID domains (“SRCR interspersed” domains), CUB (“C1r/C1s Uegf Bmp1”) domains, and ZP (“zona pellucida”) domains. The following four cDNAs have recently been identified as proteins that contain these motifs: 1) human DMBT1, a protein located in chromosome 1q25–26 that is often deleted in malignant brain tumors, suggesting that it is a tumor suppressor (17); 2) CRP-ductin α- and β-subtypes, which were isolated from cDNA libraries of mouse intestinal crypt cells (6); 3) ebnerin, a sequence identified in a rat taste bud (18). These studies raise the question whether the sequences are members of a new gene family or the product of alternative splicing of a single gene. Because the cDNAs were obtained from four different animal species it was not possible to answer that question by inspection of the sequences. Furthermore, SRCR domains are an ancient gene family that, in mammals at least, show high levels of homology among different proteins. On the basis of our analysis of hensin genomic sequences in rabbit and mouse, we conclude that the most likely interpretation of our results is that hensin, DMBT1, and CRP-ductin proteins are products of the same gene.

MATERIALS AND METHODS

Rabbit hensin cDNA and genomic cloning. Hensin protein was purified from the ECM of cultured rabbit kidney intercalated cell line, clone C (10, 29). Microsequencing of the peptide digests revealed nine novel peptide sequences. RT-PCR with clone C total RNA using degenerate primers designed from the peptides produced a 700-bp cDNA fragment of rabbit hensin as described (26). The fragment was used as a hybridization probe to screen the clone C cDNA library constructed in lambda ZAPII (Stratagene). A 4.5-kb cDNA fragment was isolated by the screen and sequencing showed that it encoded the poly-A tail. To obtain the 5’-end of hensin, rapid amplification of cDNA ends was performed (5’-RACE, version 2.0; Gibco-BRL, Life Technologies) according to the supplied protocol. The validity of the 5’-end cDNA fragment of hensin was verified by Northern blotting with clone C total RNA and by sequencing.

A rabbit hensin genomic clone was obtained by screening a rabbit genomic library (Stratagene) with a 500-bp DNA fragment from the 4.5-kb cDNA. The fragment that encodes the 5’ end of the 4.5-kb cDNA fragment was isolated by the screen and sequencing showed that it encoded the poly-A tail. To obtain the 5’-end of hensin, rapid amplification of cDNA ends was performed (5’-RACE, version 2.0; Gibco-BRL, Life Technologies) according to the supplied protocol. The validity of the 5’-end cDNA fragment of hensin was verified by Northern blotting with clone C total RNA and by sequencing.

Mouse hensin BAC clone. The SRCR1 domain of the rabbit genomic clone was amplified by PCR and used as a probe (accession no. AF078900, nt 4557–4914) for the screening of a mouse bacterial artificial chromosome (BAC) library under low stringency. The screening provided four positive clones, and they all gave the same results in all the studies described below. We also amplified CUB domains from the mouse BAC clone using PCR; the primers used were 5’-GTGGAG-GTTCCTGACCCICCTC-3’ and 5’-GAGTAGTAGTIGA-GCTIAACCCCTCTC-3’.

Southern and Northern hybridization. The DNA sample was electrophoresed in 0.7% agarose gel in 1× TAE buffer (Tris, acetate, EDTA) and then blotted onto Hybond-N (Amersham). The membrane was hybridized with the probes in the Express Hyb solution (Clontech, Palo Alto, CA) according to instructions supplied by the manufacturer. The membranes were washed and exposed to X-ray film using an intensifying screen at ~70°C. Probe DNAs were radiolabeled with [α-32P]dCTP (DuPont-New England Nuclear, Cambridge, MA) using Multiprime (Amersham). Northern blots were performed using membranes that contained RNA from different tissues or from different ages of mouse embryos (Clontech). The probes used for hybridization were made by PCR as follows: SRCR1 (GenBank accession no. AF078900, nt 4557–4914); CUB2 (AF043112, nt 3572–3887); 3’-untranslated regions (3’-UTR) of mouse CRP-ductin was amplified using the mouse small intestine cDNA library (U37438, nt 6017–6647); 3’-UTR of rabbit hensin was amplified from the 4.5-kb cDNA (AF043112, nt 4692–5108). The primers used for the amplification of the 3’-UTR were 5’-GGCTCCTACCAAG-GAAAGTGGGA-3’ and 3’-NdesT122. All PCR products were verified by sequencing.

