Molecular cloning and expression of a new rat liver cell-CAM105 isoform

Differential phosphorylation of isoforms

Ognjen CULIC,* Qi-Hui HUANG,* Donna FLANAGAN,† Douglas HIXSON† and Sue-Hwa LIN*†

* Department of Molecular Pathology, University of Texas, M.D. Anderson Cancer Center, Houston, TX 77030, U.S.A., and † Department of Medical Oncology, Rhode Island Hospital, Brown University, Providence, RI 02903, U.S.A.

An hepatocyte cell-adhesion molecule (cell-CAM105) was recently shown to be identical with the liver plasma-membrane ecto-ATPase. This protein has structural features of the immunoglobulin superfamily and is homologous with carcinoembryonic antigen proteins. We have cloned a cDNA encoding a new form of the cell-CAM105 which is a variant of the previously isolated clone. In addition to having a shorter cytoplasmic domain, the new isoform also has substitutions clustered in the first 130 amino acids of the extracellular domain. Both of these isoforms are expressed on the surface of hepatocytes with the shorter variant being the predominant form. The previously isolated cell-CAM105 (long form) has more potential phosphorylation sites than does the new isoform (short form). Both isoforms are found to be phosphorylated after incubation with [³²P]phosphate *in vitro*, with the long form being phosphorylated to a significantly higher extent. This observed differential phosphorylation could be one of the mechanisms for the regulation of isoform functions. Using anti-peptide antibodies specific for the long form and antibodies that are reactive with both isoforms, we have shown that both isoforms are localized in the canalicular domain of hepatocytes. The sequence differences between these two isoforms suggest that they are probably derived from different genes rather than from alternative splicing.

INTRODUCTION

In a previous report, we described the isolation of a full-length cDNA for rat liver ecto-ATPase which was later shown to be identical with a liver cell-adhesion molecule, cell-CAM105 (Lin & Guidotti, 1989; Aurivillius *et al.*, 1990; Lin *et al.*, 1991). Cell-CAM105 is an integral plasma-membrane glycoprotein with an apparent M_r of 105000. Two-dimensional gel-electrophoretic analysis showed cell-CAM105 to be composed of at least two acidic glycoproteins different in pI (4.1 versus 4.3) and in apparent M_r (i.e. 110000 versus 105000) (Hixson & McEntire, 1989; McEntire *et al.*, 1989). These different forms could have resulted from proteolysis during sample preparation, from variations in glycosylation or from expression of different genes. In studies with antibodies against C-terminal peptide sequences, we have found that these antibodies only immunoprecipitated the higher- M_r form (Lin *et al.*, 1991). This result suggests that at least one of the differences between the higher- and lower- M_r forms might be in the C-terminal portion of the proteins. We now report the isolation of a second clone which was shown by sequence analysis to be a variant of the previously isolated cell-CAM105, indicating that liver cell-CAM105 is composed of at least two isoforms. In addition to having a shorter cytoplasmic domain, the new isoform also has substitutions clustered in the first 130 amino acids of the extracellular domain. The sequence differences between these two isoforms suggest that they are probably derived from different genes rather than from alternative splicing.

MATERIALS AND METHODS

Materials

Peptides were synthesized according to the deduced amino acid sequence at an institutional protein core facility (M. D.

Anderson Cancer Center). Horseradish peroxidase-conjugated anti-rabbit antibodies were purchased from Bio-Rad (Richmond, CA, U.S.A.). Keyhole limpet haemocyanin (KLH) was purchased from Sigma Chemical (St. Louis, MO, U.S.A.). Protein A-Sepharose 4B was purchased from E-Y Laboratory (San Mateo, CA, U.S.A.).

Isolation and characterization of cDNAs and DNA sequence analysis

A rat liver cDNA library in λ ZAP was screened with a nick-translated fragment excised from the cell-CAM105 cDNA (*Bam*HI–*Pst*I fragment). A total of 10⁶ plaques were screened. The conditions used for hybridization and washing of filters were as described in Maniatis *et al.* (1982). Forty positive clones were obtained. After excision of phagemids *in vivo* from λ phage, the nucleotide sequence of double-stranded DNA was determined using the T₇ sequenase kit (United States Biochemical Corp., Cleveland, OH, U.S.A.) according to the procedures provided by the manufacturer.

