Molecular studies into the role of CD44 variants in metastasis in gastric cancer

H-F Hsieh, J-C Yu, L-I Ho, S-C Chiu, H-J Harn

Abstract

CD44, an integral membrane glycoprotein expressed by many cell types, serves as the principal transmembrane hyaluronate receptor and might be a determinant of metastatic and invasive behaviour in carcinomas. The generation of CD44 splice variants might be linked closely with gastric carcinoma tumorigenesis and differentiation. Some studies have reported that the magnitude of CD44 variant synthesis at the protein level correlates with lymph node metastasis. A number of studies have examined the possible mechanism of involvement of the CD44 variant in tumour metastasis. Most studies have reported that the regulation of CD44 binding to hyaluronate results from glycosylation of variably spliced exons. Direct hyaluronate binding studies of CD44 V4–V7 isoforms transfected into the human gastric carcinoma cell line, SGC-M1, have indicated that the V4–V7 isoforms themselves, in addition to glycosylation, can alter hyaluronate binding.


CD44 is a polymorphic integral membrane glycoprotein with a postulated role in lymph node homing, matrix adhesion, and T lymphocyte activation.

The traffic of lymphocytes and other leucocytes is important in maintaining the functions of the human immune system. The first step in lymphocyte recirculation is the entrance of lymphocytes into the peripheral lymph nodes and into mucosal lymph nodes through the adhesion interaction between the lymphocyte receptor and the high endothelium venule (HEV). There are a number of molecules involved in this interaction including CD44, which is recognised by Hermes antibodies, is a glycoprotein with a molecular weight of between 85 and 90 kDa. The major form of the CD44 molecule is an acidic, sulphated glycoprotein bearing both N-linked and O-linked oligosaccharide side chains. Other studies confirmed that the CD44 molecule isolated from the mucosal HEV binding B cell line (KCA) binds purified mucosal vascular addressin, a gp58–66 endothelial cell surface antigen required for the interaction of lymphocytes with mucosal HEVs. The putative protein structure of the CD44 molecule is divided into three parts: one is an extracellular domain, another is a transmembrane domain, and a third is the intracellular tail. The proximal extracellular domain, which includes the immunogenic sequence, is a potentially important structural region. The hydrophilicity, antigenicity, and probable O-glycosylation of this region shows that it is likely to be accessible to the external environment. More importantly, recognition of this domain by the Hermes-3 monoclonal antibody indicates that this portion is the receptor responsible for the binding of addressin. Interestingly, it is the proximal extracellular domain that is most variable in different cDNA clones.

The second function of the CD44 molecule is as a transmembrane link molecule between the extracellular matrix (ECM) and the cytoskeleton, a theory proposed by Carter and Wayner. Using collagen affinity chromatography, Carter and Wayner isolated a 90 kDa cell surface glycoprotein termed a class III collagen receptor (CRIII), which later proved to be the CD44 molecule. The biochemical characteristics of CRIII, its ability to interact with collagen, and its association with the detergent insoluble cytoskeleton suggests its functional role as a transmembrane link between the ECM and the cytoskeleton.

The third function of the CD44 molecule demonstrated by Huet et al is related to T cell activation. Using an anti-CD44 monoclonal antibody (H90), which is able to block the binding of lymphocytes to HEVs, Huet showed that H90 had no effect on 3H-TdR incorporation of whole peripheral blood lymphocytes stimulated by lectin allo geneic cells. In contrast, H90 strongly increased 3H-TdR incorporation of peripheral blood lymphocytes stimulated by anti-CD2 pairs of monoclonal antibodies or by anti-CD3 monoclonal antibodies linked to plastic culture plates.