Generation of domain-specific antibodies and immunoblot methods. Rabbit SRCR domains 6 and 7 (SRCR6-7) were cloned and expressed as fusion proteins in bacteria, and polyclonal antibodies were generated in guinea pigs. The antibodies recognized hensin by immunoblot, immunoprecipitation, and immunofluorescence, as described previously (26, 30). New antibodies against mouse SRCR1 were generated by first subcloning the open-reading frame of SRCR1 from the mouse hensin genomic clone followed by PCR based amplification. BamH 1 and Hind III sites were introduced at the 5’ and 3’ ends of the open-reading frame. The primers sequences were 5’-CGCGGATCCCTGTGTGTAATGCAGGTT-3’ and 5’-CCAAAGCCTTGAGCAAAAATCGGGGTGTTCTCGTGAGCC-3’. The BamH 1-Hind III fragment was cloned in frame into the corresponding sites of pQE-30 (Qiagen). Recombinant proteins were expressed in Escherichia coli strain M15 and purified using a nickel chelate affinity resin following the instructions of the manufacturer. Purified proteins were used to generate polyclonal antibodies in rabbits. Rabbit immune sera were passed through beads coupled to nonimmune rabbit serum to reduce nonspecific binding.

Kidney papilla, small intestine, and brain were freshly prepared from mouse or rabbit. Each tissue was homogenized with a Polytron in PBS with a protease inhibitors cocktail as described previously (26). The postmitochondrial fraction was obtained by centrifugation at 10,000 g for 10 min followed by one at 10,000 g for 10 min. The intercalated cell line, clone C, was cultured as described before (10). Human glioblastoma U87-MG, U118-MG, and U373-MG were obtained from ATCC (Manassas, VA) and were cultured in flasks according to the protocol provided. Confluent monolayers were solubilized in PBS with 1% Triton X-100 and used as the cell extract, and the proteins remaining on the flask were solubilized in 0.5% SDS and was used as the ECM. For SDS-PAGE, proteins (2 μg) were loaded on a 4–20% gradient acrylamide gel or a 7% acrylamide gel. Western blotting was performed as previously described (26). Signals were detected by an enhanced chemiluminescence system (ECL; Amersham International, Buckinghamshire, UK).

Chromosomal localization. Fluorescent in situ hybridization was used to localize the chromosomal position in mouse
HENSIN AND DMBT1 ARE ENCODED BY THE SAME GENE

RESULTS

Cloning of hensin cDNA. We obtained the cDNA sequence of rabbit hensin from a combination of 5′-RACE and screening of a rabbit intercalated cell cDNA library. The 4,782-bp open-reading frame was followed by a 296-bp 3′-UTR. The 5′ end of the cDNA had an ATG codon preceded by an appropriate Kozak sequence (CACCAG). Two potential polyadenylation signals (AATAAA) were located 14 and 59 bp upstream of the poly-A tail. The open-reading frame encoded a protein with 1,594 deduced amino acids and a calculated molecular mass of 172,820 Da (Fig. 1A). The deduced sequence contained all nine proteolytic peptide sequences (underlined in Fig. 1A), determined from the purified hensin protein (26). There were 17 potential N-glycosylation sites. The sequence contained a putative signal peptide, eight SRCR domains, six SID domains, two CUB domains, and one ZP domain with a short stretch of 3′-end amino acid sequence. The 5′ end of hensin consisted of seven tandemly repeated SRCR domains (13, 20). These SRCR domains had a high level of amino acid identity with each other, varying between 64% and 99% (Fig. 1B). SRCR domains are composed of ~100 amino acids containing many cysteines (13, 20). They have been found in a wide variety of secreted and membrane proteins and likely mediate protein:protein interactions (19). Each SRCR domain was separated by a short amino acid stretch previously termed SID domain (“SRCR interspersed domains”; Ref. 17) or CRP domain (6). The last SRCR was located between the two CUB domains (4). The two CUB domains shared a 39% identity at the amino acid level (Fig. 1B). CUB domains were first identified in tolloid, a Drosophila protein that is involved in the signaling pathway of decapentaplegic, the Drosophila homolog of transforming growth factor-β (TGF-β). The ZP domain (5), was present at the 3′ end of the cDNA. This domain is thought to function as sperm receptors.

The calculated molecular mass of the protein was ~173 kDa, a value different from that reported for rabbit kidney hensin, 230 kDa (26). Because there are many potential N-glycosylation sites in the predicted molecule, we examined the effects of glycosylation on the relative mobility of hensin on SDS-PAGE. Clone C cells were pulse-labeled with [35S]methionine for 5 min with [35S]methionine for 5 min in the presence or absence of tunicamycin, an inhibitor of N-glycosylation, and chased for 1 h followed by solubilization and immunoprecipitation with anti-hensin antibodies. Tunicamycin treatment decreased the relative molecular mass of the immunoprecipitated hensin from 230 to 180 kDa (unpublished results).