Preparation of isoform-specific anti-peptide antibodies

Anti-C1, anti-C2 and anti-C3 antibodies were prepared against peptide sequences within the C-terminal portion of the long-form (L-form) cell-CAM105 (Lin & Guidotti, 1989). Anti-C1 is against peptide sequence CDLTEHKPSTSSHNLP (amino acid positions 460–475), anti-C2 is against peptide sequence CDSPNKVDDVSYSVLNFNAQ (amino acid positions 478–496) and anti-C3 is against peptide sequence CQSKRPTSASSPTETVYSVVKKK (amino acid positions 497–519). Peptides were coupled to KLH using maleimido-benzoyl-*N*-hydroxysuccinimide ester (MBS) as described by Green *et al.* (1982). Briefly, 8 mg of KLH in 0.25 ml of 10 mM-sodium phosphate buffer (pH 7.2) was allowed to react with

Abbreviations used: CAM, cell-adhesion molecule; KLH, keyhole limpet haemocyanin; MBS, maleimidobenzoyl-*N*-hydroxysuccinimide ester; mAb, monoclonal antibody.

† To whom correspondence should be addressed.

1.4 mg of MBS, which was dissolved in 50 μ l of dimethyl-formamide, and stirred for 30 min at room temperature. This reaction mixture was then fractionated on a Sephadex G-25 column pre-equilibrated with 50 mM-sodium phosphate in order to remove unchanged MBS. The fractions containing KLH were combined (1 ml) and added directly to 10 mg of peptides. The reaction mixture was stirred overnight at room temperature. Antibodies to the KLH-coupled peptide were prepared by injecting subcutaneously one-tenth of the total mixture in complete Freund's adjuvant into rabbits. The rabbits received booster injections of a portion of the peptide/KLH mixture intramuscularly.

Expression of recombinant cell-CAM105 in COS cells

The full-length cell-CAM105 cDNA (L-form) in expression vector CDM8 was constructed as described previously (Lin & Guidotti, 1989). The full-length short-form (S-form) cDNA in CDM8 was constructed by insertion of full-length S-form into CDM8 vector through *XbaI*-*PstI* cloning sites. COS-M6 cells (2×10^6 /ml) were transfected with 25 μ g of plasmid DNA per 2×10^6 cells by electroporation using a Cell-Porator (Bethesda Research Laboratory, Life Technologies Inc., Gaithersburg, MD, U.S.A.). Briefly, COS-M6 cells (2×10^6 /ml) were suspended in Hepes-buffered saline containing 21 mM-Hepes (pH 7.05), 137 mM-NaCl, 5 mM-KCl, 0.7 mM- Na_2HPO_4 and 6 mM-glucose. For electroporation, COS-M6 cells in a volume of 1 ml were placed in the electroporation chamber, and the electroporation was performed at 250 V, 330 μ F. After electroporation, cells were removed from the electroporation chamber and placed in 10 ml of Dulbecco's minimum essential medium containing 10 % fetal calf serum. After incubation for 48–72 h, the culture medium was aspirated and the cells were scraped off the dishes with a 'rubber policeman'. The cells were disrupted by mixing with 0.5 ml of distilled water, followed by freezing and thawing to obtain total cell lysate.

Immunoprecipitation and immunoblot analysis of cell-CAM105 isoforms expressed in COS cells

Indirect immunoprecipitation by anti-peptide and anti-cell-CAM105 sera of isoforms expressed in COS cells was performed by solubilizing COS cells in Triton X-100 lysis solution (final concentrations: 0.5 % Triton X-100, 0.3 M-NaCl, 25 mM-sodium phosphate, pH 7.4, and 20 mM-octyl glucoside) for 30 min at 4 °C. After removal of insoluble material by centrifugation for 10 min at 13000 g, portions of detergent extract were mixed with 5 μ l of antiserum or non-immune serum. After an overnight incubation at 4 °C with constant agitation, 15 μ l of a 50 % suspension of protein A-Sepharose was added to each sample and it was further incubated at room temperature for 2 h with constant mixing. Protein A-Sepharose with bound immune complexes was collected by centrifugation (10000 g) in an Eppendorf microfuge for 15 s and washed three times (1 ml each) with 1 $\times$ lysis solution (0.5 % Triton X-100, 0.3 M-NaCl, 25 mM-sodium phosphate, pH 7.4, and 20 mM-octyl glucoside). Immunoprecipitated proteins were eluted by boiling for 5 min in SDS and 2-mercaptoethanol and submitted to SDS/PAGE. Transfer of proteins separated by SDS/PAGE on to nitrocellulose membrane filters was performed as described by Towbin *et al.* (1979). After transfer, the nitrocellulose filter was incubated for 30 min at room temperature in TBS (10 mM-Tris/HCl, pH 7.4, 150 mM-NaCl) containing 3 % non-fat dry milk powder. The nitrocellulose filter was further incubated with antiserum in the same solution overnight at 4 °C with constant mixing. The filter was then washed twice (1 min each) with TBS and incubated with horseradish peroxidase-conjugated goat anti-rabbit antibody in TBS containing 3 % non-fat dry milk powder for 2 h at room

temperature. The filter was washed twice (1 min each) with TBS and the colour was developed by using 4-chloro-1-naphthol as substrate for peroxidase.