Structure, function, and distribution of CD44, a novel class of cell adhesion molecules

Analysis of the cDNA sequence shows that the CD44 molecule, which is recognised by Hermes antibodies, is a glycoprotein with a molecular weight of between 85 and 90 kDa. The major form of the CD44 molecule is an acidic, sulphated glycoprotein bearing both N-linked and O-linked oligosaccharide side chains. Other studies confirmed that the CD44 molecule isolated from the mucosal HEV binding B cell line (KCA) binds purified mucosal vascular addressin, a gp58–66 endothelial cell surface antigen required for the interaction of lymphocytes with mucosal HEVs. The putative protein structure of the CD44 molecule is divided into three parts: one is an extracellular domain, another is a transmembrane domain, and a third is the intracellular tail. The proximal extracellular domain, which includes the immunogenic sequence, is a potentially important structural region. The hydrophilicity, antigenicity, and probable O-glycosylation of this region shows that it is likely to be accessible to the external environment. More importantly, recognition of this domain by the Hermes-3 monoclonal antibody indicates that this portion is the receptor responsible for the binding of addressin. Interestingly, it is the proximal extracellular domain that is most variable in different cDNA clones.

The CD44 cDNAs

Stamenkovic et al isolated a cDNA clone from libraries prepared from the histiocytic lymphoma cell line U373, the B lymphoblastoid line JY, the Burkitt’s lymphoma line, and the myeloid leukemia line KG-1, all of which were transfected separately into COS cells by the DEAE-dextran method. The nucleotide se-
quence of this CD44.5 cDNA clone comprises 1354 residues, with a short poly(A) site CATATAA. The ATG encoding the first methionine is present in a consensus initiation sequence and is followed by 19 hydrophobic amino acid residues resembling a secretory signal peptide. The signal peptide includes a cluster of eight neutral amino acids that could be a central hydrophobic region (Ala 7–Val 14), and five amino acid residues (Pro15–Ala 19) that could constitute a polar C-terminal region, containing Ser 17 and Ala 19, which fits the rule for signal peptidease cleavage specificity.

The N-terminal region of the putative signal peptide contains 21 residues. Cleavage of the signal peptides would generate a mature protein with 341 residues and a predicted molecular weight of 37.2 kDa. The extracellular domain consists of 248 residues followed by 21 predominately hydrophobic amino acids. The discrepancy between the predicted mass of the protein backbone and the deglycosylated forms observed in immunoprecipitates suggests that extensive O-linked glycosylation is present. The extracellular domain has six potential N-linked glycosylation sites (Asn-Xaa-Ser/Thr). In addition, this domain contains 19 serine and threonine residues, as well as numerous acidic residues and proline. These features are common in known O-glycosylation sites. The dipeptide SG that forms the minimal attachment site of serine linked chondroitin sulphate in proteoglycan proteins appears at residues 160, 170, 211, and 238 in the extracellular domain.

At the same time, the human B lymphoblastoid cell line that binds preferentially to mucosal HEV and expresses relatively high levels of gp90Hermes was used as a source of RNA by Goldstein et al to construct a lambda gt11 expression library. The size of the cDNA isolated from that source was 1.8 kb, which included the poly(A) tail and the putative signal peptides. The 5′ terminus of this clone did not include the first methionine codon, and instead started from the second amino acid. The rest of the predicted amino acids in the clone generated a 293 amino acid polypeptide.

Compared with the Stamenkovic et al clone sequence, the Goldstein et al clone lacked 67 amino acids in the 3′ terminus. This variation could be a result of differences in RNA splicing or a cloning artifact generated from pre-mRNA. A rare poly(A) site (ATTAAA) is found in the Goldstein clone, located just 12 bases upstream of the poly(A) tail. The deduced amino acid sequence predicts a mature protein with a C-terminal cytoplasmic tail, a hydrophobic transmembrane domain of 23 amino acids, and an N-terminal extracellular region of 248 amino acids. Except for the 67 amino acids not present in the Goldstein et al cDNA clone, the nucleotide sequences of the Stamenkovic et al and Goldstein et al cDNA clones are identical.

Organisation of the CD44 gene and its variant isoforms

Genomic analysis revealed the existence of 20 exons over a length of ~60 kb. The standard CD44 gene (CD44S) is composed of the following exons: exons 1s to 7s encode the extracellular part, exon 8s the transmembrane region, and exons 9s and 10s can be alternatively spliced to generate either a short (three amino acids) or a longer (70 amino acids) cytoplasmic tail, respectively. The 10 intervening, so-called variant exons 1v to 10v, positioned between exons 5v and 6v, contribute by multiple combinations to the remarkable variability of the CD44 family of molecules.