Database searches revealed four other proteins whose sequences were also composed of SRCR, CUB, and ZP domains; their structures are schematized in Fig. 2. Human DMBT1, a gene located in chromosome 10q25-26 that is frequently deleted in malignant gliomas (17), has the same domain arrangement as that of rabbit hensin except for an additional SRCR domain. Mouse CRP-ductin α- and β-subtypes were identified as cDNAs highly expressed in intestinal crypts and have seven SRCR domains, five CUB domains, and one ZP domain (6). cDNAs of CRP-ductin α-subtype and rat ebnerin (15) have a putative transmembrane domain at their 3′ end. But all protein products so far reported, including hensin, CRP-ductin, and rat ebnerin, were secreted proteins. Table 1 shows the percent nucleotide identity of rabbit hensin domains to those from other published cDNAs. Although the domains in these proteins showed a high level of homology (ranging from 63% to 93%; Table 1), it was difficult to ascertain by direct sequence comparisons whether they constituted four independent genes or were alternatively spliced versions of the same gene, because the cDNAs were cloned from different species. The large size of the DMBT1 gene (65 kb) and the presence of many homologous repeated domains precluded full-length genomic sequencing. To resolve this question, we sequenced regions of rabbit and mouse genomic hensin and concluded that the five cDNAs represent alternately spliced variants of the same gene. The nomenclature of the domains we use in this study reflects these conclusions; as will be shown below, the gene includes at least nine SRCR domains, six CUB domains, and one ZP domain. This nomenclature is tentative pending the sequencing of the whole gene.

We used RT-PCR 5′-RACE reactions on total RNA from the kidney intercalated cell line, as well as from the clone C cDNA library and rabbit kidney cDNA library to clone the 5′-end of rabbit hensin cDNA using a gene-specific primer located in SRCR2. The 5′-RACE gave three different products that had the same putative signal sequence and SRCR2, but different intervening sequences (Fig. 3, top). The results prompted us to resolve the exon arrangement of rabbit hensin gene to determine whether the three 5′-RACE products were derived from one gene or three genes. Rabbit hensin genomic sequence. A 12-kb rabbit hensin partial genomic clone was isolated from a rabbit genomic library by screening with a probe that encoded SRCR3 of hensin cDNA (Fig. 3). The complete sequencing of this genomic clone revealed that it contained two SRCRs and one SID sequence found in hensin cDNA (100% identical). The results indicate that we have indeed obtained a fragment of rabbit hensin gene. Unfortunately, the fragment did not contain the sequence of the 5′-end putative signal peptide. Instead, it had all the intervening sequences found in the three 5′-RACE products described above. The genomic clone had seven exons found in the hensin cDNA, and the potential splice donor and splice acceptor sites of these exons are shown in Table 2.

Surprisingly, the genomic clone also had an unknown SRCR (now termed SRCR1; Fig. 3, box with hatched lines) upstream of SRCR2, which had an authentic splice donor and acceptor sites (Table 2). The SRCR1 of hensin and DMBT1 resemble each other more than they do other SRCR domains (see below). Rabbit hensin...
Fig. 1. A: deduced amino acid sequence of rabbit kidney hensin cDNA. Putative signal peptide is underlined doubly. SRCR ("scavenger receptor, cysteine rich") domains are shown in green, CUB ("C1r/C1s Uegf Bmp1") domains are in gray, and ZP ("zona pellucida") domains are in yellow. Nine peptide sequences determined from the isolated protein are underlined singly, in bold (26). Sequence presented here represents splice variant type I as shown in Fig. 3 below. GenBank accession number is AF043112.

B: multiple-sequence alignment of the SRCR and CUB domains in hensin. Consensus amino acids are indicated (4, 20). Identical residues are shaded in gray; h, hydrophobic; t, turn-like; a, aromatic; 1/2, charged.
cDNA and human DMBT1 cDNA have the same domain arrangements except for one additional SRCR domain in DMBT1 (Fig. 2), and they share a high level of homology from 5' to 3' end (Table 1).

Mouse hensin BAC clone. We screened at low stringency a mouse BAC library using rabbit hensin SRCR1, one of the most divergent SRCRs, to reduce the possibility of hybridization with the many other proteins that contain SRCR domains. Four independent BAC clones were identified, all of which hybridized with SRCR, CUB, and 3'-UTR DNA probes from rabbit hensin under low stringency, suggesting that the BAC clones likely contained the full-length hensin gene (Fig. 4A).

To estimate the size of the four BAC clones, we digested each with EcoRI and HindIII. Addition of the apparent molecular mass of all the fragments gave results that varied between 113 and 150 kb. Analysis of the EcoRI digestion pattern showed that 11 bands had the same molecular mass in the four clones. Addition of the apparent molecular masses of these common fragments gave a total of 73 kb. Similarly, in the HindIII digestion studies, addition of the common bands gave an apparent molecular mass of 68 kb (data not shown). These results are compatible with the conclusions that the mouse hensin gene spans a region of ~70 kb.