Immunoprecipitation and two-dimensional gel-electrophoretic analysis

Normal adult rat hepatocytes were isolated by a modification of the collagenase perfusion technique of Bonney *et al.* (1974). Viability of isolated cells was greater than 80 % as determined by Trypan Blue dye exclusion. Surface labelling with ^{125}I was performed by the lactoperoxidase/glucose oxidase procedure of Keski-Oja *et al.* (1977). Procedures for immunoprecipitation analysis of radiolabelled antigens are described below.

To label hepatocytes biosynthetically with [^{32}P]phosphate, hepatocytes isolated by collagenase perfusion were suspended in RPMI medium supplemented with gentamicin and 10 % fetal calf serum. Hepatocytes (2×10^6) were plated into a 60 mm collagen-coated Petri dish. Cells were allowed to attach for 2 h at 37 °C. Unbound cells were removed by a single wash in medium. Cells were then incubated in phosphorus-free RPMI supplemented with SGF7 (Mowery & Hixson, 1991) and 25 mM-Hepes. After 1 h, 1 mCi of 1000 Ci/mmol [^{32}P]phosphate (Amersham) was added to each dish and cultures were maintained at 37 °C for 16 h. After two washes in labelling medium without [^{32}P]phosphate, cells were extracted on ice for 1 h in Nonidet P40 lysis buffer (0.5 % Nonidet P40, 50 mM-NaF, 50 mM-*p*-nitrophenyl phosphate, 100 μ M-sodium orthovanadate, 1 mM-phenylmethanesulphonyl fluoride, 0.5 mM-leupeptin and 0.5 mM-pepstatin). Insoluble material was removed by centrifugation at 30000 g for 15 min and extracts were precleared three times with *Staphylococcus aureus* coated with 25 μ l of normal rabbit IgG. Portions (150 μ l) of extract were immunodepleted by sequential immunoprecipitations twice each with 50 μ l of *Staphylococcus aureus* (10 % suspension) coated with the following polyclonal or monoclonal antibodies: (1) anti-C3 serum, (2) anti-C2 serum, (3) anti-C1 serum and (4) monoclonal antibody (mAb)362.50. A fifth sample was immunoprecipitated with mAb362.50 only. Bound components were eluted in first-dimension urea sample buffer and analysed by using two-dimensional PAGE as described by O'Farrell (1975).

Immunofluorescence

Staining of acetone-fixed frozen sections by indirect immunofluorescence was performed as described previously (Mowery & Hixson, 1991).

RESULTS

cDNA coding for the S-form cell-CAM105

To search for cell-CAM105 isoforms, a rat liver cDNA library was screened with ^{32}P -labelled cell-CAM105 DNA fragment derived from the middle portion of the previously isolated full-length cDNA (*Bam*HI-*Pst*I fragment, 750 bp). Forty positive clones were obtained. Analysis by restriction digestion and nucleotide sequencing showed that one of the clones (clone 20) had an open reading frame of 457 amino acids with an in-frame termination codon at position -18, indicating that the methionine codon designated as position 1 was the first in-frame initiation codon. Sequence analysis revealed that this clone was a cell-CAM105 isoform (Fig. 1b; S-form). The major differences between the present form and the previously reported cell-CAM105 sequence (Lin & Guidotti, 1989) are localized in the C-terminal intracellular domain (Figs. 1a and 1b). The new form has a shorter cytoplasmic domain (10 amino acids), whereas the previously isolated clone has 71 amino acids intracellularly. As a

Fig. 1. (a) Schematic diagram and (b) alignment of deduced amino acid sequences of the L-form and S-form cell-CAM105

In (a) the nucleotides are numbered according to those of the previously isolated clone (L-form) (Lin & Guidotti, 1989). Untranslated regions are illustrated as filled boxes. The coding regions are illustrated as boxes. Sig, signal peptide sequence; TM, transmembrane region; Cyto, cytoplasmic segment. The proposed disulphide bonds are indicated and open circles in the S-form coding region indicate the positions where amino acid substitutions were found. The sequence of the S-form contains a polyadenylation signal (AATAAA) at the 3' end. In (b), predicted membrane-spanning regions are underlined.

result, the new isoform is the shorter version (S-form) of the previously isolated clone (L-form). In addition to this major difference, there were also substitutions clustered within the first 130 amino acids of the extracellular domain (Fig. 1b). Few nucleotide substitutions were found between amino acid 130 and the end of the sequence (Table 1), and these nucleotide substitutions did not result in a change in the encoded amino acids (Fig. 1b). This sequence information indicates that the S-form and the L-form cell-CAM105 differ in both the length of the proteins and the sequences of the N-terminal domains.