CD44 expression related to tumour metastasis

Because normal lymphocytes are distinct and traffic between the lymph and bloodstream systems throughout most of their life cycle, it is reasonable to assume that the metastasis of malignant lymphoid tumours may reflect expression of their homing receptor. Bargatze et al has compared the patterns of growth of several HEV binding and HEV non-binding murine lymphomas after passage into syngeneic recipients. Although all lymphomas have access to the bloodstream, results show a significant difference between the expression of functional HEV binding ability and the in vivo metastasis patterns. CD44 expressing lymphomas that metastasised through blood showed gross, symmetric enlargement of all lymph nodes. In contrast, lymphomas not expressing CD44 showed little or no apparent change in distant lymph nodes or in Peyer’s patches. Thus, the expression of functional homing receptors for HEV correlated with the ability of cells to migrate into the lymphoid node tissue through the HEV.

Further evidence for the involvement of CD44 molecules in tumour metastasis comes from hyaluronate receptor studies. Toole and colleagues showed that carcinomas were characterised frequently by local accumulation of hyaluronate in vivo.

A subsequent report by Toole et al indicated that the ability of a rabbit carcinoma to invade surrounding tissue correlated with its ability to induce formation of hyaluronate in situ. Furthermore, Nemec et al showed that a highly invasive bladder carcinoma had high levels of hyaluronate binding activity. The CD44 molecule is the principal cell surface receptor for hyaluronate and CD44 transcript expression is raised in carcinoma tissue.11

Although all of the functions of the hyaluronate receptor are not clear, the increased expression of CD44 molecules seen in tumour tissues and proliferating epithelial cells rich in hyaluronate matrix further suggest that the CD44 molecule is involved in tumour metastasis.17 20

Gunthert et al indicated that a new variant of the glycoprotein CD44 (CD44 meta) confers metastatic potential to rat carcinoma cells.11 Gunthert et al demonstrated the presence of new variants of the glycoprotein family CD44 in two series of metastasising rat tumour cells lines. Overexpression of one of these variants, encoded by cDNA clone pMeta-1, conferred...
metastatic behaviour to one of the non-metastasising cell lines, BSp73As. Although BSp73AS cells carry the normal pattern of CD44 and remain absolutely stationary, the transfectants express both CD44 and variant CD44 and are highly metastatic in syngeneic animals. The sequences of the isolated pMeta-1 CDNA clone revealed similarities with human CD44, with an additional extracellular domain of 162 amino acids. The unique part of the sequence is the same as in CD44 V5 and CD44 V6. The variant specific rat CD44 sequences were used to detect RNA expression in human cell lines including carcinoma lines from lung, breast, and keratinocytes. Various alternatively spliced products have been detected in other human tumour cell lines and tissues, including non-Hodgkin’s lymphoma, epithelial ovarian cancer, gastric adenocarcinoma, neuroblastoma, cervical cancer, colorectal carcinoma, pancreatic adenocarcinoma, hepatoma, and prostate carcinoma. In contrast to standard CD44 (CD44s), which is expressed almost ubiquitously, the variety of CD44 isoforms (CD44v) have a much more restricted distribution—for example, on keratinocytes (exons V3–V10), epithelial cells (exons V8–V10), and activated lymphocytes and macrophages (exon V6). Splice variants containing exon V6 seem to be involved in tumour metastasis as well as lymphocyte activation.