Digestion of the BAC clone with EcoRI and hybridization with a rabbit genomic SRCR1 probe showed one strong band with several weaker bands (Fig. 4A). To estimate the size of the four BAC clones, we digested each with EcoRI or HindIII. Addition of the apparent molecular mass of all the fragments gave results that varied between 113 and 150 kb. Analysis of the EcoRI digestion pattern showed that 11 bands had the same molecular mass in the four clones. Addition of the apparent molecular masses of these common fragments gave a total of 73 kb. Similarly, in the HindIII digestion studies, addition of the common bands gave an apparent molecular mass of 68 kb (data not shown). These results are compatible with the conclusions that the mouse hensin gene spans a region of ~70 kb.
probed with a 41-bp oligonucleotide that encodes the putative signal peptide of mouse CRP-ductin, they showed only one 10-kb band (data not shown). These results suggest that the mouse BAC clone isolated likely encodes the complete mouse hensin gene.

Mouse CRP-ductin cDNA does not contain an SRCR1 domain (6). To examine whether the BAC clone containing the putative mouse hensin gene also encodes CRP-ductin, we first determined the sequence of four SRCR sequences from the mouse hensin BAC clone after HindIII digestion. They contained mouse CRP-ductin SRCR3 (97% nt identity), SRCR5 (100% nt identity), SRCR7 (99% nt identity), and SRCR8 (99% nt identity) (GenBank AF079225–AF079228). Second, CUB domains were amplified by PCR from the hensin BAC clone using primer pairs designed from CRP-ductin CUB domains (Fig. 4B). Sequencing of the 1.7- and 1.8-kb PCR products revealed that the BAC clone had CUB2 and CUB4 of CRP-ductin. We also found that both the 3'-UTR of hensin and that of CRP-ductin were present in the BAC clone by Southern blotting, implying that both genes are present in the same BAC clone (Fig. 4C).

We compared the 3'-UTR sequences of rabbit hensin and human DMBT1 and found that they were 68% identical (Table 1), further supporting the conclusion that they might be products of the same gene. The 3'-UTR of mouse CRP-ductin shares 82% identity with that of rat ebnerin, suggesting that these two cDNAs might originate from the same gene; but the 3'-UTR of rabbit hensin had minimal homology to that of mouse CRP-ductin. We used the 3'-UTR of rabbit hensin and that of mouse CRP-ductin to probe a blot of the BAC clone after EcoRI and/or Hind III digestion. Both probes recognized the same band, but when the BAC clone was further digested with EcoRI1 or PstI, the two probes gave differential hybridization patterns (Fig. 4C). This suggests that the 3'-UTR of hensin and CRP-ductin are present in distinct regions within a 6-kb fragment of the BAC clone. Furthermore, it also suggests that the genes for CRP-ductin and hensin are not present in tandem on the same large BAC clone.

An alternative explanation is that CRP-ductin and hensin/DMBT1 are present in this mouse BAC clone in tandem but in opposite directions; if so, then their 3'-UTRs could still be located in a single 6-kb DNA fragment. We think this is unlikely, based on a quantitative analysis. Using a CUB6 fragment that hybridizes with both hensin/DMBT1 and CRP-ductin domains, we found that all CUB domains and both 3'-UTR domains were present in two Hind III fragments of 10 and 15 kb in size (data not shown). As shown in Fig. 2, there are at least five CUB domains in the 3' end of CRP-ductin (6), each of which is composed of two exons and an intervening intron, which on average is ~1.5 kb in length (this study, GenBank

---

**Table 1. Domain comparison among four cDNAs**

<table>
<thead>
<tr>
<th>Domain</th>
<th>Hensin vs. DMBT1</th>
<th>Hensin vs. Ebnerin</th>
<th>DMBT1 vs. Ebnerin</th>
<th>CRP-ductin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full length</td>
<td>78</td>
<td>74</td>
<td>71</td>
<td>75</td>
</tr>
<tr>
<td>Putative signal</td>
<td>83</td>
<td>76</td>
<td>86</td>
<td></td>
</tr>
<tr>
<td>SRCR 1</td>
<td>72</td>
<td>84</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>SRCR 2</td>
<td>80</td>
<td>80</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td>SRCR 3</td>
<td>80</td>
<td>75</td>
<td>82</td>
<td></td>
</tr>
<tr>
<td>SRCR 4</td>
<td>81</td>
<td>77</td>
<td>82</td>
<td></td>
</tr>
<tr>
<td>SRCR 5</td>
<td>81</td>
<td>78</td>
<td>83</td>
<td></td>
</tr>
<tr>
<td>SRCR 6</td>
<td>81</td>
<td>78</td>
<td>78</td>
<td>83</td>
</tr>
<tr>
<td>SRCR 7</td>
<td>80</td>
<td>78</td>
<td>80</td>
<td>81</td>
</tr>
<tr>
<td>SRCR 8</td>
<td>82</td>
<td>80</td>
<td>82</td>
<td>81</td>
</tr>
<tr>
<td>SRCR 9</td>
<td>77</td>
<td>67</td>
<td>63</td>
<td>65</td>
</tr>
<tr>
<td>CUB 1</td>
<td>79</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CUB 2</td>
<td>89</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CUB 3</td>
<td>87</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CUB 6</td>
<td>81</td>
<td>76</td>
<td>75</td>
<td>82</td>
</tr>
<tr>
<td>ZP</td>
<td>81</td>
<td>81</td>
<td>78</td>
<td>81</td>
</tr>
<tr>
<td>3'-UTR</td>
<td>68</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