Expression of cell-CAM105 isoforms in liver plasma membranes

To study whether both isoforms were expressed in normal adult rat liver, anti-peptide antibodies were prepared against the three different C-terminal portions of the L-form cytoplasmic domain: amino acid positions 460–475 (anti-C1), 478–496 (anti-C2) and 497–519 (anti-C3). Since these sequences do not exist in

the S-form, these antibodies should only recognize the L-form. To test the specificity of these anti-C antibodies, we examined their reactivities with individual cell-CAM105 isoforms expressed in COS cells. Consistent with our prediction, all three anti-C terminal antibodies immunoprecipitated the L-form but showed no reactivity with the S-form expressed in COS cells transfected with cDNA coding for the L-form and the S-form respectively (result not shown). In immunoprecipitation assays, antibody no. 669 which was raised against purified cell-CAM105 (Lin *et al.*, 1991) reacted with both the L-form and S-form expressed in COS cells (results not shown). Monoclonal antibodies against cell-CAM105 (mAb362.50 and 5.4) recognized the same proteins as antibody no. 669 in immunodepletion experiments, and immunoprecipitated both the L-form and S-form of liver cell-CAM105 (Lin *et al.*, 1991). This latter observation coupled with the ability of both mAbs to label cell-CAM105 on the surface of intact cells (Mowery & Hixson, 1991) indicated that these mAbs recognized

Table 1. Nucleotide sequence difference between the L-form and S-form of cell-CAM105

The nucleotides are numbered according to those of the previously isolated ecto-ATPase (L-form) (Lin & Guidotti, 1989). The cDNA sequence available for the S-form begins at position -20. Slash indicates positions between two nucleotides where insertion was found. del indicates deletion; ins indicates insertion. Although the cDNA for the S-form was about 3.5 kb long, only the *N*-terminal 1.7 kb were sequenced.

Nucleotide no.	L-form	S-form
-20		Start
15	G	A
146-147	AG	GC
163	G	A
209	G	T
217-218	TT	AC
220-221	AA	GG
223-224	CC	TT
226	G	A
256-257	GA	AG
263	T	C
268	A	C
275	G	A
296	A	T
336	C	G
345	T	G
352	A	G
355	G	C
373	T	A
375	T	A
379	C	A
507	G	T
639	T	C
720	A	G
798	C	T
1068	T	C
1362-1414		(del)
1596/1597		C (ins)
1607/1608		CC (ins)
1608	G	C
1612/1613		G (ins)
1619	A	T
1624	G	C
1630/1631		C (ins)
1662-1663	CA	TG
1672	T	C
1680-1682	CAC	GGT
1684-1685		(del)
1688-1690	CAC	TGG
1692	C	A
1696	C	T
1698-1703	CACACA	TTTCTT
1705/1706		T (ins)
1709	A	C
1712		(del)
1714-1715	CA	TC
1717-1718	AC	TG
1720-1722	CAC	GTG
1724	C	A
1726	C	T
1728		(del)
1730	T	A
1733-1734	AT	TC
1736	C	G
1741	T	G
1743-1745	TTG	CGA
1747-1752	CTCCTT	AAAAAG
1752		end

epitopes in the extracellular domain that were shared by both isoforms.

Expression of the L-form and the S-form in liver plasma

Fig. 2. Expression of cell-CAM105 isoforms in liver plasma membranes

Hepatocytes radiolabelled with ^{125}I were solubilized, immunoprecipitated and analysed by two-dimensional gel electrophoresis as described in the Materials and methods section. (a) Total cell lysate was immunoprecipitated with mAb362.50, anti-C3, anti-C2 and anti-C1. (b) Total cell lysate was sequentially immunoprecipitated with antibodies anti-C3, anti-C2, anti-C1 and mAb362.50 as described in the Materials and methods section.