**Differential expression of the human metastasis adhesion molecule CD44v in normal and carcinomatous stomach mucosa**

In 1993, Hieder et al reported the expression of CD44 variants in gastric cancer. The analysis of RNA expression revealed dramatic differences between normal mucosa and adenocarcinomas. Whereas in normal epithelium only two of 12 CD44 variant RNAs containing exons V5 and/or V6 could be detected, intestinal-type tumours yielded a much more complex pattern of amplification products that hybridised to exons V5 and V6. In the sample of a diffuse-type tumour, expression of exon V5, but not V6, could be detected. Furthermore, using immunohistochemical staining and northern blotting, Harn et al showed that the expression of CD44 variants encoding exons V5 and V6 was upregulated in human gastric carcinomas. Most notably, CD44 V5 was expressed preferentially in poorly differentiated carcinomas and metastatic lymph nodes.

Using western blot analysis, Guo et al presented preliminary evidence that the concentration of soluble CD44 in the serum is raised in patients with advanced gastric or colon cancer. He also showed that the serum CD44 concentration correlated with tumour burden and metastasis of tumours. Surgical resection of tumours resulted in decreases in serum CD44 concentrations. Its results suggested the presence of different CD44 isoforms in the serum of patients with cancer. Serum CD44 concentrations may be an indicator of tumour burden and metastasis in patients with malignant disease. Using the enzyme linked immunosorbent assay, Harn et al showed that the concentrations of soluble CD44 V5 and V6 were raised in patients with advanced gastric carcinomas. He showed that serum CD44 V5 concentrations correlated with the extent of tumour invasion (T), the status of lymph node involvement (N), and distant metastasis (M) (TNM staging).

**Hyaluronate binding assay study of transfected CD44 V4–V7 isoforms into human gastric carcinoma cell line SC-M1**

Among the numerous ECM components, hyaluronate is of increasing interest regarding its role in cancer cell development and invasion. Because CD44 is known to be a hyaluronate receptor, Harn et al investigated whether there is a correlation between hyaluronate binding activity and CD44 V5 and V6 isoform expression. By using the Dynabead separation method and one of the well established gastric adenocarcinoma cell lines, SC-M1 was separated into V5 and V6 isoform positive and negative populations. The V5 and V6 isoform negative populations exhibited significantly higher hyaluronate binding activity than the corresponding positive cells (fig 1). The hyaluronate binding activity of V5 and V6 positive cells could be restored by pretreatment with anti-CD44 V5 and V6 monoclonal antibodies. Furthermore, transfection of an expression vector containing CD44 V5 and V6 into V5 and V6 negative cells decreased their hyaluronate binding activity to the same levels as CD44 V5 and V6 positive cells. Cells transfected with V5 and V6 recovered their hyaluronate binding activity after pretreatment with monoclonal antibodies against V5 and V6. These data suggest that cell adhesion involving hyaluronate can be regulated by multiple mechanisms, one of which involves alternative splicing of CD44 isoforms. To date, most reports have described that the regulation of CD44 binding to hyaluronate is by glycosylation of variably spliced exons. The results of Harn et al indicated that the V4–V7 isoforms themselves, in addition to the glycosylation, can alter the hyaluronate binding.

**Conclusion**

There are two important issues in CD44 variant expression related to tumour metastasis. One is how the CD44 variant is generated. Another issue is how CD44 isoform affects tumour metastasis. Regarding the first issue, there are two possible mechanisms that should be considered. One is cis element gene.
mutation, especially in the splice donor or acceptor sites. So far, there are no reports regarding gene mutation in malignant tumours that have resulted in the generation of CD44 variants. The second mechanism is that the tumour may make some regulator elements that affect alternative RNA splicing. Purification of regulator elements has been reported. 67 With regard to how the variant CD44 isoforms affects tumour metastasis, most studies have focused on glycosylation of variably spliced exons to regulate hyaluronate binding activity. An alternative explanation is that there is a new ligand present for the variant domain. However, no ligand has been identified. Recently, Harn et al suggested that tumour cell adhesion involving hyaluronate can be regulated by alternative splicing itself. Further investigations, especially in vivo, are needed before a final conclusion about how CD44 isoforms affect hyaluronate binding can be reached.

Our study was supported by grants from the National Defense of the Republic of China, Taiwan (DOD-88-39) and Armed Force Taoyuan Hospital, Research Foundation (1998–8801), Taiwan, Republic of China.