- cDNA sequences of rabbit hensin, human DMBT1, mouse CRP-ductin and rat ebnerin were compared. Domains were compared using the bestfit program of the Genetics Computer Group package. See Fig. 2 for nomenclature of the domains.
AF079229 and AF079230). The 3'-UTR of CRP-ductin is 400 nt long. This region is thus predicted to be at least 20 kb in size. Since this mouse genomic BAC clone contains the 5' end of hensin/DMBT1 (i.e., SRCR1), it is also predicted to contain all hensin CUB domains. Hence, the total amount of DNA containing the 3' ends of CRP-ductin and that of hensin/DMBT1 (in reverse orientation) would be at least 40 kb and probably much longer. Mouse BAC clones digested with Hind III produced a pattern, suggesting that there is sufficient DNA for only one gene containing CUB domains rather than two. These results, in the aggregate, suggest that hensin/DMBT1 and CRP-ductin are alternatively spliced products from the same gene.

Chromosomal localization of mouse hensin. DMBT1 was identified as a locus in human chromosome 10q25–26 that is often deleted in malignant gliomas, and this region is also implicated in breast and prostate cancer. Human DMBT1 is located between the fibroblast growth factor receptor 2 (FGFR2) and the anonymous marker D10S587. FGFR2 was localized to the terminal region of mouse chromosome 7 (3). We probed mouse chromosomes with the hensin BAC clone using fluorescent in situ hybridization (Fig. 5). We found that hensin was localized to chromosome 7 band F4. These results are consistent with the conclusion that hensin and DMBT1 are products of the same gene.

Table 2. Predicted splice sites in rabbit hensin gene

<table>
<thead>
<tr>
<th>Exon</th>
<th>Donor Site</th>
<th>Acceptor Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exon 2</td>
<td>5'-TTCTACAGGTCACTGGG</td>
<td>TTCTACAGGTCACTGGG</td>
</tr>
<tr>
<td>Exon 3</td>
<td>5'-TGCTACAGGTATGGG</td>
<td>TGCTACAGGTATGGG</td>
</tr>
<tr>
<td>Exon 4</td>
<td>5'-ACAGAAAGGTAAATGG</td>
<td>ACAGAAAGGTAAATGG</td>
</tr>
<tr>
<td>Exon 5</td>
<td>5'-ACCACAGGTAGAGTG</td>
<td>ACCACAGGTAGAGTG</td>
</tr>
<tr>
<td>Exon 6</td>
<td>5'-TGCTACAGGCTGGTGCG</td>
<td>TGCTACAGGCTGGTGCG</td>
</tr>
<tr>
<td>Exon 7</td>
<td>5'-CCACACAGGTAGCTGC</td>
<td>CCACACAGGTAGCTGC</td>
</tr>
<tr>
<td>Exon 8</td>
<td>5'-ACCCACAGGTAGCTGG</td>
<td>ACCCACAGGTAGCTGG</td>
</tr>
<tr>
<td>Exon 9</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>

A 12-kb rabbit hensin gene was isolated and sequenced (GeneBank accession no. AF007890). The arrangement of exons in the genomic DNA is shown in Fig. 3. The 5' and 3' splice sites of each exon were predicted by a neural network program provided on the world wide web (http://www-hgc.lbl.gov/projects/splice.html). ND, not determined.
distinguish between two homologous sequences. The longer the arm, the greater is the number of substitutions. The result of this analysis shows that SRCR9 is the most divergent, followed by SRCR1. In the rest of the SRCR domains, the neighbors had low bootstrap values (in DMBT varying between 21 and 51), indicating that the analytic program could not construct a tree that uniquely determined the degree of relatedness of each of the groups of two neighbors because they were too similar to each other. Interestingly, all of the SRCR domains fell into the same or similar groupings among species, suggesting that the gene contained at least nine SRCR domains. This was not the case for CUB domains, where it became clear that there were at least six recognizable domains, whereas the cDNA of CRP-ductin had five. Of these domains, CUB1 was only present in DMBT1 and hensin, whereas CUB6 was present in all four cDNAs. The results of Fig. 4A are in accord with this conclusion, since they show that the EcoR I digest of the mouse BAC clone probed with CUB6 produced six hybridizing bands.