membranes was examined by sequential immunoprecipitation of the detergent extracts of radiolabelled hepatocytes with anti-C-terminal antibodies and mAb362.50. Although both antibody no. 669 and mAb362.50 could immunoprecipitate the L-form and the S-form of cell-CAM105, the immune complexes with mAb362.50 could be more quantitatively eluted in the urea sample buffer used for first-dimension isoelectric focusing gels in two-dimensional gel analysis. For that reason, mAb362.50 was used for immunoprecipitation experiments followed by two-dimensional gel analysis. As shown in Fig. 2(a), in two-dimensional PAGE, mAb362.50 immunoprecipitated two proteins with slight differences in their M_r values and isoelectric points (a protein of M_r 110000, pI 4.1, and a protein of M_r 105000, pI 4.3). In contrast, anti-C3 antibody only showed reactivity with the 110000- M_r protein, and subsequent depletion of anti-C3-depleted extracts with anti-C2 and anti-C1 antibodies did not remove any detectable amounts of labelled proteins (Fig. 2b). Since all three anti-C-terminal anti-peptide antibodies can immunoprecipitate the 110000- M_r protein (Fig. 2a), the ability of anti-C3 to deplete extracts of anti-C1 and anti-C2 reactivity indicates that there are undetectable amounts of isoforms which only react with anti-C1 or anti-C2. However, mAb362.50 could still immunoprecipitate the 105000- M_r protein from extracts that were previously depleted with all three anti-C-terminal anti-peptide antibodies (Fig. 2b). These results indicate that the liver cell-CAM105 has two major forms: a high- M_r form containing the entire C-terminal sequence, as predicted from the nucleotide sequence of the L-form, and a low- M_r form lacking sequences recognized by the three anti-C-terminal antibodies, a pattern of reactivity consistent with the information derived from the nucleotide sequence of the S-form. From these results, we conclude that both the S-form and the L-form cell-CAM105 are expressed in the liver plasma membrane. The L- and S-forms have identical numbers of tyrosine and histidine residues in their extracellular domains and would thus be expected to incorporate identical amounts of ^{125}I after surface labelling of intact cells. The observation that the total radioactivity of the 110000- M_r protein (L-form) immunoprecipitated by anti-C3 antibodies is about half that of the 105000- M_r protein (S-form) immunoprecipitated by mAb362.50 indicates that the S-form represents about 65% and the L-form about 35% of total cell-CAM105.

Fig. 3. Phosphorylation of cell-CAM105 isoforms *in vitro*

Hepatocytes radiolabelled with [^{32}P]phosphate were solubilized, immunoprecipitated and analysed by two-dimensional gel electrophoresis as described in the Materials and methods section. (a) Total cell lysate was immunoprecipitated with mAb362.50. Gel was exposed to film for 1 week. (b) Total cell lysate was sequentially immunoprecipitated with anti-C3, anti-C2, anti-C1 and mAb362.50. Gels from anti-C3 and anti-C2 immunoprecipitates were exposed to film for 1 week and gels from anti-C1 and mAb362.50 immunoprecipitates were exposed to film for 3 weeks.

Fig. 4. Localization of cell-CAM105 in liver plasma membrane using (a) anti-C3 antibodies and (b) antibody no. 669

Phosphorylation of cell-CAM105 isoforms *in vitro*

Comparison of the deduced amino acid sequences for the two isoforms indicated that the cytoplasmic domain of the L-form contained several potential phosphorylation sites which were not present in the S-form, raising the possibility that the two forms may be differentially phosphorylated. To confirm that both isoforms were phosphorylated *in vitro*, sequential immunoprecipitation with anti-C3, anti-C2, anti-C1 and mAb362.50 was carried out on the detergent extracts of primary cultured hepatocytes labelled with ^{32}P for 2 h. As shown in Fig. 3(a), radiolabelled proteins with mobilities characteristic of the L- and S-forms were immunoprecipitated by mAb362.50, with the L-form being the major phosphorylated isoform. In the cytoplasmic domain of the L-form cell-CAM105, a cyclic AMP-dependent serine phosphorylation (KRPTS) and a tyrosine kinase phosphorylation consensus sequence (KVDDVSY) have been identified (Lin *et al.*, 1991). There are also several serine and threonine residues in the cytoplasmic domain of both L- and S-form cell-CAM105. Since the relative ratio of the phosphorylated L- to S-form was about 3 to 1 compared with a 1 to 2 ratio for the iodinated proteins, these observations suggest that the L-form is phosphorylated at more sites than the S-form and/or a larger percentage of the L-form is phosphorylated. This major phosphorylated isoform was also immunoprecipitated selectively by anti-C3 antibody (Fig. 3b). No phosphorylated protein was detected in the second immunoprecipitate with anti-C3 antibody, indicating essentially complete depletion of reactive proteins (results not shown). Subsequent immunoprecipitation of anti-C3-depleted extract with anti-C2 and anti-C1 antibodies did not result in any detectable labelled protein (Fig. 3b). As expected from results with iodinated extracts, the S-form lacking the cytoplasmic domain recognized by the three anti-C-terminal antibodies was still present in extracts depleted twice with each of the three anti-C-terminal antibodies and was thus precipitated by mAb362.50. Unexpectedly, there was still a significant amount of L-form immunoprecipitated by mAb362.50 (Fig. 3b). When this experiment was repeated with a different extract, essentially identical results were obtained, suggesting that a small percentage of the ^{32}P -labelled L-form cell-CAM105 has a C-terminal domain that lacks reactivity with all three of the C-terminal anti-peptide antibodies.

It is possible that this lack of reactivity may stem from the fact that the C-terminal antibodies were generated against non-phosphorylated peptides and consequently may not show equal levels of reactivity with the singly or multiply phosphorylated variants of the L-form. Alternatively, there may be another isoform expressed at very low levels which is detectable by ^{32}P labelling but not by iodination. In any event, this result suggests that both L- and S-form cell-CAM105s are phosphorylated *in vitro*, with the L-form being the major phosphorylated protein.

Localization of the cell-CAM105 isoforms in the rat liver plasma membrane

Since hepatocytes are polarized cells, the existence of isoforms raises questions regarding localization of these isoforms in different domains of hepatocyte plasma membranes. We have previously shown by indirect immunofluorescence with antibody no. 669 that cell-CAM105 is localized in the canalicular domain of the liver plasma membranes (Lin, 1989). Since antibody no. 669 recognizes both the S-form and the L-form cell-CAM105 and the S-form cell-CAM105 is more abundant in the liver plasma membrane, the implication of previous localization results is that the S-form may be localized in the canalicular domain. However, the localization of the L-form, which has a lower level

of expression than the S-form, cannot be certain from this experiment. In order to elucidate the localization of the L-form in the liver plasma membrane, we have carried out further indirect immunofluorescence experiments using anti-C3 antibody. As shown in Fig. 4(a), anti-C3 antibodies which react specifically with the L-form gave a pattern of immunofluorescence essentially identical with that of antibody no. 669 (Fig. 4b), with most of the reactivity localized in the canalicular domain of hepatocyte plasma membranes. This result indicates that both the L-form and the S-form cell-CAM105 are localized in the canalicular domain.

DISCUSSION

This paper reports the complete sequence of a new rat liver cell-CAM105 isoform. The sequence of this isoform is very similar to the previously isolated clone, with 86% of the amino acids being identical. The major differences between these two cell-CAM105 isoforms include substitutions in the first 130 amino acids of the *N*-terminal domains and the different lengths of the *C*-terminal intracellular domains. We present evidence that in hepatocytes both isoforms are expressed at the protein level. The relative levels of expression are 35% and 65% for the L-form and the S-form respectively, as determined by SDS/PAGE of immunoprecipitates with antibodies against the intact cell-CAM105 and anti-peptide antibodies raised against the L-form-specific sequence. Evidence at the DNA and protein levels for the existence of isoforms was further supported by results obtained from RNAase protection experiments by using RNA probes generated from the *C*-terminal and *N*-terminal portions of the L-form cell-CAM105 sequence (results not shown). We also present data showing that despite the large differences in the size of their *C*-terminal domains, both L-form and S-form cell-CAM105s were phosphorylated *in vitro* to various extents, suggesting the possibility of phosphorylation as a regulatory mechanism.

One of the methods for detecting the relative expression of isoform proteins is to use antibodies specific for each isoform. Because of the difference in the length of the *C*-terminal domain of these two isoforms, we were able to generate anti-peptide antibodies specific for the L-form. However, since the two cell-CAM105 isoforms are very similar in the rest of their primary structure, it is difficult to select a region which is different enough to produce S-form-specific anti-peptide antibody. Selection of monoclonal antibodies against synthetic peptides containing sequences unique for the short isoform may be required. The specificity of the monoclonal antibodies can be confirmed by using proteins expressed in COS cells transfected with the L- or the S-form cell-CAM105 cDNAs.

The presence of cell-CAM105 isoforms raises questions concerning the function of these isoforms. In the case of N-CAM, there are at least five isoforms generated by alternative splicing of mRNAs encoded by a single N-CAM gene. The alternatively spliced variants produce N-CAM molecules with different cytoplasmic domains and modes of attachment to the cell membrane. Although little is known about the difference in biological functions of these various isoforms (Doherty *et al.*, 1989), there is evidence that expression of variably spliced transcripts could change in response to nerve growth factor (Prentice *et al.*, 1987) and developmental regulations (Small *et al.*, 1988). In both brain and muscle, individual N-CAM isoforms differ in their patterns of expression (Edelman, 1986; Walsh, 1988), suggesting that structural variations in N-CAM may be important for its function. Another group of CAMs which has also been intensively studied is the calcium-dependent CAMs, cadherins.

Cadherins exist in the plasma membrane in two forms, detergent-soluble and -insoluble forms. The detergent-insoluble form is associated with cytoskeletal components (Nagafuchi & Takeichi, 1988; Ozawa *et al.*, 1989), possibly through interaction involving the *C*-terminal domain. The observation that *C*-terminal mutants were unable to promote cell-cell adhesion strongly suggests that the *C*-terminal domain of cadherins is essential for normal functional activity (Nagafuchi & Takeichi, 1988; Ozawa *et al.*, 1989). Studies with a closely related member of the calcium-dependent CAMs, uvomorulin, suggest that this loss of function may result from dramatic changes in their anchoring in the plasma membrane and their association with cytoskeletal components as a result of truncation of the *C*-terminal domain (Ozawa *et al.*, 1989; McNeill *et al.*, 1990). Significant in this regard is the fact that the S-form of cell-CAM105 has less intracellular domain than the L-form and thus may differ significantly from the L-form in its association with cytoskeleton. It is also possible that these two isoforms are required in order to form a functional molecule, similar to the asialoglycoprotein receptor. In the latter molecule, two related polypeptides must interact early in biosynthesis in order to assemble a functional asialoglycoprotein receptor (Shia & Lodish, 1989). Further studies are needed before the functions of these isoforms are understood.