Expression of hensin transcripts. Expression of the hensin transcript was examined by Northern blotting in multiple tissues of adult mouse. The SRCR2 probe recognized one major band of ~7 kb (Fig. 7A). This confirmed the observations of Cheng et al. (6), who detected only one band of ~7 kb as the message of CRP-ductin. The expression was highest in liver, modest in spleen, but absent in heart, brain, lung, skeletal muscle, testis, and kidney. However, we also detected minor bands at 4.5 and 2.0 kb in brain, skeletal muscle, kidney, and testis. The hensin probe hybridized with a major 7-kb band in whole embryo poly(A)⁺ RNA at day 17 (Fig. 7B). Minor bands at 5.0 and 3.0 kb were also observed from embryonic day 7 to embryonic day 17. We also performed RT-PCR in mouse kidney using specific primers and subcloned and sequenced the product. We found that it encoded CRP-ductin sequences.

We examined the various protein products of hensin gene by Western blotting in rabbit, mouse, and human tissues. Using the antibody against rabbit SRCR6 (26), we found that rabbit small intestine contained two bands of M₆ of ~230 kDa, which was the same size as that of rabbit kidney hensin (Fig. 8A). The same antibody, however, recognized a 270-kDa band in mouse small intestine, the same value reported for CRP-ductin in mouse small intestine (6). The antibody detected a strong signal, ~150 kDa, and weak signals, ~230 kDa, in both rabbit brain and mouse brain. We also tested human glioma cell lines that were used in the DMBT1 study (17). No signal was detected in the cell lysates, but the ECM fraction clearly contained a 200-kDa band (Fig. 8B). The extracellular localization of this band was similar to those obtained in rabbit kidney intercalated cell line (26, 29). These results suggest that the alternative splicing of hensin gene produce many variants in different tissues and different animals. However, species- and tissue-specific glycosylation likely affects the relative mobility of hensin on the SDS-PAGE as well, not to mention protease effects that are insensitive to the inhibitors used in the solubilization mixture.

We then compared the staining pattern of two different antibodies in immunoblots to differentiate the potential alternative spliced products. The first antibody was raised against rabbit hensin SRCR6+7 cDNA, which was expected to act as a pan-hensin antibody
(since all cDNAs contained these two domains, Fig. 2). The second antibody was generated against a fusion protein of mouse genomic hensin SRCR1. These antibodies were expected to differentiate the spliced products in various tissues, since SRCR1 was divergent from the other SRCRs among human, rabbit, and mouse (Fig. 6) and since SRCR1 is transcribed in human brain DMBT1, but not in rabbit kidney hensin, mouse small intestine CRP-ductin, or rat ebnerin (Fig. 2). Mouse SRCR1 used for generation of the antibody shared 62% amino acid identity with the SRCR6+7 fusion protein. Despite this homology, the anti-SRCR1 antibody only faintly reacted with the SRCR6+7 fusion protein. The rabbit hensin SRCR6+7 antibody recognized a 230-kDa band in the ECM from clone C (Fig. 8C) and a 260-kDa band and a smaller band of ∼160 kDa in mouse small intestine and mouse brain, respectively. Since we did not detect a positive signal in mouse kidney papilla (using up to 150 µg protein) in preliminary experiments, we enriched the hensin protein by immunoprecipitation as described previously (26). The SRCR6+7 antibody detected a positive signal in mouse kidney papilla at the same mobility as that in mouse brain. The expression level of hensin was clone C > small intestine > brain > kidney papilla. The same membrane was reprobed with the antibody against the mouse SRCR1 (Fig. 8C). The antibody recognized a strong band ∼260 kDa in mouse brain. It, however, detected no positive signal in other tissues, including the ECM of clone C. We suggest that two types of hensin variants were expressed in the mouse brain by a cell-specific splicing mechanism: a large hensin that uses the SRCR1 (likely to be the DMBT1-like variant), and a smaller variant devoid of the SRCR1. The results further suggest that these domain-specific antibodies (and other currently being produced against CUB and ZP domains) will be useful to examine the differential expression and tissue-specific localization of the spliced variants from the hensin gene.

Using immunocytochemistry, we recently found that the SRCR6+7 antibodies stained rabbit prostate (30), we performed Western blotting experiments using human prostate cancer cell lines. We found that the rabbit SRCR6+7 antibody recognized 230- and 60-kDa bands, whereas the anti-SRCR1 antibody recognized only a 50-kDa band in the same cell lines (data not shown). Table 3 summarizes the results of Northern and West-
Discussion

Hensin is a modular protein composed of SRCR, CUB, and ZP domains. The SRCR motif, originally identified in the macrophage scavenger receptor, is a 110-residue cysteine-rich domain found in many secreted and membrane proteins (2, 13, 20). The cysteines form intramolecular disulfide bonds (19). There are two types of SRCR domains: group A SRCRs contain six cysteines and are encoded by two exons, whereas group B have eight cysteines and are encoded by a single exon (2). Like all group B, the SRCRs of hensin had eight cysteines, and the domain was encoded by a single exon as shown by genomic analysis (Fig. 3). The function of the domain is still unclear, but the membrane-proximal SRCR in CD6 was shown to bind its ligand, activated leukocyte cell adhesion molecule (31). Hensin has two copies of another cysteine-rich domain, termed CUB, which was first noted in the Drosophila tolloid, a protein that binds to decapentaplegic (dpp), the Drosophila homolog of TGF-β (4, 7). Tolloid contains an astacin-type protease as well as a CUB domain. It activates dpp, and mutations in its CUB domain also results in a phenotype similar to that of dpp mutants (7, 12). The third cysteine-rich domain present in hensin is the ZP domain (5), first found in proteins of the zona pellucida. They function as presumed sperm receptors and have some similarity to the TGF-β receptor type III.

Hensin, DMBT1, CRP-ductin, and ebnerin are proteins that contain three domains (SRCR, CUB, and ZP) in different combinations (Fig. 2). We presented several lines of evidence that these proteins are likely to be encoded by the same gene. First, the putative signal sequences of hensin, DMBT1, and CRP-ductin shared a high level of homology (Table 1). Second, a partial genomic sequence of the rabbit hensin gene demonstrated the presence of an additional SRCR, upstream of the other SRCRs, but this domain was absent in the hensin cDNA. This new domain, SRCR1, was homologous to the transcribed SRCR1 domain of DMBT1 (Fig. 6). Similarly, SRCR1 was not transcribed in CRP-ductin cDNA (6). However, the mouse BAC clone that encodes CRP-ductin contained an SRCR1. Third, an antibody raised against the SRCR1 failed to recognize hensin in clone C, kidney, or hensin/CRP-ductin in small intestine, but labeled a presumed brain hensin/DMBT1 (Fig. 8C). Fourth, hensin was localized to a region of mouse chromosome 7 syntenic to chromosome 10q25–26, the locus of DMBT1. Fifth, the 3'-UTRs of rabbit hensin and human DMBT1 were remarkably similar given the interspecies distance (Table 1). These results suggest that hensin, CRP-ductin, and DMBT1 are products of the same gene.

Probing a digest of mouse hensin gene in a BAC clone with a rabbit SRCR1 domain showed seven hybridizing bands, and the sequences of four of these bands demonstrated that they were identical to four SRCR domains of CRP-ductin. Although this suggests that hensin/DMBT1 and CRP-ductin might be encoded by the same gene, it is still possible that the large BAC clone could contain the two genes in tandem. To resolve this, we examined the 3'-UTR sequences of hensin/DMBT1 and CRP-ductin, sequences which do not share any significant homology (Table 1). The two 3'-UTRs were present in the same 6-kb fragment of the BAC clone, suggesting that hensin/DMBT1 and CRP-ductin are encoded by the same gene. The 3'-UTR of ebnerin is quite similar to that of CRP-ductin as are the CUB domains and...
SRCR domains. Hence, we suggest that all four sequences are encoded by the same gene. It had already been demonstrated that this gene can readily produce several transcripts; we showed that hensin was alternatively spliced in its 5′ end (Fig. 3), and Cheng et al. (6) cloned two different forms of CRP-ductin cDNA from mouse small intestine one with a transmembrane domain in its 3′ end and another lacking it. These

Fig. 7. Expression of hensin transcript in mouse tissues. A: multi-tissue Northern blotting. B: embryo Northern blotting. Northern blotting was performed with rabbit hensin SRCR1 probe under low stringency. The 3′-UTR probes from mouse CRP-ductin and rabbit hensin gave the same results. Two micrograms of poly(A)+ RNA was loaded on each lane. Large dot around 3 kb in day 17 embryo RNA was a contamination of [α-32P]dCTP.

Fig. 8. Expression of hensin protein in various tissues. A: Western blot of small intestine and brain from rabbit or mouse. B: Western blot of human glioma cell lines. C: comparison of antibody against rabbit SRCR6+7 (αrSRCR6+7) with antibody against mouse SRCR1 (αmSRCR1). Extracellular matrix (ECM) fraction from clone C (2 µg), 100 µg of proteins from small intestine and brain isolated from rabbit or mouse, and the immunoprecipitated proteins from 9 mg of mouse kidney papilla were processed for SDS-PAGE and transferred to a nitrocellulose membrane. The same membrane was immunoblotted with antibodies against the rabbit hensin SRCR6+7, stripped and then rebotted with the anti-mouse SRCR1 antibody.
HENSIN AND DMBT1 ARE ENCODED BY THE SAME GENE

Table 3. Summary of various species of hensin/DMBT1/CRP-ductin mRNAs and proteins