The presence of isoforms also raises the question concerning the genomic arrangement of gene(s) coding for these proteins. Although the different isoforms of N-CAM are generated by alternative splicing of the message derived from a single gene, our observations suggest that the cell-CAM105 isoforms probably arise from different genes rather than alternative splicing. First, sequence comparison of the two isoforms showed that there were minor differences scattered along the entire nucleotide sequence of the two isoforms. These substitutions were observed with another five clones containing partial sequences for the S-form, indicating that these substitutions were not due to cloning artifacts. This scattered substitution pattern argues against alternative splicing as the mechanism of isoform formation. Second, by screening a genomic library, we have obtained a genomic clone containing the *N*-terminal portion of the sequence. Of the two kilobase sequences obtained, two exons were identified. Exon 1 contains the 5' untranslated sequence and the *N*-terminal 63 nucleotides. Exon 2 contains the sequence from nucleotide 64 to 424. The overall sequence of these two exons are 96% and 91% identical with the L-form at the DNA and protein level respectively. The same genomic clone is 98% and 94% identical with the S-form at the DNA and protein level respectively. This indicates that this genomic clone codes for another cell-CAM105 isoform which is very similar to the liver L-form and S-form cell-CAM105. Therefore there are multiple genes for the cell-CAM105 family. However, this conclusion needs to be further confirmed by sequencing of the genomic clones.

In the present studies, the two phosphorylated forms of cell-CAM105 displayed mobilities on two-dimensional gels that were essentially identical with their ¹²⁵I-labelled counterparts and demonstrated the same pattern of reactivity with anti-peptide antibodies and mAb362.50. On the basis of these results, we conclude that the ³²P-labelled proteins were phosphorylated derivatives of the L- and S-forms and not two distinct isoforms which were not detected in iodinated extracts. Sequence analysis of the two isoforms showed that the cytoplasmic domain of the L-form had at least two potential kinase phosphorylation sites that were not present in the S-form, raising the possibility for differential regulation of isoforms by phosphorylation. This contention is supported by earlier studies of McEntire *et al.* (1989) showing that the phosphorylated cell-CAM105 appeared

as two distinct forms in two-dimensional gel analysis, and there were differences in the pattern of ^{32}P -labelled peptides produced by cleavage of high- and low- M_r cell-CAM105 with V8 proteinase. Results from the present studies taken together with those from previous investigations suggest that the additional sites present in the extended cytoplasmic domain of the L-form are more readily or stably phosphorylated than those found in the S-form.

Phosphorylation *in vivo* of other CAMs, i.e. L-CAM (Cunningham *et al.*, 1984) and N-CAM (Sorkin *et al.*, 1984), has also been reported. However, the functional significance of phosphorylation in these molecules was not known. Recently, it was shown that efficient transcytosis of polymeric immunoglobulin receptor requires phosphorylation of a serine residue located in the cytoplasmic domain (Casanova *et al.*, 1990). According to the model presented for hepatic membrane protein trafficking, proteins that are localized in the canalicular domain are presumably transported by transcytosis from sinusoidal membranes (Bartles *et al.*, 1987). Since both cell-CAM105 isoforms are localized in the liver canalicular domain, it is possible that phosphorylation of the newly synthesized cell-CAM105 is necessary for its transport to the canalicular domain by transcytosis. More study will be needed in order to elucidate the physiological significance of cell-CAM105 phosphorylation.

Since cell-CAM105 is an hepatocyte adhesion molecule, it was expected to be found on the lateral cell surface where adhesive interactions predominate (Boyer & Thiery, 1989). However, immunofluorescence analysis of frozen sections of rat liver using L-form-specific antibody and antibody that reacts with both isoforms (669) showed that both cell-CAM105 isoforms are localized in the canalicular domain of hepatocyte membranes. Electron-microscopic examination of hepatocytes stained by the indirect immunoperoxidase technique showed that cell-CAM105 was localized in both the bile canalicular membranes and the pericanalicular region of lateral membranes (Mowery & Hixson, 1991). On the basis of these observations, it was suggested that the fraction of cell-CAM105 localized in lateral membranes was involved in the formation of cell-cell contacts. Although the functional diversities of the two isoforms have not been established at present, it is possible that one isoform is mainly involved in the formation of intercellular contacts, and the other is mainly involved in mediating intramembrane adhesion for the formation of microvilli and membrane border folds or the maintenance of the canalicular lumen (Mowery & Hixson, 1991). By using immunoperoxidase labelling and electron microscopy, we have found that cell-CAM105 is also localized in the brush-border regions of the kidney (Sabolic *et al.*, 1992). This observation is consistent with the notion that this protein may be involved in intramembrane adhesion. Immunocytochemical localization at the ultrastructural level with L- and S-form-specific antibodies should be able to provide further information to prove or disprove this hypothesis. In any event, the findings that the two isoforms have a high degree of structural similarity