<table>
<thead>
<tr>
<th>Tissue</th>
<th>Rabbit</th>
<th>Mouse</th>
<th>Human</th>
<th>Rat</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>kb</td>
<td>kDa</td>
<td>kb</td>
<td>kDa</td>
</tr>
<tr>
<td>Kidney</td>
<td>6</td>
<td>230</td>
<td>4.5, 2.0</td>
<td>160</td>
</tr>
<tr>
<td>Small intestine</td>
<td>6</td>
<td>230</td>
<td>7³</td>
<td>290³</td>
</tr>
<tr>
<td>Brain</td>
<td>200, 160</td>
<td>5, 2.0</td>
<td>200, 160</td>
<td>260, 230, 90⁶</td>
</tr>
<tr>
<td>von Ebner’s gland</td>
<td></td>
<td></td>
<td>6, 7.5</td>
<td>230, 160, 120, 60⁴</td>
</tr>
<tr>
<td>Prostate</td>
<td></td>
<td></td>
<td>8³</td>
<td>230, 60, 50⁴</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5³</td>
</tr>
</tbody>
</table>

⁴Ref. 6. ⁵Detected by mSRCR 1 antibody. ⁶Ref. 17. ⁷Cultured cell line. ⁸Ref. 15.

Observations suggest that alternative splicing occurs throughout the gene in a complex manner and results in a variety of cDNA species including hensin, DMBT1, CRP-ductin, and ebnerin. However, Northern blotting of multiple tissues in rabbit and mouse detected only one major transcript of hensin (this study and Ref. 26). Other investigators also reported one 7-kb message for mouse CRP-ductin (6) and one 5-kb species for rat ebnerin (15). In contrast, three sizes of transcripts for DMBT1 (8.0, 7.5, and 6.0 kb) were detected in human (17). The transcripts of hensin, DMBT1, and CRP-ductin are easily detected in the small intestine (6, 17, 26). DMBT1 and ebnerin were highly expressed in the lung, whereas those of hensin and CRP-ductin were not detected. In the kidney, hensin is expressed only in the collecting duct, and total kidney RNA contains very little hensin message. However, we were able to detect several minor transcripts of smaller size in many tissues (Fig. 7 and Table 3). The wide variation in the reported size and intensity of hybridization signals in different tissues is, hence, likely to be due to its low abundance in the total tissue RNA and/or species differences. Similarly, the anti-hensin antibody recognized the distinct protein bands in different tissues of the same animal and in the same tissue of different animals (Fig. 8). That the domain-specific antibodies raised against SRCR1 and SRCR6+7 detected mutually exclusive bands in mouse brain and lower-molecular-weight bands (in human glioma and prostate cell lines) reinforces the idea that the alternative splicing brought about a complex sets of hensin protein products. It further raises a new question about the function of the different spliced variants in distinct tissues. Further experiments are clearly needed to resolve the cell-specific and the species-specific splicing pattern of hensin.

The function of DMBT1, CRP-ductin, and ebnerin is unknown; however, we have previously demonstrated that hensin is a critical protein that causes terminal differentiation of the intercalated cell and perhaps other epithelia (30). Hensin was originally identified as a protein that caused reversal of the polarity of an immortalized intercalated cell line (26, 29). These cells, in the intact kidney, are responsible for acidification and alkalinization of urine. They exist in a spectrum of differentiation states with one extreme, the β-cell (i.e., low-density phenotype) was associated with a number of other changes including columnarization, i.e., an increased height and decreased cross-sectional area. Furthermore, there was reorganization of the apical cytoskeleton with the appearance of exuberant microvilli in the high-density cells as well as apical endocytosis (30). All of these changes were produced by hensin, since culturing low-density cells onto the matrix deposited by high-density cells reproduced the findings. Furthermore, anti-hensin antibodies prevented the development of apical endocytosis, columnarization, and the rearrangement of the apical cytoskeleton (26, 30).

Remodeling of the apical cytoskeleton and columnarization are hallmarks of terminal differentiation of epithelia. In the intestine, a site of high expression of hensin, crypt cells, which do not express microvillar proteins, develop exuberant microvilli as they differentiate to become villus absorptive cells (16). The roles that hensin or its other alternately spliced forms (CRP-ductin and DMBT1) play in the differentiation of the intestine and other epithelial tissues is an intriguing one and requires more direct examination. It is interesting in this regard that CRP-ductin was identified as a cDNA that was downregulated by ultraviolet irradiation, a procedure that promotes dedifferentiation (6). PCR analysis suggested that DMBT1 was widely expressed in the central nervous system. Neurons are polarized cells, and it had been proposed that axons and dendrites are analogous to the apical and the basolateral domains of epithelial cells, respectively (8). Furthermore, glia are known to control the specific functions of many neural elements (18). That deletion of DMBT1 was frequent in medulloblastoma and glioblastoma multiforme (17) raises the possibility that hensin/DMBT1 is a tumor suppressor gene. It is well-known that interruption of many terminal differentiation pathways can lead to tumorigenesis.
REFERENCES