and are localized in close proximity favour the possibility that they may have similar functional roles.

We thank Professor Guido Guidotti for his suggestions and encouragement during the course of this work. Animals were maintained in facilities approved by the American Association for Accreditation of Laboratory Animal Care, and in accordance with current United States Department of Agriculture, Department of Health and Human Services, and National Institutes of Health regulations and standards. This work was supported in part by Core Grant CA16672 from the National Institutes of Health, by grant GM43189 (to S.H.L.) and by grant CA42715 (to D.H.) from the National Institutes of Health.

REFERENCES

- Aurivillius, M., Hansen, O. C., Lazrek, M. B. S., Bock, E. & Obrink, B. (1990) *FEBS Lett.* **264**, 267–269
- Bartles, J. R., Feracci, H. M., Stieger, B. & Hubbard, A. L. (1987) *J. Cell Biol.* **105**, 1241–1251
- Bonney, R. J., Becker, J. E., Walker, P. R. & Potter, V. R. (1974) *In Vitro* **9**, 399–413
- Boyer, B. & Thiery, J. P. (1989) *J. Membr. Biol.* **112**, 97–108
- Casanova, J. E., Breitfeld, P. P., Ross, S. A. & Mostov, K. E. (1990) *Science* **248**, 742–745
- Cunningham, B. A., Leutinger, Y., Gallin, W. J., Sorkin, B. C. & Edelman, G. M. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 5787–5791
- Doherty, P., Barton, H., Dickson, G., Seaton, P., Rowett, L. H., Moore, S. E., Gower, H. J. & Walsh, F. S. (1989) *J. Cell Biol.* **109**, 789–798
- Edelman, G. M. (1986) *Annu. Rev. Cell Biol.* **2**, 81–116
- Green, N., Alexander, H., Olson, A., Alexander, S., Shinnick, T. M., Sutcliffe, J. G. & Lerner, R. A. (1982) *Cell* **28**, 477–487
- Hixson, D. C. & McEntire, K. D. (1989) *Cancer Res.* **49**, 6788–6794
- Keski-Oja, J., Mosher, D. F. & Vaheri, A. (1977) *Biochem. Biophys. Res. Commun.* **74**, 699–706
- Lin, S. H. (1989) *J. Biol. Chem.* **264**, 14403–14407
- Lin, S. H. & Guidotti, G. (1989) *J. Biol. Chem.* **264**, 14408–14414
- Lin, S. H., Culic, O., Flanagan, D. & Hixson, D. C. (1991) *Biochem. J.* **278**, 155–161
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- McEntire, K., Mowery, J. & Hixson, D. E. (1989) *Cancer Res.* **49**, 6795–6802
- McNeill, H., Ozawa, M., Kemler, R. & Nelson, W. J. (1990) *Cell* **62**, 309–316
- Mowery, J. & Hixson, D. C. (1991) *Hepatology* **13**, 47–56
- Nagafuchi, A. & Takeichi, M. (1988) *EMBO J.* **7**, 3679–3684
- O'Farrell, P. H. (1975) *J. Biol. Chem.* **250**, 4007–4021
- Ozawa, M., Baribault, H. & Kemler, R. (1989) *EMBO J.* **8**, 1711–1717
- Prentice, H. M., Moore, S. E., Dickson, J. G., Doherty, P. & Walsh, F. S. (1987) *EMBO J.* **6**, 1859–1863
- Sabolic, I., Culic, O., Lin, S. H. & Brown, D. (1992) *Am. J. Physiol.* **262**, F217–F228
- Shia, M. A. & Lodish, H. F. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 1158–1162
- Small, S. J., Haines, S. L. & Akeson, R. A. (1988) *Neuron* **1**, 1007–1017
- Sorkin, B. C., Hoffman, S., Edelman, G. M. & Cunningham, B. A. (1984) *Science* **225**, 1476–1478
- Towbin, H. T., Staehelin, T. & Gordon, J. (1979) *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4350–4354
- Walsh, F. S. (1988) *Neurochem. Int.* **12**, 263–267